

Automated Analysis of Cellular Signalling Parameters based on Images and Videos of Fluorescence Microscopy

A

Thesis submitted for the degree of DOCTOR OF PHILOSOPHY

In Computer Sciences to the School of Computing
of the University of Buckingham - United Kingdom

David Daouda TRAORE

November 2020

Abstract

Rapid changes in computer vision technologies have enabled automatic perspectives for more disciplines that tend to need heavy intervention from human experts. Computational biology, which is the application field of this thesis, is one of those disciplines where computer technologies (software and hardware systems) are applied on cell biology research, drug discovery, and disease diagnosis. The research conducted in this thesis is primarily concerned with automating the analysis of calcium imaging data obtained by two-dimensional (2D) fluorescence microscopy (FM) over living cells. The thesis also presents a theoretical and empirical analysis of the state-of-the-art object detection techniques used in Region based Convolutional Neural Network (R-CNN) and proposes a new R-CNN scheme tailored for cellular object detection in FM data. The analysed images are either individual greyscale images or image stacks of cardiac myocytes stained with DNA markers or calcium indicators. The cells are special cardiac cells found in lung veins, called Pulmonary Vein sleeve Cells (PVC) and Neo-natal Rat Ventricular Myocytes (NRVM) extracted from heart muscle. PVC stains enable the analysis of Calcium signals effect on heart physiology, whereas NRVM images allow autophagy process measuring through accurate cell counting.

In the thesis, we demonstrated that automated hotspot detection can be achieved with 79.75% of precision by a two-level segmentation procedure combining thresholding and statistical filtering of cellular regions. We also showed that photobleaching can be corrected by exponential curve fitting and baseline adjustment from normalised calcium traces with respectively a mean square error (MSE) varying between [0.09013, 6.41796] and an overall accuracy of 78.75% for a real-life dataset. Finally, based on the investigation carried over state-of-the-art object detection techniques developed in the past such as the Edge Boxes, the Selective Search, the Objectness Measure, and the Sliding Window paradigm, we demonstrated that a customised R-CNN framework based on a data-driven proposal box generation outperforms with less sampling rate (61 proposals/image) and more ground truth coverage ratio (GTCR of 99.40%).

Acknowledgements

I want to thank **Dr Hisham AL-ASSAM** and **Dr Naseer AL-JAWAD**, my internal supervisors from the School of Computing of the University of Buckingham, and **Dr Katja RIETDORF**, our external collaborator from the department of Life, Health, and Chemical Sciences of The Open University (Milton Keynes), since my knowledge on Computational Biology is the result of numerous interactions with them in a multidisciplinary environment over the course of four years.

To **my supervisors**, I own an eternal debt for their overwhelming support, guidance, patience, and believe in me more than probably deserved. May **The One** at the beginning of everything keep them blessed and healthy to face their daily obligations for ever.

Special thanks to **Professor JASSIM**, **Mr DU**, **Dr SELLAHEWA**, and **Professor LAMI** for their support and wise remarks during my research time in the school of Computing of the University of Buckingham.

To **my family** (Mom, brothers, sisters, wife, and kids), and to my friend, **Dr Nasiru IBRAHIM**, who I also considered as a brother, I would like to express my deepest gratitude for being there for me at all cost. To **my late Dad**, I hope he will be proud of his son, wherever he is. May his soul rest in peace !

Abbreviations

AP : Average Precision

APP: Application

AI : Artificial Intelligence

ANN: Artificial Neural Network

Calcium: ionized calcium molecule

CNN: Convolutional Neural Network

DNA: Deoxyribonucleic acid

FCN: Fully convolutional network

FM: Fluorescence Microscopy

FP: False Positive

FN: False Negative

GLCM: Grey Level Co-occurrence Matrix

GTCR: Ground Truth Coverage Ratio

GUI: Graphical User Interface

GPU: Graphical Processing Unit

HCM: Hypertrophic Cardiomyopathy

HOG: Histogram of Oriented Gradients

ICH: Intracellular calcium hotspots

IoU: Intersection Over Union

LBP: Local Binary Patterns

LoG: Laplacian of Gaussian

LSD: Lysosomal Storage Disorder

mAP: Mean Average Precision

MSE: Mean Squared Error

NRVM: Neo-natal Rat Ventricular Myocyte

PC: Phase Congruency

PVC: Pulmonary Vein Sleeve Cell

R-CNN: Regions with Convolutional Neural Network

R^2 : Coefficient of determination

ReLU: Rectifier Linear Unit

ROI: Region Of Interest

SVM: Support Vector Machines

TN: True Negative

TNR: True Negative Rate

TP: True Positive

TPR: True Positive Rate

VOC: Visual Object Class

2D: Two-Dimensional

Declaration of Originality

I hereby declare that all the work in my thesis entitled “Automated analysis of cellular signalling parameters based on images and videos of fluorescence microscopy” is my own work except where due references are made within the text of the thesis.

I also declare that, to the best of my knowledge, the material presented has never been submitted in the past for a degree in the University of Buckingham or any other universities.

David Daouda TRAORE - ©copyright 2020

Contents

List of Figures	xi
List of Tables	xxv
1 General Introduction	27
1.1 Biological Problems	28
1.2 Thesis Aim and Objectives	30
1.3 Frameworks for Calcium Signal Analysis and Nuclei Detection	31
1.4 List of Contributions and Publications	35
1.5 Thesis Outline	36
2 Background and Literature Review	38
2.1 Introduction	38
2.2 Computational Background	39
2.2.1 HOG Feature Representation and Description	39
2.2.2 LBP Feature Representation and Description	41
2.2.3 GLCM Feature Representation and Description	43
2.2.4 PC Feature Representation and Description	45

2.3	Review on Pre-processing Techniques	47
2.4	Review on Automated Analysis of Intracellular Calcium Signals	49
2.5	Review on Cell Segmentation Techniques Applicable to Cell Biology Images . . .	52
2.6	Summary on Literature Review	55
3	Automated Calcium Hotspot Detection and Measurement	57
3.1	Introduction	58
3.2	Calcium Hotspot Detection	62
3.2.1	Identifying Candidate Calcium Hotspots by Multi-level Thresholding . .	63
3.2.2	Filtering and Identifying the True Hotspots based on Hit-ratio Computation	68
3.3	Calcium Signal Measurements from Individual Hotspots	69
3.4	Materials & Experimental Protocols	71
3.4.1	Calcium Signal Ground Truth Datasets	72
3.4.2	Experimental Protocols	73
3.5	Results and Discussion	75
3.5.1	Calcium Signal Area or Event Segmentation in Video Frames	75
3.5.2	Ultimate Calcium Hotspot Detection in FM Videos	78
3.6	Summary on Calcium Hotspot Detection	81
4	Calcium Hotspot Signals Processing in Cardiac Myocytes	83
4.1	Introduction	84
4.2	Hotspot Calcium Signal Extraction	84
4.3	Photobleaching in Hotspot Calcium Signals	86

4.3.1	Introduction to Photobleaching	86
4.3.2	Photobleaching Detection in Hotspot Calcium Signals	88
4.3.3	Photobleaching Correction in Hotspot Calcium Signals	94
4.4	Materials & Experimental Protocols	95
4.5	Results and Discussion	96
4.6	Summary on Calcium Hotspot Signal Analysis	99
5	Investigating Standard R-CNN Variants for Nuclei Detection in Autophagy Cells	104
5.1	Introduction	105
5.2	Materials & Experimental Protocols	106
5.2.1	Ground Truth Dataset Materials	107
5.2.2	Experimental Protocols	108
5.3	Evaluating State-of-the-Art Algorithms for Region Proposal Generation in R-CNN Variants	112
5.3.1	Description of the 4 State-of-the-art Techniques	112
5.3.2	Experimental Results on Biomedical Datasets	116
5.3.3	Discussion of the Techniques Performances	123
5.3.4	Techniques Comparison	125
5.4	Summary	129
6	Customized R-CNN for Nuclei Detection based on Adaptive Region Proposal Generation	131
6.1	Introduction	132

6.2	Proposed Solution for Nuclei Detection based on R-CNN	133
6.2.1	Initial Nuclei Detection by Local Adaptive Thresholding	136
6.2.2	High-quality Nuclei Proposals Filtering by SVMs	142
6.2.3	Fine-grained Object Detection by CNN Classification	149
6.3	Materials & Experimental Protocols	150
6.3.1	Ground Truth Datasets	151
6.3.2	Experimental Protocols	151
6.3.3	SVM and CNN Training and Testing Protocols	153
6.4	Result and Discussion	155
6.4.1	Initial Nuclei Detection by Local Adaptive Thresholding	155
6.4.2	Filtering Initial Nuclei Proposals by SVM Models	159
6.4.3	Findings on Nuclei Bounding Boxes Classification by CNN	164
6.5	Summary on Nuclei Detection by Customized R-CNN	165
7	General Conclusion	179
7.1	Achievements	179
7.1.1	Calcium Signal Analysis for Understanding Arrhythmia	179
7.1.2	Nuclei Detection for Autophagy Evolution Analysis	181
7.2	Limitations	181
7.2.1	Automated Solution for Calcium Signal Analysis	181
7.2.2	Automated Solution for Cell Detection	182
7.3	Areas for Future Research	182

A Description of INCAS:Intracellular Calcium Analysis Software	196
A.1 Overview	196
A.2 Interfaces and Features Description	197
A.3 Appendix Conclusion	201

List of Figures

1.1	Importance of intracellular calcium transport during an action potential and heartbeat. AP: action potential. ATP: symbol of adenosine 5 triphosphate, principal molecule for energy transport and storage in cells. NCX: sodium calcium exchanger molecule responsible of removing calcium from cells. Na: symbol of sodium or salt nutriment. Ca: symbol of calcium. K: symbol of potassium. RyR: ryanodine receptors, major mediator for calcium-induce and release in animal cells. SR: sarcoplasmic reticulum, calcium storage in muscle cells. PLB: phospholamban protein.	29
1.2	Global framework for automated calcium signal analysis. (1): loading of calcium signal recordings into the system. (2): calcium signal segmentation by multi-level thresholding in frames. (3): calcium hotspot detection. (4): measure of calcium traces over time in hotpots. (5): photobleaching detection and correction in signals. (6): signal parameters measurements.	32
1.3	Global framework for automated nuclei detection. (1): Stage of initial proposal generation by thresholding. (2): stage of hand-crafted feature extraction. (3): stage of initial proposals filtering by SVM classifiers built upon HOG, GLCM, PC, and LBP features. (4): stage of ultimate nuclei detection through CNN classification of image patches	34
2.1	HOG feature extraction after gamma contrast enhancement. Sub-images A and B represent respectively an original NRVM nuclei image and its HOG feature representation. Nuclei samples being highlighted by yellow boxes and background contents by red boxes.	40

2.2	Uniform LBP image computation after gamma correction. Sub-images A and B represent respectively an original NRVM nuclei image and its LBP feature representation with nuclei patches highlighted by yellow boxes and background contents highlighted by red boxes	41
2.3	LBP texture vs HOG Shape information. Sub-image O shows an original NRVM nuclei image with a nuclei sample highlighted by a yellow box and a background sample highlighted by a white box. Sub-images A and B show respectively the HOG encoding of the nucleus and background patches highlighted in O. Sub-images C and D show respectively the LBP images and the derived histograms from the LBP images of both nuclei and background images computed in C. . .	42
2.4	GLCM features vs LBP features for nuclei texture encoding. Sub-image O shows an original NRVM nuclei image with a nuclei sample highlighted by a yellow box and a background sample highlighted by a white box. Sub-images A and B show respectively the LBP encoding of the of the nucleus and background patches highlighted in O. Sub-images C and D show respectively the GLCM images and four derived features extracted from the GLCM images of both nuclei and background patches computed in C.	44
2.5	PC features vs GLCM features for nuclei and background texture encoding . . .	46
2.6	Architecture of R-CNN (Region with Convolutional Neural Network).	54
3.1	Framework for automated calcium hotspots detection prior to signal measurements.	60
3.2	Calcium signal areas look like bright spots over general dark background. A: original input image sequence of PVC with twenty-two annotated regions from 1-22. First two labelled regions are outside the cell tissue and correspond to image background. Last 20 labelled regions are inside the cell tissue and correspond to calcium signal areas	64
3.3	Flowchart diagram of calcium signal segmentation by multi-level thresholding. .	66

3.4	System output at each stage of calcium segmentation by thresholding. A: original input video frame of PVC. B: reference image with ROIs highlighted in yellow. C: non-cellular regions estimated as background frame. D: background subtracted frame. E: contrast-enhanced frame. F: binary frame obtained after multi-level thresholding. G: connected components enhanced by morphological operations. H: multi-level thresholding results overlaid on original input video frame. ROIs have been coloured from blue to red according to x-coordinates.	67
3.5	Flowchart diagram of calcium hotspot detection by hit-ratio filtering of calcium signal areas. Here, objects refer to areas.	69
3.6	Calcium signal measurements output by our proposed algorithm.	71
3.7	Ground truth annotation for calcium signal segmentation. A represents original video frame showing PVC loaded with calcium indicators. B illustrates hand-free ellipsoidal labels of cellular regions showing calcium signals i.e., change of intracellular calcium concentration. C represents the binary mask of the ellipsoidal labels saved as the ground truth of calcium signals areas or events in A	73
3.8	Ground truth annotation for calcium hotspot detection. A represents an original FM video consisting of time-series images of cells loaded with calcium indicators. B shows the manual annotation made by calcium signal experts to highlight the 22 most active regions in the video. C represents the binary mask of the 22 hotspot locations across the entire video saved as the ground truth of the video in A.	74
3.9	Results of calcium signal segmentation in FM images. A represents an original PVC image with 27 ground truth calcium signal areas highlighted in yellow ellipsoids. B represents the segmentation result obtained by multilevel thresholding of A. B shows 30 connected components that indicate 30 candidate calcium signal locations. C is the comparison between A and B i.e., between the segmentation results highlighted in red and the ground truth labels highlighted in yellow . . .	76

3.10	Automated hotspot detection based on hit-ratio, size, and distance. A: reference image showing 30 ground truth hotspot locations annotated by experts for an input video of PVC. B: binary map of the ground truth hotspots in in A. C: 89 initial connected components output by the system as potential non overlapping hotspots after scanning 900 video frames of the input video and checking 23065 different calcium signals. Detection parameters based on size between [100, 300], hit-ratio between [1,100]%, and minimum distance = 1. D :30 final hotspots output by the system after applying a default minimum distance parameter of 2. The connected components are coloured according to their hit-ratio values. The colour map ranges from blue to red. The hit ratio values of the hotspots are displayed next to them. E: illustration of the overall precision of the hotspot detection method that is computed by comparing the results of the system against the ground truth hotspot location in white ellipsoids. F: Results of the automated hotspot detection method overlaid on a frame of the input image.	81
4.1	Flowchart diagram of calcium signal extraction.	84
4.2	Hotspot calcium signal plots.A and E: screenshot of original input video frame with three annotated hotspot locations. B and F: initial readouts of three hotspots from average raw pixel fluorescence over time. C and G: corrected fluorescence of the three hotspots. D and H: change of fluorescence measurements of the three hotspots	85
4.3	Hotspot calcium signal affected by photobleaching.Blue line: original calcium signal recorded for a hotspot through consecutive video frames for 13 seconds. Red line: photobleaching illustration through a baseline shift of the original calcium signal comparable to a negative exponential function. Amber line: theoretical flat baseline that should characterize the calcium signal in red if no photobleaching decay.	87
4.4	Photobleaching curves computation in Histological slices.	89

4.5	Photobleaching detection through baseline shift sign (x-axis representing time unit in second, and y-axis representing calcium change normalized units). Blue lines represent traces of hotspot calcium signals. Red lines correspond to their respective baseline shift estimation by mono-exponential fitting using non-linear least square regression technique. Amber lines represent smooth calcium signals obtained by window-averaging the original calcium signals. Curves are plotted calcium intensities over times in second. The intensities of the calcium signals have been normalized between 0 and 1 based on their maximum readouts. . . .	90
4.6	Flowchart diagram of photobleaching detection and correction in hotspot calcium signals. The correction step highlighted in red is described in Section 4.3.3 . . .	91
4.7	Photobleaching detection by mono (A, C) and bi-exponential (B, D) fitting. The blue lines represent original hotspot calcium signals measured over time in consecutive video frames. The amber lines represent smooth versions of the hotspot calcium signals by using window-averaging. The red lines represent the estimated photobleaching rate obtained by fitting exponential curves to the smooth calcium signals.	93
4.8	Photobleaching correction by baseline adjustment in 2 calcium signals by mono-exponential fitting. A: photobleaching detection and correction from a hotspot calcium signal recorded 3 minute before carvedilol treatment on cells. carvedilol is a protein that inhibits Calcium signalling in mitochondria. B: photobleaching detection and correction from a hotspot signal recorded after 10 minutes after dantrolene treatment of cells. dantrolene is a molecule used to treat some specific calcium channels disorders in cardiac cells. Blue lines: original hotspot calcium signals recorded over time. Yellow lines: estimated photobleaching decay derived from baseline shift of the original signals. Purple lines: theoretical flat baselines if the original signals were not affected by photobleaching. Amber lines: recovered calcium signals after photobleaching detection and correction of the original signals.	95
4.9	MSE report for mono and bi-exponential curve fitting of 80 hotspots. X-axis: index of 80 hotspot signals analysed. Y-axis: mean square error units.	97
4.10	R^2 report for mono and bi-exponential curve fitting of 80 hotspots. X-axis: index of 80 hotspot signals analysed. Y-axis: coefficient of determination units.	98

4.11	Exponential models comparison for photobleaching detection. Blue lines in A and B represent original calcium signals extracted from different hotspots of the same cells. Red lines in A, B, C and D correspond to smooth signals of original signals obtained by window-averaging. Blue lines in B and D correspond to mono-exponential curve fit of the smooth signals. Orange lines in B and D show bi-exponential curve fit of the smooth signals.	99
4.12	Photobleaching correction in hotspot calcium signal. Blue line: original hotspot calcium signal recorded over time. Red line: recovered signal after photobleaching detection and correction. Yellow line: estimated photobleaching decay derived from the baseline shift of the original calcium signal. Purple line: theoretical flat baseline of the correct signal if no photobleaching corruption	100
4.13	Average differences between expected and estimated baselines of corrected signals. X-axis: index of 80 hotspot signals analysed. Y-axis: differences recorded. Amber line: average differences between the baseline of the recovered signals and the theoretical flat baseline expected. Average differences below the blue line threshold means the system failure to correct the detected photobleaching effect	101
4.14	Example of a corrected calcium signal whose average difference falls below the threshold T. Blue line: original hotspot calcium signal recorded over time. Amber line: recovered signal after photobleaching detection and correction. Yellow line: estimated baseline shift of the recovered calcium signal. Purple line: theoretical expected flat baseline of the recovered signal. The average difference -0.002997 between the estimated baseline of the recovered signal and the flat expected baseline illustrates a non-successful photobleaching correction as it means the baseline of the recovered signal is still below expectation	102

4.15	<p>Example of corrected calcium signal whose average difference is above the threshold T. Blue line: original hotspot calcium signal recorded over time. Amber line: recovered signal after photobleaching detection and correction. Yellow line: estimated baseline shift of the recovered calcium signal. Purple line: theoretical expected flat baseline of the recovered signal. The average difference 0.000656 between the corrected baseline and the flat baseline illustrates a successful photobleaching correction as it means the baseline of the recovered signal is above the flat baseline.</p>	103
5.1	<p>Labelling nuclei and background patches of NRVM images. A: an original NRVM image with nuclei locations labelled by yellow bounding boxes and sample background patches labelled by red bounding boxes. B: Cropping of the labelled patches in A that illustrates the contrast, size, shape, and texture diversity in ROIs and background elements</p>	108
5.2	<p>Intersection over union metric illustration. Blue rectangles A and B illustrate boxes locations of 2 objects A and B. yellow rectangles represent their area of overlap. Both blue rectangles merged into one represent their area of union . . .</p>	109
5.3	<p>Performances of sliding window search for nuclei locations. A: overall precision of proposal boxes generated by sliding fixed-size windows over 250 NRVM images using various strides. Number of proposals generated are below each bar of the histogram. B: an original input NRVM image with 16 ground truth nuclei annotated under experts' guidance. C: 13820 proposal boxes generated by sliding fixed-size windows at step $[10 \times 10]$ over the input image. D: Comparison of the proposals with the ground truth enables performance measure in terms of TP, and FP boxes, and precision.</p>	118

5.4	Performances of random window search for nuclei locations. A: overall precision of 1000 fixed-size proposal boxes randomly generated by normal distribution over 250 NRVM images using various seeds. Number of proposals generated are below each bar of the histogram. B: an original input NRVM image with 16 ground truth nuclei annotated under experts' guidance. C: 1000 random proposal boxes generated by normal distribution at seed 1 over the input image. D: Comparison of the proposals with the ground truth enables performance measure in terms of TP, and FP boxes, and precision.	119
5.5	Performances of objectness search for nuclei locations. A: overall precision (blue line) and GTCR (amber line) of 1000 proposal boxes generated by Objectness measure tested over 220 NRVM images and trained over 30 NRVM images. B: an original input NRVM image with 16 ground truth nuclei annotated under experts' guidance. C: 1000 proposal boxes generated by Objectness Measure over the input image. D: Comparison of the proposals with the ground truth enables performance measure in terms of TP, FP boxes, GTCR, and precision	120
5.6	Precision of Selective Search for nuclei locations. A: overall precision of fixed-size proposal boxes generated by Selective Search over 250 NRVM images using various colours channels. Number of proposals generated are below each bar of the histogram. B: an original input NRVM image with 16 ground truth nuclei annotated under experts' guidance. C: 91 proposal boxes generated by Selective Search over the input image. D: Comparison of the proposals with the ground truth enables performance measure in terms of TP, and FP boxes, and precision.	121
5.7	Performances of standard edge box algorithm for nuclei proposal generation. Average precision = 15.39%.	122
5.8	Precision of customized edge box algorithm for nuclei proposal generation. GT stands for Ground Truth nuclei; BG for proposal boxes generated, BGI for average BG per image, TP for True Positive, FP for False Positive, P for precision, and GTCR for Ground Truth Coverage Ratio.	123

5.9	Faster R-CNN performance over nuclei images. A: Faster R-CNN ability at detecting weak boundary nuclei. Faster R-CNN ability at predicting the unique location of a single nucleus inside an image. C: Faster-CNN ability at ignoring misleading background with heterogeneous illumination and noise. D: predictions made by the system over 4-nuclei image. E: System failure at predicting the location of a nucleus with vague contour. F: System failure at ignoring misleading background patches.	126
5.10	Standard edge box algorithm tested on NRVM nuclei images. yellow boxes represent ground truth location of nuclei set manually by biologist experts. Red boxes represent true positive nuclei locations generated by the system	128
5.11	Objectness measure power and weakness at finding all nuclei locations. A: Original input NRVM image with reference nuclei locations highlighted by colourful rectangles. B: 1000 proposal locations computed by Objectness Measure to predict the true nuclei locations of A. C: Comparison of the proposals with the Ground returns 100% of GTCR versus 1.6% of precision.	129
6.1	Overview of customized R-CNN for nuclei detection based on adaptive proposal box generation. (A): initial proposal generation stage by local adaptive thresholding. (B): feature extraction and SVM classification stage. (C): ultimate nuclei detection stage by CNN classification.	134
6.2	Output of the different stages of R-CNN based nuclei detection framework. (A,E): original input nuclei images. (B,F): results of initial proposal box generation by local adaptive thresholding. (C,G): results of proposal box filtering by SVM models. (D,H): results of fine-grained nuclei detection by CNN.	135
6.3	Good-contrast images: nuclei regions have higher intensity profiles than background patches' ones. (A): original input nuclei image 2 annotations: a nuclei patch in red and an image background patch in blue. (B): intensity profile comparison of both nuclei and background image patches. (C) binary image of global Otsu thresholding. (D) Comparison of the Otsu-derived bounding box with ground truth	137

6.4	Low-contrast images: background contents are brighter than nuclei but handled by local adaptive thresholding. (A): original input nuclei image 2 annotations: a nuclei patch in red and an image background patch in blue. (B): intensity profile comparison of both nuclei and background image patches. (C) binary image of local adaptive thresholding. (D) Results of the local adaptive thresholding overlaid on original input image	138
6.5	Flowchart diagram for initial nuclei boxes locations from adaptive threshold. CC stands for connected component obtained after adaptive thresholding of NRVM images with kernel 161 x 161.	139
6.6	Initial nuclei proposal box generation by local adaptive thresholding. A: original grayscale input nuclei image displayed in pseudo colour. B: binary image after local adaptive thresholding. C: results of initial proposal box generation overlaid on original grayscale input image. D: results of initial proposal box generation overlaid on binary image after local adaptive thresholding.	141
6.7	SVM fusion model for filtering initial nuclei proposal boxes. L.A. stands for Local Adaptive.	143
6.8	Flowchart diagram for training SVM model based on HOG, GLCM, LBP, and PC combination.	144
6.9	HOG based SVM classifier fails to predict 2 fading cells locations. A: Original input NRVM image with 5 nuclei (red boxes) and five background patches (white boxes) labels. B: 3 TP and 5 TN predictions by the HOG based classifier C: 2 FN predictions by the HOG based system.	144
6.10	GLCM based SVM classifier fails to predict 1 cell with vague contour. A: Same original input NRVM image with 5 nuclei (red boxes) and five background patches (white boxes) labels as in 6.9. B: 4 TP and 5 TN predictions by the GLCM based classifier C: 1 FN prediction by the GLCM based system.	145
6.11	LBP based SVM classifier fails to predict 2 fading cells and a weak-boundary cell. A: Same original input NRVM image with 5 nuclei (red boxes) and five background patches (white boxes) labels as in 6.9. B: 2 TP and 4 TN predictions by the LBP based classifier C: 3 FN and 1FP predictions by the LBP based system.	146

6.12	PC based SVM classifier fails to predict three cells with vague contour and fading texture. A: Same original input NRVM image with 5 nuclei (red boxes) and five background patches (white boxes) labels as in 6.9. B: 2 TP and 5 TN predictions by the PC based classifier C: 3 FN predictions by the GLCM based system. . . .	146
6.13	SVM model based on fusion of 4 features fails to predict any nuclei locations in a challenging NRVM image showing weak-boundary and fading cells.	147
6.14	Fusion at score level of HOG, LBP, GLCM and PC SVM models with $T = -0.275$. A: output of TP and FP by the system. B: output of FP by the system. White numbers represent signed average scores output by the fusion model for patches predicted as nuclei. Black numbers represent signed average scores output by the fusion model for patches predicted as background.	148
6.15	Performance of initial nuclei proposal box generation from adaptive thresholding mask. Red line: average ground truth coverage ratio per image. Blue line: average precision of proposal boxes per image. X-axis: NRVM image index or our dataset. Y-axis: values in percentage	156
6.16	Limitations of initial nuclei proposal generation in handling cell overlaps. 3 cellular regions showing overlapping candidates are labelled with red stars in B. A: original reference image with 34 ground truth nuclei labels. B: binary mask after applying local adaptive thresholding. C: results of the initial proposal box generation. D: comparison of the initial proposals with the reference image. . . .	167
6.17	17 Finding locations of overlapping or touching cells from regional maxima centres. A, E, I: 3 original grayscale nuclei patches with overlapping or touching elements. B, F, J: foreground element marking by morphological opening-reconstruction. Regional maxima segmentation mask. Separation edges set between touching or overlapping cells.	168
6.18	Average performances of SVM models based on fusion at feature level for nuclei detection.	168

-
- 6.19 GLCM SVM classifiers outperform in presence of vague contour cells but underperform in heterogeneous image backgrounds. A: reference image with 8 ground truth nuclei locations. B: reference image with 8 ground truth sample background patches. C: 7 TP nuclei locations predicted by the system built on GLCM features. D: 8 TN background patches predicted by the system. E: no FP predicted by the system. F: only 1 FN predicted by the system. Decimal numbers represent the level of confidence of the classifier. 169
- 6.20 PC SVM classifiers outperform in presence of heterogeneous background. A: reference image with 11 ground truth nuclei locations. B: reference image with 11 ground truth sample background patches. C: 8 TP nuclei locations predicted by the system built on PC features. D: 10 TN background patches predicted by the system. E: only 1 FP predicted by the system. F: 3 FN predicted by the system. Decimal numbers represent the level of confidence of the classifier. . . . 170
- 6.21 HOG SVM may outperform LBP SVM in well-shaped nuclei images but underperform in presence of fading texture. A: reference image with 14 ground truth nuclei locations. B: reference image with 14 ground truth sample background patches. C: 14 TP nuclei locations predicted by the system built on HOG features. D: 11 TN background patches predicted by the system built on HOG. E: 3 FP predicted by the HOG based system. F: 5 FN predicted by the LBP based system. Decimal numbers represent the level of confidence of the classifiers . . . 171
- 6.22 LBP SVM may outperform HOG SVM in intense-texture nuclei images. A: All TP nuclei detected by LBP based model. B: 9/11 TN predicted by LBP based model. C: 10 TP/11 nuclei locations predicted by the HOG based system. D: 7 TN/11 background patches predicted by the HOG based system. E: 2 FP nuclei predicted by the LBP based system. F: only 1 FN predicted by the HOG based system. Decimal numbers represent the level of confidence of the classifiers. . . . 172

6.23	Optimum score fusion Threshold: trade-off between sensitivity and specificity. A: Table of sensitivity and accuracy performances recorded for filtering initial proposals by SVM fusion at score level with different thresholds applied over half of the testing dataset. B: plot of sensitivity (blue line) and specificity (orange line) performances over fusion threshold values. Optimum threshold is the point of junction of both curves i.e., the cut-off dashed line with $x = -0.275$	173
6.24	Average performances recorded after 10 Cross-validation of SVM model based on Fusion at score level for image patch classification without segmentation. . .	173
6.25	Average performances recorded after 10 Cross validation of SVM model based on Fusion at score level for image patch classification with segmentation.	174
6.26	Initial threshold nuclei location filtering by SVMs models using fusion at score level. A: original input NRVM image with 15 ground truth labelled nuclei. B: 56 initial locations generated by local adaptive thresholding. C: 18 high-quality boxes filtered by SVMs using score fusion. Confidence scores plotted in red numbers. D: 9 TN boxes filtered out by SVMs using score fusion. Confidences scores plotted in black numbers. E: 29 FP boxes output by SVMs. Confidence scores plotted in white numbers. F: SVMs model ability at filtering all background samples from the initial proposals.	175
6.27	Inception V3 based R-CNN detector is highly confident at detecting challenging nuclei and background. The label [Nc/Nc, 100%] means the classifier has predicted a nuclei class at a location labelled as nuclei patch with 100% confidence. In the same way, the label [Bg/Bg, 98.6%] means the classifier has predicted a background class at a location labelled as background patch with 98.6% of confidence.	176

6.28	CNN classifier outperforms hand-crafted feature based SVM models for nuclei patch classification. White boxes in A: represent initial known proposal boxes fed to SVM fusion model and CNN for classification. Red boxes in B: and F: represent TP predictions made by both classifiers i.e., nuclei boxes correctly classified as nuclei. Blue boxes in C: and G: represent FN predictions made by the classifiers i.e., nuclei boxes wrongly classified as background. Yellow boxes in D: and H: represent TN predictions made by the classifiers i.e., background boxes correctly classified as background. Magenta boxes in E: and I: represent FP predictions made by the classifiers, i.e., background patches wrongly classified as nuclei	177
6.29	Performance of nuclei patch classification by CNN Inception V3 without segmentation.	178
6.30	Performance of nuclei patch classification by CNN Inception V3 with segmentation.	178
A.1	INCAS main page - screenshot	198
A.2	Calcium signal processing - Screenshot	200
A.3	Statistical measurements - Screenshot	201

List of Tables

2.1	Relevant techniques for automated calcium signal analysis in FM datasets. CLS: Confocal line scan images. TIRF: Total Internal Reflection Fluorescence microscopy.	50
3.1	Average result for calcium signal detection per video. Column video lists indexes of the 20 different videos used to build the ground truth datasets and to test our proposed solution for calcium signal segmentation by multi-level thresholding. Column AP lists the average precision recorded per set of 50 video frames analysed.	77
3.2	Results for calcium signal detection using Mean + 2std at the frame level. Column video lists indexes of the 20 different videos used to build the ground truth datasets and to test the Cheng’s strategy for calcium signal segmentation. Column AP lists the average precision recorded per set of 50 video frames analysed.	78
3.3	Result for hotspots detection per video. Column video frames shows the number of video frames included in each video processed. AP stands for average precision of hotspots detected by the system for the video described in the left-hand side column. Column running time shows the time elapsed by the system to process the video in second.	79
5.1	Performances of nuclei proposal box generation by sliding window.	117

5.2 Comparison of different techniques for nuclei proposal generation in biomedical images. Column Techniques: proposal generation techniques. Column Proposals: average number of proposals generated per testing image. Column Window size: size proposal boxes generated. Column Precision: precision of the boxes generated. Column GTCR: GTCR of the boxes generated. Column Datasets: pair of numbers telling the number of images used respectively for training and testing. For instance (0,250) means no training occurs as the technique is not trainable, and the system has been tested over the entire dataset made of 250 labelled images. 127

Chapter 1

General Introduction

The discovery and development of Green Fluorescence Proteins (GFP) in 2008 [78] enabled biologists to observe and understand the inner workings of cells, thousands of crucial events that were for a long time hidden such as the development of brain cells, the spreading of cancer cells, autophagy process, excitation-contraction coupling of cardiac cells due to calcium signals, etc [14]. These advances in imaging living cells and their components have been applied to countless biological problems and led to a tremendous amount of imaging data produced by scientists that need automated processing. In this thesis, we focus on investigating image processing algorithms that can be applied to fluorescence microscopy (FM) images for intracellular calcium signal analysis and cell counting.

This thesis investigates the strengths and limitations associated with low & mid-level image processing techniques, machine learning strategies and deep learning solutions for the process and analysis of cell biology images obtained by FM. The low and mid-level techniques refer to traditional image processing techniques, including spatial domain filters, morphological operations, and intensity thresholding. The machine learning strategies in this thesis refer to image classification algorithms based on hand-crafted feature vectors such as Support Vectors Machines (SVM) and the use of deep learning techniques such as Convolutional Neural Networks (CNN) for object detection. To assess the multiple algorithms and models build and investigated during this research, we used imaging datasets that describe intracellular signalling parameters and cellular components of cardiac myocytes, such as calcium signals and nuclei. In this first section, we give some background information about intracellular calcium signalling, and nuclear detection in terms of biological problems associated with the datasets we dealt

with. Then we state the thesis aim and objectives before presenting a general framework for automated intracellular calcium analysis and nuclei detection in cells. The final section of this chapter enumerates our list of contributions and publications.

1.1 Biological Problems

The datasets used in this project were collected on cardiac myocytes: PVCs and NRVMs. However, the processes investigated are of much wider biological importance. The first biological relevance of our research study refers to intracellular calcium signalling. Changes in the intracellular calcium concentration are essential for a plethora of cellular signalling events [8], ranging from the fertilisation of an egg [23], transferring nerve impulses [14], gene transcription [31], activation of muscle contractions to causing memory formation and cell death [7].

Heart contraction is one of the most critical functions of cardiac myocytes and the heartbeats because of a synchronised combination of multiple processes involving the rise of an electrical signal called action potential. Action potential starts from a pacemaker region of the heart called the sino-atrial node and expanded over the different chambers of the heart. The action potentials cause a depolarisation of the cardiac myocytes, which in turn leads to an increase in the intracellular calcium concentration that causes the cardiac myocyte to contract. This process is called excitation-contraction coupling. A well-coordinated spreading of the action potentials over the heart causes the sequential contraction of the atrial and ventricular chambers, pumping blood through the body and lungs, and is important for the rhythmical beating of the heart throughout the lifetime [10]. Figure 1.1, adapted from [10] illustrates how calcium particles are important for the pumping function of the heart by contributing to the propagation of the electrical signal of heart muscle cells through a complex mechanism during an action potential. To know more about the complex biological process that involves sodium (Na), potassium (K), and hydrogen (H) ions, and multiple calcium pumps and channels, please read [10] for further details.

During this complex cardiac excitation-contraction coupling mechanism illustrated by Figure 1.1, calcium ions are specially important because they cause the contraction of cardiac myocytes through concentration variations. However, if the calcium concentration changes in an irregular manner, for example, in-between heartbeats, this can intervene in the propagation of AP and may inhibit the electrical pacing of cardiac myocytes through local calcium transients [31] [9]

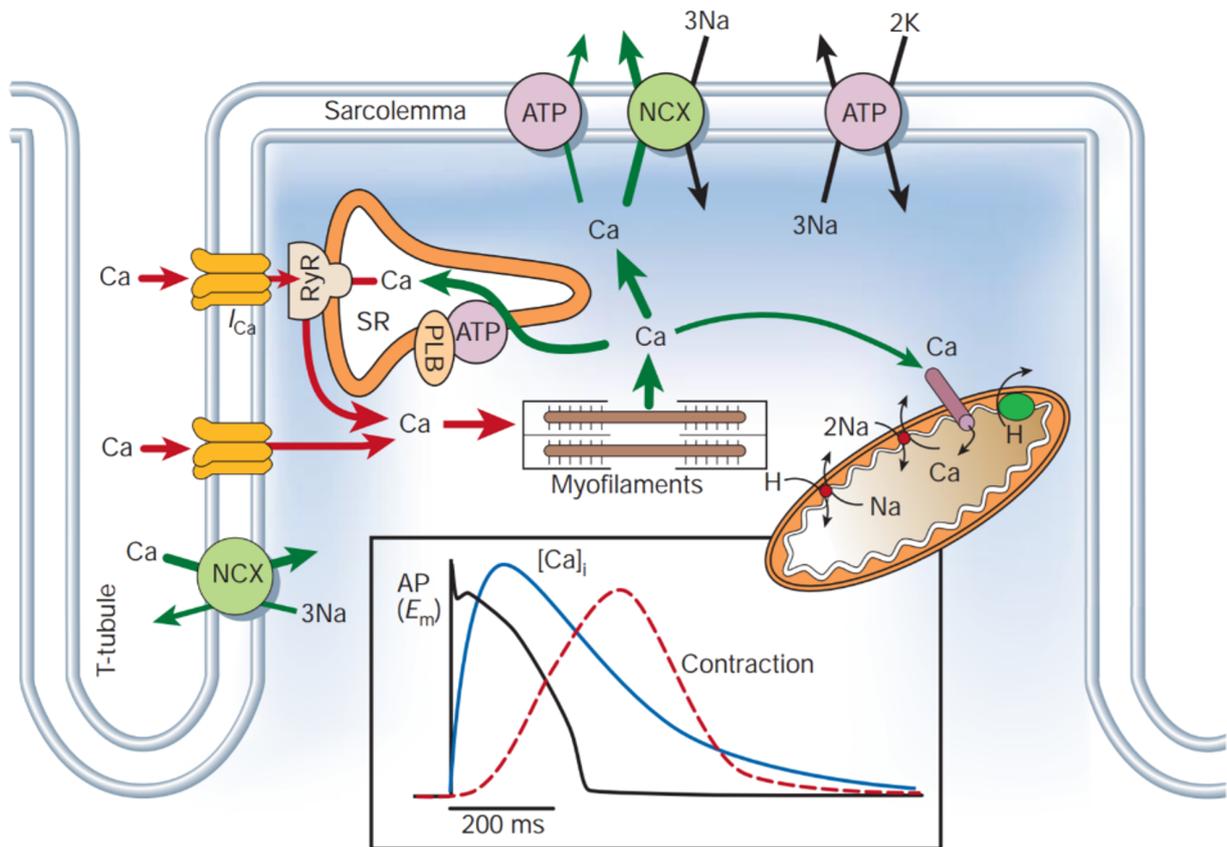


Figure 1.1: Importance of intracellular calcium transport during an action potential and heart-beat. AP: action potential. ATP: symbol of adenosine 5 triphosphate, principal molecule for energy transport and storage in cells. NCX: sodium calcium exchanger molecule responsible of removing calcium from cells. Na: symbol of sodium or salt nutriment. Ca: symbol of calcium. K: symbol of potassium. RyR: ryanodine receptors, major mediator for calcium-induce and release in animal cells. SR: sarcoplasmic reticulum, calcium storage in muscle cells. PLB: phospholamban protein.

[69] [12]. PVCs, the first group of cells whose we have dealt with their video recordings in this thesis, are one of the cell types which is known to show these spontaneous rises in the intracellular calcium concentration and that can disrupt the propagation of action potential, and thus lead to heartbeat anomalies such as atrial fibrillation, a heartbeat disturbance due to erratic signalling of atrial cells [69]. Being able to study the characteristics of calcium signalling of PVCs in detail and to identify the most active cells in calcium signalling, i.e. the hotspots is essential to understand how they contribute to the development of atrial fibrillation.

Another biological relevance of our research study is linked to cells counting in NRVM images. In the images of NRVM preparations we dealt with in this thesis, biologists are interested in

observing the autophagy process. But as autophagy, cell counting provides the foundation for multiple biomedical essays [56]. It is a fundamental task in computational biology because the automated analysis of many cellular processes requires the knowledge of how many cells are being observed. Nuclei are used to detect and count cells because every cell contains a nucleus, in which the genetic information is stored in the form of DNA. It is possible to use fluorescent markers to stain the DNA, and thus highlight one organelle in the cell. One common marker for nuclei is Hoechst, which is excited with UV light and used to produce the NRVM images analysed in this study. The number of cells can then be established by counting the number of nuclei. If researchers then perform other measurements, like measuring the ability of cells to digest and recycle macromolecules by autophagy [19] like in the NRVM cells used in this study, the results can be normalised to the level observed per cell. The second stain in a different colour can then be applied to visualise various cellular processes or stain organelles inside cells. Knowing how many cells were analysed allows normalising the signal from this second stain to the total number of cells. In addition, knowing the number of cells can also show for example if a treatment applied to the cells caused their death (e.g. X number of cells in the control treatment but only Y number of cells after treating the cells). Common diseases investigated through NRVM stains and associated to autophagy deregulation in the UK are known as lysosomal storage disorders (LSD) and hypertrophic cardiomyopathy (HCM), a cardiac disorder [58].

Based on that background information and biological relevance associated to the imaging datasets we dealt with, our research journey appears as a multidisciplinary exercise straddling digital image processing and cell biology in the sense that this thesis offers Maths, statistical and computer algorithmic solutions to understand cell-biological problems. In the next sections, we introduce our thesis aim and objectives.

1.2 Thesis Aim and Objectives

The main aim of this thesis is to develop automated routines for cellular parameters analysis through FM imaging data that provide biologist with reliable, replicable and accurate measurements for calcium imaging experiments to understand the calcium signalling in cardiac myocytes and for counting cellular nuclei. The objectives of this thesis can be summarized into five key points:

-
- a) Proposing an automated solution for calcium active cellular regions detection or hotspots detection and calcium measurement in fluorescence microscopy recordings of cardiac myocytes stained with calcium indicator.
 - b) Modelling the photobleaching decay in calcium hotspots signals as a mono-exponential function of time by non-linear least square regression.
 - c) Correcting photobleached signals of calcium hotspots by baseline adjustment.
 - d) Carrying out a theoretical and empirical analysis of R-CNN variants applicable to nuclei detection, which aims to assess the limitations of the state-of-the-art object detection techniques.
 - e) Providing practical and effective nuclei detection technique based on a customized R-CNN architecture using an adaptive generation of region proposals.

1.3 Frameworks for Calcium Signal Analysis and Nuclei Detection

Automated routines for intracellular calcium signalling analysis and for autophagy evolution via nuclei counting is very demanding among cell biologist community because the breakthroughs in imaging protocols have enabled the production of a vast quantity of imaging data to be processed in labs. Still, data analysis are often done manually or semi-automatically through multiple platforms such as ImageJ, MATLAB, Calquo, Cellprofiler, Prism, etc. To optimize the output of some of those platforms, user experiences and parameters finetuning are crucial. In this section, we list the main steps of our automated solutions to firstly analyse calcium signals through videos of fluorescent images of heart cells interiors, and secondly identify and count the number of cell nuclei present in static images of heart cells.

Figure 1.21.2 illustrates the global framework we developed to automatically analyse calcium signals in living cardiac myocytes. The process consists of 6 main stages numbered from (1) to (6).

- In stage (1) untitled “load calcium signal video”, calcium signal imaging videos are loaded into the system. These videos describe temporal screenshots of calcium release activities

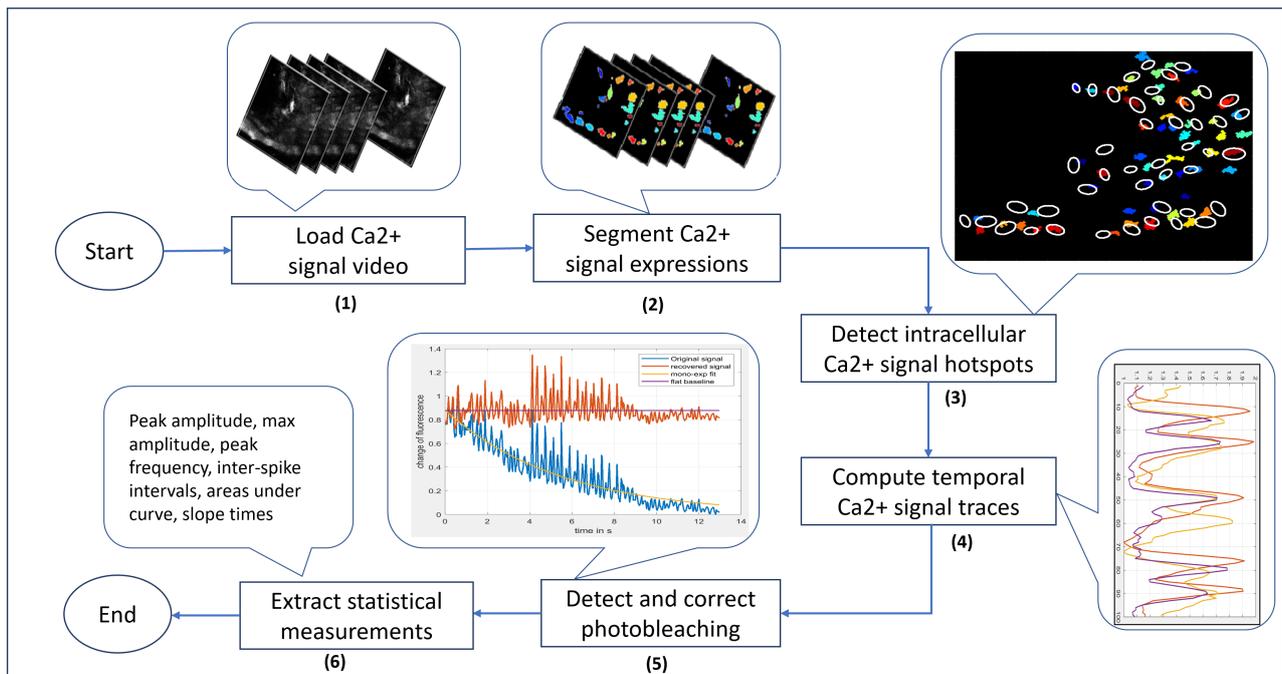


Figure 1.2: Global framework for automated calcium signal analysis. (1): loading of calcium signal recordings into the system. (2): calcium signal segmentation by multi-level thresholding in frames. (3): calcium hotspot detection. (4): measure of calcium traces over time in hotspots. (5): photobleaching detection and correction in signals. (6): signal parameters measurements.

inside living cardiac myocytes stained with calcium indicators. Therefore, to assess elementary calcium signals events at every frame time, input videos should be broken down into consecutive frames during the loading process.

- In stage (2) untitled “Segment calcium signal expressions”, the segmentation of calcium signal events is performed through individual frames. The segmentation of calcium signal events inside input frames is crucial as it tells whether cells are being active or not. This segmentation requires pre-processing of video frames prior to intensity thresholding derived from the global fluorescence of the whole video file and local fluctuations in individual frames. The pre-processing aims to reduce the Gaussian noise that may corrupt the fluorescence imaging data, and to increase the contrast of the images.
- In stage (3) untitled “Detect calcium signal hotspots”, the detection of the most active regions in calcium signal events of cells is performed. This is done by video fusion-based essentially on a hit-ratio parameter. The hit-ratio parameter is computed for all cellular regions containing calcium events and tells how often intracellular calcium concentration varies inside the given regions. The best hotspots correspond to non-overlapping areas

containing the greatest hit-ratio.

- In stage (4) untitled “Compute temporal calcium signal traces”, the average pixel intensity of the corresponding cellular areas is taken as the temporal measure of the intracellular calcium change of concentration values. The records are max normalized based on the minimum fluorescence inside each hotspot. This enabled a realistic comparison between cells observed in different biological preparations and camera settings.
- In stage (5) untitled “Detect and correct photobleaching”, our system tries to detect and correct any potential photobleaching phenomenon that might have altered the temporal measures of the intracellular change of concentration previously recorded in stage (4). Photobleaching phenomenon can be defined as an exponential decay of illumination that can affect cardiac cell preparations analysed by FM technique [90]. When calcium signals are affected by photobleaching phenomenon, their quantitative measurements get reduced and thus need to be recovered. This stage of our proposed solution for calcium signal analysis is essential as it provides data integrity for further accurate measurements of temporal calcium traces recorded from hotspots.
- In stage (6) untitled “Extract statistical measurements”, the system computes from photobleaching-free calcium signals a set of statistical measures such as peak amplitudes, maximum amplitudes, peak frequencies, inter-spike intervals, time of slopes and area under the curves. Those measurements enable formal comparison of the calcium signalling activity of various cells and active regions. In fine, they help biologists to draw conclusions related to the underlying biological problems under investigation such as the heart physiology induced by calcium signalling.

Following Figure 1.31.3 illustrates the global framework we developed for automated detection of nuclei in FM static images based on deep neural networks. The framework consists of four main steps labelled from 1 to 4.

- In step 1, as its title suggests, initial proposals boxes for potential nuclei locations are generated using local adaptive thresholding. In computer vision, object detection, and recognition tasks applied to biomedical images have been subject to numerous research papers and remain an ongoing problematic [91]. Some recent literature suggests searching for instances of objects inside images in a selective to boost performance and reduce errors

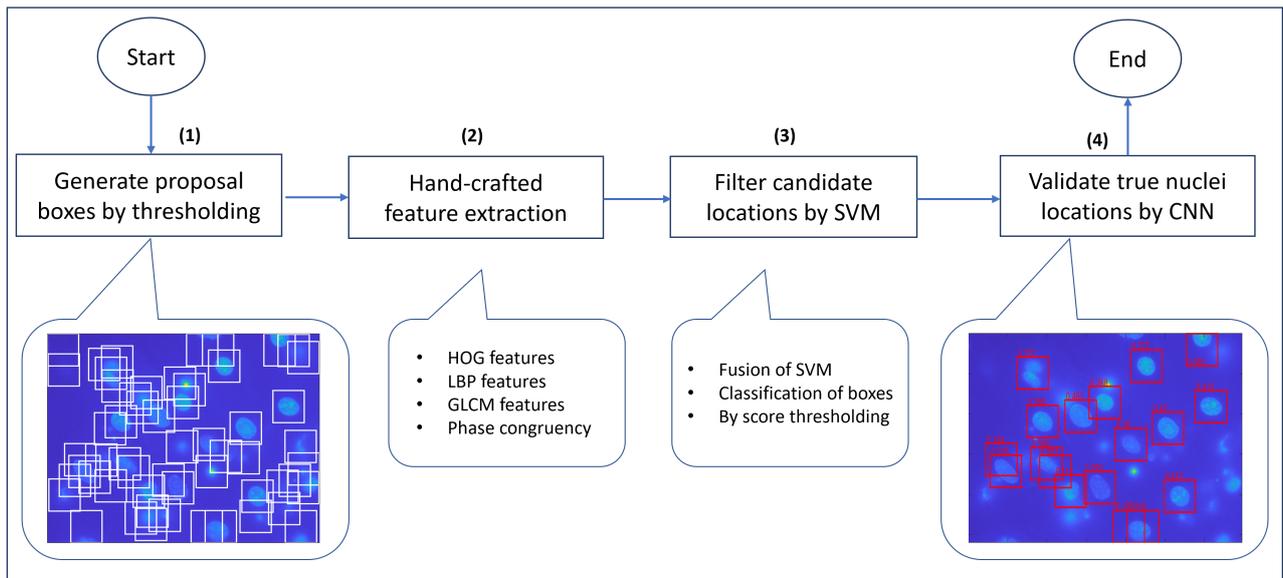


Figure 1.3: Global framework for automated nuclei detection. (1): Stage of initial proposal generation by thresholding. (2): stage of hand-crafted feature extraction. (3): stage of initial proposals filtering by SVM classifiers built upon HOG, GLCM, PC, and LBP features.(4): stage of ultimate nuclei detection through CNN classification of image patches

[87], while other authors promote an exhaustive search. We follow up the former approach by considering as potential nuclei locations, any meaningful, and connected components output by the local adaptive threshold of nuclei images.

- In step 2, untitled “hand-crafted feature extraction” the system extracts from image patches corresponding to initial candidate locations of nuclei texture-based features for further filtering. Indeed, the initial proposal boxes generated by local adaptive thresholding correspond to image patches that are more likely to contain nuclei instances. However, segmentation by intensity thresholding means that all foreground elements should be distinguished from background thanks to their pixel intensity distribution [65], which is not always the case. The hand-crafted feature extraction step included in the automated routine enables to consider extra image features for describing foreground elements and background counterparts. These features are histogram of oriented gradients (HOG), local binary patterns (LBP), grey-level co-occurrence matrices (GLCM) and phase congruency features (PC).
- In step 3 untitled “Filter candidate nuclei locations by SVM”, the system filters the initial proposal boxes based on SVM classification of the extracted features map into background and nuclei patches. We investigated various feature concatenation models,

but experiences showed that a fusion model made at score level achieved the best results.

- In step 4, untitled “Validate right nuclei locations by the convolutional neural network” a fine-grained classification of the filtered proposal boxes is applied by CNN. In fact, analysis of nuclei locations output by SVMs models built upon HOG, GLCM, LBP and PC features show a non-negligible proportion of background elements wrongly considered as nuclei contents. Which means that despite the complexity of the user-defined feature vectors, challenging background elements are yet difficult to be captured fully and comprehensively. The usage of deep learning strategy in this step aims to overcome the inherent limitations associated to hand-crafted feature vectors models: the lack of abstraction levels provided by CNN [94].

1.4 List of Contributions and Publications

The main contributions of this thesis can be summarized as the following:

- a) We proposed an automated routine for calcium hotspot detection and signal measurement based on a 2-level segmentation algorithm. At the frame level, local thresholds are computed based on Otsu’s algorithm. Then an optimum global threshold is derived from the mean and STD of the Otsu thresholds to segment calcium signal release events in individual video frames. At the video level, the best active cellular regions are identified based on a hit-ratio value derived from the frequency of calcium release events inside cells.
- b) We developed a photobleaching detection technique in hotspots calcium signals by mono-exponential curve fitting based on non-linear least square regression.
- c) We proposed a signal recovery technique after photobleaching detection in hotspot calcium signals by baseline adjustment.
- d) We built a region proposal box generation for nuclei detection in FM images using local adaptive thresholding to generate candidate locations and linear SVM fusion at score level to filter the candidates.
- e) We designed and trained a deep learning model for image patch classification to ultimately detect nuclei box using a convolutional neural network (CNNs) from the proposal boxes output by SVMs.

-
- f) We developed a Graphical User Interface (GUI) software for intracellular calcium signal analysis using MATLAB. It is a standalone application that runs on Windows PC that enables biologists to perform calcium signal segmentation, hotspots detection, photobleaching correction and signals measurements.

The following papers have been produced throughout the research:

- a) D. Traore, K. Rietdorf, N. Al-Jawad, and H. Al-Assam, “Automatic hotspots detection for intracellular calcium analysis in fluorescence microscopy videos,” in *Communications in Computer and Information Science*, 2017.
- b) D. Traore, K. Rietdorf, N. Al-Jawad, and H. Al-Assam, “Photobleaching detection and correction in cardiac cell calcium signals by negative exponential curve fitting and baseline adjustment,” in *Communications in Computer and Information Science*, 2020. [in preparation]
- c) D. Traore, K. Rietdorf, N. Al-Jawad, and H. Al-Assam, “Customized R-CNN with adaptive proposal box generation for nuclei detection in fluorescence microscopy images” in *Proceedings - International Symposium on Biomedical Imaging*, 2020. [in preparation].

1.5 Thesis Outline

Chapter 2 presents background information and literature about techniques for cellular parameters analysis through fluorescence microscopy images.

Chapter 3 proposes a 2-level segmentation technique for intracellular calcium hotspot detection and signal measurement in cardiac myocytes. It also presents a comparison between our proposed approach and a pioneering technique.

Chapter 4 provides a machine learning approach for photobleaching detection and correction in calcium hotspots signals using non-linear regression.

Chapter 5 describes a theoretical and empirical analysis of objection detection techniques that are applicable to nuclei detection in images obtained by fluorescence microscopy.

Chapter 6 proposes a feasible and efficient solution for nuclei detection in fluorescence microscopy images using R-CNN and adaptive region proposal generation.

Chapter 7 highlights the thesis achievements alongside, its limitations and potential areas for future research.

The Appendix A presents the biomedical image analysis toolbox we developed in MATLAB to automate the analysis of intracellular calcium imaging through video recordings of cardiac myocytes stained with calcium indicators.

Chapter 2

Background and Literature Review

The background study and literature review reported in this chapter emphasizes on relevant and recent solutions proposed for and applicable to the automated analysis of biomedical images obtained by fluorescence microscopy (FM). In general, we shall investigate texture related features that can help encode and distinguish biomedical image contents such as cellular objects and background elements and review pre-processing techniques to enhance image quality and to deal with specific problems associated with FM images such as noises and illumination artefacts. Moreover, we shall discuss segmentation algorithms applicable to calcium signal analysis in calcium imaging videos. Finally, we shall review some advanced techniques developed for cellular and/or nuclear detection in biomedical images.

2.1 Introduction

In the last decade, a considerable amount of literature has been published on computational biology applicable to the analysis of fluorescence microscopy cell images. These studies can be classified into two main categories. First, there is a set of traditional methods based on standard image processing techniques such as intensity-based segmentation [61], frequency domain component filtering [73], mathematical morphology [89], edge and region-based segmentation algorithms [24]. Then, the group of machine learning techniques using state-of-the-art classifiers that can be either deep networks [80] [25] or hand-crafted feature extraction models for mainly image segmentation, object detection and image classification [59] [1] [26].

Despite the number of previous works and the relative performances of various state-of-the-art techniques over specific datasets such as natural images and high-resolution medical images, multiple instance object detection in low-contrast FM images remains problematic. Challenges associated to FM images are essentially image quality, movement of cells, potentially unwanted background due to unspecific staining, variability in image illumination, and variability in regions of interest (ROI) sizes, shapes, colours and textures which makes most of the techniques tested on specific datasets to fail on unknown images i.e. on images with different characteristics [71] [43].

In the following sections, we firstly report a computational background study and investigation on four popular texture related features that may help encode and distinguish image contents in FM datasets such as Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), Grey Level Co-occurrence Matrix (GLCM) and Phase Congruency features (PC). Then, we present a quick review on the literature about pre-processing techniques that deal with some of the challenges associated to FM cell images. Finally, we review techniques for intracellular calcium imaging analysis, and cellular objects detection in 2-D greyscale biomedical images.

2.2 Computational Background

To develop robust machine learning techniques for automated cell detection in FM images, we investigated multiple hand-crafted features vectors such as HOG, LBP, GLCM and PC. In the next subsections we describe these features ability to encode cellular contents and background elements and their visual representation in FM images stained with DNA markers.

2.2.1 HOG Feature Representation and Description

HOG features represent local shape information within images patches [82]. Here, we found out that computing HOG within a cell window of [10 x 10] discriminates well background patches from their counterpart nuclei in many images, after gamma transform. Gamma transform is a simple image contrast enhancement technique that converts a low contrast input image into a visually pleasant image without significant alterations [66] according to the formula $A = c * I^\gamma$, where A is the output image, c a constant, I the input image, and γ the power. Figure 2.1 describes shape information encoded by HOG in NRVM patches.

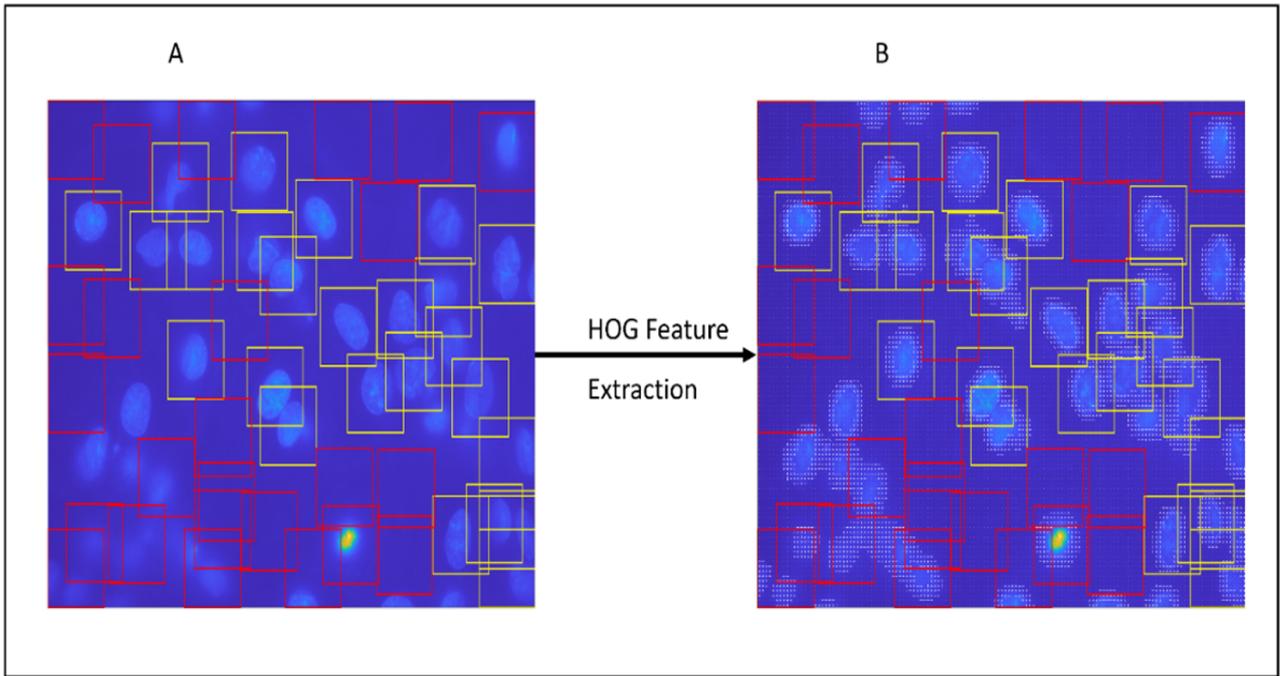


Figure 2.1: HOG feature extraction after gamma contrast enhancement. Sub-images A and B represent respectively an original NRVM nuclei image and its HOG feature representation. Nuclei samples being highlighted by yellow boxes and background contents by red boxes.

Figure 2.1 demonstrates visually that HOG features can capture various object shapes in image patches. Figure 2.1(A) illustrates an original NRVM image with nuclei and background patches labelled by bounding boxes. Red boxes correspond to background labels, whereas yellow boxes highlight nuclei locations. The input image is originally greyscale, but is displayed in pseudo colour for visual purpose. Figure 2.1(B) illustrates the visual representation of the HOG features extracted from patches in A. To implement the HOG feature extraction, we used the dedicated MATLAB proprietary function presented here [22].

In conclusion, as illustrated in Figure 2.1, HOG features help encode various shapes of objects inside images. Nuclei can look like egg, heart, peanut, circle while background elements may show fancy patterns ranging from flat surfaces, elongated forms, circular shapes, to complex patterns. This means that a background element may also look like a nucleus, as illustrated in Figure 2.1(B) mid-bottom region, where a bright noisy dot is encoded as a circular object by HOG. Such visual confusion related to HOG features may indicate that more information such as texture encoding may be necessary to distinguish nuclei from background elements in addition to shape descriptors. Moreover, in some lower-contrast images, nuclei contours may appear as shadow and therefore difficult to be captured by HOG descriptor. This is generally due to staining deficiency. In such cases, HOG may also fail to recognize shapes, and hence to

distinguish accurate background contents from nuclei. The performance of SVM models built on HOG and tested over splits of our nuclei datasets is reported and discussed in Section 6.4.2. Filtering Initial Nuclei Proposals by SVM Models.

2.2.2 LBP Feature Representation and Description

The fact that background elements may exhibit the same visual shape as foreground elements led us to investigate additional texture-based descriptors. The first texture-based features, we analysed are LBP features. LBP are famous in the field of computer vision for achieving great success in image classification tasks [26] over various medical datasets.

To learn and differentiate nuclei and background textures, we extracted from uniform LBP image regions, 59-bins histograms. Figure 2.2 illustrates the LBP features' ability to encoding rich texture details in NRVM images. Here, for instance, the bottom background element with misleading HOG shape observed in Figure 2.1 can be better distinguished as non-nucleus by the LBP texture encoded (Figure 2.2(B)).

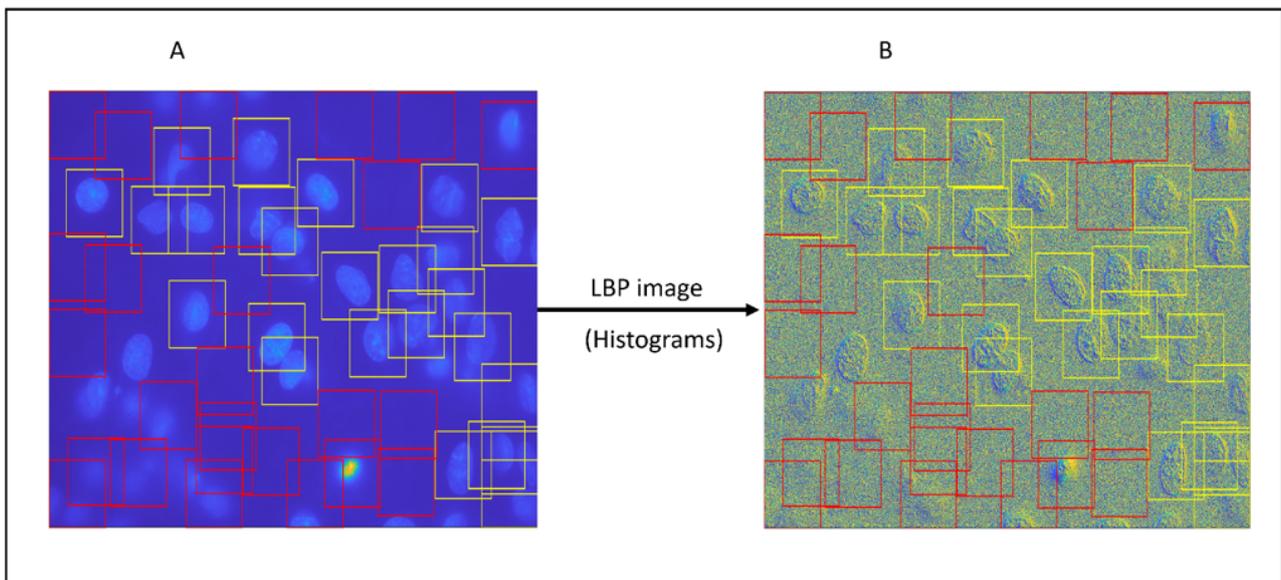


Figure 2.2: Uniform LBP image computation after gamma correction. Sub-images A and B represent respectively an original NRVM nuclei image and its LBP feature representation with nuclei patches highlighted by yellow boxes and background contents highlighted by red boxes

Figure 2.2 illustrates the texture information encoded by LBP features when applied to NRVM nuclei images. Figure 2.2(A) represents an original NRVM image with nuclei ground accurate labels (in yellow) and background ground truth labels (in red). Figure 2.2(B) illustrates the

LBP version of A, where 59-bin histograms will be later extracted as LBP feature vectors to describe individual patches. Visually it can be noticed that right nuclei regions are encoded as subtle prominence over general flat valleys surface corresponding to background elements.

However, few background elements exhibit some peak pattern like the misleading background dot element located at the bottom of Figure 2.2(A-B). A close look shows that the LBP texture representation of the problematic background element is well different than the texture representation of nuclei regions. Moreover, the 59-bin histogram extracted from the misleading background region shows some differences when compared to the histogram extracted from a true nuclei patch, as highlighted in Figure 2.3(C-D).

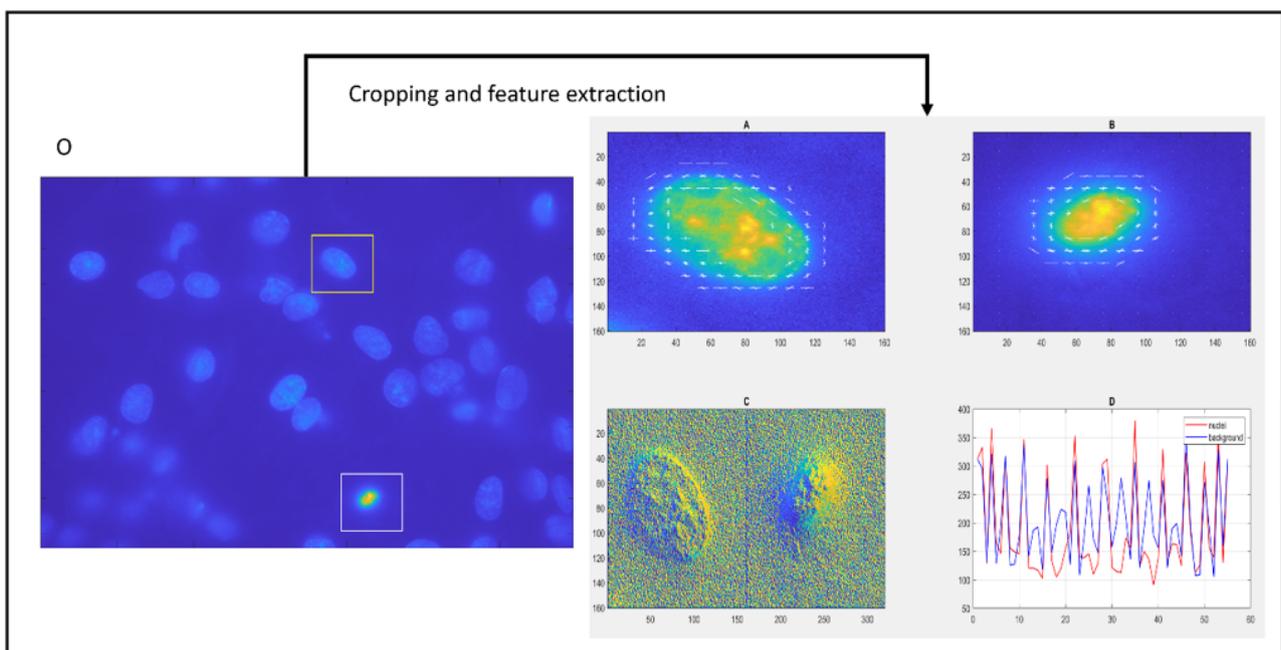


Figure 2.3: LBP texture vs HOG Shape information. Sub-image O shows an original NRVM nuclei image with a nuclei sample highlighted by a yellow box and a background sample highlighted by a white box. Sub-images A and B show respectively the HOG encoding of the nucleus and background patches highlighted in O. Sub-images C and D show respectively the LBP images and the derived histograms from the LBP images of both nuclei and background images computed in C.

Figure 2.3 illustrates a qualitative comparison between shape information encoded by HOG features and texture information encoded by LBP features for nuclei and background patches classification. Figure 2.3(O) represents an original NRVM nuclei image with two labelled regions nucleus (in yellow) and background (in white). Figure 2.3(A) and Figure 2.3(B) represent respectively, the HOG features of the nucleus and the background regions. Visually it is evident

that the nucleus in A and the background element in B exhibit the same circular shape which highlights the limitations of HOG shape descriptors.

Figure 2.3(C) shows the corresponding LBP image computation of the patches (A) and (B) cropped from (O). Figure 2.3(D) represents the 59-bins LBP histogram of the nuclei patch in (A) (red line) and the background patch in (B) (blue line). From those LBP histogram plots, it is noticeable that nuclei feature distribution spans in larger value intervals than background counterparts even if they intersect sometimes.

Based on that visual comparison of HOG and LBP feature representation, we argued that LBP could be an added value to use alongside HOG features to filter nuclei patches. This assumption related to shape descriptors will be later confirmed when comparing the output of SVM models based on separate HOG and LBP feature features. The overall performance of SVM models built on LBP and tested over splits of our nuclei datasets is reported and discussed in Section 6.4.2.

2.2.3 GLCM Feature Representation and Description

GLCM is another textural feature that we investigated after LBP features to encode nuclei and background clues in NRVM images. GLCM derives from the spatial distribution at different angles or offsets of pixel pairs [88] that we investigated. Here, we found out that extracting statistics such as the contrast, the correlation, the energy, and the homogeneity derived from GLCM computation with a 59 x 2 offset window gives a richer textural description of nuclei and background patches. The co-occurrence matrix we computed for each NRVM image of our dataset is a matrix that tells how often pairs of pixels with specific values and in a specified spatial relationship occur inside NRVM images. This matrix is defined to be the distribution of co-occurring pixel values at a given offset or distance according to the following formula [88]:

Consider:

$$N = \sum_i \sum_j G_d(i, j); \quad (2.1)$$

which is the total number of occurrence pairs in G_d

Let:

$$GN_d(i, j) = \frac{1}{N} G_d(i, j).GN_d; \quad (2.2)$$

$GN_d(i, j)$ is called the normalized co-occurrence matrix, where $(i, j)^{th}$ entry of $GN_d(i, j)$ is the joint probability of co-occurrences of pixels with intensity i and pixels with intensity j separated by a distance k , in a particular direction d .

Figure 2.4 below shows that energy features (solid blue line) extracted from GLCM nuclei images are different than those extracted from GLCM background patches(Figure 2.4(C-D)). In contrast, for the same cell and background patches, LBP histograms exhibit little differences (Figure 2.4 A-B). In other terms, where LBP could fail, GLCM might offer a healthy alternative to recognize nuclei and background patches.

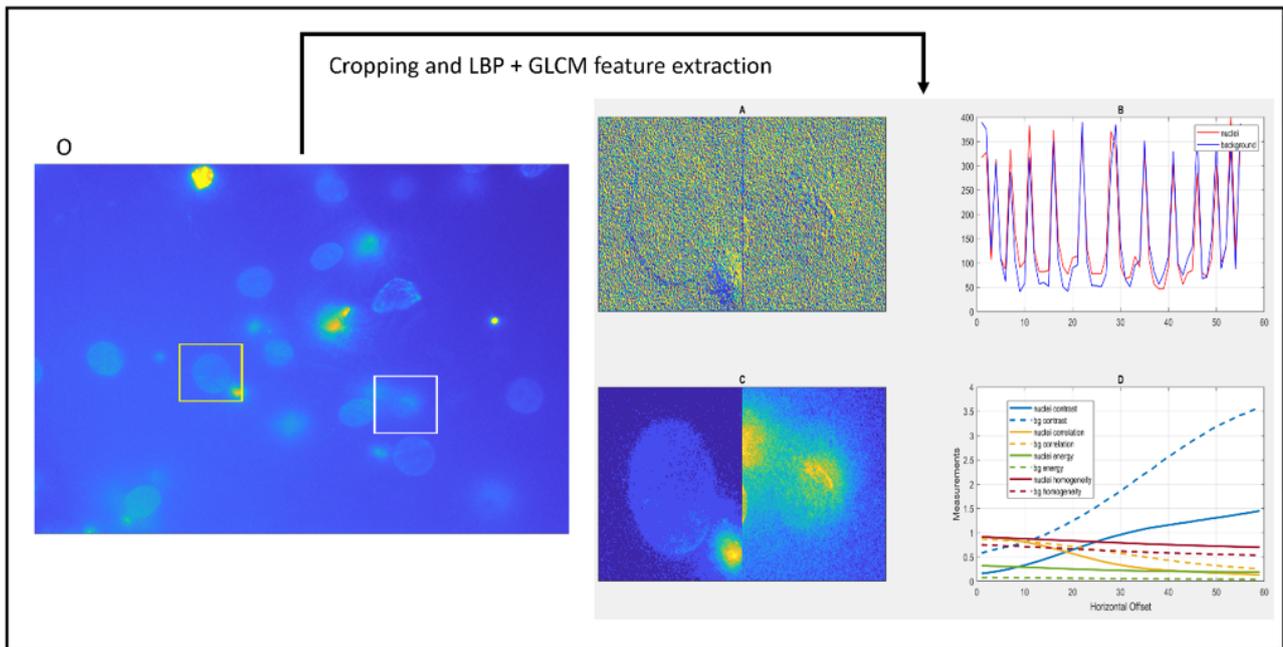


Figure 2.4: GLCM features vs LBP features for nuclei texture encoding. Sub-image O shows an original NRVM nuclei image with a nuclei sample highlighted by a yellow box and a background sample highlighted by a white box. Sub-images A and B show respectively the LBP encoding of the of the nucleus and background patches highlighted in O. Sub-images C and D show respectively the GLCM images and four derived features extracted from the GLCM images of both nuclei and background patches computed in C.

In Figure 2.4, (O) represents an original NRVM image with one nucleus fade region labelled in yellow and one background region labelled in white. (A) represents LBP computation of the cropping versions of those image regions highlighted in (O). (B) represents the LBP histograms plotting of both LBP images displayed in A. So, visually, it is easy to notice that both regions exhibit similar LBP texture with no obvious cell texture standing out or background prominence as in Figure 2.2. This is further verified in (D) where their corresponding GLCM features

extracted are plotted. Solid lines represent GLCM features of the nuclei patch, whereas the dashed line corresponds to the background patch's ones.

To sum up, GLCM may outperform LBP visually as it enables encoding fine details texture from fading nuclei regions that will not be possible with LBP. Therefore, we made the hypothesis that GLCM based SVMs models may also outperform LBP counterpart in challenging cases where images exhibit low contrast cells. The performance of SVM models built on GLCM and tested over splits of our nuclei datasets is reported and discussed in Section 6.4.2.

2.2.4 PC Feature Representation and Description

PC is defined as a low level invariant property of image features [50] that tries to capture edges and corners of objects from frequency domain representation of image contents. PC is known for its performance at detecting object edges and corners in challenging environments such as local image illumination variations and low contrast images [50] [49] [59]. For these reasons, PC might be interested to test on FM images which exhibit similar challenges. Multiple theories behind the calculation of phase congruency have been released in the past. Here, the Kovessi's formula based on a new measure of the phase deviation using sin and cos functions in 2D is adopted [50].

$$PC(x) = \frac{\sum_n W(x) [A_n(x) \Delta \phi_n(x) - T]}{\sum_n A_n(x) + \epsilon} \quad (2.3)$$

Where $PC(x)$ is the PC value at x .

$$W(x) = \frac{1}{1 + e^{\gamma(c-s(x))}} \quad (2.4)$$

$A_n(x)$: amplitude of filter responses at x ; $\phi_n(x)$: a phase angle at x ; $\bar{\phi}(x)$: mean phase angle; ϵ : a small constant to avoid division by zero such that the phase deviation measure is given by:

$$\Delta \phi_n(x) = \cos \phi_n(x) - \bar{\phi}(x) - \|\sin \phi_n(x) - \bar{\phi}(x)\| \quad (2.5)$$

To encode nuclei features based on PC, we adopted the improved MATLAB implementation of Kovessi, which consists of a set of wavelet filters convolutions over input images prior to Gaussian smoothing. For more details, please read [50]. Thus, to learn relevant nuclei features

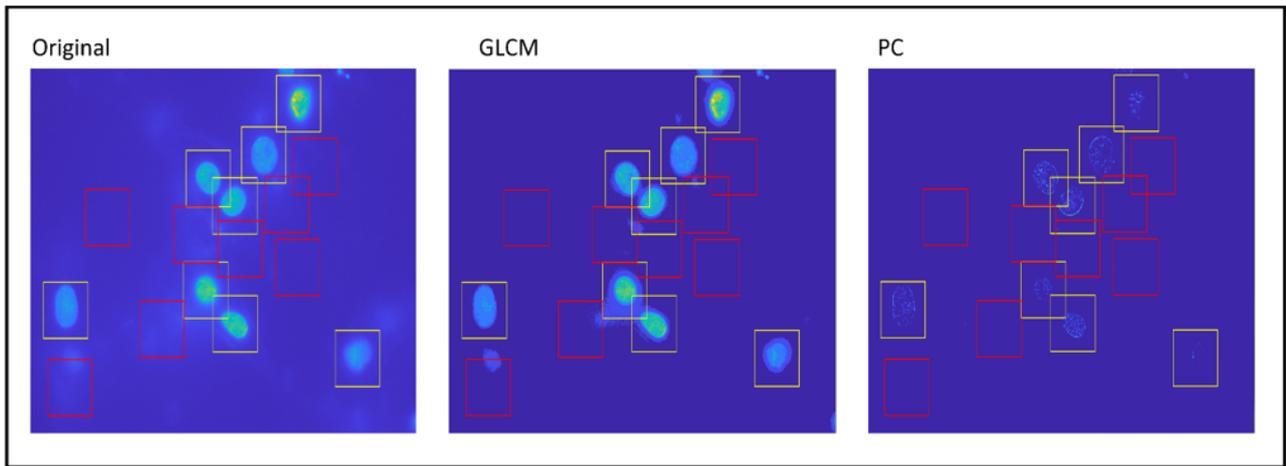


Figure 2.5: PC features vs GLCM features for nuclei and background texture encoding

in NRVM images, we extracted from PC versions of nuclei and background patches, 59-bins histograms as feature descriptors.

Figure 2.5 illustrates that PC can often outperform GLCM features and can better help discriminate heterogeneous background patches. Heterogeneous background patches are NRVM image windows that contain numerous illumination artefacts that could be mistaken as nuclei. This is illustrated in Figure 2.5 by the red labelled background region located at the bottom-left corner of the image labelled “original”. Visually it can be noticed that in this background region, PC did not rightly encode any strong edges (image labelled “PC”). In contrast, the corresponding GLCM image produced a cellular texture (image labelled “GLCM”). Therefore, we made the hypothesis that PC could be an added value to GLCM feature vectors for SVM filtering of nuclei and background patches due to their sensibility at denoting or not object edges in image regions.

In details, Figure 2.5 demonstrates that PC and GLCM could be used in parallel for nuclei and background patches classification in NRVM images by SVM models. As the titles suggest, the three images displayed in Figure 2.5 represent respectively, an original NRVM input image, a GLCM image version of the original image and a PC image version of the original image. As it can be denoted visually, texture information encoded by GLCM features and those encoded by PC features seems proportional in all patches except for a background patch located at the bottom left corner of the three images and for a nucleus patch situated at the bottom right corner.

In the seven nuclei patches located at the top of the images and highlighted in yellow, GLCM

texture is visible and prominent, while PC texture is strong. It is the same for the seven background patches located at the top of the images and highlighted in red. In those background regions, GLCM texture and PC edges are logically null or hardly visible.

The only disparity happens inside the two lowest image patches located respectively at the bottom-left corner (highlighted in red) and the bottom-right corner (highlighted in yellow) of Original image of Figure 2.5. In the first region, which exhibits a heterogeneous background, GLCM feature encodes some tiny nuclear texture as illustrated in the GLCM image of Figure 2.5. In contrast, the PC does not exhibit any edge map. In the second problematic region, which exhibits a vague-contour cell, GLCM feature encodes some strong nuclear texture, while PC still shows slight edges fragments challenging to visualize.

This dissimilarity of the two types of feature vectors at encoding a heterogeneous background patch and a weak boundary cell made us investigate confidence levels associated to the object classes output by SVM models built upon unique feature vectors such as PC and GLCM. We expanded this analysis to SVM models based on HOG and LBP. We also investigated SVM fusion models based on multiple combinations of the four individual feature vectors: HOG, LBP, GLCM, and PC. All these investigations about SVM models for nuclei and background patches filtering are presented and discussed in Sections 6.3 and 6.4. In next Section 2.3, we introduce the relevant literature about pre-processing techniques, automated calcium signal analysis, and cellular segmentation in FM images and videos.

2.3 Review on Pre-processing Techniques

The advent of FM and its application in cell biology has been a fantastic mean for visualizing and understanding complex sub-cellular and molecular mechanisms involving thousands of proteins that reside in living organisms. Fluorescent labelling through the usage of dyes or contrast-enhanced products also called GFP had enabled easy tracking, in vitro and in vivo, of tissues, cells, proteins, ions, and various macro-molecular structures. To learn more about the groundbreaking contributions of GFP in cellular biology, please read Roger Tsien's paper that won the 2008 Nobel Prize for Chemistry [78]. From these optical advances, numerous biomedical studies and research are getting benefits. But, like often, computer scientists who aim to provide hardware and software solutions to automate the analysis of such medical data, are yet to overcome all the computational challenges associated to the imaging power.

Some common problems associated with FM imaging datasets are illumination artefacts, low resolution images, photobleaching, Gaussian and Poisson noises [61] [71] [54] [92]. To deal with illumination artefacts, poor resolution and noises in fluorescence images, many solutions have been proposed such as contrast enhancement techniques based on greyscale histogram manipulation, morphological reconstruction, gamma transform, background subtraction techniques, point operators, low-pass filtering using Fourier Transform or wavelet transform, median filtering and Anscombe transform [36] [63] [20] [64] [45]. Those techniques can be remarkably successful on specific datasets but dramatically fail on others due to wrong parameter setting.

Besides, robust noise removal techniques may cause blurring effects on images and crucial information contained within the images might be lost. This could be the case when applying noise removal in calcium signalling images where intracellular calcium changes are captured by the difference in the intensity (brightness) of a calcium indicator. The direct manipulation of pixel values in such images could alter the calcium concentration derived from cellular regions' brightness.

Another reason that makes it difficult to apply standard techniques such as morphological operations to unseen datasets is related to the fact that the performances of morphological operations rely on the technique adopted, and the type and size of the structuring elements chosen. For instance, the rolling ball technique used for background estimation, and based on morphological top-hat operations described in [60] would need prior knowledge about input images, for example, the size (in μm) of the signals the researchers are interested in, if default parameters do not work.

Therefore, instead of developing robust pre-processing techniques for noise removals and image contrast enhancement, we adopted cautious strategies according to the phenotype under investigation. For the calcium signalling image datasets, a set of routines consisting of adaptive background subtraction, a default [3x3] median filtering, and a simple morphological reconstruction were investigated and adopted. For the counting of nuclei cells, we applied gamma correction over image patches prior to feature extraction to increase the brightness and adjust the contrast of fading cells. In following Section 2.4, we report a literature review on automated solutions for intracellular calcium signal analysis in FM datasets.

2.4 Review on Automated Analysis of Intracellular Calcium Signals

Calcium ions are truly one of the most important minerals in human body [40]. They help solidify human bones and teeth. They can also be found free in many soft tissues and intracellular domains where they are used as secondary messengers or triggering signals to multiple cellular functions [14]. The automated analysis of such intracellular calcium signals remains a challenging task due to their spatio-temporal diversity and staining protocols i.e., the spatial extent, duration, and amplitude of calcium signals, and the microscopy technique and camera settings or calcium indicators used to stain the calcium signals. Table 2.1 list the existing and relevant techniques of calcium signal analysis in FM images found in the literature.

One of the earlier investigations for automated analysis of such intracellular calcium signals was published in February 1999 by Cheng and Co [21]. In their findings, the authors demonstrated that local calcium transients or calcium sparks observed in confocal line-scan images of rat ventricular myocytes are outliers relative to the mean and standard deviation of the whole image fluorescence. They believed that calcium spark signals could be segmented above a relative fluorescence threshold equalling to the mean plus two times the standard deviation of the total image fluorescence. That was brilliant at the time but difficult to sustain in conditions of illumination variation inside image sequences, or for different cell types or cells with different intracellular calcium intensities or for experimental conditions using different calcium indicators or different camera settings that may affect the overall fluorescence measured.

Later in 2012, a plugin developed in ImageJ was proposed for automated ROIs analysis of dynamic calcium signals in images sequence of rat arterial cells [34]. The technique combines statistical noise filtering and arbitrary ellipse fitting of high-intensity image contents to identify and measure local calcium signal events in cells within an arbitrary radius of 15 pixels per ROI. The technique also applies an arbitrary of $p < 0.05$ to segment the events. Using a fixed threshold value and assuming all calcium signal events circular make this approach limited in the sense that calcium signals are diverse in shape, frequency, lasting time, intensity and spread according to the cells involved and experimental conditions in place. A technique based on dynamically computation of such parameters of size and intensity thresholding might be theoretically more efficient.

Methods	Features	Dataset	Performance	Limitation
[21]	Raw pixel data	CLS images of skeletal muscle cells and cardiac myocytes.	Specificity = 90% Accuracy = 50%	Fails in Background fluctuation and local variation of calcium intensities.
[34]	Normalized pixel data	CLS images of pulmonary cells and arterial cells	Sensitivity = 99%	Fails to detect heterogeneous calcium signals
[28]	Raw pixel data	TIRF images of neuroblastoma cells	Speed: 30 seconds for 4000 frames.	Only dedicated to puff calcium signals.
[76]	Raw pixel data	TIRF images of plasma membrane of <i>Xenopus laevis</i>	Accuracy = 100% for long calcium signals.	Only dedicated to puff calcium signals.
[52]	Fourier transform spectrum	Confocal microscopy images of Jurka T-cells	classification Precision = 100%	No calcium signal segmentation.

Table 2.1: Relevant techniques for automated calcium signal analysis in FM datasets. CLS: Confocal line scan images. TIRF: Total Internal Reflection Fluorescence microscopy.

Another approach using dynamic user-defined spatio-temporal filters, threshold detection level and selection of the background level to be subtracted was proposed in 2014 for local calcium puffs detection and measurement in Total Internal Reflection Fluorescence microscopy (TIRF) [28]. This approach written in MATLAB and Python is interesting but the requirement of user input of crucial parameters such as the size of the ROIs of the algorithm makes it prone to same user-bias as a manual procedure would be.

Developed and tested on the same type of microscopy data (TIRF) in 2018, the CellSpecks software solution written in Java, came as a more autonomous and robust GUI application for small calcium channels detection and activity measurement of oocytes plasma membrane at high frame rate ($> 500fps$) [76]. Although, the Java environment provides flexibility and portability to CellSpecks software, its proven performances over only local calcium signals do not guarantee same performance over global signals i.e., calcium transients that repeat frequently or spread in bigger spaces and last longer such oscillatory signals or calcium waves of cardiac cells for instance. These limitations of the CellSpecks solution are partially addressed by the CalQuo2 approach for calcium signal classification described here [52].

In CalQuo2, a frequency domain filter based on Fourier transform [52] is designed to detect and differentiate single calcium response and global calcium oscillations of Jurka T-cells. In their approach, authors of CalQuo2 demonstrated that the Fourier spectrum of the intensity-based curve of calcium signals could help classify cells into triggering and non-triggering cells of global calcium signals according to a certain radius threshold set over parts of their Fourier spectrum. Calquo2 is an interesting solution for cells classification based on their calcium signals. However, it does not answer the calcium signal detection problematic itself i.e., it is not an automated framework for the detection and measurement of calcium signals in cells images.

All in one, to the best of our knowledge, several solutions for intracellular calcium signals analysis exist but none of them propose an automated solution for the detection and measurement of the best active regions in cells that is data driven in two-dimensional images i.e., where the ROIs parameters are inferred from the series of FM images input and analysed. That is why in this thesis we aim to investigate an automated solution for intracellular hotspot detection and measurement through FM videos of cardiac myocytes without prior knowledge or user input of crucial parameters of size and threshold value of the calcium signal under investigation. Next Section 2.5 reports our literature review on the cell segmentation techniques applicable to cell biology images obtained by FM.

2.5 Review on Cell Segmentation Techniques Applicable to Cell Biology Images

Cell or and nuclei detection techniques applied to FM images range globally from hand-crafted algorithms and machine learning techniques to deep learning strategies. For an extensive review on nuclei detection and segmentation techniques please read [42], [61],[93], and [45]. In this thesis, we aim to investigate the limitations of hand-crafted algorithms and some machine learning techniques and to propose better practical and working solutions that can be easily adapted to other datasets.

By hand-crafted algorithms, we mean traditional image processing techniques where computers are specifically programmed to perform cell and/or nuclei detection and segmentation tasks from end-to-end without any learning strategy. These traditional methods for nuclei detection and/or segmentation in microscopy images may consist of threshold-based techniques [65] [57], Laplacian of Gauss (LoG) filtering [47], morphological operations [89], region growing solutions [44], and marker-controlled watershed segmentation techniques [89] [85].

Thresholding techniques that rely on global or local thresholds to separate foreground element from the background are said to perform well when the signal-to-noise-ratio of images is high, and when there is a clear a separation of the pixel intensity distribution between nuclei and the remaining image contents [59]. This might not be the case in many real-world scenarios, where Gaussian and Poisson noises may corrupt FM images. Moreover, such images may also exhibit non-uniform background, weak cell boundaries and varying illumination artefacts sharing same colour intensities like the ROIs.

To overcome the limitations related to thresholding, people use LoG filters. LoG filters enable better handling of background disparities and noise reduction. However, LoG performance is still critically dependant to the bandwidth choice [74] relevant to ROIs' sizes. Another popular traditional image processing technique used to detect nuclear components in microscopic images refers to mathematical morphology such as Top Hat and h-dome operators where structuring elements with tuned sizes help differentiate foreground elements from image background [45].

However, as the structuring element tuning suggests, morphological solutions assume prior knowledge about the dataset, i.e. the size of the nuclei and the undesirable artefacts. In other terms, it would be a happy coincidence that default parameters of algorithms tested over

different datasets give same results over new datasets without prior knowledge and finetuning of user-input parameters.

It is to reduce this need of tuned input parameters that machine learning (ML) detectors have been proposed for nuclei detection like in the reviews described in [61],[42], [43]. Those ML techniques include unsupervised clustering algorithms and supervised ML classifiers such as random forest model, Gaussian mixture model, Bayesian graph cut, binary graph cut, spatial fuzzy c-means, fuzzy level set algorithms [4] and Markov random fields model [61] [95]. The advantage of using these ML techniques is related to the fact that the knowledge about ROIs are inferred from features within images rather than from users' inputs. Clustering and object detection are data-driven accordingly even if crucial inputs are still needed in some cases such as seed points notification in graph cut techniques.

The main limitations related to the ML techniques described above refers to feature selection paradigm i.e. how to make sure that the selected features are the best to describe any kind of nuclei in most fluorescence images knowing that there are no perfect features or a set of features to describe all kind of nuclei regions. This is a trivial question that may never get answered soon as it is even believed that the selection of the best features may depend only on someone's experience and luck rather than any rationality [96]. For example, detection models based on HOG features which are popular in object recognition task [82] for capturing shape information of rigid objects may fail to recognize deformable objects or objects showing various shapes within same images. In the same sense, it can be argued that classification models based on HOG features would make wrong predictions about different objects showing similar shape information.

In our case, the observed nuclei may have various shapes or forms within the same images, and exhibit weak boundaries. Moreover, in some extreme conditions, microscope images might show background elements with circular, ellipsoid, elongated shapes that can be mistaken as nuclei when shapes are the only descriptors chosen to distinguish foreground and background elements. The conclusion we can draw from such conditions is that theoretically, HOG features alone may not be enough to discriminate nuclei. The same analysis can be applied to other state-of-the-art ML techniques available in the literature using only hand-crafted feature selection. To sum up, the generalization is still an issue for ML techniques due to the inherent limitations of the feature vector selection to encode image contents.

To overcome the limitations associated with ML models that are dependent to their feature extraction type, supervised deep learning (DL) strategies seem better alternatives. The DL strategies based on CNNs for object detection, make the feature vector extraction more systematic, richer, hierarchical, and portable to any kind of object as long as there are enough image patch/label pairs of data to learn from as described in [35].

The most popular CNNs applicable to object detection in FM images are regions with convolutional neural network (R-CNN) [35], faster RCNN [67], mask RCNN [38], and fully convolutional networks (FCN) [77]. In those general ANN architectures, the aim is dual: either to assign labels to image patches or to assign labels to image pixels. The two strategies are respectively called box and pixel-wise classifications.

With regards to R-CNN, for instance, the architecture we have adopted in this thesis, the object detection is performed in three phases: first, around 2000 candidate object proposals are generated by sliding windows [41], or edges boxes [97], or by selective search using SIFT features combined with a SVM classifier [87]. Then rich hierarchical feature extraction is performed by CNN layers from each proposal box flattened into a fixed size using a popular technique called affine image warping. In geometry, an affine transformation is a transformation function between affine spaces that preserves points, straight lines, and planes. Finally, fine-grained object detection is performed at the output layer of the CNN using image patch classification by SVM and bounding box refinement by a box regressor. Figure 2.6 describes the standard architecture of R-CNN as described in [97].

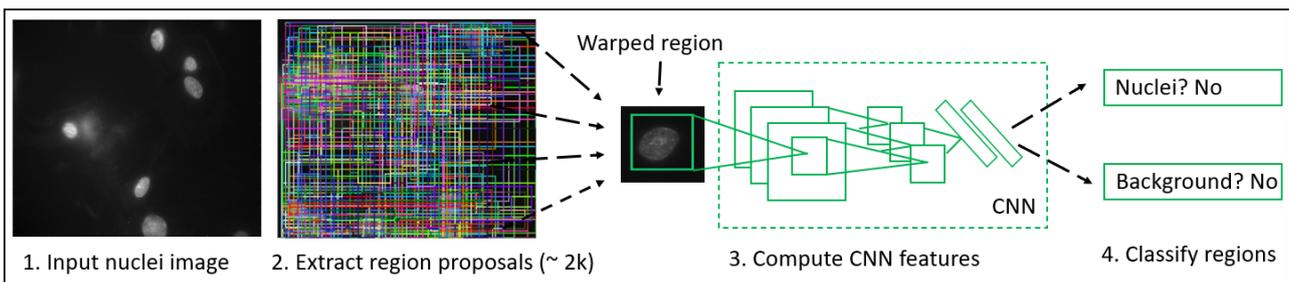


Figure 2.6: Architecture of R-CNN (Region with Convolutional Neural Network).

With regards to pixel-wise classification, for instance, authors in [72] proposed the U-Net solution, a variant of FCN applied to light microscopy cell images. This U-Net consists of both down-sampling and up-sampling convolutions over input images while preserving region context to segment Hela cells and neurons with respectively 92% and 77% intersection over union (IoU) accuracy. Another novel deep learning framework presented in [39] for nuclei segmentation

uses a combination of Mask R-CNN and U-net architectures to outperform 739 techniques presented at the 2018 Data Science Bowl competition. Authors of this technique described in [39] achieved this performance over unknown testing datasets due to the artificial data augmentation strategies they developed and applied to imitate multiple image conditions.

With regards to similar biological ROIs analysed by FM to the nuclei images we dealt with in this thesis, authors in [96] report experiments of a faster R-CNN based approach for detecting six categories of microscopic cells such as red blood cells, white blood cells, yeast, crystal, cast, and epithelium cells. Although those cells differ in shape, texture, size, area, colour intensity, and spatial distribution within images, authors demonstrated that a CNN classifier (detector fine-tuned upon VGG16 [18]) could outperform traditional SVM models trained with hand-crafted features. The different cell types detected by the technique described in [96] with a mean average precision of 91.6% indicated how well deep networks could generalize well across other datasets compared to hand-crafted algorithms [6].

Faster R-CNN is an evolved version of R-CNN where the proposal box generation is performed by a CNN module called region proposal network (RPN) instead of a separate object detection technique that can be Edge Boxes [97], Selective Search [87], Objectness Measure [1] or Sliding Window search [41]. In this thesis, we aim to investigate solutions based on the basic R-CNN as it is simpler and requires less computation power in terms of double training CNN for proposal box generation and final box classification.

2.6 Summary on Literature Review

To sum up, a considerable amount of literature has been published on automated analysis of intracellular parameters through images and videos of FM. For pre-processing FM images, many techniques can be applied with relative outstanding performances according to the testing datasets such as spatial and frequency domain filters, histogram-based operations, and morphological operations. However, human interactions are needed for parameters setting as filtering may cause unwanted blurring effects on input images.

For calcium signalling segmentation or the segmentation of calcium imaging events, the detection of calcium hotspots in videos of cardiac myocytes recorded using FM, and the photo-bleaching correction in calcium signals recorded in cardiac myocytes, no previous study has fully

investigated automated solutions for the whole procedure, to the best of our knowledge without prior knowledge of the spatial spread, the duration, and the intensity of the calcium signals under investigation. That is why, in Chapter 3, we propose an automated solution for calcium hotspot detection and measurement in cardiac myocytes based on a multi-level thresholding method derived from the input data themselves, and in Chapter 4 an automated routine to deal with photobleaching in temporal calcium signals based on non-linear least square fitting.

For cellular segmentation techniques in biological images, there have been several solutions proposed for multiple cell types and imaging protocols. Still, the generalisation of much-published research on cell or nuclei detection is problematic. We aim to investigate that gap in Chapter 5 and Chapter 6 by arguing that it is feasible to build an efficient nuclei detection strategy based on a customized R-CNN using data-driven region proposal generation.

Chapter 3

Automated Calcium Hotspot Detection and Measurement

This chapter presents our first investigations for the automation of intracellular signalling parameters analysis through fluorescence microscopy (FM) videos. The primary study reported here consists of three stages:

1. The segmentation of calcium imaging events or calcium signals objects by combining local and global thresholding to identify candidate hotspots (active regions)
2. Locating and identifying the true active regions inside cells, i.e. calcium hotspots by employing a novel hit-ratio computation.
3. Extracting calcium signal measurements from individual hotspots to be used at the analysis phase.

Many UK biology labs study calcium signalling processes using FM in which the analysis is usually done manually. In this chapter, we argue that an automated procedure for hotspot detection and measurement in fluorescence videos of calcium imaging can provide scientists with experiment replicability, speed and can prevent from inter-intra observer variability. Thus the contributions of the chapter is twofold: firstly we demonstrate that a two-level segmentation technique based on thresholding and video fusion using region hit-ratio parameter may outperform manual procedure at detecting calcium hotspots in cardiac cells; then secondly by developing a working solution for hotspot analysis, we provide biologist with a graphical user

interface application essential for the automation of their daily calcium data analysis and also applicable to multiple biological problems.

This chapter begins by outlining the biological relevance of the study and an overview of our proposed solution for automated detection of calcium hotspots in cardiac myocytes (in Section 3.1). Then in Sections 3.2, and 3.3, we describe in detail our proposed solution for hotspot detection and measurement. Section 3.4 of the chapter presents the experiments conducted for calcium signal object segmentation and hotspot detection. Finally, in Section 3.5, results of calcium imaging experiments carried out in PVC and NRVM calcium images are reported and discussed prior to concluding remarks presented in Section 3.6.

3.1 Introduction

As mentioned in Chapter 1, calcium signals trigger and inhibit a vast number of cellular functions through various signalling pathways and spatio-temporal patterns, i.e. according to their arise time, space and amplitude [14] [15] [12] [46]. Invariably, in animal cells, intracellular calcium dynamics involve the activation of various calcium channels and other processes inside multiple cellular organelles such as nuclei, sarcoplasmic or endoplasmic reticulum, mitochondria, and/or lysosomes [68][14] [19].

It is therefore essential in many laboratory experiments for biologists to be able to detect, localize and measure intracellular calcium signals of living cells to infer better their effects on the related cellular functions under investigations. For instance, in PVCs, biologists use calcium indicators to analyse the frequency of the spontaneous cell of calcium signals that can modify, beyond a certain threshold, cells' natural ability to follow electrical pacing and cause atrial fibrillation, the most common cardiac arrhythmia [68]. Another cell type used to investigate cardiac signalling is NRVMs. Calcium signalling in these cells can also be studied using fluorescent calcium indicators. The cells can also be used to investigate how calcium signalling regulates other cellular processes, like for example, autophagy [15]. During autophagy, calcium signals are said to act both as, triggering and inhibitory factors according to the calcium channels activated and the signal amplitudes, i.e. the magnitude of the change of calcium concentration [15].

The current chapter consists of investigating automated procedures for the analysis of calcium

imaging data in two types of cardiomyocytes PVCs and NRVMs for the understanding of the biological correlation between changes in the pattern of changes in the intracellular calcium concentration and the development of cardiac dysfunctions like arrhythmia. In these calcium imaging experiments; biologists are interested in identifying calcium hotspots and measuring the changes in the calcium concentration inside the hotspots over time.

As their name suggests, intracellular calcium hotspots (ICH) correspond to areas in the cells where frequent and repeated changes of intracellular calcium concentration are observed over time. In other words, any active cellular regions in calcium signalling events can be considered as ICH. The accurate detection of ICH is needed because it is essential to understand the timing/frequency and amplitude of calcium signals to understand why cellular signalling might be disturbed and affect the function of cardiac myocytes [9][69]. Analysing ICH responses to different stimuli such as treatment with pharmacological modulators (drugs) usage may also be highly beneficial for diagnostic formulation and treatment discovery [13] [14].

We aim to automate the detection of ICH in cardiac myocytes through fluorescence microscopy. Since it is hard to find worldwide established routines for ICH detection and measurement in fluorescent cardiac myocytes, researchers rely on their own experience and computing knowledge to perform calcium imaging data analysis using different programs. Popular software programs among calcium signalling community are Image J, MATLAB environment, Prism and Microsoft Excel. For those familiar with these software, Image J or MATLAB are essential for hotspot setting and calcium signal segmentation, whereas Prism or Microsoft Excel are used for data plotting and statistical analysis.

A typical data analysis workflow of calcium imaging experiment is like the following: a video of cardiomyocytes stained with calcium indicator is loaded into Image J. Users navigate through the input video frames and manually set the active regions in calcium release-events as hotspot locations. From the pixel regions highlighted by the manual annotations, initial readouts of changes in the calcium concentration are extracted, corresponding to the average pixel intensities of the selected regions in all consecutive frames. After the initial readouts, background brightness values are subtracted and the raw pixel intensities normalized, then curve and statistical measurements are computed. Signal normalisation is based on the estimation of the baseline fluorescence that corresponds to the minimum fluorescence of the calcium measurements recorded in hotspots. All in one, computing tasks performed for calcium signal analysis are essentially calcium signal object segmentation, ICH detection, time-series ICH signal ex-

traction, and graph-derived measurements of signal amplitude, width, frequency, the area under a curve, inter-spike interval and time of slope.

Although elaborated, this manual procedure for calcium signals analysis described above and adopted by many UK-based experts in calcium signalling shows many limitations. These are related to speed, accuracy, and replicability of the procedure. Indeed, due to the lack of automated standards, analysis is time-consuming, prone to errors, human biased and difficult to replicate.

In this chapter, we aim to address those issues related to the manual or semi-automated approach by proposing a framework which relies on a two-level segmentation algorithm for ICH detection and measurement in videos of cardiac myocytes obtained by fluorescence microscopy. We argue that contrast enhancement techniques combined with thresholding at the frame and video level using hit-ratio computation outperform manual procedures and provide speed, accuracy and replicability for the detection and measurement of intracellular calcium traces in videos of cardiac myocytes. The flowchart in Figure 3.1 describes the whole process for automated calcium hotspots detection in fluorescence microscopy videos prior to calcium signals processing.

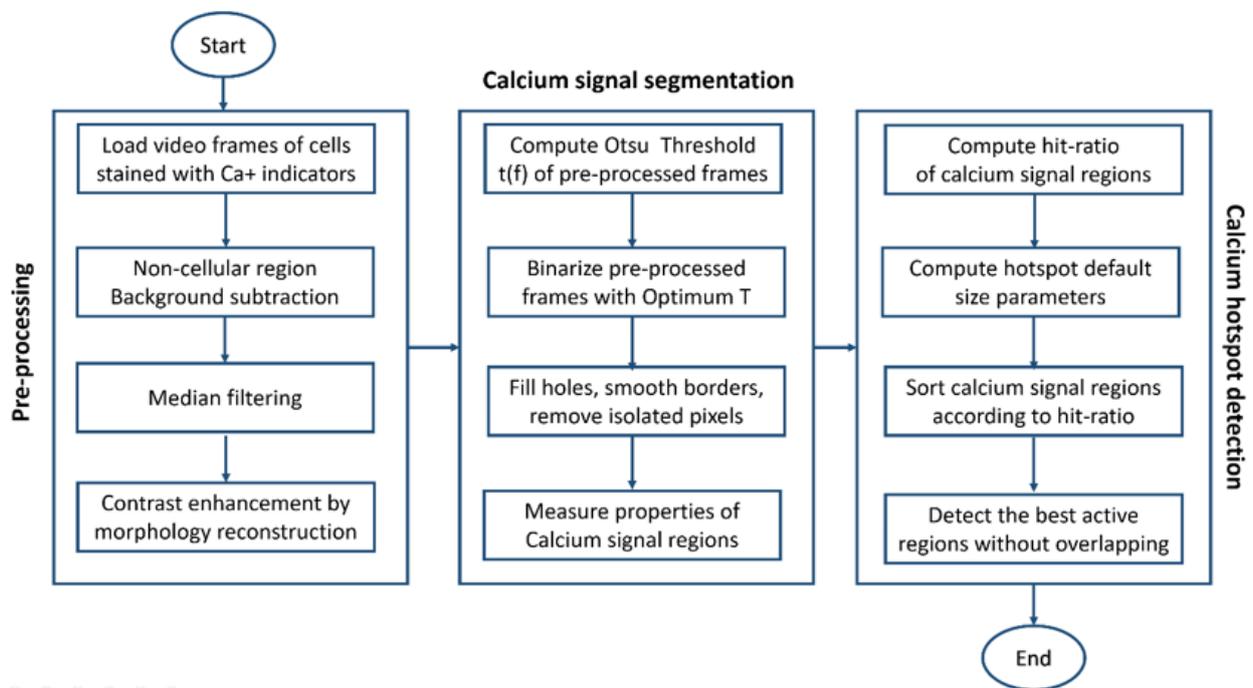


Figure 3.1: Framework for automated calcium hotspots detection prior to signal measurements.

Figure 3.1 describes our automated routine for ICH detection in fluorescence microscopy videos of cardiac myocytes based on a pre-processing stage followed by two separate segmentation

tasks. The pre-processing stage is made of four sub-routines: loading of individual video frames, non-cellular region background subtraction from frames, median filtering of background-free images, and contrast enhancement of the filtered video frames by morphological reconstruction. The calcium signal segmentation stage is also made of four sub-routines. Computing Otsu threshold from all individual pre-processed video frames is followed by the segmentation of all individual pre-processed frames by applying an optimum threshold T derived from the mean and standard deviation of individual frames Otsu thresholds. Then a post-processing routine is applied over the output of the optimum thresholding by using morphological operations such as hole filling, border smoothing, and isolated pixel removal of the binary objects representing calcium signal areas. Finally, properties of the binary calcium signal regions are measured for further analysis such as centroid locations, bounding boxes, pixel element points, size, ellipsoid fitted radius. As shown in the flowchart, the calcium detection stage which follows the calcium signal segmentation stage also consists of four main sub-tasks: the hit-ratio computation of all detected calcium signal regions, the calculation of hotspot size default parameters based on the mean and standard deviation of calcium signal region sizes, the sorting of calcium signal regions based on hit-ratio and the detection of the best active regions based their hit-ratio without overlapping hotspots selection.

In the past, many automated solutions for intracellular calcium imaging analysis have been proposed in proprietary software solutions, but few were free [45][71][46]. Most of them rely on intensity thresholding and enable the detection of local calcium signals, also called intracellular calcium sparks in time-lapse recordings using confocal microscopy in skeletal and cardiac myocytes. Those techniques are based on the general assumption that connected components whose normalized pixel intensities are above a specific threshold value represent calcium spark or wave expressions. This value is roughly equalled to the sum of average pixel intensity of the whole image + n times its standard deviation[71][45][72]. A popular algorithm using this strategy has been implemented in Java as an Image J plugin called Spark master with a graphical user interface [73].

Others methods rely on Fourier transforms of temporal pixel value fluctuation of cellular regions for the detection of global intracellular calcium responses of non-excitabile immune cells such as Jurka T-cells [46]. CALQUO, the standalone MATLAB application with graphical user interface described in [46] for global calcium response detection in individual cells, considers a discriminated radius of value 0.34 as a threshold for classifying Fourier spectrum of cell

fluorescence into global calcium oscillation or not.

Here, we adopted a strategy similar to [71] [72]. Still, instead of relying only on the mean and standard deviation of individual frames pixel intensity to detect calcium signal areas at a given time, we first considered Otsu threshold values of each time frame as a starting point. Then, the optimal threshold value for detecting the calcium signal in the whole time-series of images is derived from the average and standard deviation of all Otsu thresholds. In such a way, our algorithm takes into consideration illumination variation across consecutive frames and keep detected calcium signals equivalent and not necessarily of one type, i.e. only local or only global. In certain biological essays, it is essential to distinguish whether calcium signals are spread locally inside individual cells or globally across multiple cells, and to analyse them separately [2], but that was beyond the scope of this research because of a lack of enough ground truth images stained both with calcium and cell membrane indicators. In following Section 3.2, we describe our automated routine for ICH detection in fluorescence microscopy videos of cardiac myocytes.

3.2 Calcium Hotspot Detection

As stated in Section 3.1, hotspots correspond to dynamic intracellular active regions where calcium concentrations are likely to vary the most with time. In biological terms, they represent cellular spaces where calcium release events happen the most. This may be caused by the presence of specific calcium channels or of organelles acting as calcium stores, which can release calcium after stimulation [14]. When the cells' calcium activities are monitored using calcium-sensitive fluorescent indicators, such as Fluo-4 or Oregon-green BAPTA, calcium release events can be visualised as bright spots with dynamic sizes and located on the photo-chemical principle that the fluorescence increases with an increase in the calcium concentration.

So, knowing that changes of the intracellular calcium concentration are captured in the form of increased brightness with irregular shapes popping up randomly or regularly in various or same cellular regions, the identification of the hotspots in cytoplasmic calcium fluctuation can be directly denoted from the frequency of bright areas observed in video frames patches. In other terms, the higher is the frequency of the bright regions within a cellular region, the more active the cellular region in showing changes of intracellular calcium concentration. Therefore, from the segmentation of the calcium signal areas described in next Subsection 3.2.1 we derive the

location of the most active regions in showing changes of calcium concentration inside cardiac myocytes based on the frequency of calcium signals within regions, i.e. their hit-ratio.

We defined as hit-ratio the percentage metric that tells how often calcium fluctuations happen inside sub-cellular locations. In that sense, by default, any cellular regions showing at least one calcium change are potentially calcium hotspot locations. Thus, the next question to be answered might be how to automatically determine the hit-ratio of intracellular regions based on the fluorescence fluctuation associated with them. To do so, we broke down the calcium hotspot detection procedure into two phases: first, is the segmentation of bright areas in time-lapse microscopic images corresponding to the presence of calcium signals and second, is the filtering of the most active regions based on the calcium signal hit-ratio metric.

3.2.1 Identifying Candidate Calcium Hotspots by Multi-level Thresholding

In this study, we refer to calcium signal areas, any presence of a change in intracellular calcium concentration captured in fluorescent video frames produced at a usual rate of 30 fps with a low-resolution camera under Oregon Green Bapta or Fluo4 loading. To read more about the protocol used for calcium imaging in cardiac myocytes, please refer to [11].

To detect intracellular calcium hotspots in fluorescent video frames, we built our strategy based on a simple but powerful image processing technique called intensity thresholding. Thresholding was the obvious solution to consider because in such fluorescent videos, calcium signal areas have a common characteristic: they should look brighter than everything else (Figure 3.2(B)). In other terms, when scanning individual video frames, white pixels are more likely to belong to calcium signal areas, whereas darker pixels are more likely to fall into image background. Figure 3.2 below illustrates our hypothesis that calcium signal regions show higher pixel intensity values in images of cardiac myocytes loaded with calcium indicator.

3.2 illustrates the fact that regions where calcium signals occur exhibit higher pixel fluorescence than non-active regions. In fact, 3.2 (B) represents fluorescence measurement or intensity profiles overtime of 2 different groups of image regions highlighted in Figure 3.2 (A). The image regions highlighted in Figure 3.2 (A) are numbered in white for areas showing calcium signals and red for background samples. In Figure 3.2 (B) the intensity profiles of the calcium signal

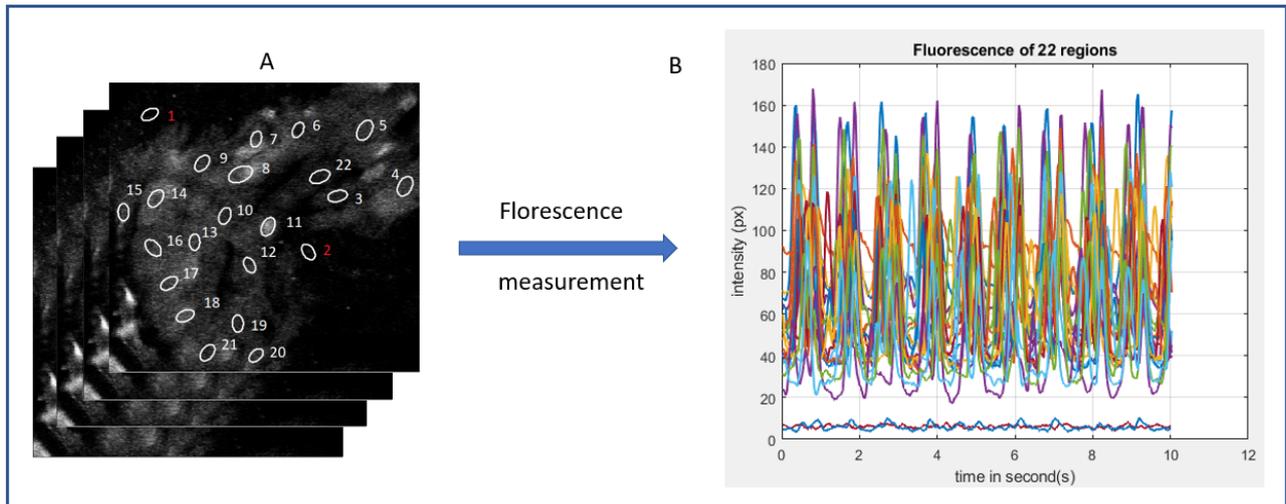


Figure 3.2: Calcium signal areas look like bright spots over general dark background. A: original input image sequence of PVC with twenty-two annotated regions from 1-22. First two labelled regions are outside the cell tissue and correspond to image background. Last 20 labelled regions are inside the cell tissue and correspond to calcium signal areas

areas correspond to the plots above 20 y-axis value. In contrast, the intensity profiles of the two background samples labelled (1,2) correspond to the plots below 20 y-axis value. Visually, it is evident that image regions showing calcium signals exhibit higher pixel intensity values than the areas without calcium change of concentration or signals. Therefore, in such case, it is obvious to conclude that a thresholding value around 20 can help distinguish calcium signal regions from the image background.

So, based on clear intensity differences between image patch samples with and without calcium signal areas, we adopted a strategy based on intensity thresholding to segment calcium signal areas. The next fundamental task to be solved was the search for the optimal threshold pixel value that can truly separate the ROIs, i.e. calcium signal areas. To do so, in the literature, many strategies have been developed. Some of them rely on global thresholds computed from grey-level histograms distributions of individual images like the Otsu method described in [62]. Others are based on local thresholds derived from image patches' statistics like in [65].

In this study, our strategy combines both local and global approaches in such a way that the best threshold value for separating calcium signal areas in fluorescent videos is assessed at two levels: locally and globally. Locally, by first considering grey-level pixel distribution of individual frames, and globally by considering the whole video file as a unique entity. Indeed, since our ROIs are changes of calcium concentration in cellular regions which can easily be perceived

when comparing consecutive video images of same stained cells, best thresholds should not be only dependant to individual frames pixel intensities but all the pixel fluctuations in the recording of the imaging experiment. In that way, calcium signal areas would stay proportional no matter their rising time in videos.

To validate this hypothesis, we compared the pixel intensity profiles of image regions with and without the presence of change of calcium concentration of different cardiac cell preparations. We found out that, at the frame level, the average pixel intensities inside areas with calcium signals are effectively greater than in background regions, i.e. in areas without calcium signals. The average brightness can be up to sixteen-fold above the average pixel intensities of non-active regions depending on the calcium signals amplitudes, the bit-depth of the camera, and the background brightness (Figure 3.2). This means that estimating calcium signal thresholds only at frame level would be hazardous because their fluorescence amplitudes vary with time, i.e. from frame to frame.

However, at the video level, we found out that the average intensity profiles of image regions with and without calcium signals demonstrated a constant pattern: an optimum threshold equalling to the double of the sum of the average and the standard deviation of Otsu thresholds computed for all individual frames can help separate calcium signal areas from image background in 95% of the videos according to the formula:

$$OptimumT = 2 * (averageOtsuT + stdOtsuT) \quad (3.1)$$

with $averageOtsuT$ = the average threshold value computed for all frames using Otsu method; and $stdOtsuT$ = the standard deviation of all threshold values calculated for all frames using Otsu method.

To sum up, the dynamics of intracellular calcium signalling in the cardiac myocytes used in this study are very intriguing because calcium signal areas may start as small dots, grow, become brighter, move around, and merge with each other at any time before disappearing over a heterogeneous background. However, intensity profiling of such bright areas shows that local thresholding may be an excellent alternative to localize local increases in calcium concentrations in the fluorescence time-series of images. The flowchart diagram below describes the different steps we followed for calcium signal areas segmentation by thresholding through fluorescence video frames.

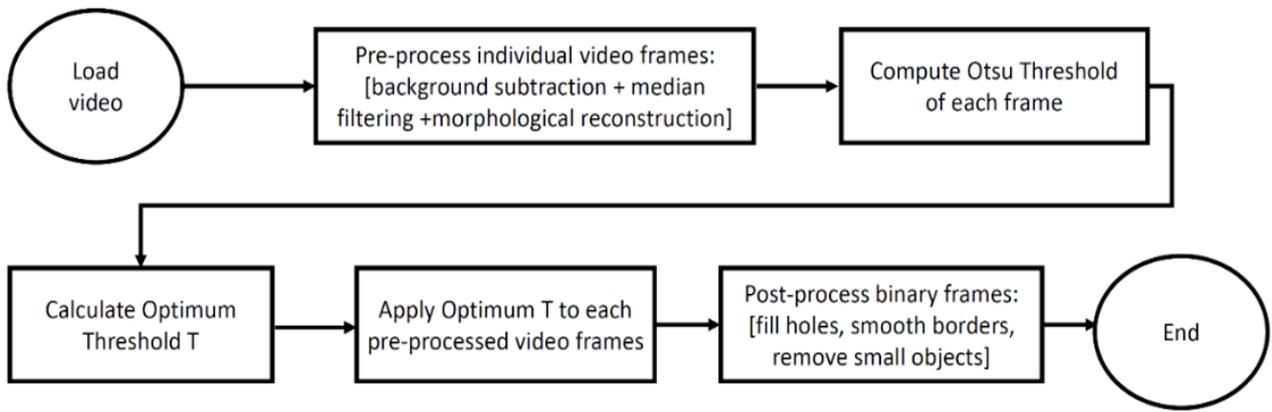


Figure 3.3: Flowchart diagram of calcium signal segmentation by multi-level thresholding.

First, a video or time-series image stack is loaded, and frames illuminations are normalized according to [92]. Then video background is roughly estimated and subtracted to obtain a background-free image. We defined as a background element, any pixel point that is not part of cellular regions. Further analysis demonstrated that the minimal fluorescence of cellular regions can be roughly deducted from frame pixel average and standard deviation. Each background free images was later smoothed by median filtering [63] to reduce Gaussian noises that usually corrupts fluorescent images.

After the smoothing operation, morphological reconstruction is applied over the resulting images for contrast enhancement. Morphological reconstruction is performed using mask and marker images of video frames. Mask images are derived from background-free frames whose contrast has been adjusted by adaptive histogram equalization [98]. Marker images are obtained by eroding the mask images.

After morphological reconstruction, Otsu thresholds of the resulting video frames are computed and recorded. Then, from those calculated frame-level thresholds, a single global threshold at video level is estimated according to the optimum threshold formula above (Equation 3.1). Figure 3.4 below highlights the results output by our automated routine for calcium signal segmentation at every stage over an example image of PVCs.

The image in Figure 3.4(A) shows an original video frame displayed in MATLAB 2019a environment. It represents a timely screenshot of the intracellular frequency of calcium signals produced by a set of PVCs stained with Oregon Green BAPTA. The actual image has been recorded by a digital camera 8 seconds after the beginning of the calcium imaging experiment. Six different Calcium signal areas in which the calcium concentration has increased have been

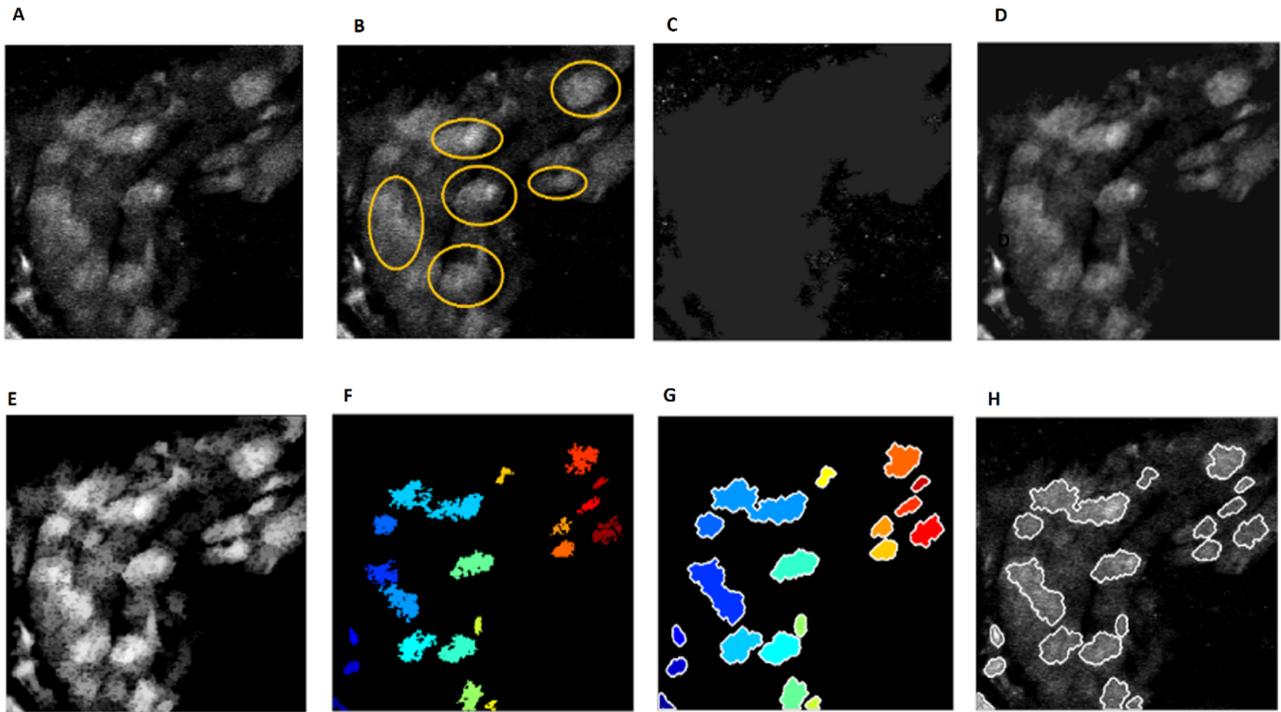


Figure 3.4: System output at each stage of calcium segmentation by thresholding. A: original input video frame of PVC. B: reference image with ROIs highlighted in yellow. C: non-cellular regions estimated as background frame. D: background subtracted frame. E: contrast-enhanced frame. F: binary frame obtained after multi-level thresholding. G: connected components enhanced by morphological operations. H: multi-level thresholding results overlaid on original input video frame. ROIs have been coloured from blue to red according to x-coordinates.

manually highlighted in yellow for visual purpose (Figure 3.4(B)).

As it can be denoted visually, bright spots correspond mostly to calcium signals, whereas dark pixels belong to background elements. When we feed such an image into our system, a rough background image is computed (as shown in Figure 3.4(C)) corresponding to non-cellular regions using $\text{mean} - \text{std}/2$ as a minimum fluorescence value for cellular pixel contents. Figure 3.4(D) shows the output of our system after median filtering of the background-free image of Figure 3.4(C). Visually it is clear that the resulting image in Figure 3.4(D) is smoother than the original input image (Figure 3.4(A)).

Figure 3.4(E) represents the output of our system after contrast enhancement application to the smooth image obtained in Figure 3.4(D) by using morphological reconstruction. It can also be noticed that visually high-intensity objects corresponding to calcium signals regions became more emphasized compared to ordinary cellular regions without significant calcium signals.

Figure 3.4(F) and (G) represent respectively the output of global optimal thresholding applied to Figure 3.4(E) and morphological operations applied to Figure 3.4(F). Finally, Figure 3.4(H) shows the final output of calcium signal areas detected by our system from the original input PVC image. The performance of the procedure for detecting calcium increases in individual frames is presented in the result section.

3.2.2 Filtering and Identifying the True Hotspots based on Hit-ratio Computation

As stated in Section 3.1, calcium signal hit-ratio is a percentage metric that tells how often certain intracellular domains show variations in their calcium concentration. The hit-ratio parameter is derived from the number of occurrences of calcium signal objects inside a frame location. Knowing that calcium signal objects are characterised by bright spots with varying sizes (that may grow or shrink over time) in recordings of imaging experiments, counting the number of bright objects inside frame regions is equivalent of counting the number of overlapping calcium objects inside same cellular areas through all video frames.

The actual locations of calcium signal areas are obtained by thresholding the video frames, the output of the calcium signal areas segmentation procedure. The resulting binary frames are used as input of the proposed method for calcium hotspot detection based on the hit-ratio computation of calcium signal events. The flowchart diagram below describes our proposed algorithm's steps for the final hotspots' detection through filtering of calcium signal locations without overlapping.

The flowchart diagram above shows that, by default, any cellular regions exhibiting calcium signal areas are potentially hotspot locations. Then, the system computes for each potential hotspot location identified in each frame, the number of overlaps with other calcium signal areas expressed in other consecutive frames. The hit-ratio is equalled to the number of overlaps divided by the total number of frames. Two binary objects are said overlapping when the separating distance between their closest borders is null or negative or when bounding boxes overlap according to the technique implemented in [86].

Finally, the system filters the best hotspot regions based on their hit-ratio, size, and separating distances. Default parameters for the minimum hit-ratio are set to 15%, the separating distance

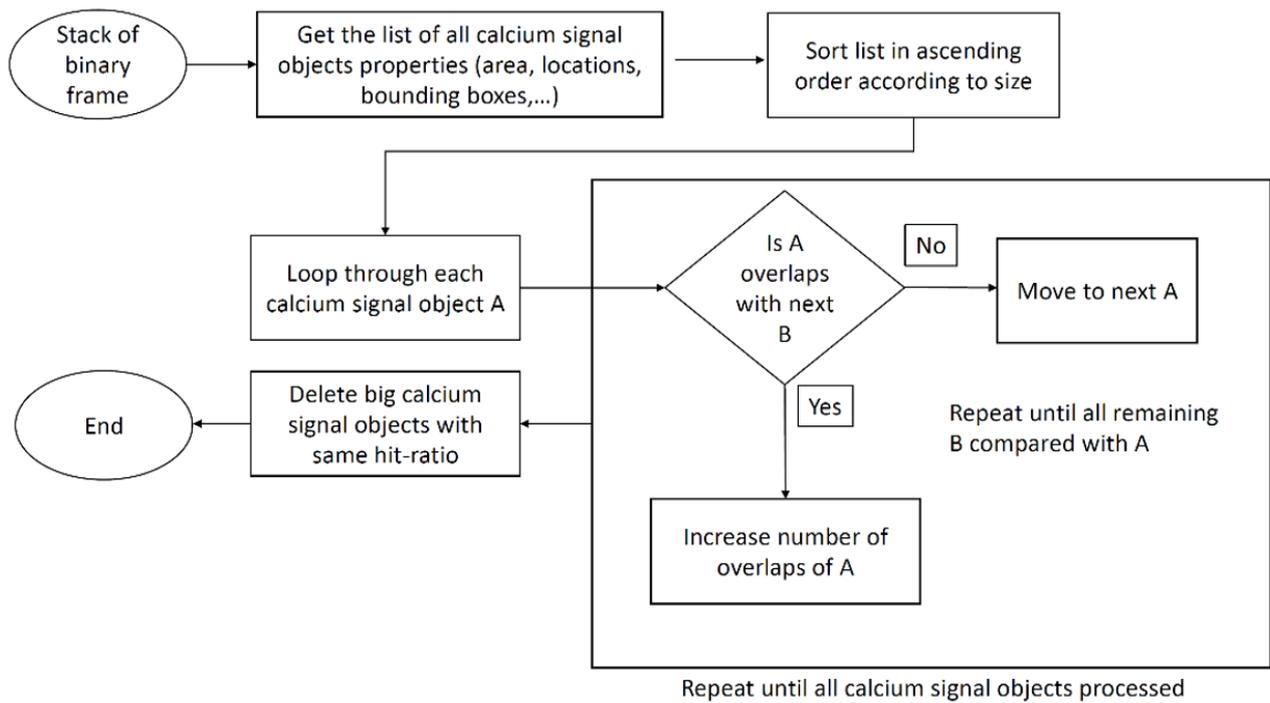


Figure 3.5: Flowchart diagram of calcium hotspot detection by hit-ratio filtering of calcium signal areas. Here, objects refer to areas.

to 5 pixels (the distance between the two borders of two separate calcium signals). In contrast, the default size is dynamically set according to the average size of calcium signal areas of the whole video file. The performance of the hotspot detection method is presented in Section 3.5.

3.3 Calcium Signal Measurements from Individual Hotspots

After the detection of hotspots, we proceed to the analysis of the calcium signal inside each of them. It consists of the temporal measurement of the change of calcium concentration inside the active regions detected as hotspots. The goal here is to provide useful measures that characterise any given active regions inside cardiac myocytes and enable biological comparisons between cells' calcium responses across different experiments.

- Initial readout of calcium signals of individual hotspots The first measurements our algorithm extracts for the detected hotspots are their average greyscale pixel intensity values. Because in each frame time, we may have different values of pixel intensity per hotspot, the plotting of those initial traces is a direct first indication of cells responses to calcium

events.

- **Non-cell tissue background fluorescence** The second measurement that can be useful for biologists to report is the estimation of the average background fluorescence inside a given frame, i.e. the non-cellular region fluorescence. Non-cell tissue regions fluorescence is a rough estimation of pixels which do not belong to any cellular areas. It is useful to make that distinction because the actual fluorescence of any calcium signal needs to be differentiated from the overall fluorescence of the cells. Non-cell tissue image corresponds to cell tissue image fluorescence subtracted from the original input image. For the cells used in this study, we found that We found that any pixel below the mean $- \text{std}/2$ falls to the background fluorescence in each frame.
- **Corrected fluorescence F** By correcting fluorescence F of a given hotspot, we intend to obtain a more accurate estimation of the value of calcium changes being released by cells. The fixed fluorescence value (f) is calculated by subtracting from the initial readouts of individual hotspot traces, the background value of the corresponding frame where the calcium signal has been detected.

$$f_i = m_i - \text{background}, \quad (3.2)$$

where $m_i = \text{initialreadouts}$

- **Change of fluorescence F/F_m** The change of fluorescence of a hotspot at a given frame time corresponds to the ratio of its corrected fluorescence at the given frame time over a particular minimum fluorescence inside the hotspot. The minimum fluorescence itself is a constant value corresponding to the pixel intensity level below which there are no events of calcium signalling captured in the hotspot. Since we assume that frame backgrounds are snapshots of calcium signalling experiments without the actual calcium expressions, averaging all backgrounds of the entire video should give us a close indication of the minimum fluorescence value if we minimize the effect of photobleaching on the whole video. So, to get the minimum fluorescence, we take the mean of the estimated frame background previously calculated. The reason for the change of fluorescence estimation is to enable reliable comparison of hotspots between different videos. When plotting F/F_m , a straight line of value 1 acts like a baseline for the change of fluorescence.
- **Statistical measurements** From the change of fluorescence, we extracted statistical measurements such as locations of the peak of the calcium signal, their amplitudes and frequencies, the area under the curve, inter-spike intervals, time of slopes. The peak is

defined to be above a minimum threshold of 20% of the maximum signal amplitude recorded per hotspot. The inter-spike intervals value corresponds to the elapsed time between two consecutive calcium signal peaks. The time of slopes refers to the elapsed time between a peak point and the time when the hotspot signals get back to the normal baseline. All the measurements output by our system are illustrated in Figure 3.6.

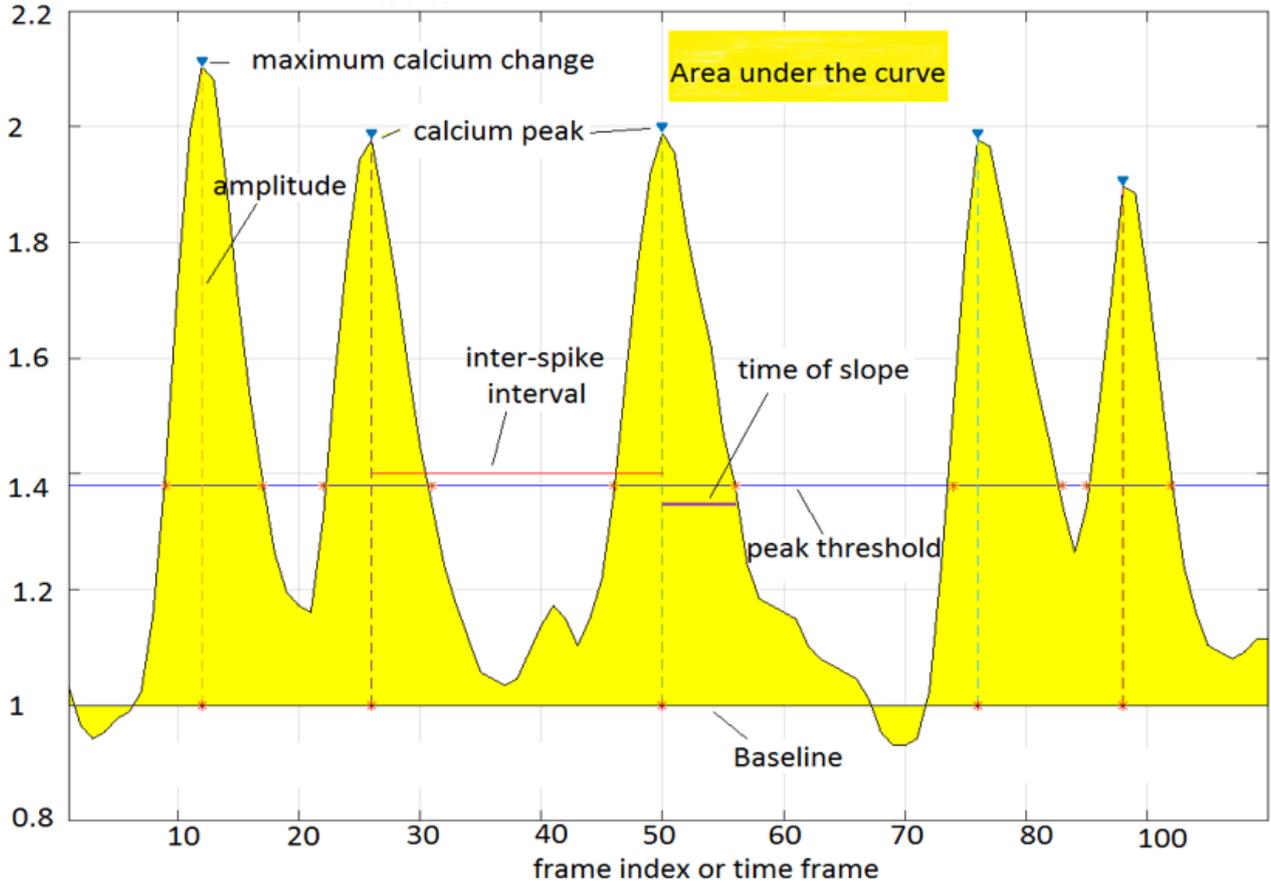


Figure 3.6: Calcium signal measurements output by our proposed algorithm.

3.4 Materials & Experimental Protocols

Since the proposed solution for automated calcium hotspot detection and measurement in FM videos consists of inferring calcium hotspots or active cellular regions in calcium change of concentrations from the frequency of calcium signal events, materials and experimental protocols were designed accordingly.

In terms of materials, two ground truth datasets were developed. The first ground truth dataset describes calcium signal events or areas present in individual video frames of PVC and NRVM

cells' preparations. It enables to evaluate and investigate the performances of the proposed multi-level thresholding algorithm for calcium signals segmentation presented in Section 3.2.1 against the Cheng's pioneering work published here [21]. The second ground truth dataset also developed over FM videos of PVC and NRVM describes true location of calcium hotspots in cells preparations after scanning all videos. It allows performance measure and the validation of the ultimate calcium hotspot detection subroutine presented in Section 3.2.2. The following subsections give an insight on the labelling procedure performed to build the different calcium signal ground truth datasets and the experimental protocols carried to assess the performance of the different techniques investigated.

3.4.1 Calcium Signal Ground Truth Datasets

As highlighted in Section 3.1, various signalling pathways and spatio-temporal patterns make intracellular calcium signals difficult to shape. In cells preparations loaded with calcium indicators, they may look like dot points or calcium sparks i.e., localized calcium signals with little spread within a cell, then gradually grow, and move to calcium waves before shrinking and disappearing as clouds or smokes in cellular domains.

Therefore, to develop a ground truth dataset of such calcium signals, a hand-free labelling procedure was performed under the supervision of three calcium expert biologists in a user-friendly and interactive application coded in MATLAB. After loading calcium imaging videos in the system, the experts were asked to highlight in a frame basis, cellular areas showing calcium signals or intracellular calcium change of concentration as the smallest ellipsoids containing them. Then the pixel locations of the different ellipsoidal labelled areas were saved as ground truth of the corresponding intracellular calcium signals manually marked. For each video analysed, 50 consecutive frames were hand-free labelled by experts. This makes a final ground truth calcium signal datasets of 1000 labelled frames derived 20 different calcium imaging videos that was used in this analysis. Figure 3.7 illustrates the hand-free labelling procedure conducted under biologists' supervision where cellular regions showing calcium signal events are highlighted in ellipsoids.

In Figure 3.7, 25 cellular regions highlighted in yellow show the locations of calcium signals recorded at the 30th frame of a video of PVC. These locations were identified by calcium signalling experts. The procedure is applied on all 1000 frames.

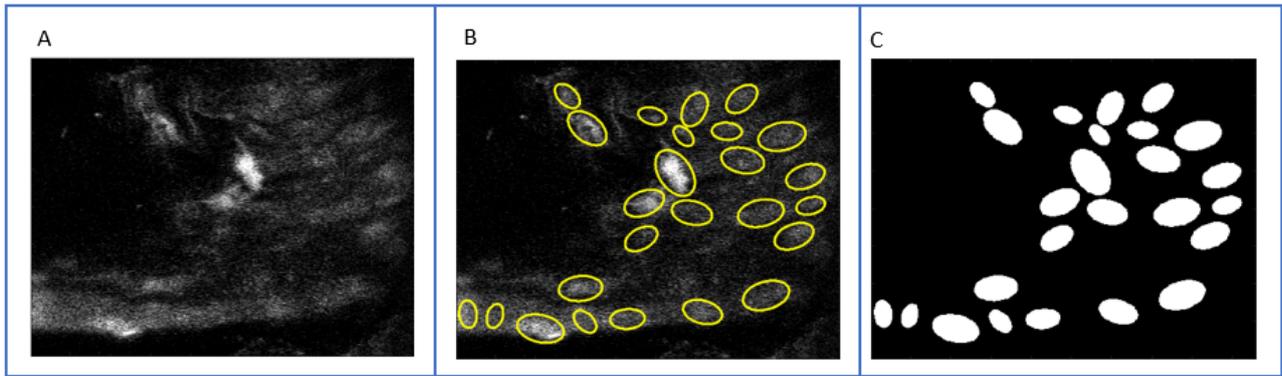


Figure 3.7: Ground truth annotation for calcium signal segmentation. A represents original video frame showing PVC loaded with calcium indicators. B illustrates hand-free ellipsoidal labels of cellular regions showing calcium signals i.e., change of intracellular calcium concentration. C represents the binary mask of the ellipsoidal labels saved as the ground truth of calcium signals areas or events in A

Since calcium hotspots correspond to the locations of the cellular regions that exhibit the highest frequency of intracellular calcium signals over time, the same hand-free labelling approach was adopted as for the signals themselves. After loading a calcium imaging video into a GUI app designed in MATLAB, 3 expert biologists were asked to play the video then to manually labelled by ellipsoids the most active regions in calcium release events during the entire video recording. In such a way, for each video analysed, they drew in average 20 hand-free hotspots locations inside a single reference image. Thus from 20 labelled videos in total, 20 different hotspot maps were obtained. The final ground truth for calcium hotspots was made of 20 hotspot image masks as illustrated in Figure 3.8.

Next subsections of this chapter summarize the experimental protocols performed for investigating automated solutions to segment calcium signal in individual video frames and detect calcium hotspots in videos.

3.4.2 Experimental Protocols

According to the two main steps of the proposed solution for automated detection and measurement of ICH, two types of experiments are performed. First calcium signals are segmented from FM video frames by using the Cheng's method [21] and the multi-level thresholding introduced in Section 3.2.1 . Then, ultimate calcium hotspots are detected from the video fusion technique

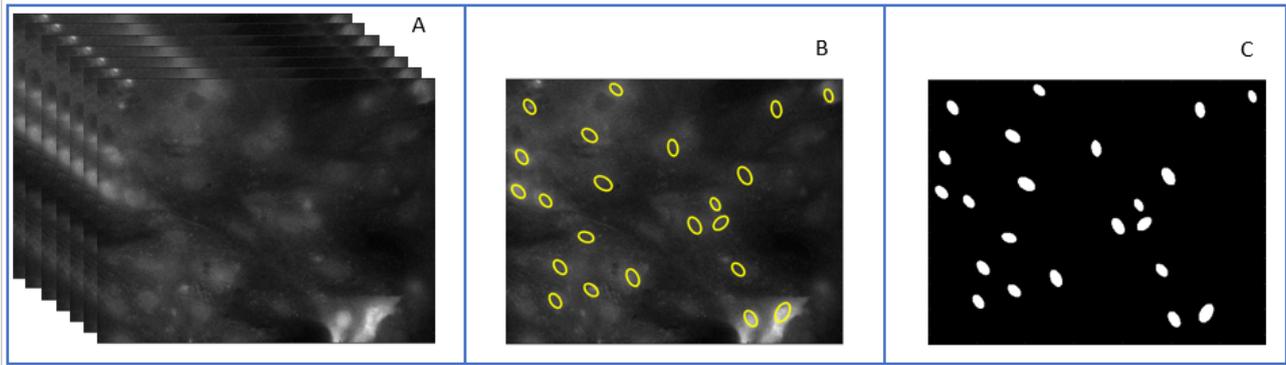


Figure 3.8: Ground truth annotation for calcium hotspot detection. A represents an original FM video consisting of time-series images of cells loaded with calcium indicators. B shows the manual annotation made by calcium signal experts to highlight the 22 most active regions in the video. C represents the binary mask of the 22 hotspot locations across the entire video saved as the ground truth of the video in A.

based on hit-ratio filtering of cellular regions introduced in Section 3.2.2 . The calcium imaging videos used were captured by Zeiss cameras with fluorescence confocal microscopy setup to record videos at an average frame rate of 30 fps. Each video frame exhibits a bit-depth varying from video to video between 8 to 16 bits.

To evaluate the performance of both types of experiments, outputs of the investigated algorithms are compared against their corresponding ground truth in terms of precision metric. In classification problems, the precision metric is defined as the proportion of relevant instances among the retrieved ones. Here, it corresponds to the fraction of calcium signal area or hotspot's locations that are detected by the system and which overlap with their corresponding ground truth set up manually by biologist experts. The formula of the precision metric is the following:

$$P = \frac{TP}{TP + FP} \quad (3.3)$$

where P = precision, TP = true positive, FP = false positive.

The TP calcium signal areas in video frames and the TP calcium hotspots respectively correspond to the calcium signal areas and hotspots output by automated systems and which overlap with the ground truth datasets. Thus, the FP calcium signal areas and hotspots in images and videos correspond to those ones that do not overlap with the ground truth. Two cellular regions are said overlapping when they share common pixel points (x, y) . In Section

3.5 the results and discussions on the findings of the experiments carried are reported.

3.5 Results and Discussion

According to the types of experiments conducted, the results and discussions on the findings are classified into two groups: the ones related to calcium signal area segmentation and those related to hotspot detection experiments.

3.5.1 Calcium Signal Area or Event Segmentation in Video Frames

For calcium signal area segmentation in individual frames, the proposed multi-level thresholding i.e., an optimum threshold value estimated both at frame and video levels achieved a mean average precision of 65.5% over the entire labelled calcium signal dataset (Table 3.1). However, it is worth mentioning that these performances are relative to the number of ground truth locations experts limited themselves to per FM video frames labelled. Figure 3.9 illustrates the result of the multi-level thresholding method over an example PVC video frame where ROIs or targeted calcium signal areas are highlighted in yellow (Figure 3.9 (A)), calcium signal areas segmented by thresholding are represented by white binary connected components (Figure 3.9 (B)), and the system output of calcium signal areas in red (Figure 3.9 (C)).

As it can be noticed in Figure 3.9. (C), over 30 predictions, 28 calcium signal areas truly match with ground truth labels i.e., 0.93 of precision. Table 3.1 lists the average results obtained by multi-level thresholding 1000 FM video frames of cardiac myocytes. Next Table 3.2 reports the average results obtained by applying the Cheng's method over the same datasets.

Intracellular calcium dynamics of the cardiac myocytes used in this study are very intriguing across different experimental conditions. Still, our results shows that applying an optimum threshold derived from local Otsu thresholds computed for every frame and their standard deviation enables the segmentation of areas showing change of calcium concentration with high precision than thresholds derived only from Cheng's method based on the mean plus two time the standard deviation of individual frames as suggested in [21] and [86]. Table 3.2 presents the results obtained from Cheng method.

Despite the various spatio-temporal patterns that characterize intracellular calcium signals of

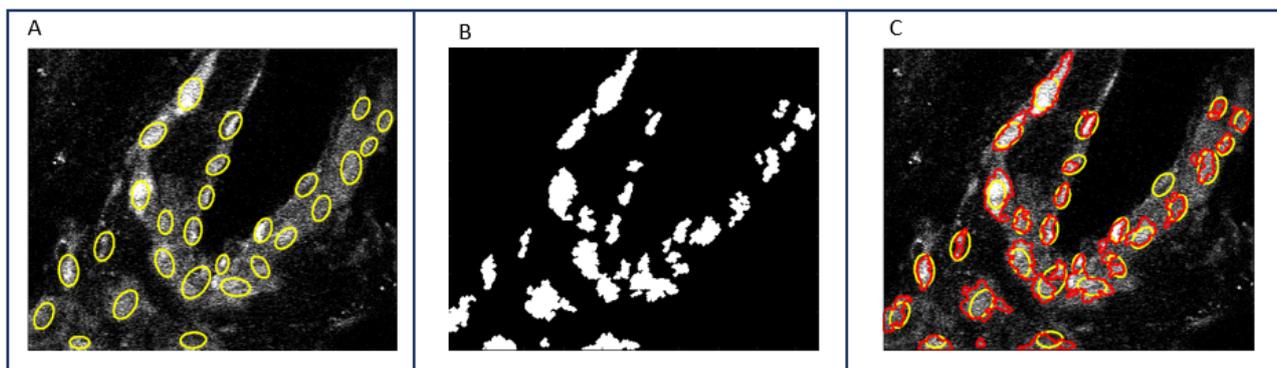


Figure 3.9: Results of calcium signal segmentation in FM images. A represents an original PVC image with 27 ground truth calcium signal areas highlighted in yellow ellipsoids. B represents the segmentation result obtained by multilevel thresholding of A. B shows 30 connected components that indicate 30 candidate calcium signal locations. C is the comparison between A and B i.e., between the segmentation results highlighted in red and the ground truth labels highlighted in yellow

cardiac myocytes, these results show that an optimum threshold derived at two levels and using mean and standard deviation of Otsu thresholds enables a better precision (65%) than the Cheng’s method that relies on thresholds computed solely at frame level (53% precision).The difference could be explained by the fact that Cheng’s approach assumes that calcium signal areas correspond to outlier pixels whose intensities are above a constant threshold equalling to the mean + 2 times the standard deviation of individual cell video frames. This assumption implies that any FM video frames showing cells loaded with calcium indicators there should be calcium signal events no matter the cell types and states except for extreme scenario of fully black images or fully bright images.

This assumption cannot stand in any cases because intracellular calcium signalling is a dynamic procedure that may show some pauses, increases, or decreases over time depending on the cells or calcium channels being activated or not. Therefore, it is not always right to expect calcium signals while monitoring life cells stained with calcium indicators every moment of the imaging experiment. In other terms, bright pixels above a certain threshold should not be considered as part of calcium signal areas in individual images without comparing with other consecutive images.

Indeed, the changes in calcium concentration can only be denoted when spotting fluorescence variation between consecutive frames or in cellular areas at different time points. This is the

video	AP	video	AP
1	61	11	50
2	74	12	63
3	83	13	61
4	57	14	63
5	79	15	79
6	74	16	56
7	50	17	80
8	70	18	64
9	59	19	49
10	86	20	52
Mean	65.5		
STD	11.83		

Table 3.1: Average result for calcium signal detection per video. Column video lists indexes of the 20 different videos used to build the ground truth datasets and to test our proposed solution for calcium signal segmentation by multi-level thresholding. Column AP lists the average precision recorded per set of 50 video frames analysed.

fundamental difference between our method and Cheng’s strategy for calcium signal detection in FM videos i.e., the search for calcium signal locations through all video frames as a whole set instead of individual frames as independent entities.

Although the two-level thresholding advantages, it shows two major drawbacks. First it wrongly identifies some noisy illumination artefacts common to FM data as calcium signal areas. Second the system fails to handle FM videos with high illumination variances across consecutive frames. The noisy artefacts generally correspond to persistent bright spots or areas which are present across consecutive frames but with high pixel intensities. To enable the system to identify such false friends, we introduced a post-processing step which filters those persistent bright areas before the ultimate hotspot detection.

The system failure to process video frames with varying fluorescence spectrum may be explained by severe photobleaching, change in camera settings during experiment recordings, or significant variances in intercellular calcium waves of the cells. To overcome such problems, illumination normalization techniques of video frames , multiple optimum thresholds estimation for a group

video	AP	video	AP
1	71	11	74
2	53	12	33
3	30	13	40
4	44	14	26
5	63	15	60
6	33	16	68
7	72	17	58
8	37	18	50
9	85	19	65
10	39	20	65
Mean	53.3		
STD	17.17		

Table 3.2: Results for calcium signal detection using Mean + 2std at the frame level. Column video lists indexes of the 20 different videos used to build the ground truth datasets and to test the Cheng’s strategy for calcium signal segmentation. Column AP lists the average precision recorded per set of 50 video frames analysed.

of frames, and photobleaching correction could be investigated. This is done in Chapter 4 for illumination normalization and photobleaching correction . In next Section 3.5.2, results, and discussion on the findings from experiments of filtering cellular areas into calcium hotspots are reported.

3.5.2 Ultimate Calcium Hotspot Detection in FM Videos

From the calcium signal areas output by multi-level thresholding, the system performs ultimate calcium hotspot detection by filtering the intracellular regions showing calcium signals based on the hit-ratio metric associated with two other parameters such as region size, and minimum distance separating them. The default size parameter is a size interval based on the average size of calcium signal areas. The default hi-ratio is set between [1-100%]. In contrast, the minimum distance separating two hotspots is set to two pixels. This approach enables to develop a novelty technique for automated calcium signal analysis. In fact, to the best of our knowledge, there was no previous fully automated framework available for intracellular calcium hotspots

detection through FM videos of cardiac myocytes.

The proposed framework is unique and enables to achieve ICH detection with an overall accuracy of 80% over twenty testing video datasets by using default parameters of size, hit-ratio and minimum distance. Table 3.3 summarizes the results we obtained over expert manually labelled ground truth.

video frames	AP	running time (s)	video frames	AP	running time (s)
900	71	359.51	390	97	112.43
301	68	67.10	900	82	427.33
1799	64	860.19	390	98	105.74
900	85	571.66	206	62	225.44
900	69	403.35	206	83	231.15
900	72	387.13	206	85	243.27
900	79	552.40	389	64	103.58
900	73	418.63	350	90	72.88
300	91	85.04	441	93	83.91
899	78	224.11	441	89	74.25
Mean AP	STD AP		Mean Running Time		
79.65	11.34		280.45 seconds 4.67 minutes		

Table 3.3: Result for hotspots detection per video. Column video frames shows the number of video frames included in each video processed. AP stands for average precision of hotspots detected by the system for the video described in the left-hand side column. Column running time shows the time elapsed by the system to process the video in second.

Further investigation on the effect of the three parameters onto our automated hotspot detection algorithm shows strong correlations between regions size, separating distance, the system performance, and the total number of hotspots retrieved. Large region size parameters tend to produce few hotspot locations with big hit-ratio. But those large hotspots are not significant for biologists as they may be covering multiple cells or multiple active regions inside different cells or same cells. Significant separating distance parameters also tend to reduce the number of hotspots detected but do not affect the overall system performance if smaller than 5 pixels. Besides, increasing the minimum distance between ROIs reasonably enables to avoid redundancy, i.e., picking up multiple hotspots inside the same cells, which is a matter of concerns for biologists.

Indeed, biologists are seeking to identify the most active cells responsible for some physiological patterns involving calcium signalling without being distracted by same cells hotspots calcium measurements. This is valid in cases where they want to know if a treatment increases or decreases the number of active regions in a cell or if it affects the activity of certain or all cellular regions, for instance. So, in an ideal world, the system should enable full detection of the best active hotspots inside distinct cells without overlapping. However, a calcium hotspot detection technique considering cell boundaries was difficult to investigate in this analysis as cell membranes are not stained nor visible in the FM datasets used.

The three parameters of desired hotspot size, minimum separating distance and hit ratio that have been introduced in this analysis for automated hotspot detection in FM videos by connected components filtering enables rich, consistent, and unbiased selection of active regions compared to the manual procedure described in Section 3.1. Figure 3.10 illustrates this strength of the proposed solution where from scanning 900 video frames, the system picks up only 30 hotspots from a total of 23065 calcium signal areas previously identified by multi-level thresholding of the video frames. The automated procedure enables speed, accuracy, and replicability for the following reasons:

- It takes only 280 seconds i.e., 4 minutes 40 seconds in average for the system to process the datasets against 30 minutes in average by biologists when performing a manual procedure on a video made of 300 frames.
- The selection of hotspots or active regions is more consistent and less human-biased nor subjective as it relies on a hit-ratio that tells how often the concentration of intracellular calcium varies within cells. The accuracy of the automated system is derived from automated procedure with a precision of 79.65% that will stay consistent despite human factors such as fatigue, and errors.
- Results are replicable because the same parameters of size, hit-ratio and separating distance will always retrieve the same hotspots over the same cell preparations no matter how many times the analysis is repeated whether by the same biologists or not. Replicability of hotspot measurements is essential for scientists because it enables them to observe cells' behaviour at different times and under different circumstances. For instance, to understand the effect of lysosomal calcium signals in NRVM cells undergoing autophagy, same cells preparations are loaded with calcium indicators and monitored at different days (1,

8, 10). During these experiments, cells can be excited with caffeine or not or electrically field stimulated.

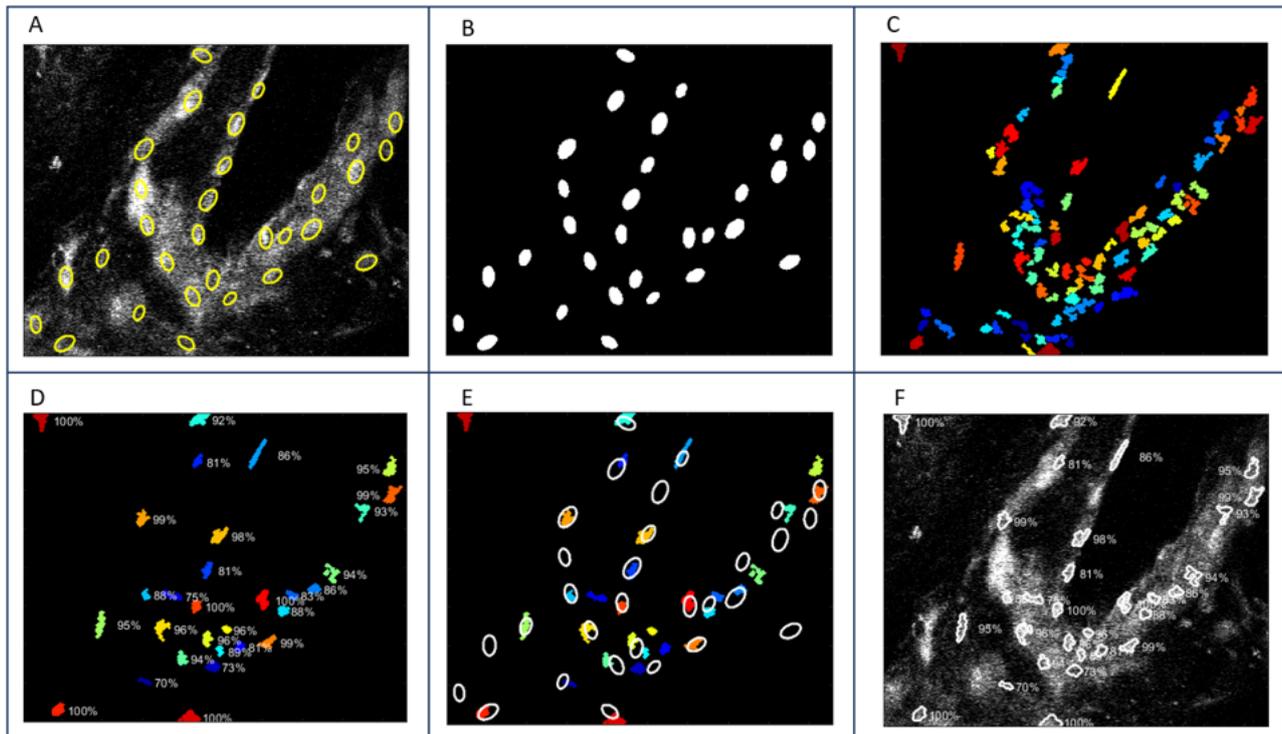


Figure 3.10: Automated hotspot detection based on hit-ratio, size, and distance. A: reference image showing 30 ground truth hotspot locations annotated by experts for an input video of PVC. B: binary map of the ground truth hotspots in in A. C: 89 initial connected components output by the system as potential non overlapping hotspots after scanning 900 video frames of the input video and checking 23065 different calcium signals. Detection parameters based on size between $[100, 300]$, hit-ratio between $[1,100]\%$, and minimum distance = 1. D :30 final hotspots output by the system after applying a default minimum distance parameter of 2. The connected components are coloured according to their hit-ratio values. The colour map ranges from blue to red. The hit ratio values of the hotspots are displayed next to them. E: illustration of the overall precision of the hotspot detection method that is computed by comparing the results of the system against the ground truth hotspot location in white ellipsoids. F: Results of the automated hotspot detection method overlaid on a frame of the input image.

3.6 Summary on Calcium Hotspot Detection

To sum up, our automated routine for hotspots detection and measurements in FM video has shown that optimum thresholding derived from individual frames pixel histograms and calcium

fluctuation hit-ratio in cellular regions outperformed standard routines developed by Cheng in [21] and an early automated technique we developed in [86] with more than 26.35% accuracy margin. Our new framework, which relied on a two-level segmentation algorithm enables fast, accurate and reproducible analysis of intracellular calcium signals within certain types of cardiac myocytes loaded with fluorescence indicators.

Automated algorithms for intracellular calcium signal analysis are essential issues for future research due to their biological relevance. Our proposed technique has contributed to a rapid, accurate and better understanding of clinical research data performed by biologists about cardiac arrhythmia through the analysis of intracellular hotspot calcium measurements.

However, when analysing the calcium change of concentration curve over time for measurements extraction inside detected hotspots, we noticed that some signals exhibit photobleaching effects (progressive fading of the fluorescence signal over time). In quantitative microscopy, it is believed to take into consideration photobleaching effects because they can reduce the precision of the derived measurements, and therefore lead to false conclusions [79]. Photobleaching detection and correction are discussed in the next chapter.

Chapter 4

Calcium Hotspot Signals Processing in Cardiac Myocytes

In quantitative microscopy, photobleaching is a degenerative photo-chemical phenomenon that severely alters measurements derived from raw pixel data such as intracellular calcium signal traces in cardiac myocytes. In this chapter, we proposed automated routines that guarantee data integrity in terms of calcium signals measurement extracted from intracellular calcium hotspots despite eventual photobleaching decay that may corrupt measurements. Therefore, the main contributions of this chapter are:

1. Extracting and normalising intracellular calcium signals in cellular hotspots
2. Detecting photobleaching in calcium signals of cardiac myocytes and demonstrating that it can be modelled as a mono-exponential function of time and corrected by baseline adjustment according to the exponential decay initially estimated by non-linear least square regression.
3. Correcting calcium signals affected by photobleaching and validating the results by measuring the goodness of the recovered signals against the expected flat baselines.

Structurally, the chapter is made of 6 sections. Section 4.1 briefly presents the goal of calcium hotspot signal processing in terms of signal extraction and normalisation, and photobleaching handling. In Section 4.2, an automated framework for extracting and normalising intracellular calcium signals in calcium hotspots is described. Section 4.3 presents our investigations and

proposed solutions for photobleaching detection and correction in calcium signals of cardiac myocytes' hotspots. This framework is based on exponential curve fitting and signal baseline adjustment. In Section 4.4, the materials and experimental protocols used to evaluate the proposed techniques for both photobleaching detection and correction are described and the results and discussion on the findings are reported in Section 4.5. Finally, the chapter ends by concluding remarks in Section 4.6.

4.1 Introduction

Calcium hotspot signal processing in cardiac myocytes refers to a set of measurement tasks performed by biologists after identifying calcium hotspots inside cellular regions from fluorescence microscopy (FM) videos. These tasks are; therefore, an extensive follow up of the automated hotspot detection technique discussed in Chapter 3, i.e. the signal processing behind the measurement of calcium change of concentration inside ROIs.

Our goal here is to make sure that calcium traces captured for individual hotspots are accurate and can be used to compare calcium signalling in different cells or in the same cells observed in different circumstances, regardless of experimental differences and photobleaching loss.

4.2 Hotspot Calcium Signal Extraction

Calcium signal extraction from hotspots (see Section 3.3 for detailed description) corresponds to the measurement of calcium concentration changes inside hotspot regions over time. Technically it derives from the average fluorescence measured in consecutive video frames inside cellular areas labelled as hotspots. To do so, for a video of N frames, we measure as hotspot calcium traces N average greyscale pixel values of their corresponding location through each frame. The flowchart diagram below schematizes our automated routine for hotspots calcium signal extraction if hotspots regions are detected:

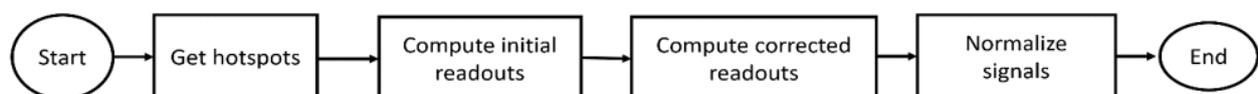


Figure 4.1: Flowchart diagram of calcium signal extraction.

Figure 4.1 lists three main tasks: initial calcium readout computation, corrected readouts computation and definitive calcium signals normalization. The initial readouts computation corresponds to averaging the raw pixels values of hotspots regions through video frames. The corrected readouts computation corresponds to the first correction applied to the initial readout by subtracting the non-cellular region fluorescence. The reason for such operation is to make sure that the measurements extracted reflect essentially the fluorescence attributed to calcium change of concentration rather than something else (noisy artefacts) in video frames. The definitive calcium signal normalization corresponds to a ratio of the corrected fluorescence over an estimated minimum fluorescence for every hotspot region identified.

Figure 4.2 illustrates the importance of normalized readouts against initial ones for comparison between hotspots detected from different cardiac myocytes observed in various experimental conditions.

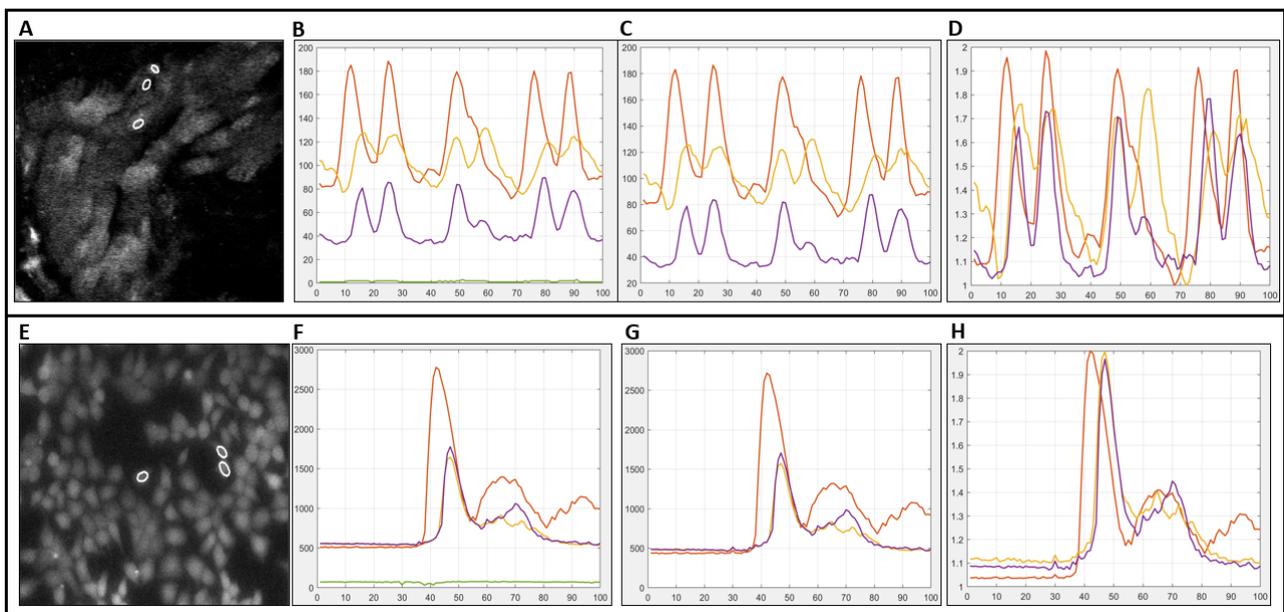


Figure 4.2: Hotspot calcium signal plots. A and E: screenshot of original input video frame with three annotated hotspot locations. B and F: initial readouts of three hotspots from average raw pixel fluorescence over time. C and G: corrected fluorescence of the three hotspots. D and H: change of fluorescence measurements of the three hotspots

Figure 4.2(A) and Figure 4.2(E) represent images of two preparations of cardiac myocytes where three hotspot locations have been highlighted in white ellipses. For the first 100 frames, the initial readouts of the hotspots from the different cells are plotted in Figure 4.2(B) and Figure 4.2(F) with the green curves representing the non-cellular region fluorescence, i.e. the background fluorescence. The amplitudes of the initial readouts range from [30,200] for cells

(A) whereas the amplitudes of the initial readouts for cells (E) range from [500, 3000]. The variation of the ranges is due to differences in experimental conditions, i.e., camera image resolution (8 bits for the first one, and 12 bits for the second).

Therefore, strong conclusions from the initial readouts of the frequency of calcium signals of those two types of cells might be misleading. Figure 4.2(C) and Figure 4.2(G) represent the corrected estimation of the fluorescence in the hotspots inside cells (A) and (E). But they are still dependant to experimental conditions such as the camera bit-depth. The final Figure 4.2(D) and Figure 4.2(H) show the normalized calcium signals extracted from hotspots in Figure 4.2(A) and Figure 4.2(E). The range of the normalization procedure, being set between [1,2], enables appropriate comparison to both cells' frequency of calcium signals without experimental conditions interferences according to the formula:

$$F_i = \frac{f_i}{\max(f_i)} + 1 \quad (4.1)$$

where F_i is the normalized fluorescence value of the hotspot at time i scaled between [1,2],

f_i the corrected fluorescence value of the hotspot at time i ,

And $\max(f_i)$ the maximum value of the corrected fluorescence values of the hotspot.

4.3 Photobleaching in Hotspot Calcium Signals

As stated in the introduction, measurements of calcium changes of concentration in cardiac myocytes can be altered quantitatively by photobleaching loss. In following subsections 4.3.1, 4.3.2 and 4.3.3, we define what is meant by photobleaching in quantitative microscopy, and then we present our investigation for its detection and correction in calcium signals recorded in calcium hotspots of cardiac myocytes.

4.3.1 Introduction to Photobleaching

In FM, photobleaching is defined as the lost ability of fluorescent molecules to change contrast, fluoresce or to re-emit light at an expected wavelength when exposed to some controlled excitation light at the presence of binding targets [90]. In simple terms, there is photobleaching when

fluorochromes or fluorophores that are supposed to change colour when exposed to light and at the presence of target molecules, stop doing so because of the death of the fluorescent molecules [90] [92]. It is a progressive phenomenon which shows fluorescent molecules fade the more they are exposed to light [55]. As highlighted in Figure 4.3, the change of fluorescence measured inside an active region of cardiomyocyte (plotted in blue) can exhibit a negative baseline shift corresponding to an exponential decay (plotted in red). Ideally, its baseline should show a flat horizontal baseline as plotted in amber if there was no photobleaching phenomenon affecting the original signal plotted in blue.

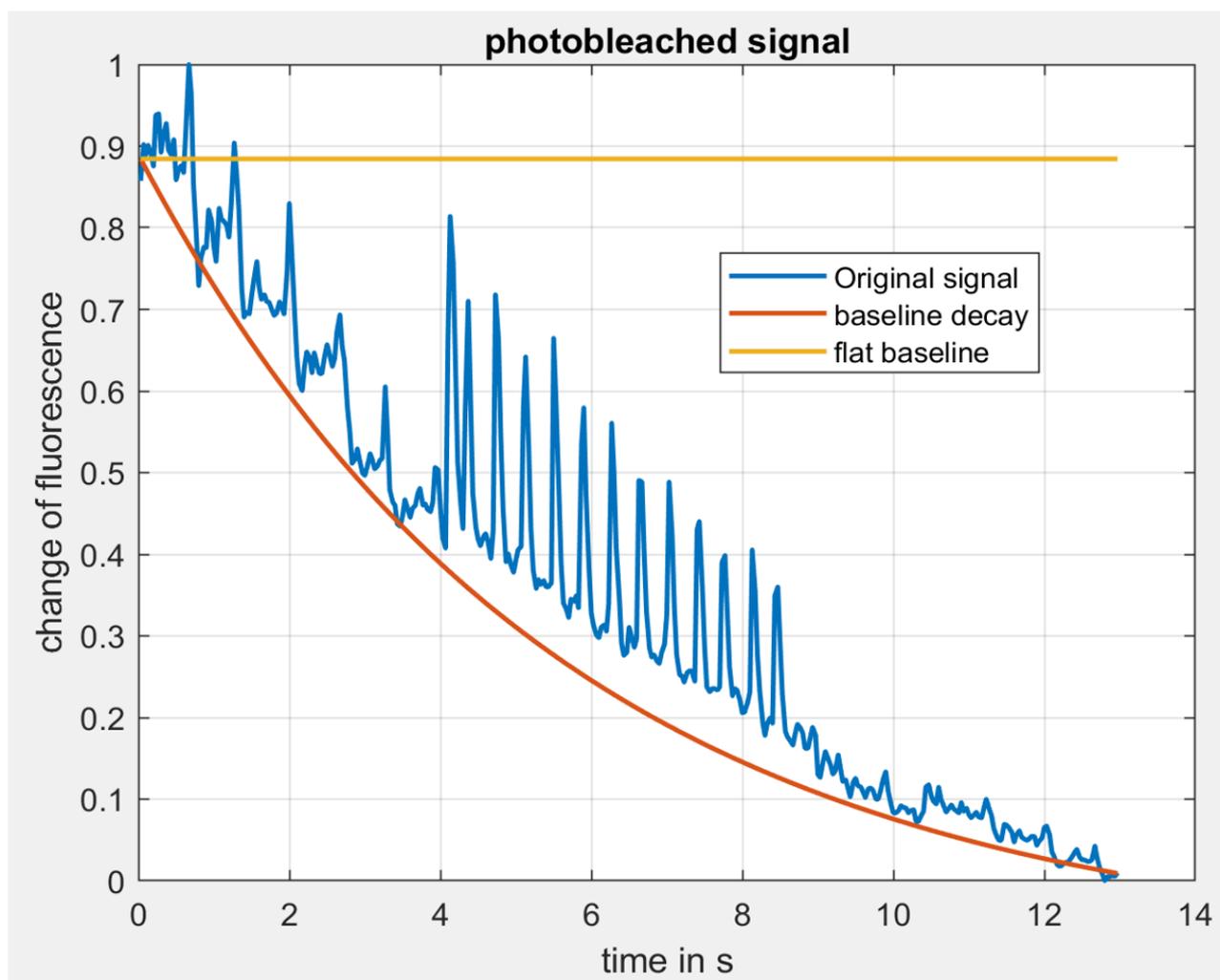


Figure 4.3: Hotspot calcium signal affected by photobleaching. Blue line: original calcium signal recorded for a hotspot through consecutive video frames for 13 seconds. Red line: photobleaching illustration through a baseline shift of the original calcium signal comparable to a negative exponential function. Amber line: theoretical flat baseline that should characterize the calcium signal in red if no photobleaching decay.

Many factors contribute to the advent of photobleaching in FM. These can be the nature of dyes

used, the chemical composition of the observed cells, the oxygen concentration inside them, the amount of excitation light, and the camera settings [79].

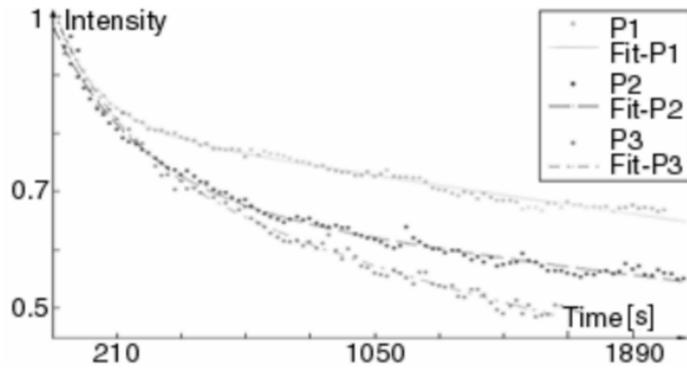
Although these diverse factors make it difficult to design general mathematical models applicable to different imaging experiments, research have not stopped investigating automatic solutions to deal with photobleaching in raw data measurements. Indeed, it is crucial in quantitative microscopy, to deal with the progressive decay of pixels values within images due to photobleaching when the aim is to infer directly from their raw image data the measurements of biological phenomenon such as intracellular calcium signals through FM [92].

To deal with the photobleaching loss that may alter calcium signal measurements in cardiac myocytes analysed in this thesis, we adopted a machine learning approach based on negative exponential models described here [90] [79]. These models based on mono and two exponential functions have been successfully applied to various types of cells and image modalities but not on biological signals observed in cardiac myocytes. In this chapter, we aim to investigate the same mathematical hypothesis over intracellular calcium signals of PVC and NRVM cell preparations obtained by FM that can be severely altered by photobleaching decay [68] [11] when exposed to excitation light for some time.

4.3.2 Photobleaching Detection in Hotspot Calcium Signals

Ideally, when the imaging experiments are supposed to exhibit constant fluorescence spectrum over time as described in [90] and [70] for histological slices of *Bufo Arenarum* Embryos stained with E-cadherin indicator, photobleaching effects can be directly measured from changes of raw pixels fluorescence in time-sequential images. Since the concentration of the protein E-cadherin under investigation does not fluctuate over time, any decay of the fluorescence associated with its expression can be considered as due to photobleaching loss. Figure 4.4 adapted from [90] illustrates the computation of photobleaching curves from three slices of *Bufo Arenarum* Embryos. Exponential curves (solid lines Fit-P1, Fit-P2, Fit-P3) are fitted directly to average pixel values of E-Catherin regions (dot lines P1, P2, P3) to estimate photobleaching curves affecting the images obtained by immunofluorescence.

However, in the case of cardiac myocytes stained with intracellular calcium indicators, the change of raw pixel intensities in time-sequential images could reflect the intracellular calcium changes of concentration rather than direct fluorescence decay in cellular preparations. In such



Photobleaching curves. Three negative exponential curves are shown (Fit-P1, Fit-P2 and Fit-P3). These were fitted to the data extracted from selected pixels of three different histological slices (P1, P2 and P3, respectively).

Figure 4.4: Photobleaching curves computation in Histological slices.

biological experiments, the fluctuation in cell regions due to intermittent and transient calcium signals with different amplitudes make it hard to distinguish normal fluorescence decays and abnormal ones induced by photobleaching. Normal fluorescence decays may be observed in cellular regions where calcium signals are gradually decreased. Abnormal fluorescence decays may occur in calcium hotspot regions undergoing photobleaching phenomenon but difficult to detect as inhibited by the cardiomyocytes' calcium signalling behaviour. Cardiac myocytes calcium signalling typically shows large transient increases of the calcium concentration, caused by calcium release from the endoplasmic reticulum, which can have various spatio-temporal patterns [68], resulting in fluorescence intensities spectrum ranging from dark to bright when the cells are imaged.

To distinguish the fluorescence decay due to photobleaching in cellular regions from a real decrease in the intracellular calcium concentrations, we built a photobleaching detection strategy based on calcium signals baseline estimation in such a way that the global trends of calcium signals are captured in a window-averaging (the window size is 6.6% of video length) manner. By doing so, our algorithm ignores the effect of fluorescence fluctuations inherent to calcium signals intakes and releases in cells and focus on filtered versions of calcium signals that may better represent the global flow of the original signals, i.e. the baseline or the resting calcium concentration.

Our strategy assumes that if there is no photobleaching affecting calcium signals, cells should roughly show positive-shifts or flat baselines when analysing their general signal trends. Otherwise, negative shifts of the baseline of calcium signals would mean the occurrences of photobleaching loss. To compute the sign of the baseline shift of hotspot calcium signals over time, we adopted a machine learning approach using a non-linear least square regression algorithm

that has been successfully implemented in [90]. Non-linear least square regression is a data fitting solving problem that minimizes the sum of squares of auxiliary functions defined by a user-defined objective function f according to the following formula [37]:

$$\text{minimize}_x f(x) = \sum_{i=1}^m f_i(x)^2 \equiv \frac{1}{2} F(x)^T F(x); \quad (4.2)$$

Where $f(x)$ is the objective function, $f_i(x)$ the auxiliary function, and $F(x)$ the vector-valued function. To read more about the algorithm mathematical concept please refer to [37] [81]. Here we used exponential models as the objective function.

Figure 4.5 illustrates our automated procedure for photobleaching detection in calcium hotspot signals by baseline shift' data fitting using nonlinear least square regression.

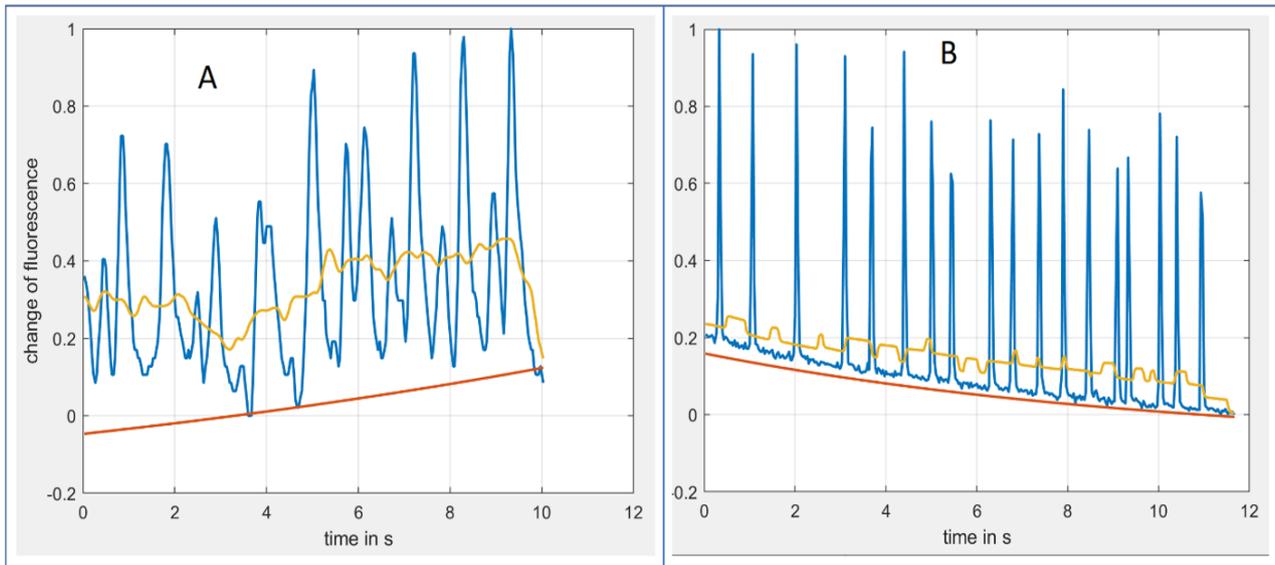


Figure 4.5: Photobleaching detection through baseline shift sign (x-axis representing time unit in second, and y-axis representing calcium change normalized units). Blue lines represent traces of hotspot calcium signals. Red lines correspond to their respective baseline shift estimation by mono-exponential fitting using non-linear least square regression technique. Amber lines represent smooth calcium signals obtained by window-averaging the original calcium signals. Curves are plotted calcium intensities over times in second. The intensities of the calcium signals have been normalized between 0 and 1 based on their maximum readouts.

Figure 4.5 describes original calcium signals for two hotspots (in solid blue lines) monitored in two different cell types. Blue lines represent traces of hotspot calcium signals. Red lines correspond to their respective baseline shift estimation by mono-exponential fitting using non-linear

least square regression technique. Amber solid lines represent smooth calcium signals obtained by window-averaging the original calcium signals. Curves are plotted calcium intensities over times in second. The intensities of the calcium signals have been normalized between 0 and 1 based on their maximum readouts. The first hotspot in Figure 4.5(A) does not show any sign of photobleaching artefact as it has a positive-shift baseline highlighted in the solid red line. The baseline is estimated from the smooth signal of the original signal highlighted in amber by using non-linear least square regression with the exponential model (Equation 4.2). The second hotspot in Figure 4.5(B) in contrast demonstrates an evident photobleaching decay affecting the overall original calcium signal depicted by a negative shift of the estimated baseline of the signal (solid red line in Figure 4.5(B)).

The flowchart diagram in Figure 4.6 below describes our automated routine steps for photobleaching detection by non-linear least square regression using the exponential model and photobleaching correction by baseline adjustment in hotspot calcium signals.

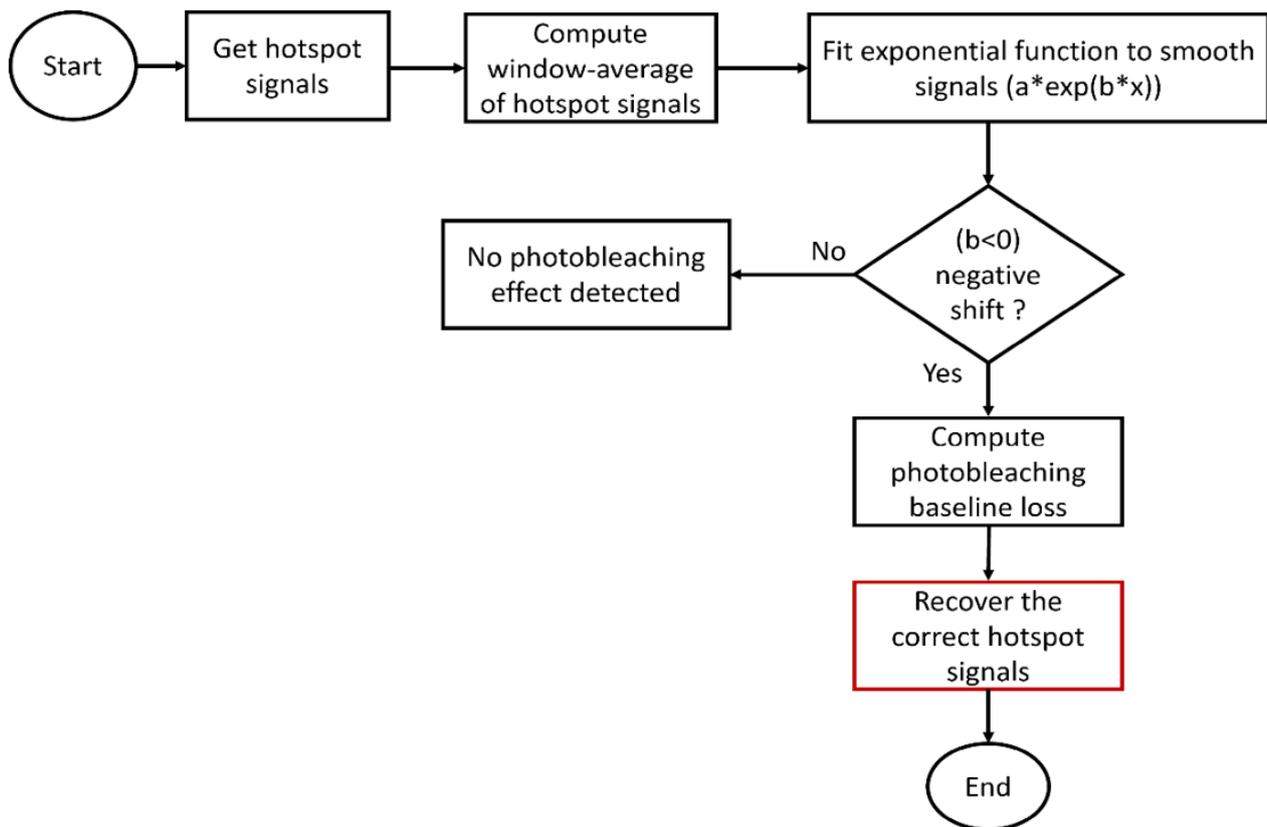


Figure 4.6: Flowchart diagram of photobleaching detection and correction in hotspot calcium signals. The correction step highlighted in red is described in Section 4.3.3

Figure 4.6 shows four main steps for photobleaching detection and correction: window-averaging

of hotspot signals, exponential curve fitting of hotspot calcium signals, baseline loss estimation by photobleaching, and signal correction after photobleaching. The window-averaging of hotspot signals consists of a trend estimation by mean filtering in one dimension where each output signal amplitude contains the mean value in a window of k neighbourhood around the corresponding amplitude in the input hotspot signal. The greater is the neighbourhood size, the smoother the output signal becomes. The smaller is the neighbourhood size, the closer the output signal will be to the initial input. We dynamically set the neighbourhood size to the nearest odd number roughly corresponding to 6.6% of the input video length. For example, a video of 301 frames will be smoothed within a window of 21 frames. By smoothing the original calcium signal in a relatively large window (21 frames in average), we aim to capture the baselines of calcium signals, i.e. their general trends with little interferences related to inherent cellular calcium dynamics.

The exponential curve fitting of the smooth signals obtained after mean filtering corresponds to a maximization procedure that looks randomly for the best exponential function parameters to minimize the mean square error between observed data and predictions modelled (Equation 4.2). We investigated two exponential models (functions with one and two terms) and opted for the best fit. For mono-exponential fitting, we searched for two parameters (a, b) whereas for bi-exponential fitting, we searched for four parameters (a, b, c, d) according to the following equations:

$$f_1 = a * exp(b * x) \tag{4.3}$$

$$f_2 = a * exp(b * x) + c * exp(d * x) \tag{4.4}$$

where f_1 is a mono exponential function , f_2 , a bi-exponential function of time x modelling the photobleaching loss of fluorescence inside hotspot signals. In mono-fitting, $b < 0$ would mean a negative shift of the calcium signal baseline. In bi-exponential fitting, $b < 0$ and $d < 0$ would mean a negative shift of the calcium signal baseline. Otherwise, the system assumes there is no photobleaching effect on the input hotspot signal processed. The corresponding photobleaching loss affecting the calcium signal is derived from the parameters of the exponential function found using curve fitting of the smooth calcium signals.

Although we are dealing with different ROIs (Calcium signals) and different fluorescent indicators (e.g. Oregon Green BAPTA and fluo-4) than the methods described in [90] and [37]

which are static cells populations loaded with membrane stains, our proposed photobleaching detection solution argues that it is possible to model the photobleaching loss affecting intracellular Calcium traces of cardiac myocytes as a negative exponential function of time. Figure 4.7 illustrates the result of the two exponential fittings using Equation 4.3 and Equation 4.4. The performance of our method for photobleaching detection based on exponential curve fitting of the baseline shift of calcium signals is presented in Section 4.5.

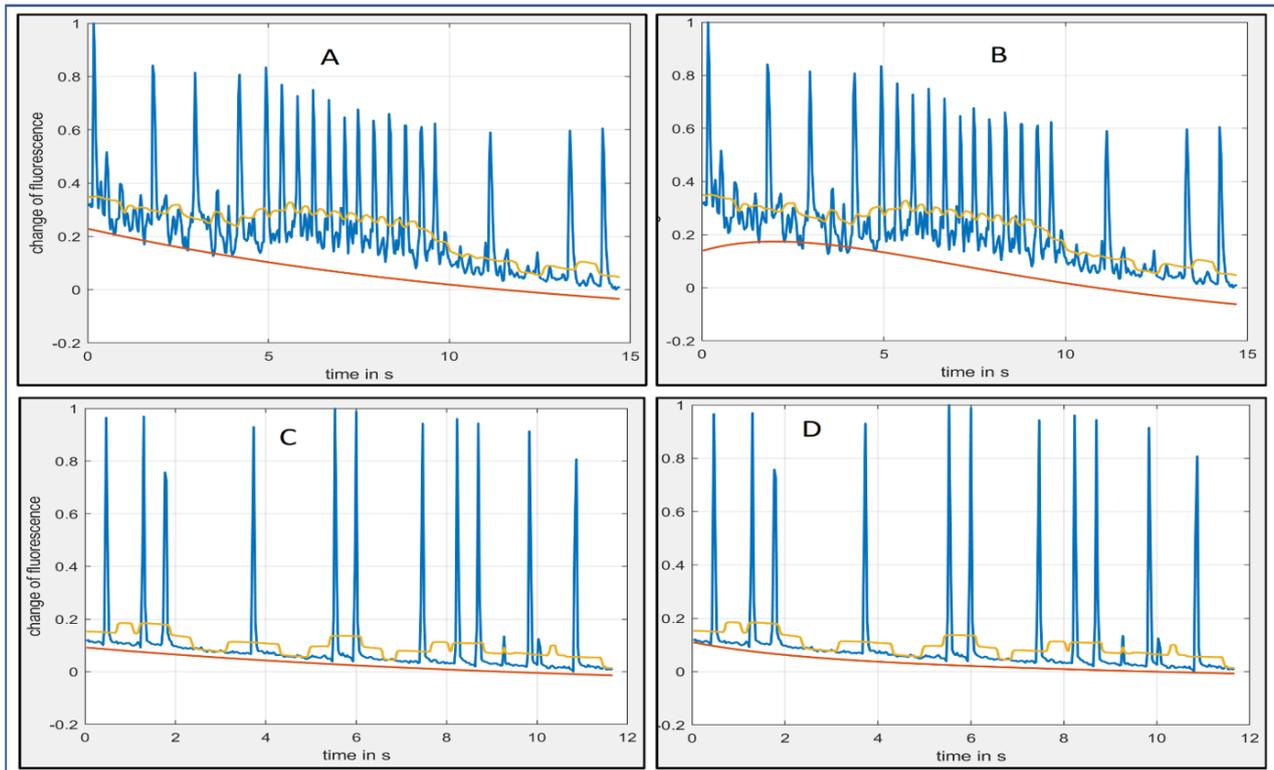


Figure 4.7: Photobleaching detection by mono (A, C) and bi-exponential (B, D) fitting. The blue lines represent original hotspot calcium signals measured over time in consecutive video frames. The amber lines represent smooth versions of the hotspot calcium signals by using window-averaging. The red lines represent the estimated photobleaching rate obtained by fitting exponential curves to the smooth calcium signals.

As Figure 4.7 shows, models based on mono-exponential curves (red curves (A) and (C)) seem more appropriate at capturing photobleaching losses described as negative exponential functions of time rather than bi-exponential curves (red curves (B) and (D)). Bi-exponential curves are more complex (4 parameters) and tend to overfit the observed data in smooth signals (red curve B). In best cases, they exhibit similar shapes as mono-exponential curves (red curves (C) and (D)). Moreover, the visual inspection demonstrates that the mono-exponential fits show consistent exponential decays in the photobleaching rate, whereas the bi-exponential fits can be

inconsistent. The bi-exponential fit in Figure 4.7(B) does not show a progressive decrease of the fluorescence, i.e. that does not exhibit photobleaching effect. Therefore, we based the correction of photobleaching in calcium hotspot signals on the mono-exponential fitting of baseline shifts. Next section 4.3.3 presents our investigation for calcium signal recovery once photobleaching loss has been inferred by exponential curve fitting of signal baseline.

4.3.3 Photobleaching Correction in Hotspot Calcium Signals

If negative baseline shifts are detected from initial calcium signals instead of expected flat baselines, we correct the corresponding photobleaching effects by adding to the original hotspot signals the amplitude differences between the expected flat baseline and the estimated shifted baseline obtained by exponential curve fitting, as shown in Figure 4.8.

If negative baseline shifts are detected from initial calcium signals instead of expected flat baselines, the correct calcium signal is recovered by adding to the altered hotspot signals the absolute amplitude differences between the expected flat baseline and the estimated shifted baseline corresponding to the photobleaching loss and obtained by exponential curve fitting, as shown in Figure 4.8.

Figure 4.8 shows the photobleaching effect and correction of the calcium signals in the two types of cardiac myocytes used in this study: PVCs (A) and NRVMs (B). The two cell types exhibit different levels of photobleaching effects plotted in amber lines (estimated baselines by non-linear regression using a mono-exponential model over smooth signals). The blue lines (in both cells' plots) represent the original calcium signals showing or affected by photobleaching loss. The purple lines (in both cells' plots) represent the expected horizontal baselines of correct calcium signals, i.e. the flat baselines of signals not affected by photobleaching. The red lines (in both cells' plots) correspond to the corrected calcium signals obtained by baseline adjustment. The baseline adjustment is performed by taking the absolute differences between the expected flat baseline (in purple) and the estimated shifted baseline (in amber) due to photobleaching. Then the absolute differences are added to the initial calcium signals (in blue) according to the following formula:

$$CorrectSignal = InitialSignal + abs(FlatBaseline - ShiftedBaseline) \quad (4.5)$$

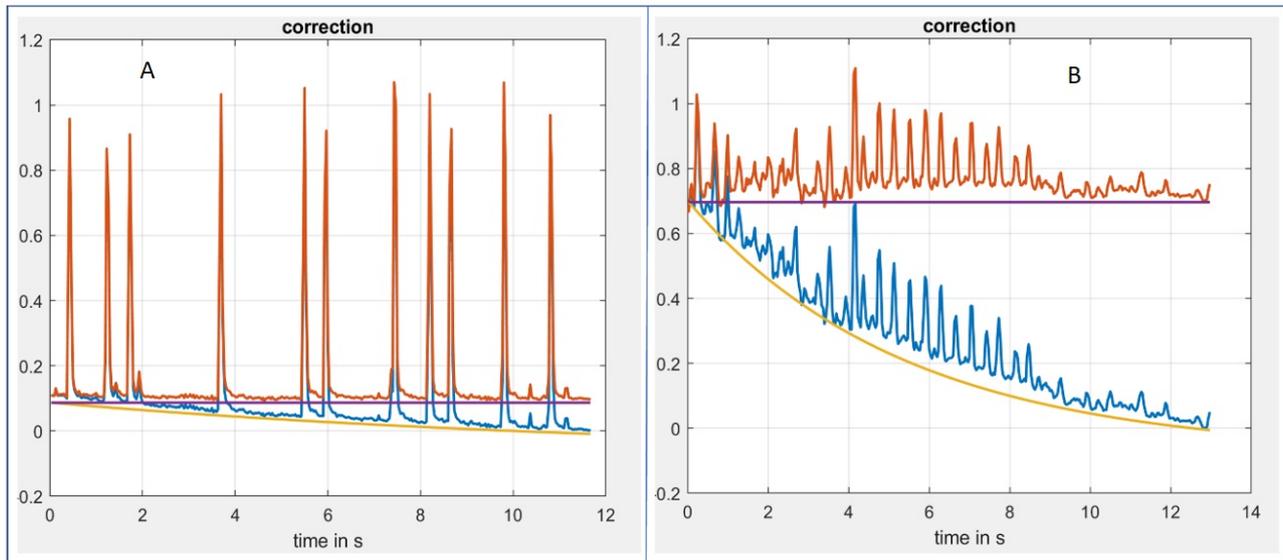


Figure 4.8: Photobleaching correction by baseline adjustment in 2 calcium signals by mono-exponential fitting. A: photobleaching detection and correction from a hotspot calcium signal recorded 3 minute before carvedilol treatment on cells. carvedilol is a protein that inhibits Calcium signalling in mitochondria. B: photobleaching detection and correction from a hotspot signal recorded after 10 minutes after dantrolene treatment of cells. dantrolene is a molecule used to treat some specific calcium channels disorders in cardiac cells. Blue lines: original hotspot calcium signals recorded over time. Yellow lines: estimated photobleaching decay derived from baseline shift of the original signals. Purple lines: theoretical flat baselines if the original signals were not affected by photobleaching. Amber lines: recovered calcium signals after photobleaching detection and correction of the original signals.

where *CorrectSignal* is the expected calcium hotspot signal without photobleaching effect; *FlatBaseline* is the ideal expected horizontal baseline that should characterizes the *InitialSignal*; and *ShiftedBaseline* is the abnormal baseline caused by photobleaching and estimated by exponential curve fitting.

4.4 Materials & Experimental Protocols

The experiments we performed in calcium hotspot signals processing correspond to the two main tasks related to our handling of photobleaching in cardiac myocytes, i.e., photobleaching detection by non-linear least square regression with exponential models and photobleaching correction by baseline adjustment. To obtain time-series calcium signals of hotspots, we used

the output of our automated routine for hotspots detection over PVC and NRVM calcium imaging datasets described in Chapter 3.

To assess the performances of our automated algorithm for photobleaching detection in hotspot calcium signals by curve fitting, we performed qualitative comparison and measured the goodness of the fits through two metrics: the mean square error of and the coefficient of determination. The mean square error (Equation (4.6)) describes the difference between the estimator (the original baseline obtained by smoothing the original calcium signal) and the estimated value (the predicted value output by exponential function models). The coefficient of determination (Equation (4.7)) is a percentage metric that tells how much the original baseline is closed to the exponential function whose parameters have been estimated by non-linear least square regression.

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - f_i)^2; \quad (4.6)$$

where MSE is the mean square error, n the number of observations, y_i the estimator or baseline dataset point at position i , and the f_i the corresponding estimated value or baseline value by exponential function at position i .

$$R^2 = 1 - \frac{SS_{res}}{SS_{tot}}; \quad SS_{res} = \sum_i (y_i - f_i)^2; \quad SS_{tot} = \sum_i (y_i - y')^2; \quad (4.7)$$

Where R^2 is the coefficient of determination, SS_{res} is the residual sum of squares, SS_{tot} is the total sum of squares, y_i is the observed data at position i , f_i the predicted data, and y' the observed data mean.

To assess the performance of our automated routine for photobleaching correction in hotspot calcium signals by baseline adjustment, we considered the proportion of hotspot calcium signals correctly adjusted by detecting photobleaching in the corrected signals.

4.5 Results and Discussion

Mono-exponential and bi-exponential functions were fitted to sets of data derived from window-averaging of calcium signals extracted from hotspot regions in consecutive video frames. Statistical analysis of MSE (Figure 4.9) and R^2 (Figure 4.10) of the two models shows that mono and

bi-exponential models output similar performances. MSE of mono-exponential models varies between $[0.09013, 6.41796]$, whereas MSE of bi-exponential models varies between $[0.03444, 6.41796]$. R^2 of the mono-exponential models varies between $[0.014009, 0.999998]$, while R^2 of bi-exponential models have also got the same maximum similitude of 0.999998 but a smaller minimum of 0.006958. So, all in one, based on the MSE and the R^2 , a conclusion about the best fits cannot be drawn. Therefore, we visually compared the outputs of both models.

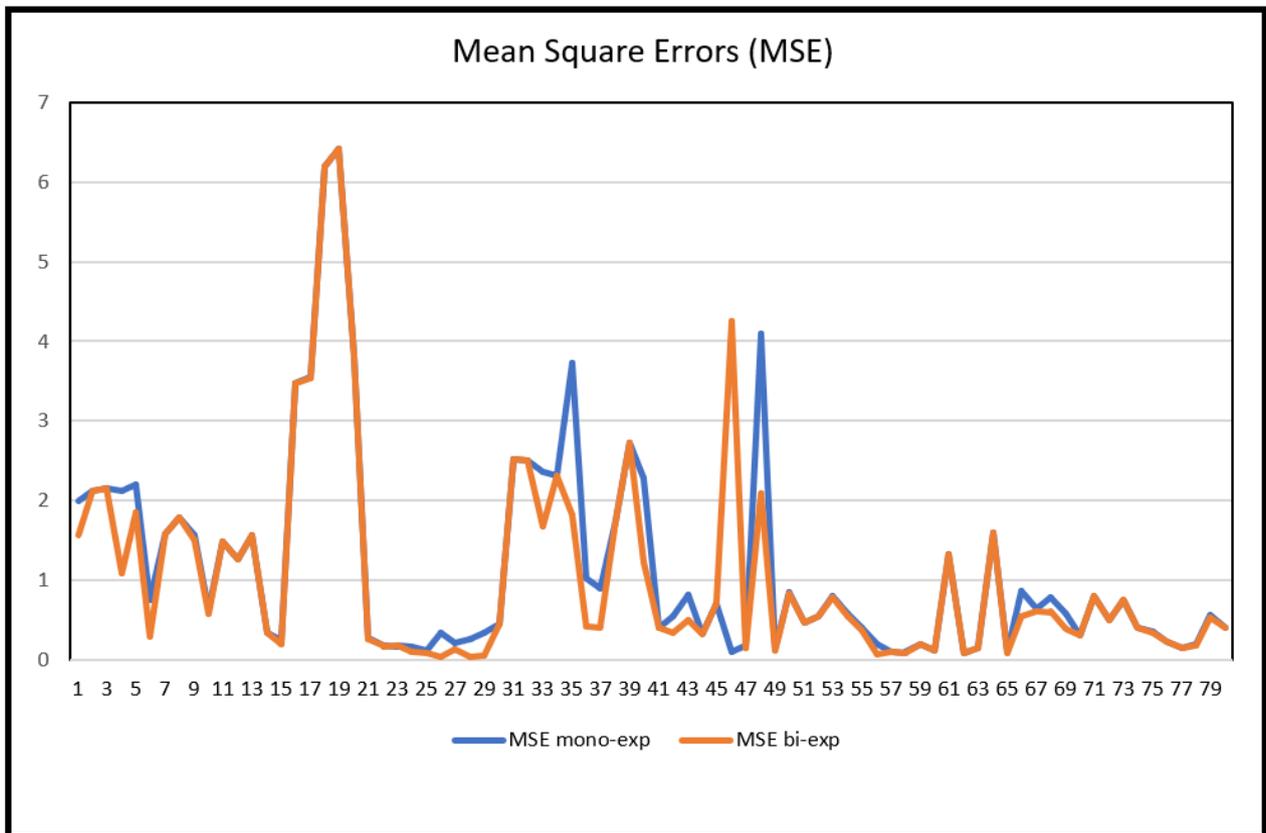


Figure 4.9: MSE report for mono and bi-exponential curve fitting of 80 hotspots. X-axis: index of 80 hotspot signals analysed. Y-axis: mean square error units.

A visual analysis of the exponential curves shows that bi-exponential models tend to overfit baseline shift datasets. Bi-exponential curves can exhibit complex shapes where function derivatives change signs at some points. In contrast, mono-exponential curves usually show smooth shapes where function derivatives never change signs and remain negative.

For instance, in Figure 4.11, it is evident that mono and bi-exponential curves can capture similar photobleaching decline of fluorescence (B) as well as different shapes (D). In the case of D, the mono-exponential model represents the best fit for capturing photobleaching effect of the original calcium signal in (C), because it shows a consistent decrease in fluorescence value, the bi-exponential model describes a non-continuous decline of fluorescence values. Therefore,

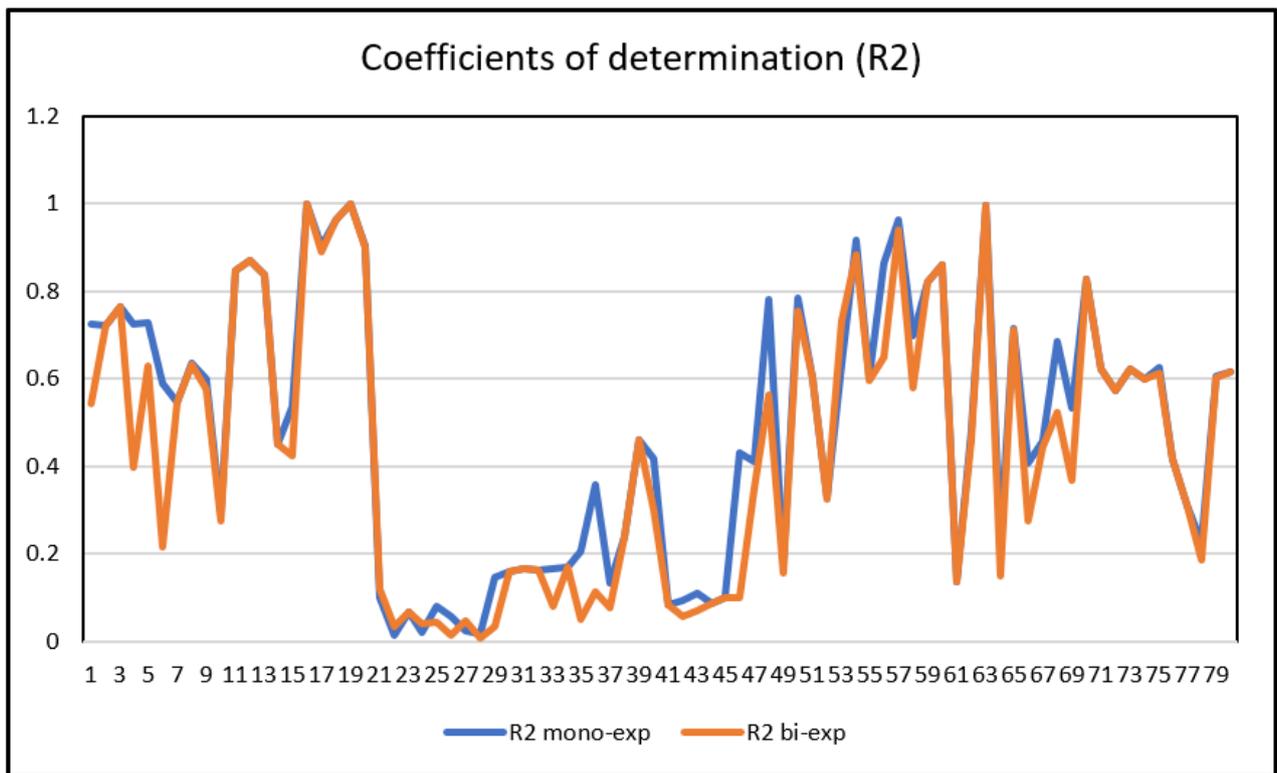


Figure 4.10: R^2 report for mono and bi-exponential curve fitting of 80 hotspots. X-axis: index of 80 hotspot signals analysed. Y-axis: coefficient of determination units.

bi-exponential fails to capture (in some cases) the properties of photobleaching decay in hotspot calcium signals.

Based upon on the statistical and visual analysis of both models investigated, mono-exponential curve models were employed in our automated routine for photobleaching detection. For photobleaching correction in calcium imaging recordings, we added to the original calcium signals, the fluorescence differences between the expected flat baseline and the mono-exponential fitted baseline, as shown in Figure 4.12.

Corrected calcium signals are considered truly adjusted when the average difference between the flat expected baseline and the baseline of the signals is greater than a threshold T . The threshold T corresponds to $1/100^{th}$ of the average fluorescence decay observed in our hotspot calcium datasets made of 80 elements. $T = -0.0014$. Figure 4.13 shows the average differences between estimated baselines of corrected signals and expected baselines of 80 hotspots analysed.

Comparing the average differences of corrected signals with the threshold (as shown in Figure 4.13), we found 63 of the 80 hotspots calcium to be above the threshold, i.e. 78.75% of hotspots have been correctly adjusted. Figure 4.14 illustrates a situation where our system has failed

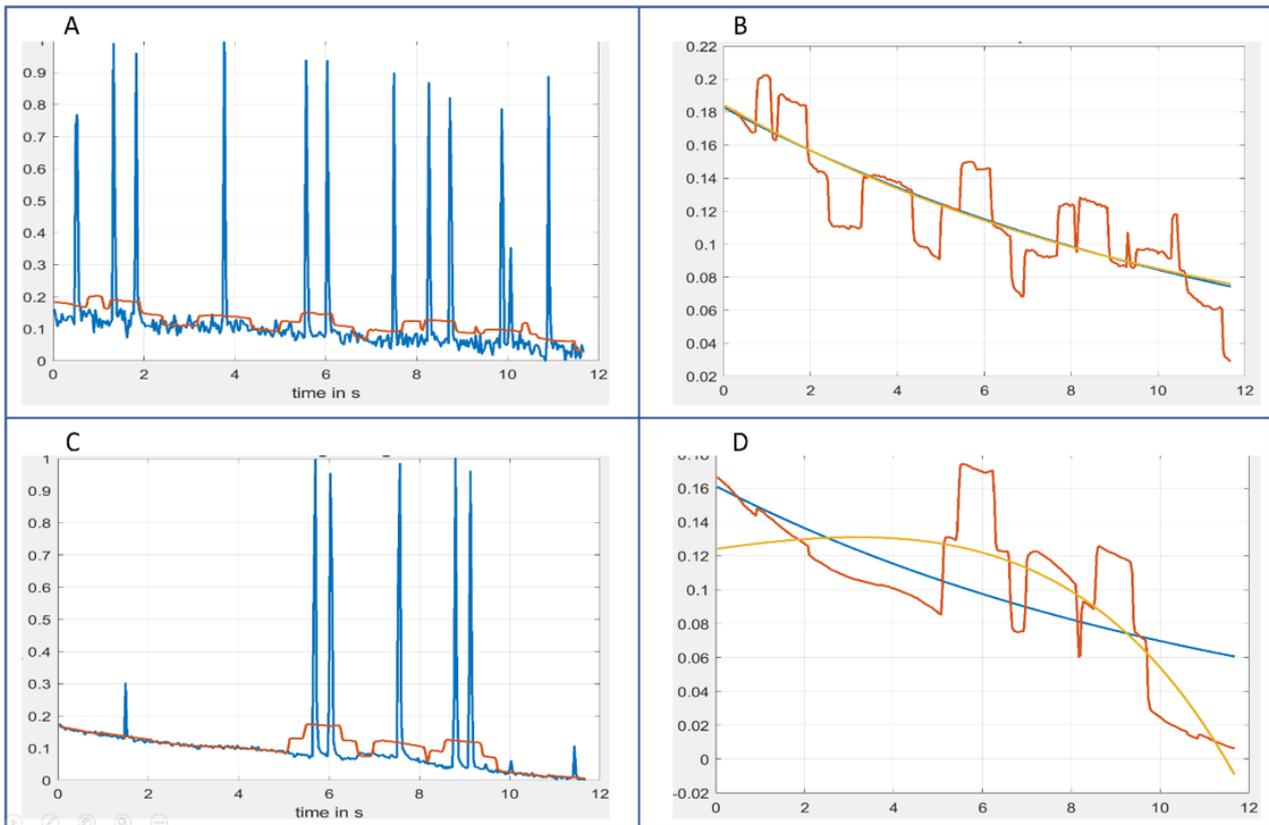


Figure 4.11: Exponential models comparison for photobleaching detection. Blue lines in A and B represent original calcium signals extracted from different hotspots of the same cells. Red lines in A, B, C and D correspond to smooth signals of original signals obtained by window-averaging. Blue lines in B and D correspond to mono-exponential curve fit of the smooth signals. Orange lines in B and D show bi-exponential curve fit of the smooth signals.

to adjust a photobleached signal to the accepted threshold we set up, whereas Figure 4.15 illustrates a successful detection and correction outcome.

4.6 Summary on Calcium Hotspot Signal Analysis

To sum up, the detection and correction of photobleaching effects in calcium imaging experiments provide biologists with more precise measurements that can be used to compare different cells signalling activity regardless of experimental conditions. The chapter shows that photobleaching can be detected with two exponential curve fittings using nonlinear least square regression in cardiac myocytes datasets. The two models show similar performances in terms of MSE and R^2 . However, upon a visual investigation, calcium baseline shifts estimated by

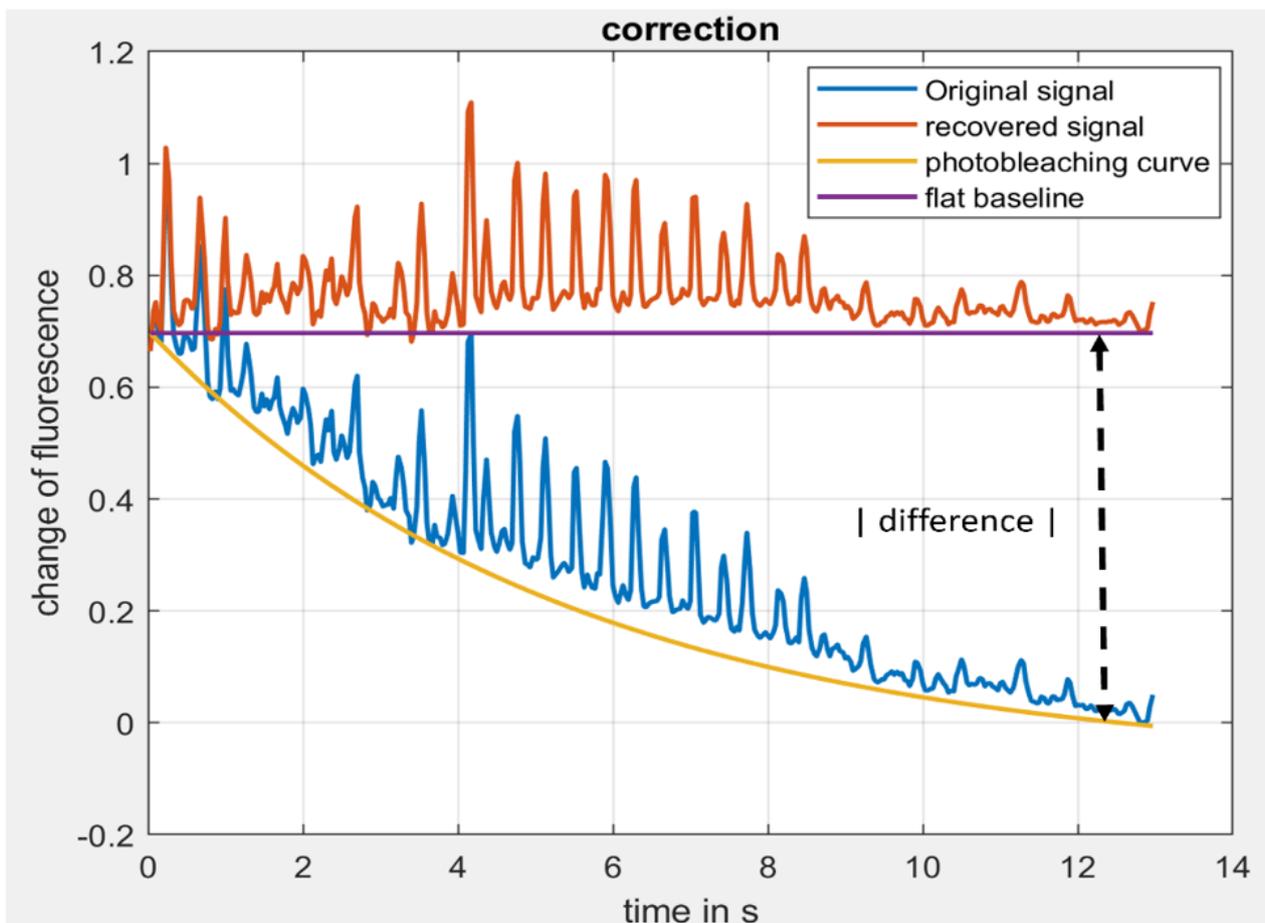


Figure 4.12: Photobleaching correction in hotspot calcium signal. Blue line: original hotspot calcium signal recorded over time. Red line: recovered signal after photobleaching detection and correction. Yellow line: estimated photobleaching decay derived from the baseline shift of the original calcium signal. Purple line: theoretical flat baseline of the correct signal if no photobleaching corruption

mono-exponential models exhibit the properties of photobleaching rates, i.e. the exponential continuous decline of fluorescence over time. This can be expanded to other experiments that observe changes in fluorescence over time.

The chapter also shows that the detected photobleaching declines in calcium signals can be corrected by baseline adjustment of the original signals. The adjustment is made by adding to the original signals the differences between the expected flat baselines and the estimated baselines exhibiting photobleaching loss. With this correction, 78.75% of the detected hotspots exhibiting photobleaching effects were adjusted correctly. For the remaining 21.25% of the detected hotspots that the system failed to correct, the photobleaching detection and correction could be applied to them multiple times until an acceptable threshold is reached.

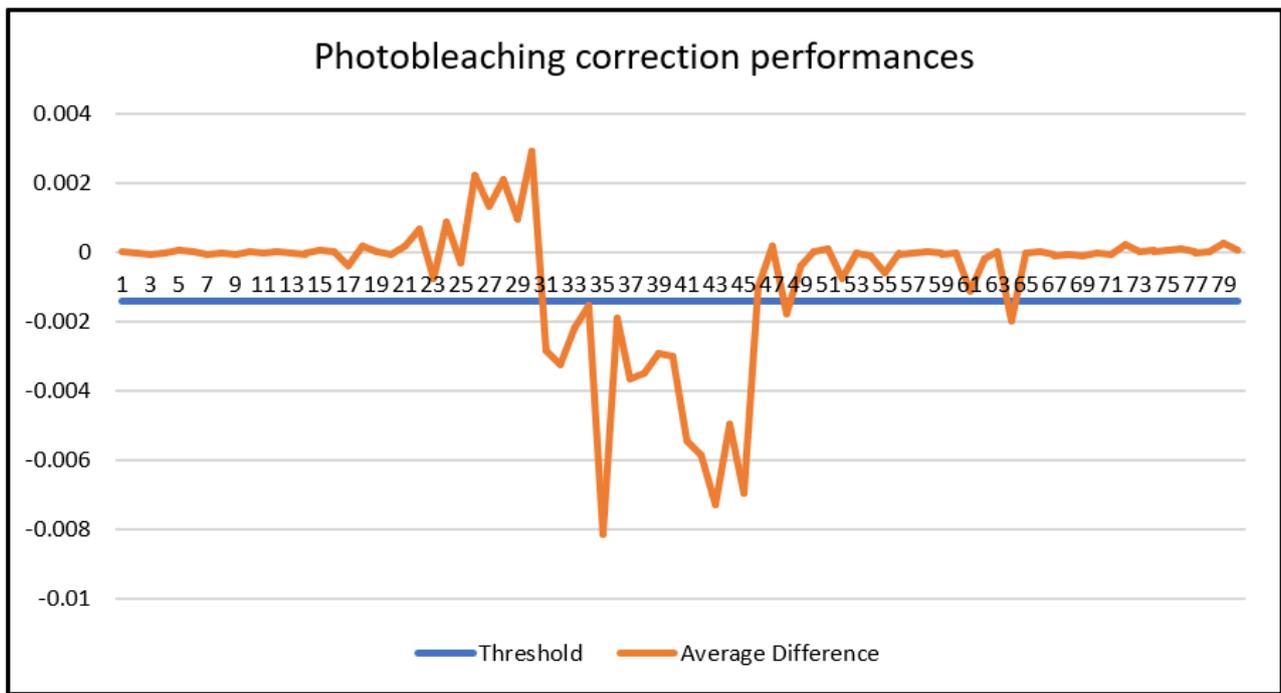


Figure 4.13: Average differences between expected and estimated baselines of corrected signals. X-axis: index of 80 hotspot signals analysed. Y-axis: differences recorded. Amber line: average differences between the baseline of the recovered signals and the theoretical flat baseline expected. Average differences below the blue line threshold means the system failure to correct the detected photobleaching effect

Measuring calcium signals accurately helps to understand how cells function. Calcium signals regulate a plethora of other processes inside cells, some of which can also be observed in living cells using fluorescent indicators. By observing two processes in parallel, one can draw important conclusions on how, for example, calcium signals affect the regulation of autophagy. For many of these other cellular processes, it is critical to know how many cells are observed. Detecting and counting the nuclei in FM images is the central theme of Chapter 5.

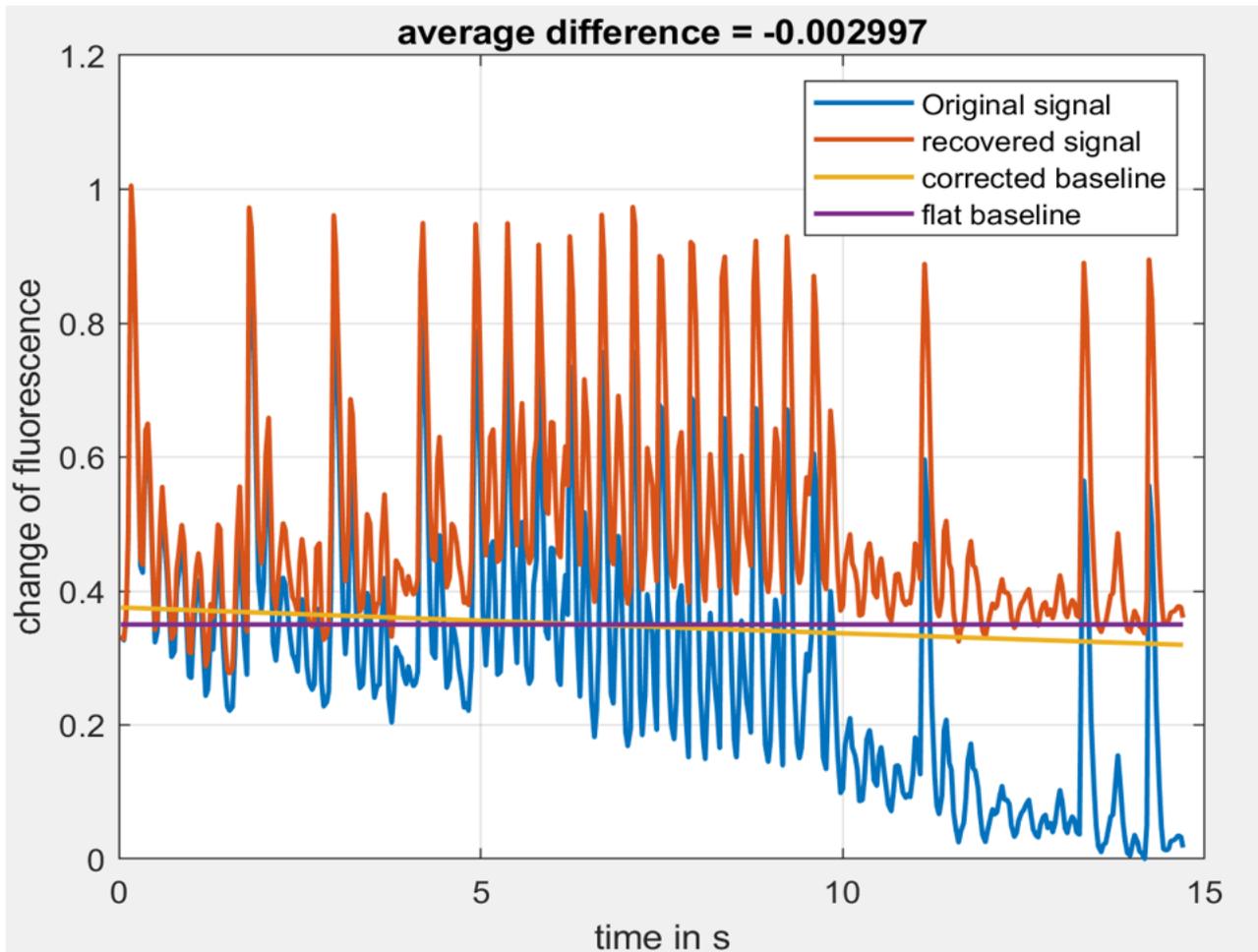


Figure 4.14: Example of a corrected calcium signal whose average difference falls below the threshold T . Blue line: original hotspot calcium signal recorded over time. Amber line: recovered signal after photobleaching detection and correction. Yellow line: estimated baseline shift of the recovered calcium signal. Purple line: theoretical expected flat baseline of the recovered signal. The average difference -0.002997 between the estimated baseline of the recovered signal and the flat expected baseline illustrates a non-successful photobleaching correction as it means the baseline of the recovered signal is still below expectation

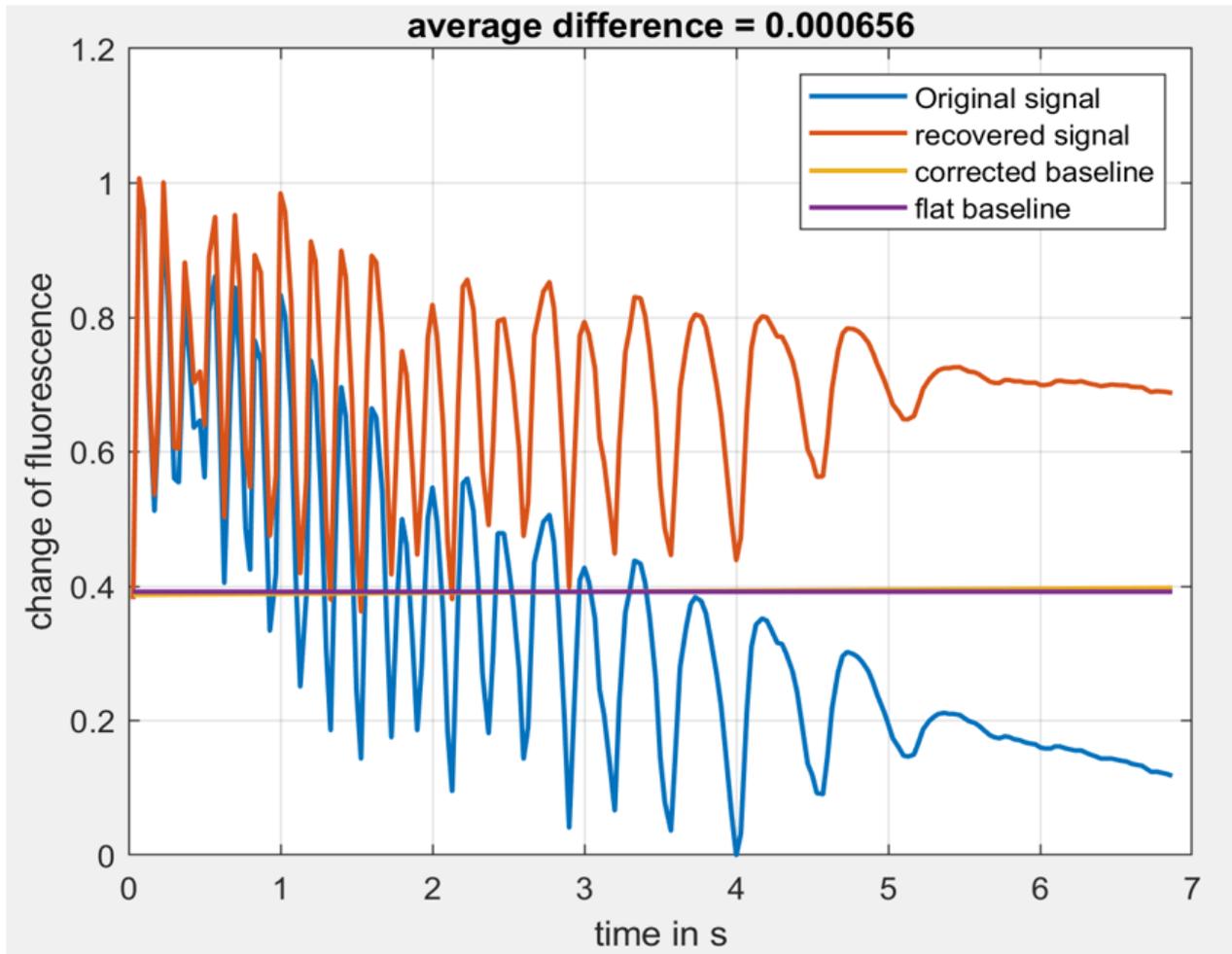


Figure 4.15: Example of corrected calcium signal whose average difference is above the threshold T . Blue line: original hotspot calcium signal recorded over time. Amber line: recovered signal after photobleaching detection and correction. Yellow line: estimated baseline shift of the recovered calcium signal. Purple line: theoretical expected flat baseline of the recovered signal. The average difference 0.000656 between the corrected baseline and the flat baseline illustrates a successful photobleaching correction as it means the baseline of the recovered signal is above the flat baseline.

Chapter 5

Investigating Standard R-CNN Variants for Nuclei Detection in Autophagy Cells

This chapter presents a theoretical and empirical analysis of state-of-the-art object detection techniques based on R-CNN architecture for nuclear object proposal generation in FM images. In R-CNN based solutions, object detection techniques rely on different strategies to identify “proposal boxes” to be fed into a CNN classifier. The performances of such strategies were largely demonstrated over natural images, but not yet rigorously done on biomedical images at a similar scale. This chapter evaluates four popular strategies for proposal box generation in R-CNN, namely, exhaustive search by sliding windows, edge box search, selective search, and region proposal network (RPN). The four techniques are compared, and their strengths and limitations are reported over NRVM images stained with a DNA blue fluorescent dye called Hoechst. The evaluation reported in this chapter enabled us to make recommendations for nuclei detection in biomedical images based on a customized R-CNN architecture using adaptive region proposal generation. The contribution of this chapter is twofold:

1. Evaluating the strengths and limitations of four state-of-the-art object detection algorithms for object proposal generation in biomedical images.
2. Arguing that a customized R-CNN based on adaptive region proposal box generation using the data-driven technique may produce a great performance for nuclei detection in

biomedical images.

The chapter begins by giving an overview of R-CNN in Section 5.1. Then in Section 5.2, the materials and experimental protocols used to evaluate four state-of-the-art object proposal techniques over biomedical images are described. In Section 5.3, the theoretical and empirical evaluations of the four techniques are reported in terms of techniques description, experiment results and discussion, and techniques comparison. The chapter ends by concluding remarks in Section 5.4.

5.1 Introduction

R-CNN was first introduced in 2014 by Ross Girshick, a research scientist at Facebook AI Research. Ross Girshick is also well-known for developing variants of R-CNN called Fast-R-CNN, Faster R-CNN, [67], and Mask R-CNN [38] used respectively for object recognition, object detection, and semantic segmentation. Those deep networks have been successfully tested on computer vision benchmark databases made of natural images such as Pascal Visual Object Class (VOC) [30], and COCO [53], but are yet to achieve the same success in biomedical image analysis such as the detection of nuclei in cardiac myocytes obtained by FM.

The focus of this chapter is on the core landmark R-CNN architecture introduced in 2014 [35]. Since, challenges associated to our datasets of NRVM nuclei are ROI variability (in size, colour, and texture), and illumination artefacts, we argue that R-CNN based strategy might enable us to build a nuclei detector object that can generalize well across different ROIs and images characteristics, providing that the proposal box generation of nuclei locations is efficient. In fact, many researchers in the field of computer vision and object recognition believe that the overall accuracy of supervised object detectors based on image region analysis depends on the efficiency of the proposals initially computed by the systems [65]. Efficient object proposal methods should be fast and achieve high recall or sensitivity i.e., True Positive Rate (TPR). Thus, in this chapter we seek to mainly investigate multiple object detection techniques used as core initial proposal generation techniques in R-CNN models for nuclei detection in biomedical images.

Typical R-CNN architecture is interesting as it is simple and relies on three basic modules. First, an initial region proposal module generates candidate object locations that may be present in

input images. Then, a feature extraction module using CNN layers computes rich hierarchical features from the candidate locations proposed initially. Finally, a classification module outputs the actual locations of objects after image patch classification using linear SVM and box refinement using regression built upon the CNN feature maps as described in [35].

Despite the simplicity of R-CNN models, legitimate research questions tend to focus on the computational cost of their deployment, the size of network training datasets, and the accuracy of the proposal box generation technique used. To overcome the deployment burden, systems built on GPU hardware solutions and using Faster R-CNN architectures have been proposed as credible alternatives [67] because the proposal box generation and the final object detection are performed simultaneously.

About the issues linked to network training datasets, it is tough to find in the literature a “one size fits all cases ” answer because the amount of sufficient training data required to build performing models can depend on multiple parameters such as the differences between object classes to be identified, and the quality of the dataset someone may be using [27]. To overcome these limitations associated with the lack of sufficient training data, data augmentation, and transfer learning solutions have been proposed by many researchers [54]. For more information about data augmentation and transfer learning in deep neural networks, please read [5].

About the problems related to the quality of candidate object locations generated by R-CNN models, state-of-the-art object detection techniques are yet to be widely adapted and tested over cellular domains images. That is why, in this chapter, we present a comprehensive investigation of the performances and limitations of four non-exhaustive proposal box generation algorithms over biomedical image data. These techniques include object proposal generation by exhaustive search with or without objectness measure [1], selective search [87], edge boxes [97], and Region Proposal Network (RPN) [67].

5.2 Materials & Experimental Protocols

In some object detection and recognition systems, the development of efficient ground truth labelled datasets are crucial for assessing models’ performance and adjusting training protocols [33]. That is the case for instance, in supervised learning, where it is based on the ground truth that researchers refine algorithms and perform a sort of reality check of their output against

what they know as the truth or the reality. Therefore, in this case study, we developed a ground truth nuclei dataset under the supervision of 2 biologist experts. The dataset was made from nuclei images of NRVM cells stained with blue Hoechst dye and used to investigate and evaluate four different proposal box generation algorithms that were initially built, trained and/or tested over natural images. Following subsections describe the labelling procedure we performed and the experimental protocols we followed to compare the different techniques investigated, i.e., object proposal generation by exhaustive search with or without objectness measure, selective search, edge boxes and RPN.

5.2.1 Ground Truth Dataset Materials

Since the techniques investigated output bounding boxes as the locations of object instances inside testing images, we performed a bounding box labelling of nuclei images under supervision of biologist experts to set the ground truth dataset. To this end, from 250 NRVM images, the same number of locations of nuclei and background samples were manually indicated in terms of $[x, y, w, h]$ vectors. The vectors $[x, y, w, h]$ represent respectively (x, y) geographical coordinates of the upper-left corner of the bounding box containing an object class with width ' w ' and height ' h '. Only two object categories were set: "nuclei" and "background" as illustrated in Figure 5.1. An arbitrary fixed size of $[160 \times 160]$ as width and height of sample patches was chosen to enable feature vector length consistency of techniques based on feature extraction and to conform with the size of the smallest window that can contain the most significant NRVM nucleus observed within the dataset. In total, the labelled dataset consists of 250 labelled NRVM images made of 2734 nuclei locations and 2734 background locations.

Figure 5.1 describes the labelling procedure adopted to generate the ground truth dataset used to evaluate performances of state-of-the-art proposal box generation techniques. In Figure 5.1, 15 nuclei locations and 15 background locations have been respectively labelled into yellow boxes or positive samples (top left illustration) and red boxes or negative samples (bottom left illustrations) on an NRVM image. The right-hand side illustration of Figure 5.1 represents individual crops of the nuclei and background regions labelled.

Visually, it can be noticed that the different object categories samples exhibit heterogeneous illumination, texture, and contents sizes. Such variability in the positive and negative samples of our labelled database comes as extra arguments in favour of generalized approach toward

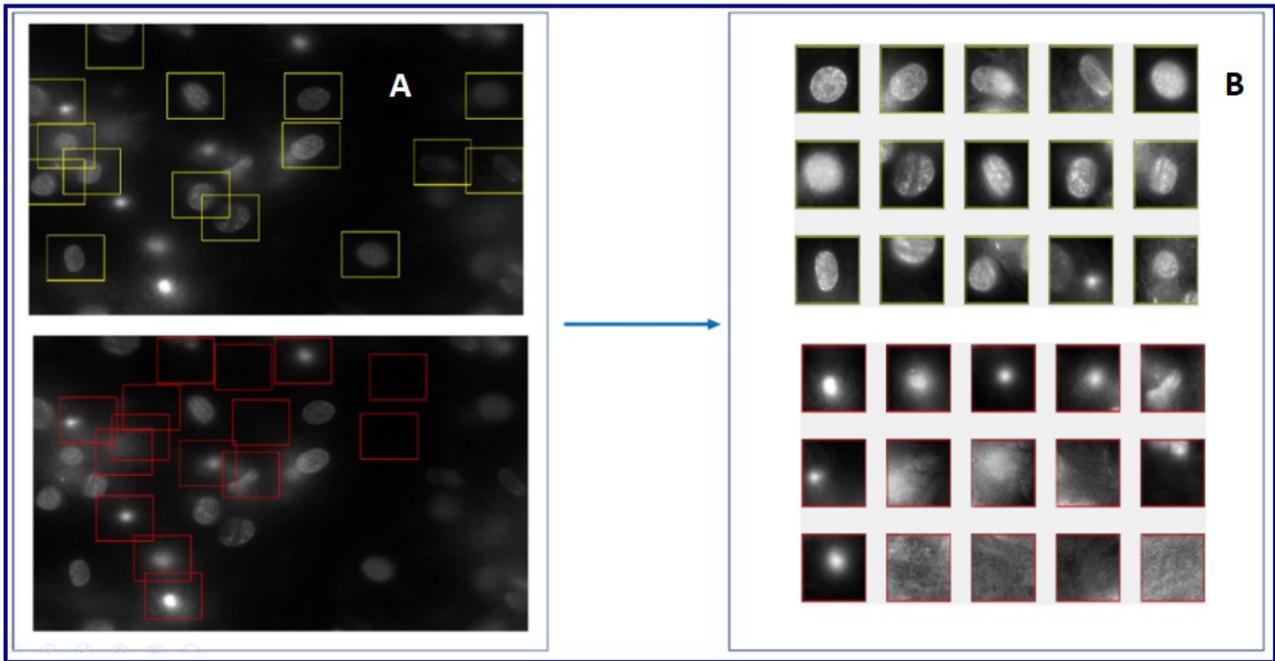


Figure 5.1: Labelling nuclei and background patches of NRVM images. A: an original NRVM image with nuclei locations labelled by yellow bounding boxes and sample background patches labelled by red bounding boxes. B: Cropping of the labelled patches in A that illustrates the contrast, size, shape, and texture diversity in ROIs and background elements

nuclei detection in FM images instead of hand-crafted algorithms or traditional image processing techniques that rely on prior knowledge of the datasets as discussed in Chapter 2. It also illustrates the ROI variability challenges faced by algorithm designers for nuclei instance counting in FM images. The next section of this chapter summarizes the experimental protocols performed for investigating nuclei proposal box generation by exhaustive search, selective search, edge boxes and RPN.

5.2.2 Experimental Protocols

To evaluate the performance of the four state-of-the-art object detection algorithms chosen over the ground truth nuclei datasets described in Section 5.2.1, experiments of nuclei proposal box generations were conducted for each technique. Then, the results output by the different techniques were analysed in terms of TP proposal boxes generated, FP proposal boxes, boxes' precision, and their ground truth coverage ratio (GTCR).

Performance Metrics

We defined as TP boxes, proposal boxes which overlap at least 50% with ground truth nuclei boxes according to a computed intersection over union metric (IoU). FP boxes in opposite correspond to proposal that do not overlap at least 50% with any ground truth nuclei locations. The IoU metric enables to measure the overlap ratio between proposal boxes and ground truth. It is a standard metric used in object recognition systems for evaluating the bounding box accuracy of object detection algorithms made popular by the well-known benchmark Pascal VOC challenge [30]. The IoU is defined by the following formula:

$$IoU = \frac{area(A \cap B)}{area(A \cup B)} \quad (5.1)$$

Where $area(A \cap B)$, area of overlap between boxes A and B is defined by:

$area(A \cap B) = width * height$ and $area(A \cup B)$, area of union of boxes A and B is defined by:
 $area(A \cup B) = area(A) + area(B) - area(A \cap B)$

Such that *width*: the width of the area of overlap corresponds to the closest right corner to the y-axis minus the farthest left corner to the y-axis of both rectangles A and B , and *height*: the height of the area of overlap corresponds to the closest bottom corner to the x-axis minus the farthest top corner to the x-axis of both bounding boxes A and B as illustrated in Figure 5.2.

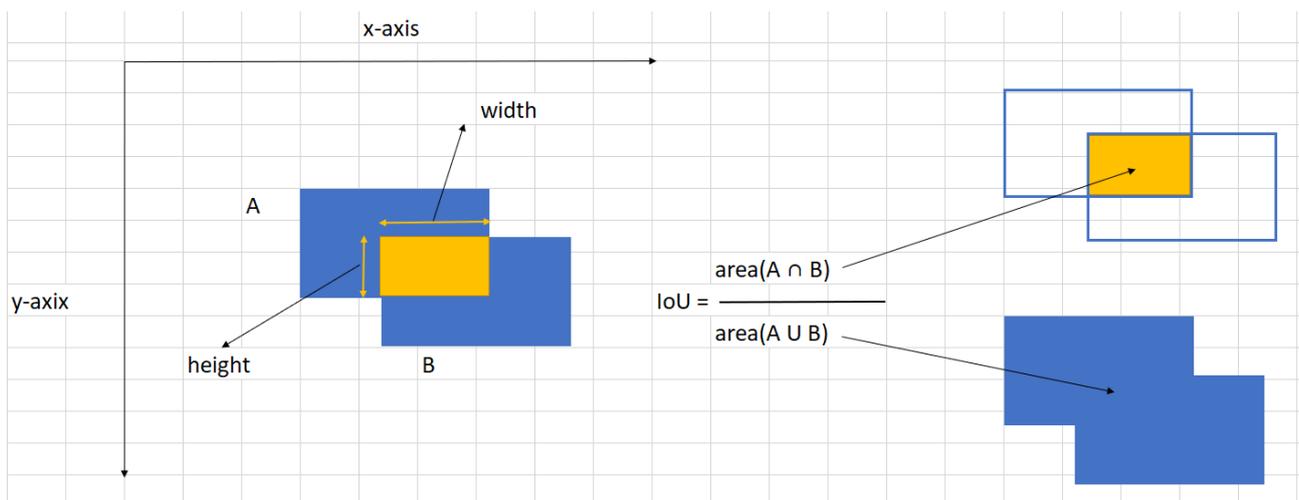


Figure 5.2: Intersection over union metric illustration. Blue rectangles A and B illustrate boxes locations of 2 objects A and B. yellow rectangles represent their area of overlap. Both blue rectangles merged into one represent their area of union

Therefore, based on the IoU overlapping criterion, total number of TP and FP boxes were counted, and models' precision was derived according to Equation 3.3. Equally, the GTCR of the different models were measured like the following:

$$GTCR = \frac{GTPB}{TGTN} \quad (5.2)$$

where $GTPB$ = the number of Ground Truth nuclei covered by Proposal Boxes of an image, and $TGTN$ = the Total number of Ground Truth Nuclei of an image.

NB: At this stage, the performance of the different systems at predicting correct background patches were not assessed as the aim for proposal box generation is to make sure that all true nuclei locations are covered by the candidate proposals as much as possible prior to adequate object detection by category-specific classifiers.

Set of Experiments

For each object detection technique investigated, specific set of experiments were conducted. These experiments consist of proposal box generation by sliding window, random generation, objectness measure, training classifiers based multiple image cues over splits of labelled datasets, and fine-tuning algorithm parameters.

- Experiments of proposal generation by Objectness Measure

To investigate proposal box generation by objectness measure over biomedical images, three experiments were conducted: (i) the generation by sliding window and evaluation of fixed-size bounding boxes to predict nuclei locations using different strides in horizontal and vertical directions such as (1,2,10,20,50,60,80,100,160); (ii) the generation and evaluation of random fixed-size bounding boxes over biomedical images to predict nuclei locations following a normal distribution and using 1000 proposals per image 10 times; (iii) the generation by objectness measure and evaluation of 1000 fixed-bounding boxes to predict nuclei locations after training a Bayesian framework over 30 nuclei images. These experiments enabled us to access and compare the performance of blind searching of nuclei objects inside images by sliding window and random boxes versus a search guided by objectness criterion. Results of the three experiments are reported in Section 5.3.2.

-
- Experiments of proposal generation by Selective Search

To investigate the quality of object hypothesis based on a Selective Search using graph segmentation and hierarchical grouping, five types of proposal box generation experiments were conducted over NRVM nuclei images according to the five different colour channels from which the technique is derived. These experiments enabled us to assess the performance of the selective search technique and identify the best parameters applicable to nuclei images. Results of the five experiments are reported in Section 5.3.2.

- Experiments of proposal generation by Edge boxes

To investigate the performance of Edge-boxes based strategy for nuclei proposal generation in biomedical images, we adapted Piotr Dollar's codes [97] that describe a random Forest framework for edge strength learning over natural images to build, train, and test a similar framework applicable to nuclei datasets. Originally, the available codes were set to work on 8-bit depth images. We adapted them to work on 16-bit images and generate fixed-size nuclei locations based on 3 default parameters (α , β and $minimumscore$). The ground truth dataset was split into 75% for training and 25% for testing. Performances of a standard technique and a customized framework trained with biomedical datasets were reported and in results Section 5.3.2.

- Experiments of proposal generation by RPN

To analyse the quality of candidate nuclei locations generated by RPN over biomedical images, a Faster R-CNN nuclei detector based on Resnet50 CNN was implemented, fine-tuned, and tested. For specification restriction on our computer system (Dell XPS, Intel Core i7, RAM 16 GB, NVIDIA GPU GeForce GTX 1050 Ti, 8GB, OS windows 10), we were able to train the deep artificial neural network (ANN) only on 40 images instead of 75% of 250 the labelled nuclei dataset described in Section 5.2.1. The main hyperparameters of the network were the following: Stochastic Gradient Descent (SGD) as optimizer, a step-size of 0.004, and 40 epochs in total. We used 2 anchors instead of 9 by default. The restricted ground truth dataset was split into 75% for training and 25% for validation and the remaining 210 nuclei images were used for testing. The training procedure enabled us to optimize by Stochastic Gradient Descent (SGD), convolution filters of ResNet50, RPN weights, and last fully connected layers weights for nuclei detection in pre-processed nuclei NRVM images.

During the pre-processing stage, input biomedical images with size 1040 x 1932 were

resized into $(224 \times 224 \times 3)$ according to the ResNet 50 input layer requirement. Labels of negative and positive samples of the Ground Truth dataset were updated accordingly. A set of data augmentation techniques such as horizontal-flipping and Poisson noise adding were also applied to increase the training size and quality.

Technically, performances reported in Section 5.3.2 describe a full object detector rather than RPN alone. However, the analysis of all categories of bounding boxes output by the system gives an insight on the power of CNN. Following Section 5.3 reports our theoretical and experimental investigations of the four state-of-the-art techniques for nuclei proposal generation over biomedical images.

5.3 Evaluating State-of-the-Art Algorithms for Region Proposal Generation in R-CNN Variants

In the past decade, several researchers reported techniques for generating category-independent region proposals based on various segmentation strategies and image cues. These techniques include the sliding window paradigm that can be enhanced by an objectness measure in image windows [1], the selective search based on graph segmentation [75] [29], constrain parametric min-cut [16], multi-scale combinatorial grouping [3], edge boxes [97] and RPN [96]. These state-of-the-art object recognition algorithms have recorded great performances over natural images they have been trained and tested on such as the Pascal VOC datasets. Here, we present a theoretical and empirical evaluation of those techniques over biomedical images for candidate object detection.

5.3.1 Description of the 4 State-of-the-art Techniques

The four object detection techniques we investigated differ essentially in two ways: a) the sampling procedure of candidate object locations and b) the probability measure of bounding boxes to contain objects. For some of them, the sampling procedure is carried exhaustively, and the scoring probability requires learning a classifier based on multiple features (objectness measure, and RPN), whereas, for others, the sampling procedure is more selective, and the scoring probability does not always involve any leaning function (selective search and edge boxes). The following subsections detail the sampling procedure and the probability measure

of bounding boxes generated by objectness measure, selective search, edge boxes and RPN.

Region Proposal Generation by Objectness Measure

The objectness criterion technique is built on the assumption that anything inside an image that stands out from its surrounding with a well-defined closed boundary, and a different appearance is an object [1] [2]. It is an exhaustive search, where the goal is to capture all possible object locations that are opposed to background elements like grass and sands. To output object hypotheses, the objectness measure technique randomly generates a maximum number of 100 000 bounding boxes at different scales and locations according to an objectness heat map computed all over images. The sampling procedure follows a random normal distribution in image regions showing high scoring probabilities. A non-maximum suppression (NMS) filter can be applied to control the sampling ratio based on proposal overlaps and objectness score. The objectness heat map i.e. the scoring probability of candidate object locations is performed by image patch classification through a trained Bayesian framework that combines four images cues: multi-scale saliency (MS), colour contrast (CC), edge density (ED) and superpixels straddling (SS).

The MS descriptor is derived from Fourier Transform spectrum of images and enables to capture image regions with a specific appearance. The CC descriptor enables to encode the dissimilarity characteristic of image regions with their surrounding based on the Chi-square distance between their Lab histograms. The ED descriptor allows measuring the closed boundary characteristic of object bounding boxes based on Canny edge detector. The SS descriptor plays the same role as the ED descriptor. Still, it relies on superpixel segmentation, i.e. the segmentation of image contents into small regions of similar colour or texture. For more details about these four image cues, please read [2].

Region Proposal Generation by Selective Search

As the objectness criterion, the selective search for guessing object locations in images, is data-driven and uses images cues to guide the sampling procedure, i.e. to identify images patches that are more likely to contain objects [87]. However, in contrast to the objectness measure that uses a trainable Bayesian framework, the selective search adopts an unsupervised learning approach that combines graph segmentation, k-mean clustering, and region growing or merging

techniques. When the selective search algorithm is run, image contents are first subdivided into small sets through graph segmentation; then the multiple regions are merged into bigger sets to form full object parts thanks to multiple similarities functions and random seed points selection.

In details, authors first apply Felzenszwalb and Huttenlocher segmentation algorithm [94] to generate initial object locations by graph representation of pixel points, then perform hierarchical grouping of candidate object locations based on four similarity functions computed at five different colour channels: Hue Saturation Value (HSV), Intensity or greyscale image, Lab, RG channels of normalized RGB plus Intensity (rgI), and Red Green Blue standard colour (RGB). The scoring probabilities of object hypotheses are denoted from their order of generation from the merging strategy. The four similarity functions consist of four different measures in image patches such as colour, texture, gap filling and size. The colour similarity derives from the one-dimensional histogram of each colour channel. The texture similarity is computed from the histograms of Gaussian derivatives of the different channels. The filling gap similarity enables to measure how close two regions can be based on their sizes and their bounding boxes.

All in one, those features associated to different similarity functions and random seed points are meant to enable diversification in the search for objects that will work well in different lighting conditions and image modalities. To know more about these image cues implemented in selective search for object detection, please read [87]. In the next section of this chapter, we shall describe a powerful data-driven technique for proposal box generation in R-CNN that is only based on a single image feature: object edges.

Region Proposal Generation by Edges Detection

As the selective approach, the edge-based technique for generating candidate object locations in images aims to narrow down the sampling procedure based on image structure. However, rather than segmenting images using a graph representation of pixel points prior to a hierarchical grouping of the different segments, authors of Edge Boxes perform Sobel-edge deconvolution over images and generate initial object locations only in places where strong object edges or contours have been denoted and enhanced. The object contours are enhanced by structured learning [24], and the candidate object locations are computed by sliding window at multiple scales and orientations. The objectness ranking of the candidate locations is built on the assumption that the number of contours that are fully contained in bounding boxes strongly

determines their probability to contain objects [97].

User input parameters associated with the edge box-based algorithm are the following: alpha (α), beta (β) and minimum score (ms) parameters. α is a percentage that controls the density of window samples. The higher is α , and the bigger is the number of proposal boxes generated. β represents a threshold value that controls the Non-Maximum Suppression filtering (NMS) of boxes that overlap, i.e. the parameter that greedily reduces the sampling density. The last parameter ms is also a percentage value that determines the minimum score of output boxes to contain an object. The default parameters ($\alpha = 0.65, \beta = 0.75, ms = 0.1$) mean for a window density of 65%, all boxes with minimum scores and that overlap 75% are disregarded, whereas each output box should have at least an object score of 10%.

All in one, though the Edge boxes technique is a non-exhaustive strategy solely based on a single low-level image feature, it has achieved high performances over natural images in Pascal VOC challenges. The next section of this chapter presents RPN, another state-of-the-art object detection technique that relies on a richer and more complex feature extraction model called CNN feature maps for region proposal generation in R-CNN variants.

Region Proposal Generation by RPN

In contrast to the first three techniques presented above that can be detached from R-CNN variants, RPN is a joint module that is hardly detachable from the R-CNN models it is built-in. As its name suggests, RPN is a neural network designed for computing fast and accurate object locations from complex feature maps through multiple convolutional filters. Here the sampling procedure does not perform any edge detection, nor graph segmentation and objectness measure through multiple image cues over input images. In contrast, the sampling is more exhaustive and essentially relies on a principle called anchor boxes. Anchor boxes are nine rectangular locations that are generated around central points at various scale and orientations. In RPN, candidate object locations are generated in the following manner:

- Compute anchor boxes by sliding multiple windows around pixel points at a stride of 16. By default, for each anchor point, nine different boxes are computed according to three scale values [8, 16, 32] and three aspect ratios [0.5, 1, 2].
- Classify each anchor box into foreground or background and learn the offsets for the

foreground boxes to be adjusted. The adjustment is then achieved by cross-entropy and smooth L1 loss functions through respective classification and regression layers. The location and score differences between ground truth boxes and anchor boxes are the inputs used to train the model. To learn more about the whole RPN training and deployment process, please read [67].

To sum up, region proposals generation by RPN combines the strength of an exhaustive search through multiple anchor boxes to the power of deep neural network through box classification and regression of convolutional features. In the following sections, we report and discuss the performances of the four state-of-the-art object detection techniques for nuclei proposal box generation over biomedical datasets described above.

5.3.2 Experimental Results on Biomedical Datasets

The four state-of-the-art techniques for region proposal generation have different sampling procedures and scoring probabilities, i.e. they follow different techniques to identify and rank object of interest in images as explained above. This section systemically evaluates their performances over biomedical images datasets.

Performances of Region Proposal Generation by Objectness Measure

As explained in Section 5.3.1, the objectness measure assumes that objects can appear everywhere inside images at different scales and orientation according to four essential cues that are multi-scale saliency, colour contrast, edge density and superpixels straddling. Therefore, to assess the quality of candidate nuclei locations generated by objectness measure over our nuclei datasets, we first measured the performances of a sliding window search with different strides at predicting nuclei locations in biomedical images, and those of a random search following a normal distribution. Then we measured the performances of a customized objectness measure search trained and tested over splits of our nuclei datasets for comparison.

For these first two experiments, we recorded the following performances: a) The sliding window strategy generates in average 155238 boxes per image with an average precision of 4.12% and an average GTCR of 85.81% at multiple strides values of (1,2,10,20,50,60,80,100,160) as reported in Table 5.1 and Figure 5.3. b) The analysis of a fixed number of 1000 of proposal boxes randomly

generated by normal distribution shows an average precision of 5.88% and an average GTCR of 3.33% per image as reported in Figure 5.4.

Stride	Windows	Run-time	TP	FP	Precision%	GTCR%
1	1 087 824	0.08138	47899	1 039 925	4.40	100
2	274 881	0.032032	12099	255 782	4.40	100
10	13 820	0.016686	597	13 223	4.32	100
20	5 574	0.016193	234	5 340	4.20	100
50	3 234	0.016125	129	3 105	3.99	99.73
60	3 099	0.016201	123	2 976	3.97	98.15
80	2 976	0.016289	118	2 858	3.97	86.47
100	2 901	0.016102	114	2 787	3.93	63.59
160	2 832	0.016137	111	2721	3.92	24.36

Table 5.1: Performances of nuclei proposal box generation by sliding window.

Each row of Table 5.3.2 illustrates respectively the step value taken in both horizontally and vertically to generate $[160 \times 160]$ fix-candidate locations, the number of proposals generated per image, the runtime of the algorithm, the number of true positive predictions (TP), the number of false-positive predictions (FP), the precision and the GTCR of the boxes in percentage.

Figure 5.3 illustrates the performance of the sliding window strategy at predicting nuclei locations in biomedical images. The bar histogram shows in x-axis coordinates, the number of proposal boxes generated per stride per image, and in y-axis coordinates, the average precision recorded over 250 nuclei images for each stride value. It can be noticed the total number of proposal boxes generated varies between 1 087 824 to 2832 while the precision of the boxes stagnates between 4.4% to 3.9%. The other illustrations represent respectively a ground nuclei image with 16 nuclei labelled in yellow boxes, then 13820 proposal boxes generated by sliding windows in a step of $[10 \times 10]$ and labelled in black boxes, and their comparison with the ground truth which gives 840 TP, 12980 FP and 6% of precision.

Figure 5.4 illustrates the performance of the random search strategy at predicting nuclei locations in biomedical images. The bar histogram shows in x-axis coordinates, the number of proposal boxes generated per stride per image, and in y-axis coordinates, the average precision recorded over 250 nuclei images for each stride value. It can be noticed that for a fixed number of 1000 of random proposals generated by normal distribution ten times; the average precision

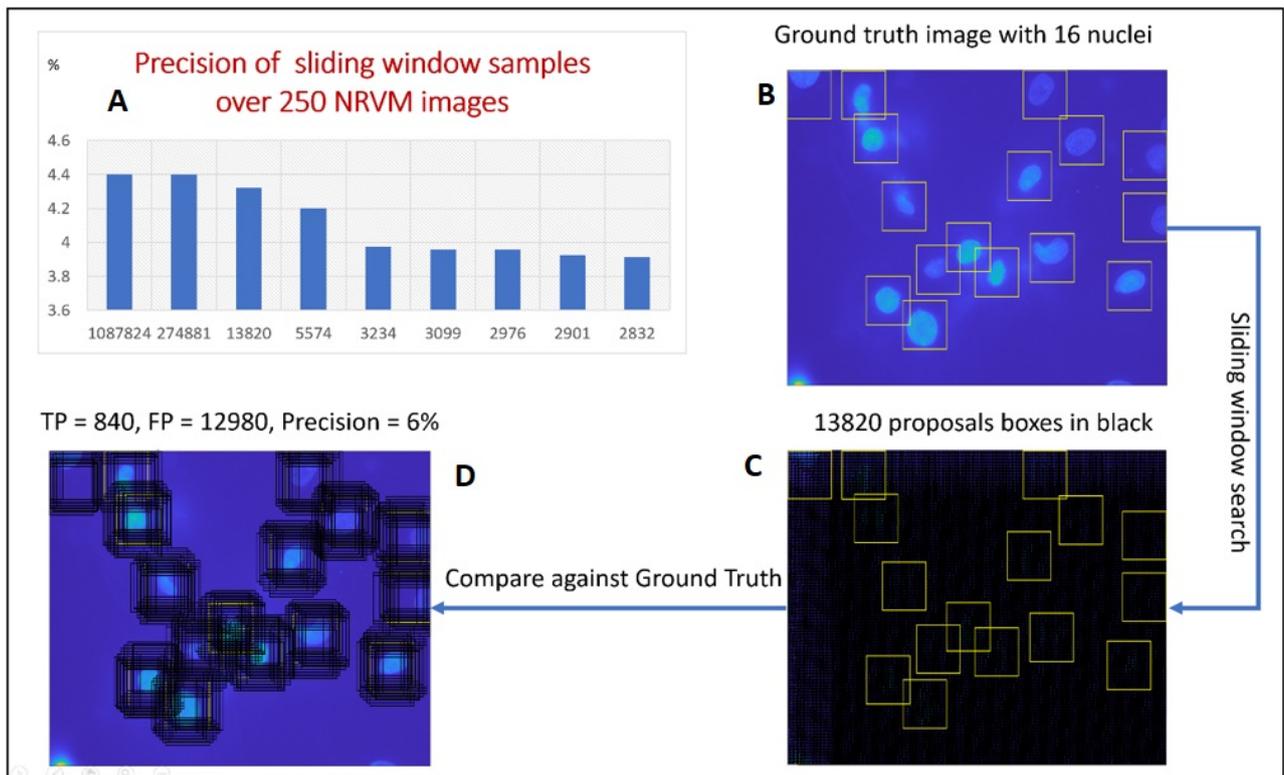


Figure 5.3: Performances of sliding window search for nuclei locations. A: overall precision of proposal boxes generated by sliding fixed-size windows over 250 NRVM images using various strides. Number of proposals generated are below each bar of the histogram. B: an original input NRVM image with 16 ground truth nuclei annotated under experts' guidance. C: 13820 proposal boxes generated by sliding fixed-size windows at step $[10 \times 10]$ over the input image. D: Comparison of the proposals with the ground truth enables performance measure in terms of TP, and FP boxes, and precision.

of the boxes stagnates between 5.75% to 6%. The other illustrations represent respectively a ground nuclei image with 16 nuclei labelled in yellow boxes, then 1000 proposal boxes randomly generated by a normal distribution and labelled in black boxes, and their comparison with the ground truth which gives 75 TP, 925 FP and 7.5% of precision.

After training an objectness model described in [1] and testing over splits of our biomedical datasets, we recorded an average precision of 1.5% and a GTCR of 73.4% when generating a fixed number of 1000 proposal boxes per image as reported in Figure 5.5. When generating an increased fix number of 2000 and 3000 of proposal boxes per image respectively, the average precision dropped down to 0.9% and 0.6% while the GTCR remains at 84.98%. For memory restriction, the dataset was split into two separate sets: 30 images allocated to training and 220 images allocated to testing as described in the experimental protocols in Section 5.2.

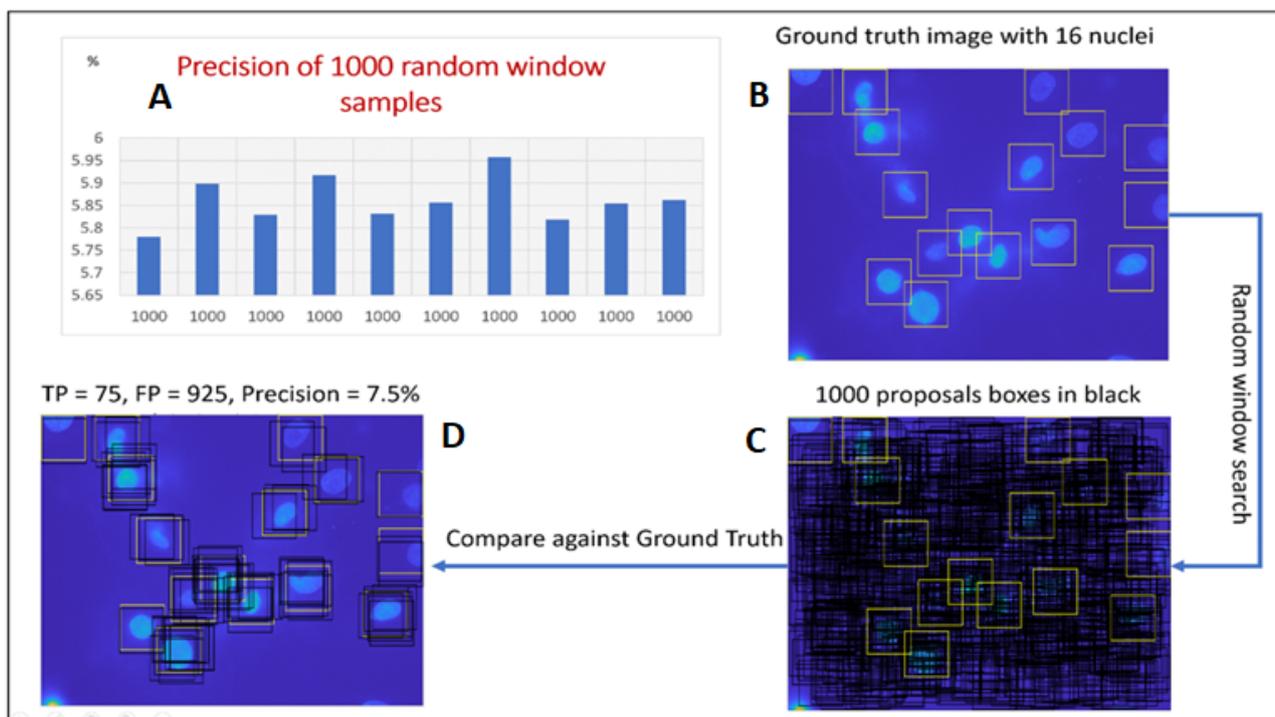


Figure 5.4: Performances of random window search for nuclei locations. A: overall precision of 1000 fixed-size proposal boxes randomly generated by normal distribution over 250 NRVM images using various seeds. Number of proposals generated are below each bar of the histogram. B: an original input NRVM image with 16 ground truth nuclei annotated under experts' guidance. C: 1000 random proposal boxes generated by normal distribution at seed 1 over the input image. D: Comparison of the proposals with the ground truth enables performance measure in terms of TP, and FP boxes, and precision.

Figure 5.5 illustrates the performance of the objectness measure search strategy at predicting nuclei locations in biomedical images. The lines chart shows in x-axis coordinates, the indexes of the 220 images used to test the system, and in y-axis coordinates the average precision and GTCR of the proposal boxes recorded for each image. It can be noticed that for a fixed number of 1000 of candidate locations generated by objectness measure; the average precision of the boxes is less than 2% while the GTCR is set to 73.4%. The other illustrations represent respectively a ground nuclei image with 16 nuclei labelled in yellow boxes, then 1000 proposal boxes with various sizes generated by objectness measure and labelled in black boxes, and their comparison with the ground truth which gives 15 TP, 985 FP and 1.5% of precision.

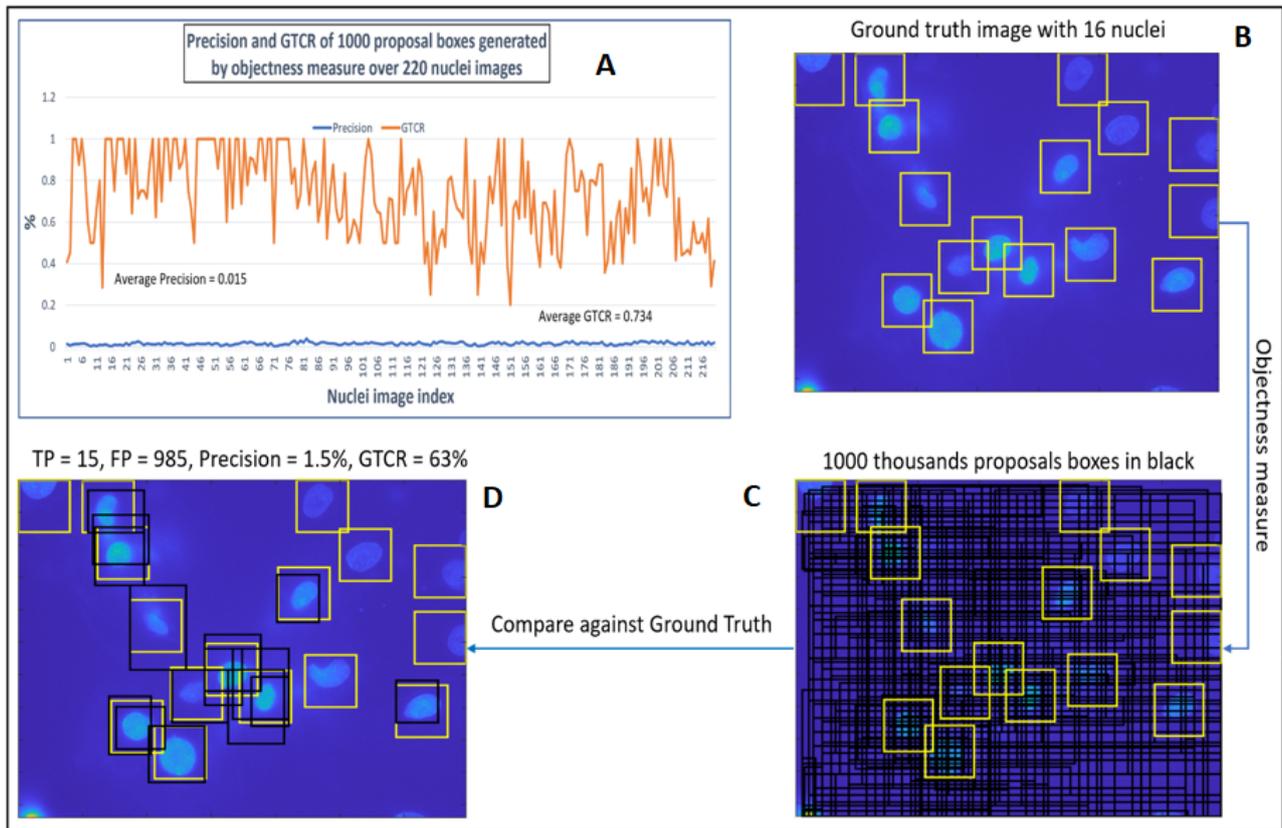


Figure 5.5: Performances of objectness search for nuclei locations. A: overall precision (blue line) and GTCR (amber line) of 1000 proposal boxes generated by Objectness measure tested over 220 NRVM images and trained over 30 NRVM images. B: an original input NRVM image with 16 ground truth nuclei annotated under experts' guidance. C: 1000 proposal boxes generated by Objectness Measure over the input image. D: Comparison of the proposals with the ground truth enables performance measure in terms of TP, FP boxes, GTCR, and precision

Performances of Region Proposal Generation by Selective Search

As stated in Section experimental protocols, we performed five different types of proposal generation experiments to analyse the quality of object hypotheses based on selective search according to the five different colour channels-based merging techniques used by the method (HSV, RGI, H, Intensity, Lab). Since the selective search algorithm requires two user-input parameters: the minimum segment (k) size and the graph threshold (σ), a grid-search was later performed to fine-tune those parameters according to the best combination of similarity functions and colour type applicable identified earlier.

In that sense, we found out that the best combination of the similarity functions and colour type parameters applicable to our NVRM dataset consists of four similarity functions (two

textures, and two-size variant criteria) and RGI colour type with fine-tune parameters $k = 80$ and $\sigma = 0.6$. With those parameters, the selective search-based technique achieved a maximum average precision of 4.25% and a GTCR of 60.44% as reported in Figure 5.6.



Figure 5.6: Precision of Selective Search for nuclei locations. A: overall precision of fixed-size proposal boxes generated by Selective Search over 250 NRVM images using various colours channels. Number of proposals generated are below each bar of the histogram. B: an original input NRVM image with 16 ground truth nuclei annotated under experts' guidance. C: 91 proposal boxes generated by Selective Search over the input image. D: Comparison of the proposals with the ground truth enables performance measure in terms of TP, and FP boxes, and precision.

Figure 5.6 illustrates the performance of the selective search strategy at predicting nuclei locations in biomedical images. The bar chart shows in x-axis coordinates, the average number of proposal boxes generated per colour type and using four similarity functions with the fine-tuned parameters $k = 80$, $\sigma = 0.6$, and in y-axis coordinates the average precision recorded for each colour type. It can be noticed that for an average number of 149 of candidate locations generated by Selective Search over biomedical images; RGI colour-based segmentation strategy outperforms with an average precision of 4.25%. The other illustrations of Figure 5.6 represent

respectively a ground nuclei image with 16 nuclei labelled in yellow boxes, then 91 proposal boxes with various sizes generated by Selective Search and labelled in black boxes, and their comparison with the ground truth which gives 13 TP, 78 FP and 14.28% of precision.

Performances of Region Proposal Generation by Edge Boxes

As reported in the experimental protocols Section, two types of experiments for nuclei proposal generation were performed on biomedical images to assess performances of Edge boxes at finding nuclei locations. In the first type of experiment, a standard MATLAB implementation of edge boxes was used to generate nuclei proposal boxes from 250 images stained with DNA marker. The average precision recorded was equalled 15.4% for around 21 proposals per image with a GTCR of 45,36% on the entire dataset as illustrated in Figure 5.7.

In the second type of experiment, a customized fast edge detection technique using structured forests [22] was used to generate nuclei proposals on splits of biomedical images stained with DNA marker. The average precision recorded for this second experiment increased to 30.61% with a GTCR of 95.43% as illustrated in Figure 5.8. However, the average number of proposal boxes per image is drastically high (8364), i.e. too big for efficient object detection.

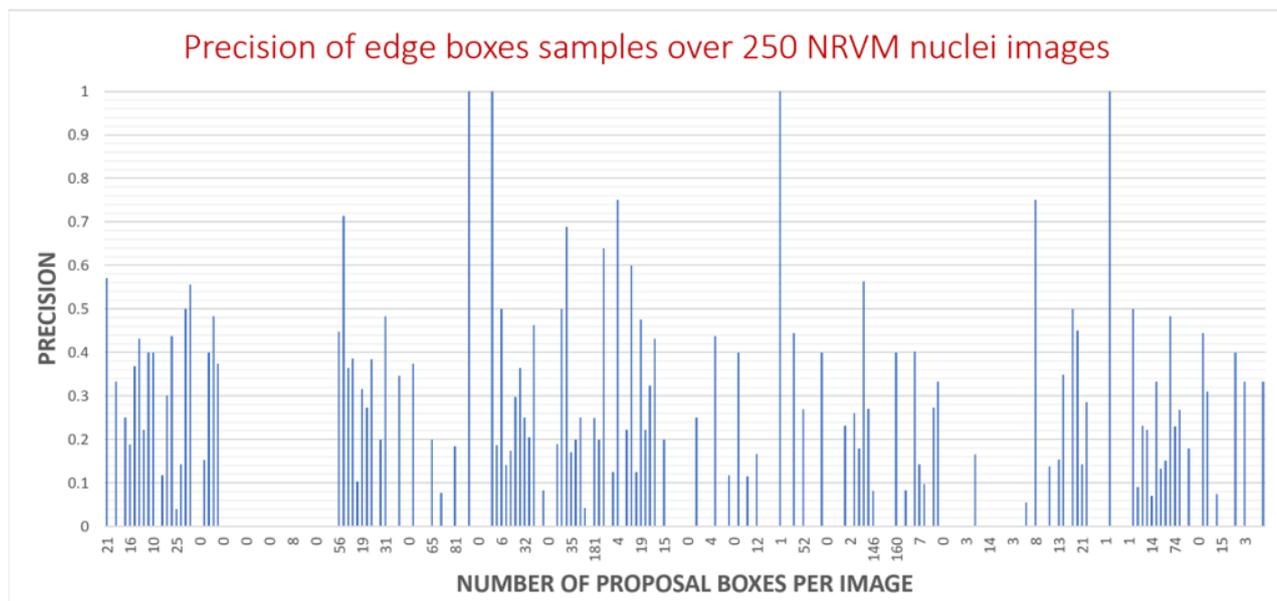


Figure 5.7: Performances of standard edge box algorithm for nuclei proposal generation. Average precision = 15.39%.

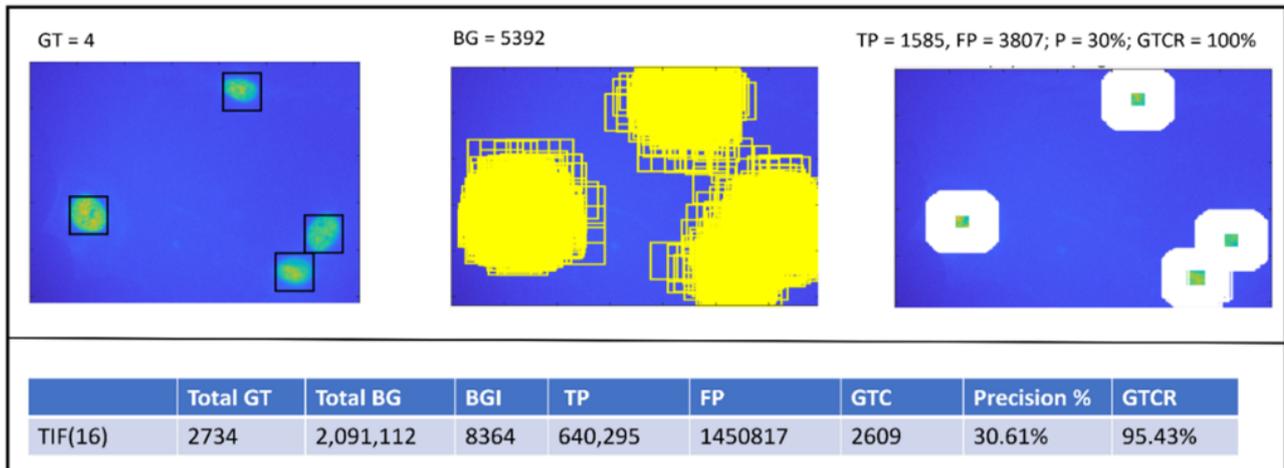


Figure 5.8: Precision of customized edge box algorithm for nuclei proposal generation. GT stands for Ground Truth nuclei; BG for proposal boxes generated, BGI for average BG per image, TP for True Positive, FP for False Positive, P for precision, and GTCR for Ground Truth Coverage Ratio.

Performances of Region Proposal Generation by RPN

To assess the performance of RPN at finding nuclei proposals in biomedical images stained with DNA marker, a faster R-CNN model was built and tested over splits of our labelled dataset like described in Section 5.2. Early training of 10 epochs achieved 30% of average precision. Longer training sessions made of 40 epochs enabled 66% of average precision on a reduced validation dataset made of 30 images, and 57% of precision over 220 unseen test images. Next Section 5.3.3 reports our discussions on the findings denoted from the experimental results.

5.3.3 Discussion of the Techniques Performances

As expected, the results of the different experiments of nuclei proposal box generation by exhaustive search or guided search (sliding windows, random boxes, objectness measure, edge boxes, selective, anchors and CNN boxes) enabled us to draw the five following observations and conclusions regarding to the four state-of-the-art techniques investigated.

- a) The sliding window search for biomedical objects in every location is secure but not realistic because of the number of image regions to check. The sliding window search is secure because it assumes that objects can appear everywhere in any size inside images. However, if images are big and the stepping size is not guided, looking at every potential

location inside images may quickly become an impossible task due to the number of locations to visit. In our case study, experiments with different strides values of 1, 2, 10, and 20 reported in Section 5.3.2 show that millions of proposal locations of fix-size boxes can be generated with similar GTCR (100%) and precision (4%) as thousands of locations. These observations mean a waste of energy or computing resources as a system built on such a sliding window approach will imply checking a huge number of false positive locations, i.e. 96% of proposal generated. Moreover, experiments of generating 1000 fix-size windows by random normal distribution demonstrate to be more precise than an exhaustive search by sliding windows (5.88% versus 4.12%) despite a smaller average GTCR of 3.33% versus 85.81%. These poor results of an exhaustive search by sliding windows or randomness argue for a sampling rate guided by image contents such as the objectness measure strategy.

- b) The Objectness Measure strategy for finding nuclei proposals in biomedical images reduces the sampling rate but shows damaging low precision in overall to build an efficient object detector. The experiments reported in Section 5.3.2 demonstrate that with only 1000 proposals of various sizes, the Objectness Measure strategy achieves a GTCR of 73.4%. However, drawbacks remain at the precision of the proposals. The recorded precision of 1.5% means that only 15 locations are genuine over 1000 proposals. A proportion that keeps getting smaller as the sampling rate increases (0.9% of precision for 2000 proposals). These promising results led us to investigate another object detection technique based on image contents such as the Selective Search.
- c) The generation of candidate nuclei locations by Selective Search using graph segmentation and hierarchical grouping exhibits a smaller number of proposals (149 in average) with more precision (4.25%) but lower GTCR (60.44%) as reported in Section 5.3.2. Since our testing datasets was encoded in greyscale format, we argue that the Selective Search technique may under-perform on them because the technique relies on multiple colour channels to compute similarity functions and generate object proposals. These contrasting results argue in favour of a further investigation of object detection technique based on image structure such as edge detection using structured forests.
- d) The generation of nuclei hypotheses by Edge boxes using structured Forests shows the most promising results over biomedical datasets with an average precision of 30.61% and a GTCR of 95.43% as reported in Section 5.3.2. However, the huge number of proposals

generated per test image (8364) jeopardizes the usage of the technique into a R-CNN model to build an object detector taking as input the proposal boxes generated. This huge number of proposals boxes generated by the Edge Boxes approach is essentially due to overlapping tiles, a sampling ratio that might be controlled by fine-tuning user-input parameters α and β of the algorithm. α and β stand respectively for window density parameter and overlap threshold value.

- e) CNN filters are powerful at capturing biomedical images features. The experiments of nuclei detection by Faster-RCNN model based on Resnet show that, despite a relatively small training dataset (28 images), CNNs enable to detect cells in 210 unseen images with a precision of 57% as reported in Section 5.3.2. A close analysis of the behaviour of the deep neural network detector on some problematic images showing weak boundary cells, various cellular shapes and heterogeneous background elements shows that CNNs generalize well across multiple image contents variability as illustrated in Figure 5.9(A) and Figure 5.9(C). The number of proposal boxes output by RPN is very low on average as illustrated in Figure 5.9(B), where for a single nuclei location, the system effectively proposes one single true location.

5.3.4 Techniques Comparison

Based on the results of the testing of the different proposal box generation techniques over splits of our nuclei ground truth datasets, the following qualitative comparison can be drawn as illustrated in Table 5.2.

Although the techniques have been tested on different splits of our dataset due mainly to computer system restriction, Table 5.2 demonstrates that Edge Boxes based strategy for finding nuclei proposals in biomedical images is the outstanding performer because it achieved the highest average GTCR (30.61%) and Precision (95.43%). This technique enables to narrow down the sampling rate of proposal boxes according to object contours learned by random Forest classification of image pixel points. Providing that images exhibit strong edges that can be easily detected, the system will show some limitations for images showing the extreme weak boundary and vague-contour cells like illustrated in Figure 5.10. Another drawback of the Edge Boxes approach is related to the huge number of proposals generated per biomedical images. Here, the average number of proposal boxes is equalled to 8364, which makes it hard

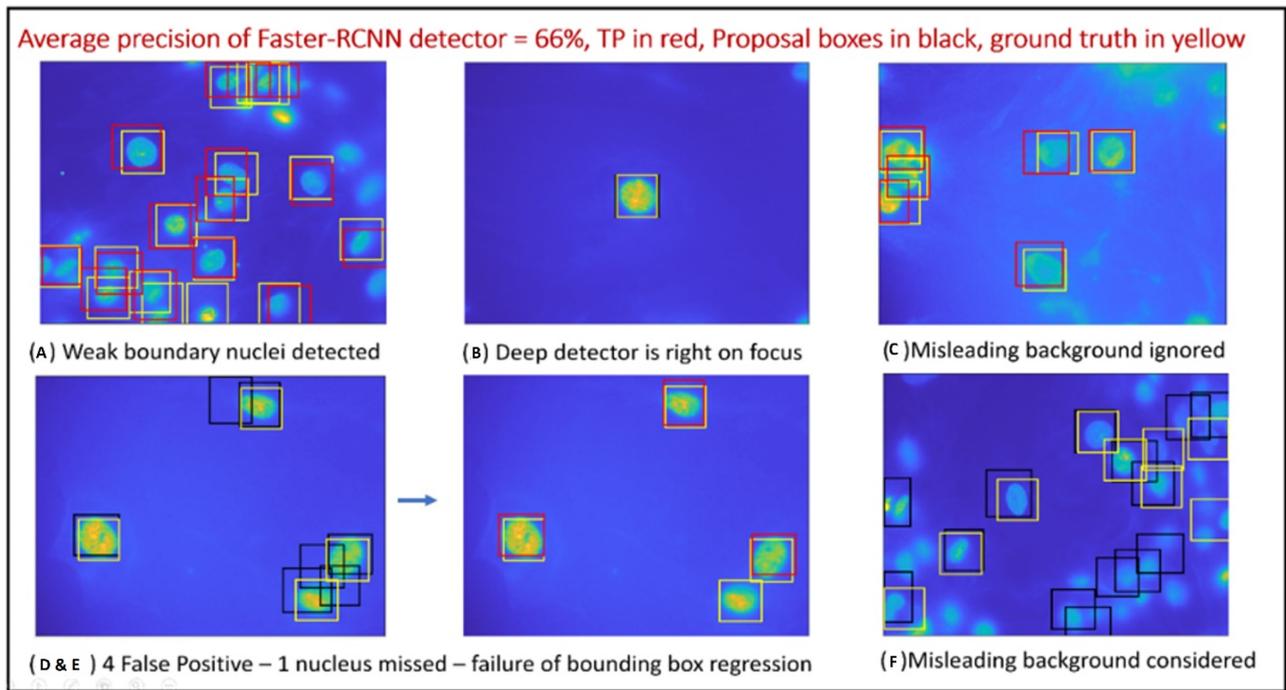


Figure 5.9: Faster R-CNN performance over nuclei images. A: Faster R-CNN ability at detecting weak boundary nuclei. B: Faster R-CNN ability at predicting the unique location of a single nucleus inside an image. C: Faster-CNN ability at ignoring misleading background with heterogeneous illumination and noise. D: predictions made by the system over 4-nuclei image. E: System failure at predicting the location of a nucleus with vague contour. F: System failure at ignoring misleading background patches.

to implement efficient object detection framework.

Figure 5.10 shows the output of the edge box method for nuclei proposal generation over two cells images. The first column of the two rows of the Figure represents the input NRVM images with nuclei ground truth locations highlighted in yellow boxes. The second column represents the initial edges maps output by the edge-based method for the generation of nuclei object proposals. For visual comparison ground, the yellow boxes are overlaid to show the location of the nuclei ground truth. The final column represents the output of the edge-based proposal method highlighted in red boxes against the yellow boxes ground truth nuclei. The proposal boxes derived from the edge strength obtained by Sobel filtering and classified by random forest. Variable performances are observed, as shown in the last column. In the first image the system identified only 6 nuclei over 16 ground truth through 60 boxes over a total number of 165 proposals. In the second image, the system fails to identify any single nuclei location as it did not generate proposal boxes. Visually, it is evident that the edge-based method performs

Techniques	Proposals	Window size	Precision %	GTCR%	Datasets
Sliding Windows	155238	[160 x160]	4.12	85.81	(0, 250)
Random Boxes	274881	[160 x 160]	5.88	3.33	(0, 250)
Objectness Measure	1000	variable	1.5	73.4	(30, 220)
Selective Search	149	variable	4.25	60.44	(0, 250)
Edge Boxes	8364	variable	30.61	95.43	(175, 75)
RPN/ Faster R-CNN	69	variable	57	57	(30, 220)

Table 5.2: Comparison of different techniques for nuclei proposal generation in biomedical images. Column Techniques: proposal generation techniques. Column Proposals: average number of proposals generated per testing image. Column Window size: size proposal boxes generated. Column Precision: precision of the boxes generated. Column GTCR: GTCR of the boxes generated. Column Datasets: pair of numbers telling the number of images used respectively for training and testing. For instance (0,250) means no training occurs as the technique is not trainable, and the system has been tested over the entire dataset made of 250 labelled images.

better over good contrast nuclei images (first-row output) than low-contrast one (second-row output). But overall, the method is not ideal because all the ground truth should be covered even if with a high number of proposal boxes.

The second-best proposal generation technique investigated according to Table 5.2 is the Objectness Measure approach which achieves an average GTCR of 73.4% against 1.5% of Precision. In contrast to the Edge Boxes that uses a single image feature, the Objectness Measure derives from multiple image cues (saliency, colour contrast, edge density and superpixel straddling) combined in a Bayesian Framework to find object hypotheses. As illustrated in Figure 5.11, though the Objectness Measure can exhibit great GTCR (100% in Figure 5.11(C)) over biomedical images, the average precision is still very low, i.e. the proportion of FP too high (984 FP over 1000 proposals Figure 5.11(B-C)).

Figure 5.11 illustrates both the power and weakness of Objectness Measure at searching for nuclei locations in biomedical images stained with DNA-marker. The illustration labelled (A) represents a FM image that shows 6 nuclei to be identified. The illustration labelled (B) represents 1000 proposal boxes generated by Objectness Measure at different scales and orientations to predict the 6 Ground Truth nuclei shown in (A). The illustration in (C) represents the correct

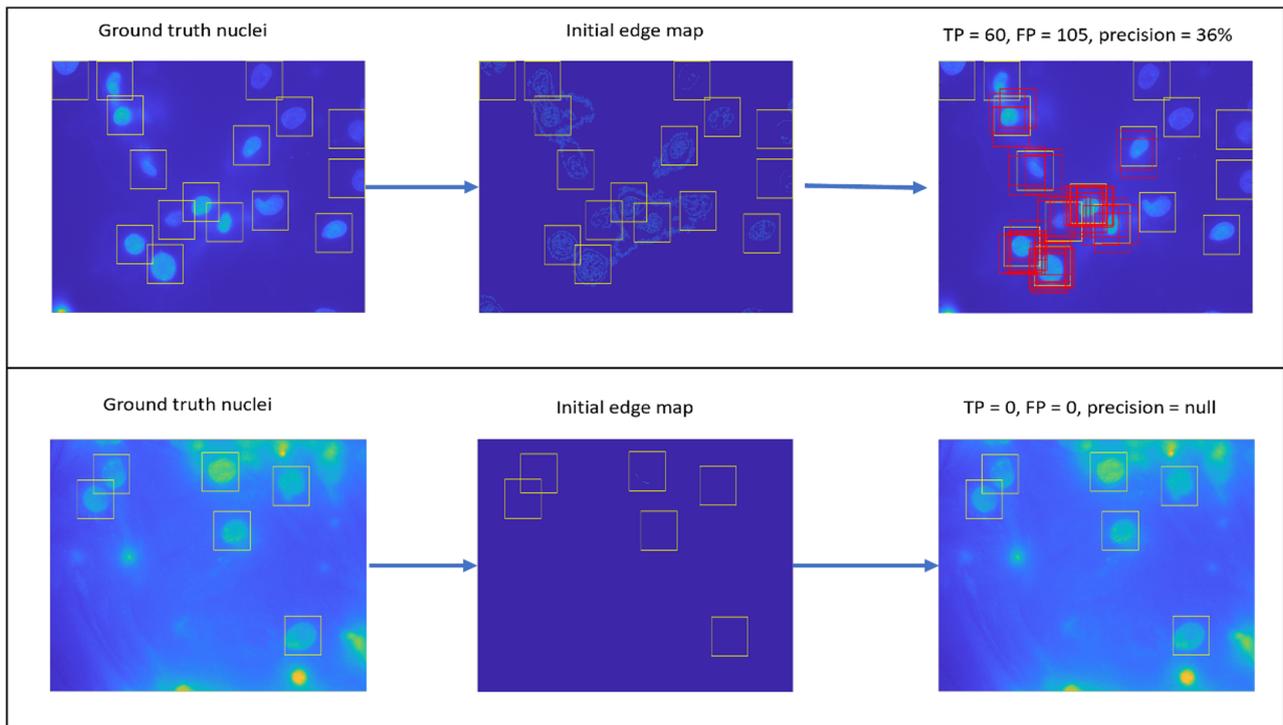


Figure 5.10: Standard edge box algorithm tested on NRVM nuclei images. yellow boxes represent ground truth location of nuclei set manually by biologist experts. Red boxes represent true positive nuclei locations generated by the system

predictions made by the Objectness Measure and the resulting Precision (1.6%) and GTCR (100%) recorded. Visually, it can be noticed despite 16 proposals that overlap with the six ground truth nuclei; there are still 984 proposals that can be classified as FP. This proportion of FP ratio is massively disadvantageous to Objectness Measure approaches when someone aims to build R-CNN object detector upon their outputs.

The remaining state-of-the-art object detection techniques investigated such as the Selective Search and RPN come respectively at the third and fourth position of Table 5.2. The former is a data-driven approach that combines the power of exhaustive search and the precision of image segmentation using graph representation and hierarchical grouping of image segments. The latter is also data-driven but relies essentially on anchor boxes and deep convolutional feature vectors to regress and classify image content into the foreground and background elements. Closed analysis of the Faster-RCNN model developed to assess RPN behaviour shows that CNN generalizes well across challenging image contexts. Indeed, despite hardware limitations to train the a ResNet 50 based Faster-RCNN, our system enabled to predict nuclei locations in images showing vague contour and weak boundary cells.

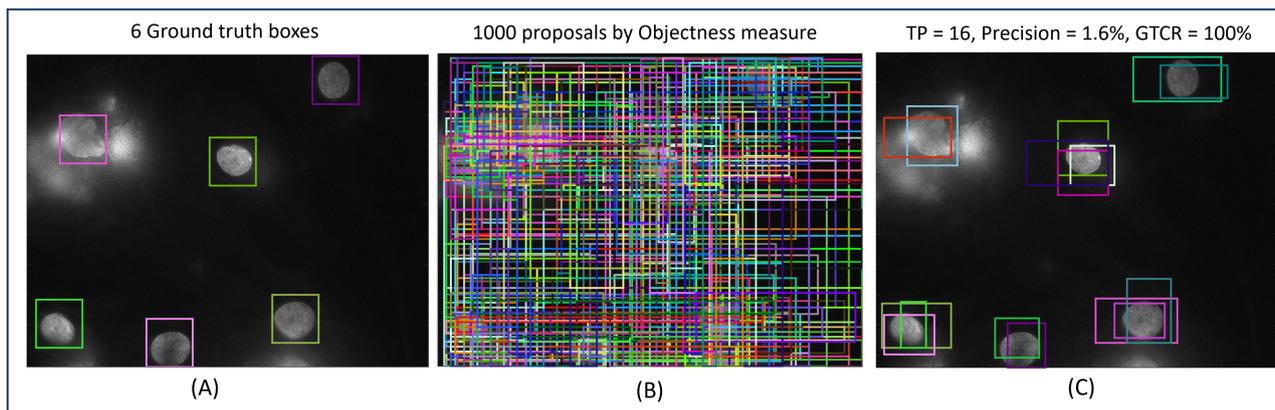


Figure 5.11: Objectness measure power and weakness at finding all nuclei locations. A: Original input NRVM image with reference nuclei locations highlighted by colourful rectangles. B: 1000 proposal locations computed by Objectness Measure to predict the true nuclei locations of A. C: Comparison of the proposals with the Ground returns 100% of GTCR versus 1.6% of precision.

Thus, knowing the fact that data-driven techniques such as Edge Boxes and Objectness Measure can achieve promising results, we made the following postulate: a customized R-CNN framework using an image-content based region proposal generation technique might be an efficient solution for automated detection of cell/nuclei in biomedical images.

5.4 Summary

Our investigations about various proposal box generation techniques used in R-CNN architecture for object detection in biomedical images show that the sliding window and the random techniques are time-consuming, exhaustive and in practice not realistic to implement because of the huge number of potential locations to be examined in high-resolution images. Objectness Measure, Selective search and Edge box-based algorithms, which were proposed as better alternatives to the sliding windows had limitations in terms of the number of FP ratio of nuclei proposals and the sampling rate when applied on biomedical images. For Objectness measure, the average precision is very low (1.5%). For selective search, the GTCR and the precision are both low (4.25% versus 60.44%). For Edge Boxes, the sampling rate per image is too high (8364 proposals in average) despite great GTCR of 95.44%. The analysis of RPN through Faster-RCNN model enabled us to measure the ability of CNN at capturing fine details of nuclei elements in challenging image context such as background heterogeneity, and illumination artefacts.

All in one, having in mind the fact that edge box strategy based on edge detection and random forest classification of edge strength can generate promising candidate locations for nuclei detection (30.61% of average precision), and the fact that CNN detectors are more precise in detecting nuclei objects in various conditions (66%), we formulated the hypothesis that a data-driven strategy for nuclei proposal generation combined with a deep classifier should enable us to achieve better performances for nuclei location proposals in FM images.

In order terms, a customized variant of R-CNN based on candidate bounding box generation algorithm by supervised learning should enable us to reduce the number of false-positive, i.e. the proportion of sampling windows where to search for nuclei objects while a deep classifier will ultimately detect the accurate nuclei locations. Custom R-CNN technique based on adaptive nuclei region proposal generation for nuclei detection in NRVM images obtained by FM is the main theme of Chapter 6.

Chapter 6

Customized R-CNN for Nuclei Detection based on Adaptive Region Proposal Generation

The main conclusion from the theoretical and empirical analysis reported in the previous chapter is that the variants of R-CNN based on proposal box generation algorithms tested on natural images are limited and inadequate to detect nuclei in biomedical images. Hence, one can argue for the need to take a step forward and investigate ways in which candidate nuclei locations could be generated and enhanced.

This chapter proposes a new scheme for adaptive region proposal generation that would enhance the accuracy of identifying candidate locations and reduce the number of sampling windows prior to ultimate image patch classification by CNN. The proposed scheme works in three steps. First, initial nuclei locations are generated by local adaptive thresholding, followed by a filtering step using linear Support Vectors Machines (SVM) fusion model. The SVM model is built upon four different features: Histograms of Oriented Gradients (HOG), Local Binary Patterns (LBP), Grey-Level Co-occurrence Matrix (GLCM) and Phase Congruency (PC). Finally, the filtered proposal boxes are classified into nuclei and background elements through a deep learning approach. This chapter extensively reports and discusses the experimental results about the three steps of the proposed scheme. Thus, its contributions are twofold:

1. Demonstrating that data-driven object proposal generation technique using local adap-

tive thresholding and SVM filtering based on hand-crafted features can help reduce the sampling rate of candidate objects while keeping a high level of sensitivity.

2. Demonstrating that a deep object detector for image patch classification integrated into a customized R-CNN model can help build a robust nuclei detection framework applicable to FM images.

The chapter consists of five sections. Section 6.1 is an introductory set of paragraphs that emphasize on the need for machine and deep learning approaches for nuclei detection in FM images. Section 6.2 presents an overview of our proposed solution for nuclei detection based on a customized R-CNN and describes in detail all the steps of the algorithm. In Section 6.3, the nuclei ground truth dataset, and the experimental protocols used to validate each step of our proposed solution are presented. Then the results and discussion on the findings are reported in Section 6.4. Finally, the chapter ends by a summary in Section 6.5.

6.1 Introduction

In recent literature, various techniques including deep semantic segmentation methods (U-Net [48]) and traditional methods such as thresholding [61], watershed [61], and LoG [57] have been proposed for nuclear detection and segmentation over several images modalities: histological tumours sections stained with haematoxylin and eosin (H&E), breast cancer images obtained by immunochemistry, and light microscopy Hela cells stained with cytoplasm markers. One major of the bottlenecks of these traditional methods is related to the fact that algorithms performances are dependent to the imaging datasets used for training and/or testing the techniques. For instance, thresholding, LoG filters and watershed ridges are sensitive to noise and therefore would be difficult to apply over new imaging datasets without prior knowledge as discussed in the literature review (Chapter 2);.

In contrast, deep neural networks approach for nuclear detection or segmentation [54], [25],[48] are widely believed to be robust against cell variability (size, shape, texture) and noisy artefacts, because of neural networks ability to learn rich image feature maps gathered at different levels of abstraction. Moreover, we demonstrated in Chapter 5 that RCNN models using data-driven technique to generate object proposal boxes could be an interesting path to follow as it would reduce the number of sampling rate and increase the accuracy of a nuclei object detector using

CNN features for classification [35].

The R-CNN based approach adopted in this chapter considers the nuclei detection task as a general object recognition problem where object locations are represented by their bounding boxes, i.e. the smallest rectangle that contains an object. To locate nuclei, we need to find the smallest bounding box that fully contains a nucleus which is different from the pixel-level classification method described in [48]. Based on the investigations presented in Chapter 5, we developed an R-CNN based strategy where proposal box generation is done in an adaptive way using thresholding and SVM classifiers and the final object detection is performed by CNN classification. Next Section 6.2 describes the overall scheme of our proposed solution for nuclei detection by adaptive R-CNN.

6.2 Proposed Solution for Nuclei Detection based on R-CNN

Our proposed solution is inspired from the guided search approaches investigated in Chapter 5 for object proposal generation, i.e. the Objectness measure [1], the Selective Search [87] [32] and the Edge Box strategy [97]. However, instead of relying on the same features or sampling procedure such as a Bayesian framework built on superpixel straddling, edge detection by Sobel filtering or a graph segmentation technique combined with a hierarchical grouping of image segments, our method generates initial object locations from adaptive thresholding and a combination of four image texture features, namely HOG, LBP, GLCM and PC to generate reliable candidate locations based on an SVM fusion framework.

The above four features were chosen to integrate variable texture related features into a nuclei proposal generation algorithm in order to maximise our chances to encode nuclei features that can be hard to capture in FM images due to illumination artefacts, noises, cell variability, and image modalities. The combination of HOG, LBP, GLCM and PC features makes our approach different than the selective search described in [87] [32] which used colour SIFT descriptors to perform graph-based segmentation and the edge box strategy that relies on raw pixel values to build a Bayesian framework [97]. After filtering the initial nuclei locations, candidate nuclei patches are fed into a CNN model for final classification as in the methodology described in [35] -using specific domain fine-tuning of R-CNN for natural objects detection in Pascal VOC

dataset. Figure 6.1 highlights the three main steps of our overall scheme for nuclei detection using customized R-CNN. Figure 6.2 that follows illustrates the outputs of the different stages of our algorithm.

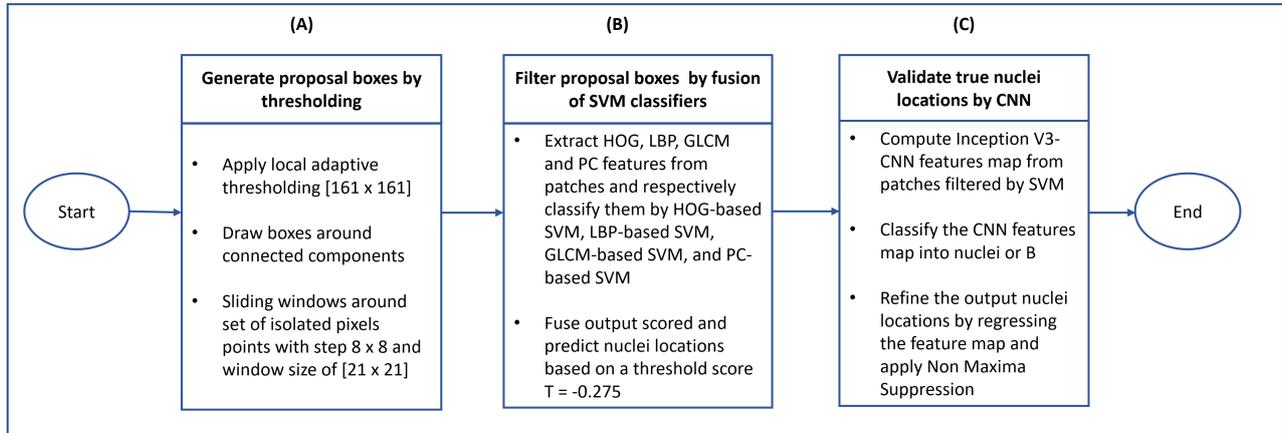


Figure 6.1: Overview of customized R-CNN for nuclei detection based on adaptive proposal box generation. (A): initial proposal generation stage by local adaptive thresholding. (B): feature extraction and SVM classification stage. (C): ultimate nuclei detection stage by CNN classification.

Figure 6.1 illustrates the overall R-CNN scheme adopted to detect nuclei in FM images based on three main steps. The first step (1) consists of generating from input images, initial nuclei locations based on intensity thresholding. This is done in two ways. First, local adaptive thresholding with a kernel of [161 x 161] is applied to images and proposal boxes are drawn around connected components with size greater than 16. Second, around sets of isolated pixels, further nuclei candidate locations are generated by a sliding window with a kernel of [21 x 21] and a step of [8 x 8]. Outputs of the first step (1) is illustrated by Figure 6.2(B and F) where from two input images Figure 6.2(A and E) are detected respectively 128 and 122 nuclei by Local Adaptive Thresholding.

The second step (2) of the proposed solution for nuclei detection by customized R-CNN consists of filtering the initial proposals by using hand-crafted features and SVM fusion. The initial proposal boxes generated by local adaptive thresholding correspond to image patches that are more likely to contain nuclei instances. However, segmentation by intensity thresholding means that all foreground elements should be distinguished from background thanks to their pixel intensity distribution [65], which is not always the case. The hand-crafted feature extraction step we included in our automated routine enables us to consider extra image features for describing foreground patches and background counterparts such as HOG, LBP, GLCM and PC.

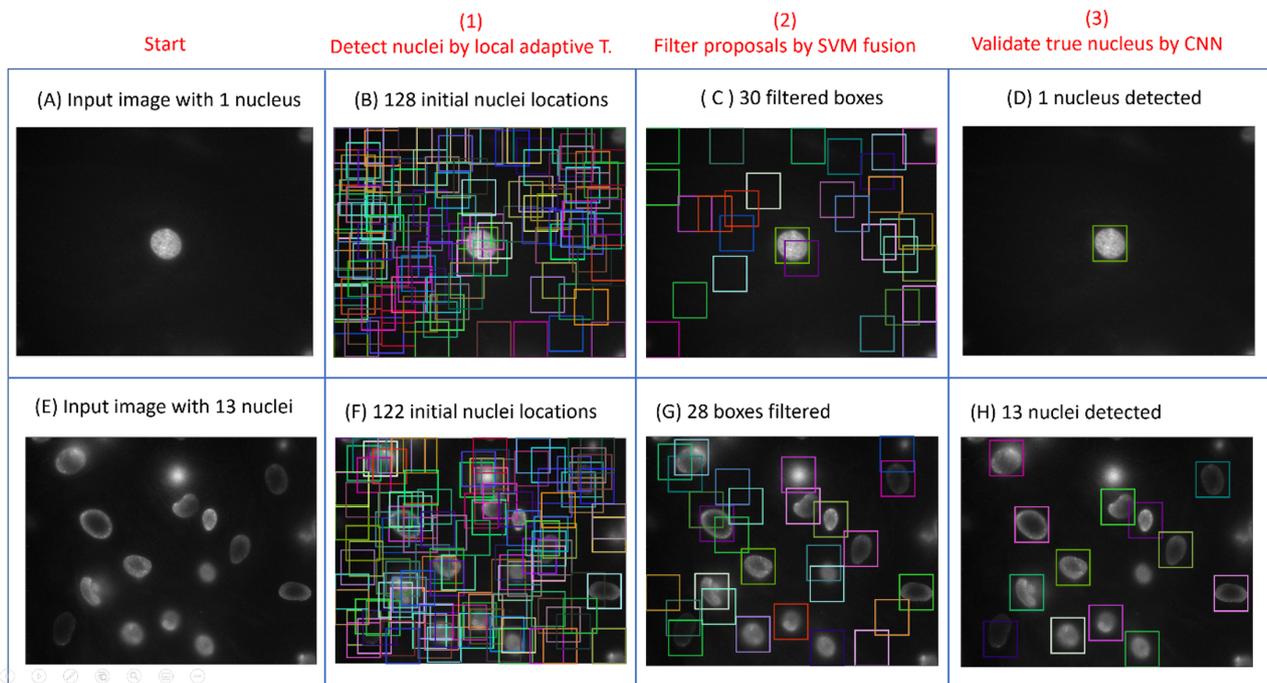


Figure 6.2: Output of the different stages of R-CNN based nuclei detection framework. (A,E): original input nuclei images. (B,F): results of initial proposal box generation by local adaptive thresholding. (C,G): results of proposal box filtering by SVM models. (D,H): results of fine-grained nuclei detection by CNN.

Once features vectors are extracted from initial candidate locations, the corresponding proposal boxes are filtered into nuclei and background classes thanks to trained linear SVM models (HOG-based SVM, GLCM-based SVM, LBP-based SVM, and PC-based SVM). We investigated various feature concatenation models, but experiences reported in Section 6.4 showed that a fusion model made at score level and built upon the four above individual models achieved the best results. This is illustrated in Figure 6.2(C and G) where respectively 30 and 28 boxes have been filtered out from 128 and 122 nuclei candidates initially detected by Local Adaptive Thresholding.

The last step of our proposed solution consists of validating the right nuclei locations previously generated by local adaptive thresholding and SVM classification by using CNN fine-grained classification. Compared to hand-crafted feature extraction, CNN enables transfer learning and to extract new rich feature maps with different levels of abstraction from image patches and thus may guaranty ultimately accurate object detector. Something that is hard to implement in hand-crafted feature extraction models without prior knowledge of the datasets or intensive feature engineering. Finally, to avoid duplicate locations of the same nuclei objects, a customized Non-Maxima Suppression algorithm is applied over the output of the CNN clas-

sification. NMS consists of keeping the bounding boxes with the highest objectness score, i.e. more likely to contain objects by ignoring those which overlap with them at a certain IoU ratio. Outputs of this last step (3) are illustrated by Figure 6.2(D and H) where from sets of proposals filtered by SVM fusions (Figure 6.2(C and G)), ultimate nuclei objects are detected by CNN classification. In the following subsections, we describe in detail the above three steps of our proposed solution based on R-CNN for nuclei detection in FM images.

6.2.1 Initial Nuclei Detection by Local Adaptive Thresholding

As demonstrated in Chapter 5, the guided search approach for generating object proposals in images is the most efficient way to build robust object detector based on R-CNN architecture because it may help reduce the sampling rate and prevent someone searching for objects everywhere inside images in an exhaustive manner. Based on such observation, we investigated a simple way of generating initial nuclei locations from the local variation of pixel intensity profiles in image regions, i.e. by local adaptive thresholding.

The principle of generating initial nuclei locations from intensity thresholding of NRVM images derived from two series of observations. First, when analysing intensity profiles of nuclei regions and non-nuclei regions, we noticed that in most images, there is a clear separation between background and nuclei patches in such a way that a global threshold value based on pixel intensity can help separate them. This is usually observed in images of cells showing the solid boundary and great contrast that demarcate them from their surrounding background like described in Figure 6.3.

Figure 6.3 illustrates good contrast nuclei images where nuclei regions are brighter than background elements. Figure 6.3(A) represents an example NRVM nuclei image with two ground truth labels: a background patch highlighted in blue and a nucleus patch highlighted in red. The two labelled regions are crossed by horizontal blue and red lines. Figure 6.3(B) represents the pixel intensity profiles of the labelled regions where the respective blue and red line cross the regions. As it can be noticed visually, the greyscale intensity of the nuclei region is brighter than the one of the corresponding background patches. Figure 6.3(C) represents the output of a global threshold using Otsu to segment the regions. Figure 6.3(D) shows the yellow bounding box drawn around the connected component obtained by global Otsu threshold and corresponding to the ground-truth nuclei labelled in Figure 6.3(A). Thus, in good contrast images with

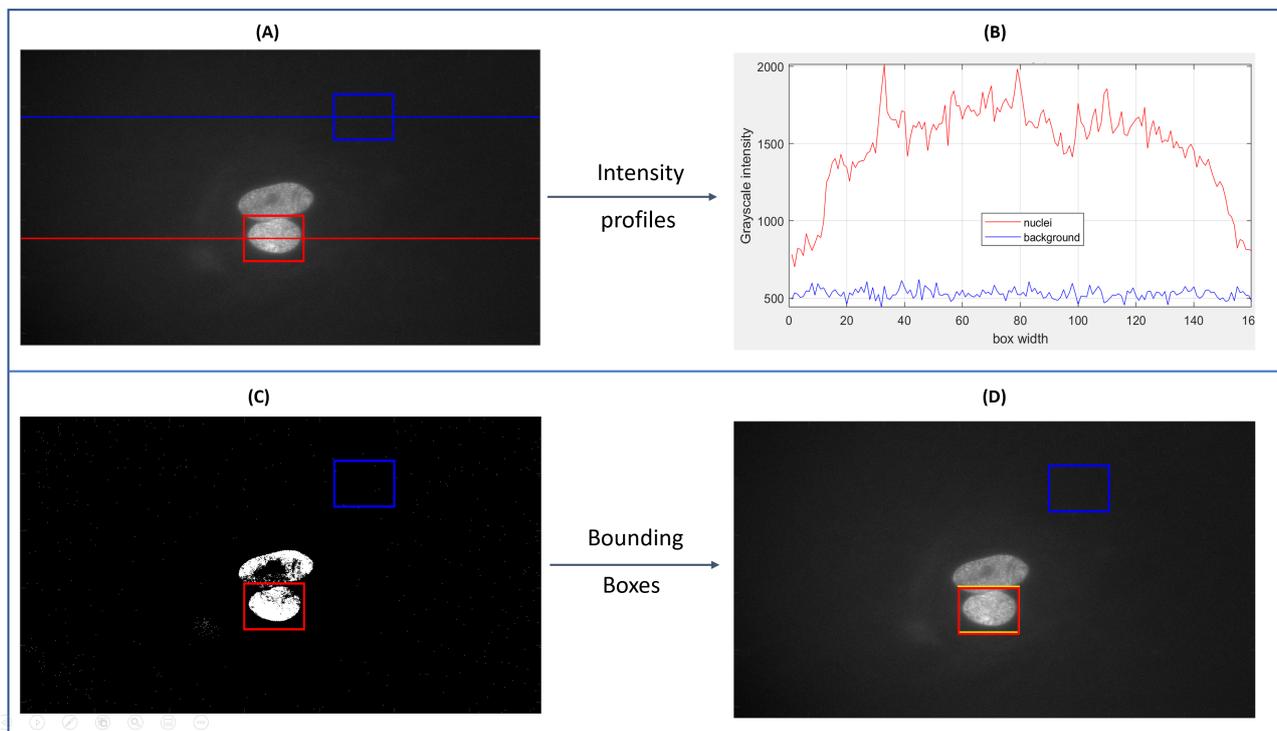


Figure 6.3: Good-contrast images: nuclei regions have higher intensity profiles than background patches' ones. (A): original input nuclei image 2 annotations: a nuclei patch in red and an image background patch in blue. (B): intensity profile comparison of both nuclei and background image patches. (C) binary image of global Otsu thresholding. (D) Comparison of the Otsu-derived bounding box with ground truth

clear pixel intensity demarcation between background and nuclei patches, a global threshold can be helpful. But this is not always the case in FM images.

The analysis of the intensity profile of some intriguing sample patches of our datasets that show little contrast between background and foreground elements revealed different patterns. In low-contrast images with heterogeneous backgrounds, some illuminations artefacts tend to be brighter than regions containing cells, as illustrated in Figure 6.4. Moreover, in some extreme cases, nuclei and background elements may share the same pixel intensity fluorescence. Those complex scenarios are denoted by crossing intensity profiles curves. In such cases, it is evident that global thresholding based on pixel intensity will fail to segment cells and background regions into two distinct sets.

That is why we opted for a local adaptive threshold approach as these techniques have shown great ability to capture local contrast variations in image regions regardless of their relative illumination [65]. Thus, we argued that nonetheless the illumination fluctuation of image

windows containing nuclei or background elements, nuclei regions are still brighter than their immediate neighbouring pixel elements corresponding to the relative image background in those image patches. In other terms, local adaptive thresholding can be an interesting starting point to guess rough locations of nuclei in FM images as illustrated in Figure 6.4(C).

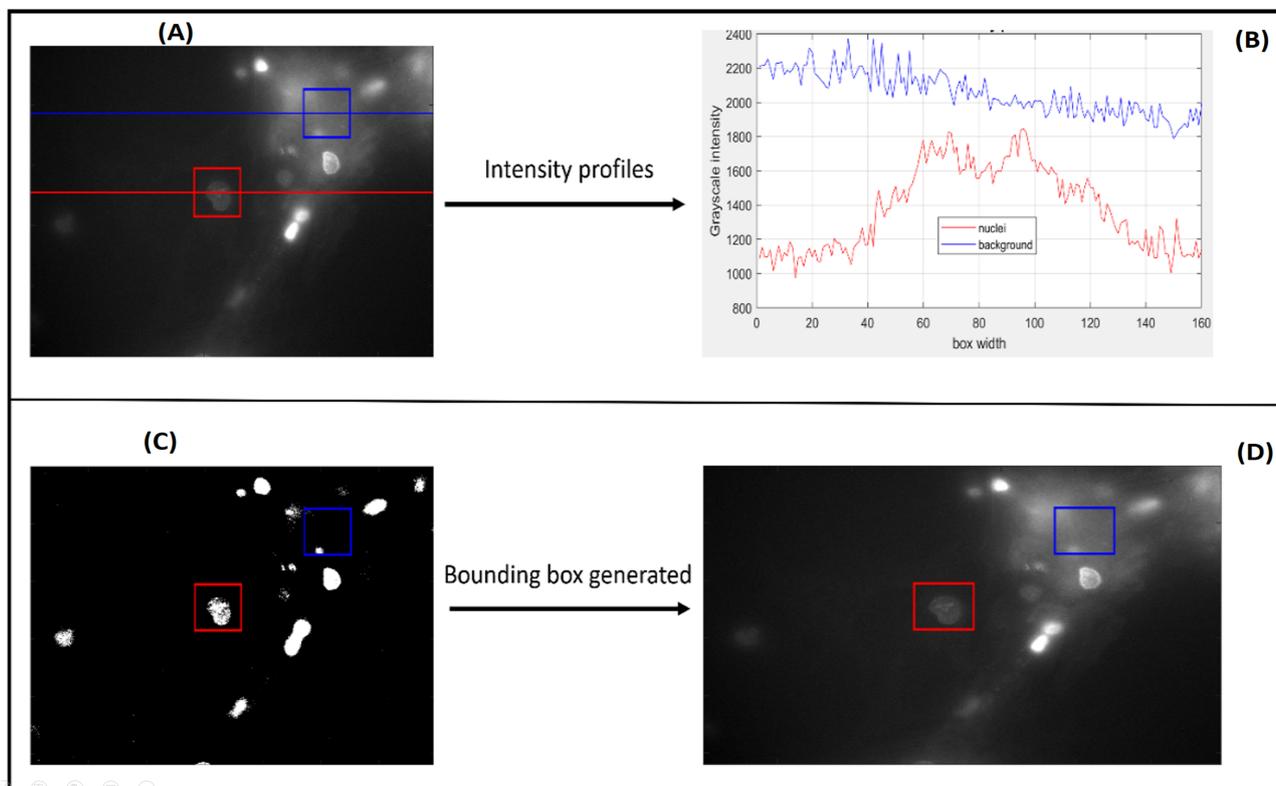


Figure 6.4: Low-contrast images: background contents are brighter than nuclei but handled by local adaptive thresholding. (A): original input nuclei image 2 annotations: a nuclei patch in red and an image background patch in blue. (B): intensity profile comparison of both nuclei and background image patches. (C) binary image of local adaptive thresholding. (D) Results of the local adaptive thresholding overlaid on original input image

Figure 6.4 illustrates the ability of local adaptive thresholding to segment nuclei in low contrast images where background patches look brighter than cellular regions. Figure 6.4(A) represents an example NRVM nuclei image with two ground truth labels: a background patch highlighted in blue and a nucleus patch highlighted in red. The two labelled regions are crossed by two horizontal lines: blue and red. Visually it can be noticed that the nucleus is fade contrasted and darker than the selected background region. Figure 6.4(B) represents the pixel intensity profiles of the labelled regions where the respective blue and red line across the selected regions. As it can be noticed visually, the greyscale intensity of the nuclei region is effectively darker than the corresponding background patch. Figure 6.4(C) represents the output of a local adaptive

threshold using the arithmetic mean of neighbourhood pixels (161x161). Here it is clear that local adaptive thresholding has wrongly detected a potential nucleus inside a background patch due to the heterogeneous illumination within that area of the input image. Figure 6.4(D) shows that bounding boxes are generated for each labelled region which means a false background element is segmented as nucleus. Therefore, local adaptive thresholding may be better at capturing fading nuclei regions than global thresholding, but bright heterogeneous background may mislead to false nuclei proposals.

To generate initial nuclei locations from local adaptive thresholding, our system performs the following steps:

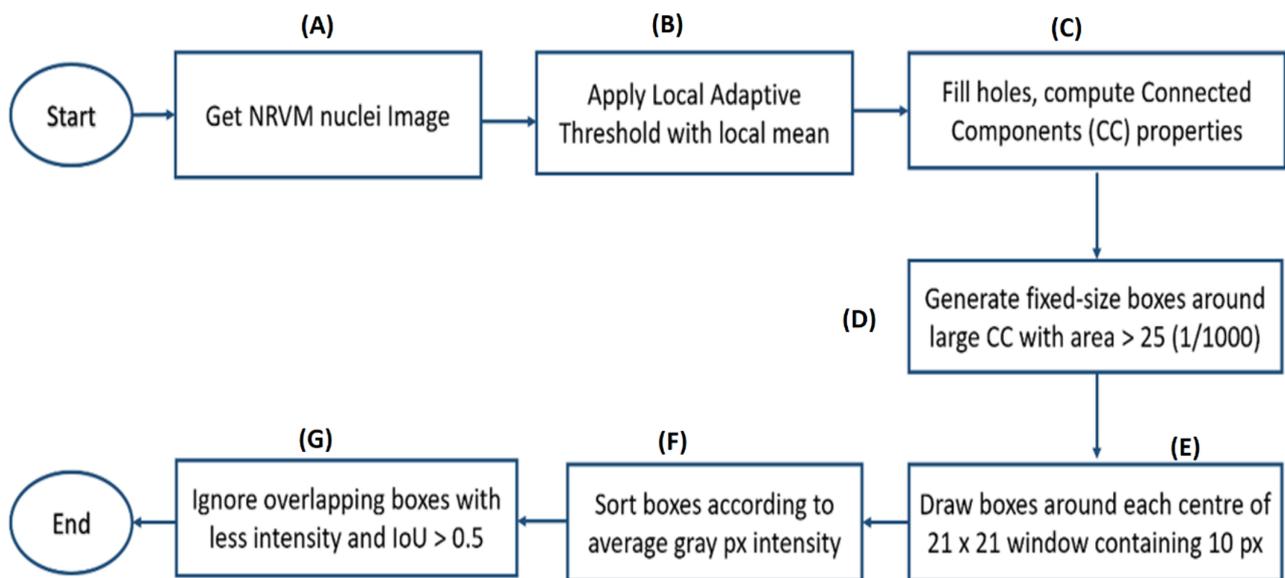


Figure 6.5: Flowchart diagram for initial nuclei boxes locations from adaptive threshold. CC stands for connected component obtained after adaptive thresholding of NRVM images with kernel 161 x 161.

Figure 6.5 represents the flowchart diagram of our routine for generating initial nuclei proposal boxes by adaptive thresholding. The routine follows six main steps labelled tasks B to G. First, over an individual greyscale image of NRVM nuclei, local adaptive thresholding using the mean statistic in a kernel of 161 x 161 is applied (Figure 6.5(B)). We selected this pair of odd numbers (161, 161) because our technique should deal with nuclei that cannot exceed 160 x 160-pixel size.

After local adaptive thresholding, the resulting binary images are further processed by morphological operations such as filling holes. Then, connected components are computed, and their

properties such as area and average greyscale pixel intensity are extracted (Figure 6.5(C)).

At the next step illustrated by Figure 6.5(D), the system begins generating initial nuclei locations by drawing fixed-size boxes of 160 x 160 around each connected component whose areas are greater than 25 (i.e. $160 \times 160 / 1000$).

At stage Figure 6.5(E), further nuclei locations are generated around groups of isolated pixels whose number of pixel elements are greater than 10 and fits within a window of 20 x 20. This second routine for nuclei proposal box generation around small connected components is performed in a sliding window manner with a step of 8 x 8. We chose to draw potential nuclei bounding boxes around groups of isolated pixels in a controlled-sliding-window manner because we wanted to make sure that no potential nuclei locations are left behind due to low illumination or low image contrast.

At stage Figure 6.5(F), the system sorts all generated nuclei proposal boxes in descending order according to the average greyscale pixel intensity of the region they refer to in the original input image. This is done to prepare for the final step of the routine corresponding to Figure 6.5(G), i.e. the filtering of all initial proposal boxes into non-overlapping windows according to their pixel intensity.

By default, when two image windows overlap, the brightest one is kept. This filtering of image windows assumes that in contiguous image regions brighter pixel points are more likely to belong to cellular regions. The image windows that are ignored correspond to the ones which overlap with the brighter image regions. The IoU ratio between regions is set to 0.5.

To sum up, our proposed technique for the generation of initial nuclei locations from local adaptive thresholding argues that despite disparities in background and nuclei illuminations, it is possible to generate relevant guesses of nuclei locations from connected binary components and set of white isolated pixels. The connected components selected are the binary objects whose areas are larger than $1/1000^{th}$ of the maximum size of nuclei within NRVM images. From sets of ten isolated white points located in a neighbourhood of 21 x 21 window, the system also generates potential nuclei locations. The performance of our method for generating initial nuclei locations from local adaptive thresholding is presented and discussed in Subsection 6.4. Figure 6.6 illustrates initial candidate locations output by our procedure based on local adaptive thresholding over an example nuclei image.

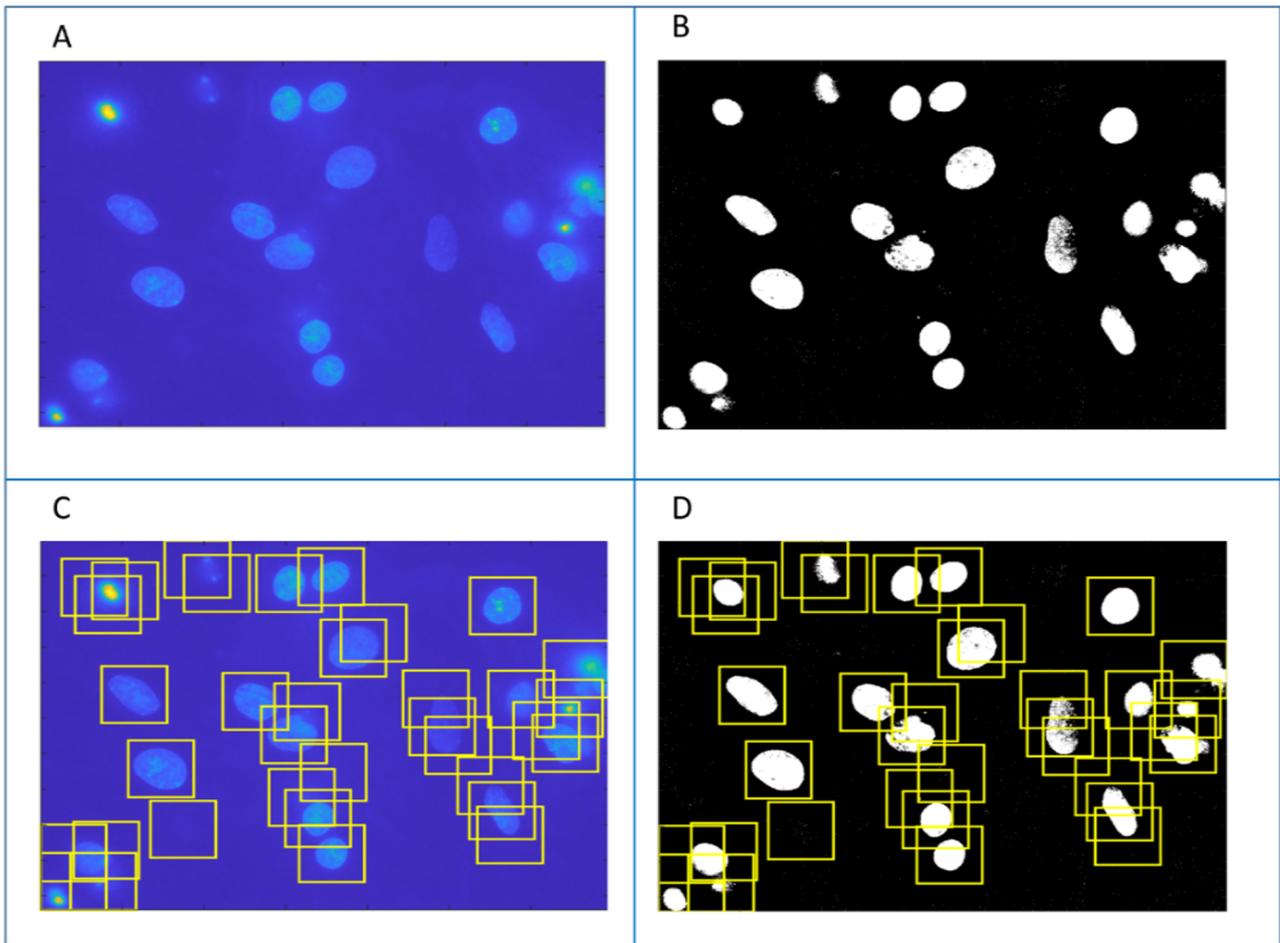


Figure 6.6: Initial nuclei proposal box generation by local adaptive thresholding. A: original grayscale input nuclei image displayed in pseudo colour. B: binary image after local adaptive thresholding. C: results of initial proposal box generation overlaid on original grayscale input image. D: results of initial proposal box generation overlaid on binary image after local adaptive thresholding.

Figure 6.6(A) represents an original grayscale image containing 14 right nuclei. The image has been displayed in pseudo-colour for visual purpose. Blue shades correspond to smaller pixel intensities, whereas warm or red shades represent higher pixel values. Figure 6.6(B) represents the binary map obtained when a local adaptive threshold is applied to Figure 6.6(A). Figure 6.6(C) represents 35 nuclei proposal boxes generated from the binary map in Figure 6.6(B) and overlaid on the original input image in Figure 6.6(A). Figure 6.6(D) represents 35 initial guesses of nuclei locations output by our system and overlaid on the binary map in Figure 6.6(B). Visually, it can be noticed that all the ground truth nuclei locations of the input image in Figure 6.6(A) have been covered by the initial guesses.

However, unnecessary locations were also generated around small bright background regions as

highlighted in the top-left corner and the middle-right border of Figure 6.6(C). These limitations linked to background fluctuations and unnecessary proposal boxes argue that pixel intensity features alone may not be enough to generate high-quality nuclei proposals in FM images. In Next Subsection 6.2.2, we introduce four powerful image cues: HOG, LBP, GLCM and PC features that may help produce better nuclei hypotheses than intensity features alone in image regions in an SVM fusion framework.

6.2.2 High-quality Nuclei Proposals Filtering by SVMs

To filter nuclei locations output by local adaptive thresholding, we investigated four types of feature (HOG, LBP, GLCM, and PC) and built SVMs models upon them. SVM was chosen due to its simplicity and effectiveness for binary classification. The aim of filtering the initial locations is to make sure that all patches that are more likely to contain nuclei are kept whereas those that are more likely to represent background patches are ignored. The filtering is performed by estimating the probability of candidate patches to be nucleus or not. In that way, our model aims to achieve a better sampling rate than the state-of-the-art object detection techniques investigated in Chapter 5 such as the objectness measure, the selective search, the edge boxes, and the RPN.

The flowchart diagram in Figure 6.7 illustrates our SVM filtering approach to generate high-quality nuclei proposals after initial generation by local adaptive thresholding. It implies to build a fusion model consisting of four linear SVM classifiers based respectively on HOG, LBP, GLCM and PC feature vectors and to perform hand-crafted feature extraction from image patches corresponding to the initial nuclei locations previously generated by local adaptive thresholding. As it can be seen from the diagram to output nuclei hypotheses, the fusion model first extracts HOG, LBP, GLCM and PC features from initial proposal boxes and normalised them. Then, each SVM model assigns one category and a score to the input patches' feature vectors. Scores of nuclei categories are set to positive sign, while scores of background categories are set to negative sign. Finally extracted feature vectors are classified into nuclei and background through the four different linear SVMs based on the fusion of all scores output by the four classifiers. Experiments demonstrated that the comparison with an optimum score threshold value of $T = -0.275$ enables the binary classification as illustrated in Figure 6.7 and demonstrated in 6.23.

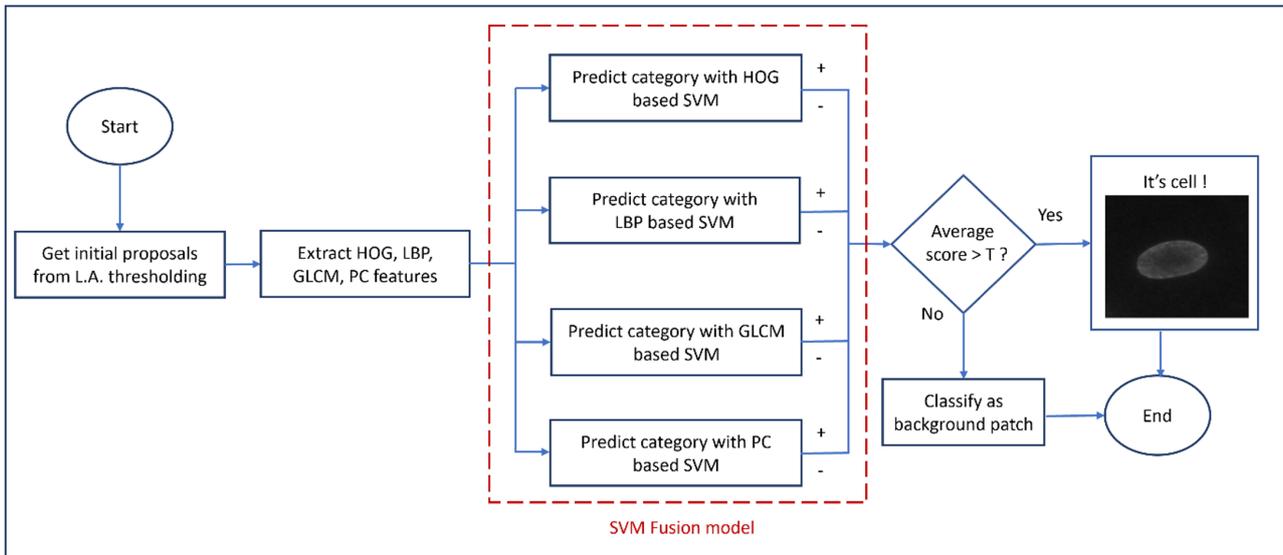


Figure 6.7: SVM fusion model for filtering initial nuclei proposal boxes. L.A. stands for Local Adaptive.

The performance of our method for generating nuclei proposals after SVM filtering of initial nuclei locations generated by local adaptive thresholding is presented and discussed in Subsection 6.4.2. In next Subsections 6.2.2 and 6.2.2, we describe the two types of SVM fusion models we built based on HOG, LBP, GLCM and PC.

SVM Models based on Fusion at Feature Level

The disparity of HOG, LBP, GLCM and PC feature response to challenging image regions such as weak-boundary cells, fading-texture cells, and heterogeneous background elements made us argue that models based on a single feature type may not be enough to distinguish nuclei patches from background elements. A single feature type does not seem to capture all relevant characteristics of nuclei objects in NRVM images due to non-exhaustive factors such as: image quality, illumination artefacts, and noises. The flow-chart diagram in Figure 6.8 describes the standard training routine we adopted to build 14 SVM models based on all possible combinations of the 4 basic feature types: HOG, LBP, GLCM, and PC.

As the flowchart diagram suggests, the combination of the four basic features enabled us to build different linear SVM models by 10-cross validation i.e., 10 random splits with the usage of 10 predefined different seeds. Knowing the seeds of the different round of validation allowed us to compare the behaviour of different models on a particular testing image, and therefore to perform qualitative and visual analysis. Systematic analysis was performed by computing and

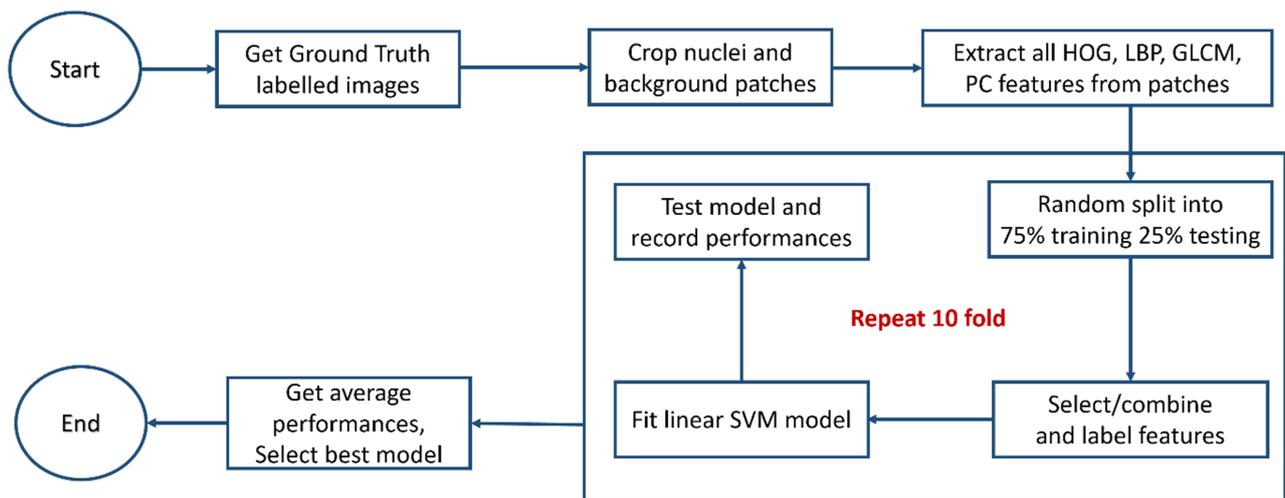


Figure 6.8: Flowchart diagram for training SVM model based on HOG, GLCM, LBP, and PC combination.

comparing all models performance metrics such F1 scores, sensitivity, specificity, and accuracy. These performance metrics adopted are described in Section 6.3.

Following Figures 6.9, 6.10, 6.11, and 6.12 illustrate the output of four basic SVM models built respectively from HOG, LBP, GLCM, and PC feature vectors and tested over example NRVM images. As it can be denoted visually, the four classifiers perform differently in the presence of fading cells vague-contour or weak-boundary and heterogeneous background and can even output opposing results - another reason in favour for investigating SVM models based on feature fusion at score level.

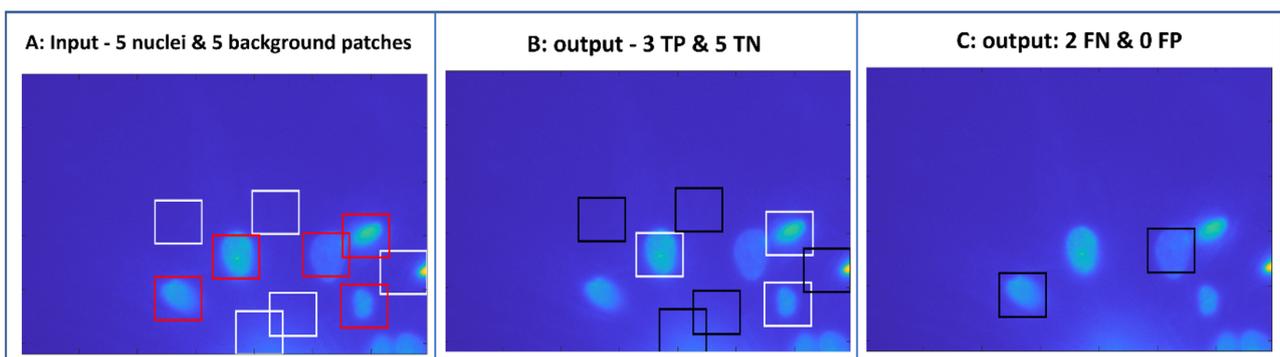


Figure 6.9: HOG based SVM classifier fails to predict 2 fading cells locations. A: Original input NRVM image with 5 nuclei (red boxes) and five background patches (white boxes) labels. B: 3 TP and 5 TN predictions by the HOG based classifier C: 2 FN predictions by the HOG based system.

Figure 6.9(A) represents five nuclei (red boxes) and five background patches (white) that are

fed into a HOG SVM classifier for object class prediction. Figure 6.9(B) illustrates the output of the HOG classifier. 3 TP & 5 TN means that the classifier correctly predicts three nuclei patches (white) over 5 and 5 background patches over 5. In Figure 6.9(C), 2 FN predictions are recorded, which means that the HOG based classifier predicts wrongly two nuclei patches as part of the image background.

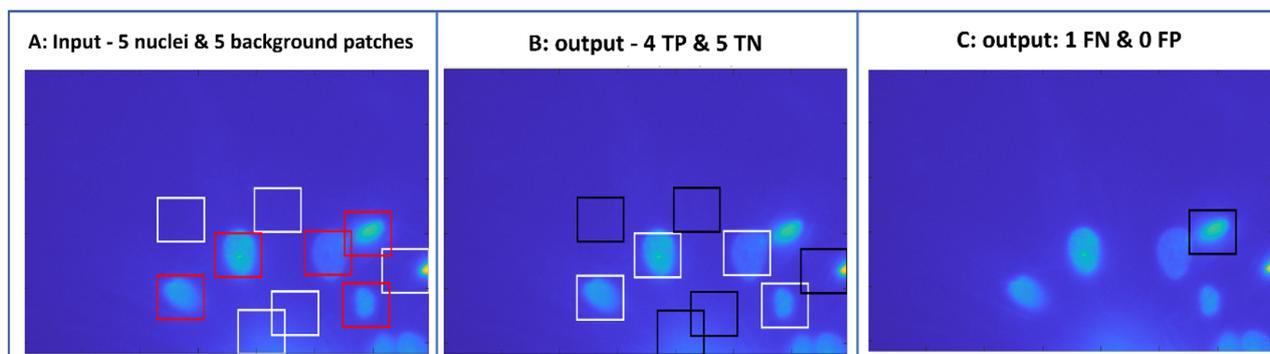


Figure 6.10: GLCM based SVM classifier fails to predict 1 cell with vague contour. A: Same original input NRVM image with 5 nuclei (red boxes) and five background patches (white boxes) labels as in 6.9. B: 4 TP and 5 TN predictions by the GLCM based classifier C: 1 FN prediction by the GLCM based system.

Figure 6.10(A) represents the 5 same nuclei (red boxes) and 5 background patches (white) early used in Figure 6.9. Here, their GLCM features are fed into a GLCM-based SVM classifier for prediction. Figure 6.10(B) illustrates the output of the GLCM classifier. 4 TP & 5 TN means that the classifier correctly predicts four nuclei patches (white) over 5 and 5 background patches over 5. In Figure 6.10(C), 1 FN prediction is recorded, which means that the GLCM based classifier predicts wrongly 1 nuclei patch with vague contour as part of the image background. Compared to HOG, GLCM is outperforming in overall, but the missing of one nucleus region showing vague contour suggests that HOG and GLCM should be fused at a certain level to be determined.

Figure 6.11(A) represents the same 5 nuclei (red boxes) and five background patches (white) used for HOG and GLCM SVM classifiers in previous Figure 6.9 and Figure 6.10. Figure 6.11(B) illustrates the output of the LBP classifier over the ten patches. 2 TP & 4 TN means that the classifier correctly predicts only two nuclei patches (white) over 5 and 4 background patches over 5. In Figure 6.11(C), 3 FN and 1 FP predictions are recorded, which means that the LBP based classifier wrongly predicts 3 nuclei patches as part of the image background and one background patch as a nucleus. Compared to HOG and GLCM, LBP based model is

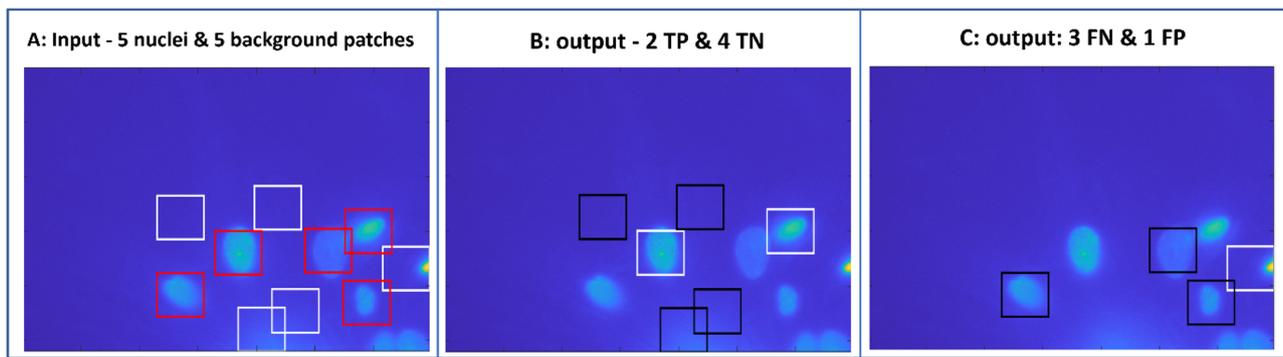


Figure 6.11: LBP based SVM classifier fails to predict 2 fading cells and a weak-boundary cell. A: Same original input NRVM image with 5 nuclei (red boxes) and five background patches (white boxes) labels as in 6.9. B: 2 TP and 4 TN predictions by the LBP based classifier C: 3 FN and 1FP predictions by the LBP based system.

under-performing in overall. However, the missing of a nuclei patch showing a vague-contour cell by GLCM and captured by LBP, suggests a fusion of LBP and GLCM feature at a certain level to maximize the detection of all challenging cellular patches. The problematic nuclei patch is the furthest window at the top-right side in Figure 6.11(B).

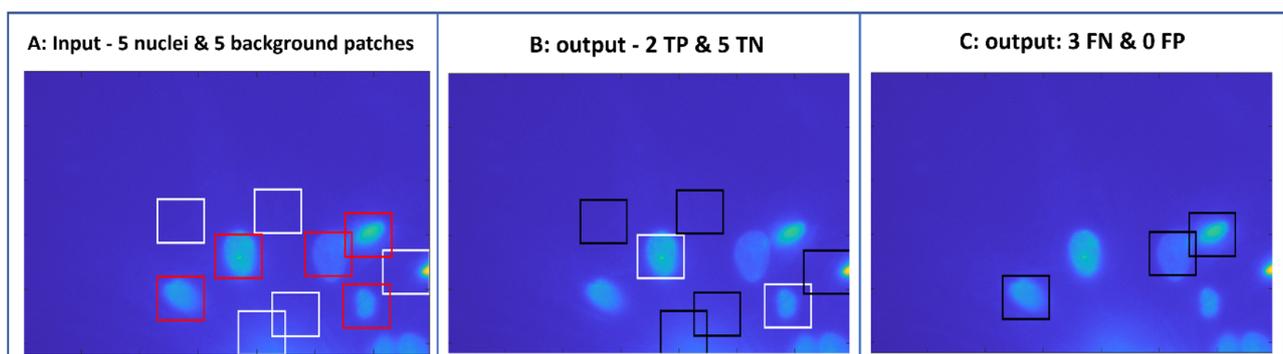


Figure 6.12: PC based SVM classifier fails to predict three cells with vague contour and fading texture. A: Same original input NRVM image with 5 nuclei (red boxes) and five background patches (white boxes) labels as in 6.9. B: 2 TP and 5 TN predictions by the PC based classifier C: 3 FN predictions by the GLCM based system.

Figure 6.12(A) represents the same five nuclei (red boxes) and five background patches (white) used for HOG, GLCM and LBP SVM classifiers. Figure 6.12(B) illustrates the output of the PC classifier over the ten patches. 2 TP & 5 TN means that the classifier correctly predicts only 2 nuclei patches (white) over 5 and 5 background patches over 5. In Figure 6.12(C), 3 FN predictions are recorded, which means that the PC based classifier predicts wrongly 3 nuclei patches as background patches. Compared to LBP, the PC based model seems identical in

terms of a number of right nuclei predictions. Still, each model has identified a cell that has been ignored by the other one (nuclei located in far-right hand side of images 6.11(B) and 6.12(B)).

HOG-LBP-GLCM-PC based SVM model output: 0 TP & 5 TN

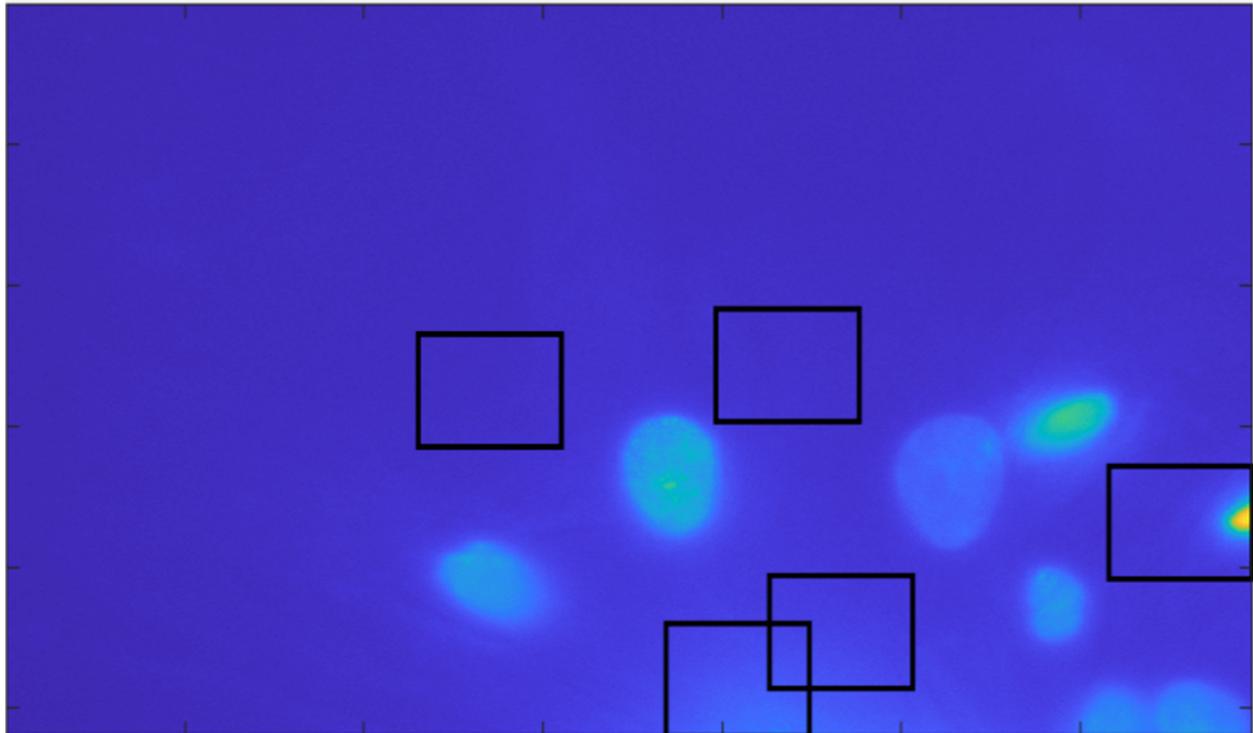


Figure 6.13: SVM model based on fusion of 4 features fails to predict any nuclei locations in a challenging NRVM image showing weak-boundary and fading cells.

After investigating the behaviour of SVM models built upon individual features over challenging image regions, we also analysed the output of SVM models based on the plain concatenation of HOG, LBP, GLCM and PC features. Figure 6.13 represents the output of an SVM model using a combination of the four basic features vectors over the same cells' locations discussed in Figure 6.9 - Figure 6.12. 0 TP and 5 TN means that the fusion model could only accurately predict five background patches over 5, but no nuclei locations over the five ground true locations checked.

Visually, it is evident that despite encouraging results for SVM models based on unique features, a model based on the concatenation of HOG, LBP, GLCM, and PC at features level may not achieve better performances. The visual analysis of results output by different SVM fusion models using various concatenations of HOG, GLCM, LBP and PC features demonstrated that a fusion at score level might be necessary to increase the ratio of TP predictions.

The overall performance of the procedure for initial nuclei locations filtering by linear SVMs

fusion at feature level over the NRVM nuclei dataset is presented in result Section 6.4. The next section presents our investigation on fusion SVM models at score level for nuclei patch filtering.

SVM Models based on Fusion at Score Level

To perform SVM fusion at score level, we analysed the scores of the classes output by the 4 SVM models built upon individual feature vectors such as HOG, LBP, GLCM and PC. From the output scores of the unique-feature-based classifiers, we computed a signed average score and performed score fusion in the following manner: 1) output scores for nuclei classes are set positive, while output scores for background classes are assigned negative; 2) compute the average of all signed scores predicted by the 4 SVMs and for all signed average scores greater than -0.275, we consider the corresponding predictions as of class nuclei otherwise background. The entire procedure of the fusion model at score level we built to predict nuclei and background patches is schematized in Figure 6.7. The analysis that led to the identification of the optimal fusion score threshold -0.275 is presented in Section 6.4.2 and demonstrated in Figure 6.23. An illustration of the system's output on an example NRVM image is also illustrated by Figure 6.14.

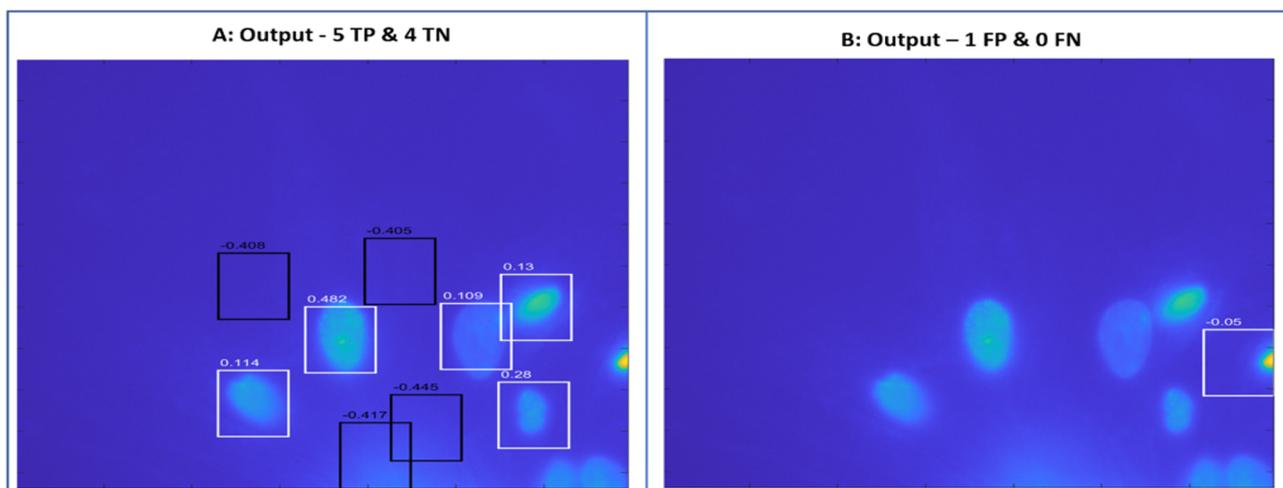


Figure 6.14: Fusion at score level of HOG, LBP, GLCM and PC SVM models with $T = -0.275$. A: output of TP and FP by the system. B: output of FP by the system. White numbers represent signed average scores output by the fusion model for patches predicted as nuclei. Black numbers represent signed average scores output by the fusion model for patches predicted as background.

Visually, it can be noticed that the model based on fusion at score level outperforms the feature

concatenation model illustrated previously in Figure 6.13. With a score threshold of -0.275, all potential nuclei locations are correctly predicted (as white-highlighted in Figure 6.14(A)). However, Figure 6.14(B) shows that a background patch has been wrongly predicted as nuclei with a signed average score of -0.05 (as white-labelled), which means that further processing is still required for ultimate right nuclei detection with less errors.

The overall performance of the procedure for initial nuclei locations filtering by linear SVMs fusion at score level over the NRVM nuclei dataset introduced in 6.2 and detailed in Section 6.3.3 is presented in result Section 6.4.1. In next Section 6.2.3, our final procedure for image window classification by the deep neural network is fully described.

6.2.3 Fine-grained Object Detection by CNN Classification

Once high-quality proposals are obtained by local adaptive thresholding, and SVM filtering, ultimate nuclei patch detection, the last step of our customized R-CNN model for automated nuclei detection can resume. Here, the aim is to perform a fine-grained object detection through a deep learner by feeding into the network, high-quality proposal boxes for classification and refinement. We argue that neural networks provide rich feature maps with different levels of abstraction that can help distinguish spurious background contents that was misclassified as nuclei locations by SVM models based on hand-crafted features. The overall performance and accuracy of the three different CNN networks investigated for fine-grained object detection (AlexNet, GoogLeNet, and Inception V3) are explored in Subsection 6.4.3. Before then, next Subsections give a brief overview of their architecture.

AlexNet model

AlexNet is a convolutional neural network made of eight layers and trained on more than a million natural images of ImageNet database. The eight main layers consist of three fully connected layers and five convolution layers separated by max-pooling layers and using a rectifier linear unit (ReLU) activation function. It is one of the most famous pioneering CNN that won the annual software contest in computer vision and machine learning, the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) in 2012 [51]. AlexNet requires input images of $227 \times 227 \times 3$. The network is powerful as it has learnt a rich feature representation of about 1000

categories of objects whose it has been trained to classify the images [51]. However, AlexNet was outperformed quickly by deeper CNNs due to its relatively small depth (a series of eight layers), and a huge number of parameters (61 million weights).

GoogLeNet model

The second CNN we investigated for nuclei proposal patch classification is a deeper CNN that consists of twenty-two layers organized in parallel or in a direct acyclic graph called GoogLeNet. GoogLeNet was also trained on ImageNet datasets and won the ILSVRC 2014 contest due to its architectural complexity with the introduction of the concept of inception module but with fewer parameters to train (7 million). The inception module consists of multiple convolutions filters applied to the same input, with some pooling to prevent redundancy and overfitting. Thus, GoogLeNet was designed to be deeper and lighter than AlexNet. Besides, the network only requires $224 \times 224 \times 3$ as input images [83] versus $227 \times 227 \times 3$ for AlexNet.

Inception V3 model

The third CNN backbone model we investigated for proposal patch classification is called Inception V3 and goes deeper than GoogLeNet. Its architecture is made of forty-eight layers organized in parallel and trained on ImageNet datasets. Inception V3 requires 229×229 input images and contains more than 23 million parameters to train [84]. Inception V3 achieves greater classification accuracy on the ImageNet dataset with higher relative prediction time than the GoogLeNet and AlexNet. For positive performances, we argue that Inception V3 might be the best network backbone for building our deep neural network strategy to classify nuclei proposal boxes. In next sections, we describe the materials and experimental protocols followed to measure the performance of the different steps of our proposed solution based on R-CNN for nuclei detection in images of FM.

6.3 Materials & Experimental Protocols

Since our proposed solution for nuclei detection in FM images involve initial generation of nuclei proposals by local adaptive thresholding, SVM filtering of high-quality candidate locations

based on HOG, LBP, GLCM and PC features, and image patch classification by CNN, we developed materials and experimental protocols accordingly. In terms of materials, ground truth datasets based on full input images and image patches were necessary to assess the performance of each step of our proposed algorithms. Performance metrics such as F1 score, sensitivity, specificity, the mean average precision (mAP) and GTCR were used to measure the efficiency of the system at each step. Following Subsections give an insight on the ground truth datasets we relied on and the experimental protocols we followed to assess the performances of our proposed solution for nuclei detection at every steps, i.e. initial nuclei proposal generation by local adaptive thresholding, SVM filtering of high-quality nuclei proposals based on HOG, LBP, GLCM and PC features, and image patch classification by deep neural networks.

6.3.1 Ground Truth Datasets

Two datasets have been used as ground truths of our deep learning model for nuclei detection in NRVM images. The first datasets correspond to the 250 labelled nuclei images described in Chapter 5. The second datasets consist of 2734 labelled individual images of nuclei, and 2734 labelled individual images of background patches obtained by the cropping of size (160 x 160). We used the first datasets made of 250 labelled images to build a customized region proposal algorithm for nuclei detection in FM images. The second ground truth dataset made of 2734 nuclei labels and 2734 background labels was used to develop and fine-tune a CNN detector for nuclei patches classification. Both datasets were split into 70% for training and 30% for testing.

6.3.2 Experimental Protocols

According to the three main steps of our proposed solution for nuclei detection by customized R-CNN based on adaptive region proposal generation, three types of experiments are conducted. First initial nuclei locations are generated from local adaptive thresholding map over the 250 NRVM images as described in 6.2.1, then the initial nuclei proposals are filtered by SVMs models using different feature concatenation models, i.e. fusion at feature and score levels of HOG, LBP, GLCM and PC as described in 6.2.2. Finally, the filtered proposal boxes are classified by different fine-tuned CNNs (AlexNet, GoogLeNet, Inception V3) into true nuclei or background patches as reported in Section 6.2.3.

Assessment of initial proposal generation technique

To assess the performance of the method for initial nuclei proposals boxes generation by local adaptive thresholding, we used two performances metrics: the GTCR and the mAP. First, the Ground Truth Coverage Ratio (GTCR) which is equivalent to the percentage of ground truth nuclei covered by the proposal boxes; and the mean average precision (mAP) which corresponds to the ratio of the number of nuclei proposal boxes that overlap with the ground truth nuclei over the total number of proposals boxes generated.

$$mAP = \frac{NPGT}{TNPG} \quad (6.1)$$

where $NPGT$ = the number of nuclei proposals boxes overlapping with Ground Truth, and $TNPG$ = the total number of nuclei proposal boxes generated.

Two boxes A and B are said overlapping when their IoU is greater than 0.5, according to the formula described in Figure 5.2. The higher are the GTCR and the mAP metrics, and the more accurate are the initial nuclei proposals generated. However, at this stage, only the GTCR matters because the aim is to reduce the sampling rate of the proposal generation technique by guiding the search to image locations that are more likely to contain nuclei rather than searching at every location and orientation inside images.

Assessment of high-quality proposal locations by SVM classification

To assess the performance of the different SVM filters built upon HOG, LBP, GLCM and PC and the different deep neural networks investigated, and to compare them; four metrics were used: the True Positive Rate (TPR or sensitivity or recall), the True Negative Rate (TNR or specificity), the accuracy (AC) and the F1score according to the following equations:

$$TPR = sensitivity = \frac{TP}{TP + FN} \quad (6.2)$$

$$TNR = specificity = \frac{TN}{TN + FP} \quad (6.3)$$

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN} \quad (6.4)$$

$$F_1 = 2 * \frac{\textit{precision} * \textit{sensitivity}}{\textit{precision} + \textit{sensitivity}} \quad (6.5)$$

TP correspond to all nuclei predictions that overlap with ground truth nuclei locations with $\text{IoU} \geq 0.5$; *TN* correspond to all background predictions that overlap with ground truth background patches (or do not overlap with any ground truth nuclei) with $\text{IoU} \geq 0.5$; *FP* correspond to all nuclei predictions that do not overlap with ground truth nuclei with $\text{IoU} \geq 0.5$; *FN* correspond to all background predictions that do not overlap with any ground truth background patches (or do overlap with a nuclei ground truth patch) with an $\text{IoU} \geq 0.5$. These same metrics were used to assess the performance of the ultimate nuclei detection by CNN. Next Subsection 6.3.3 describes the training and testing protocols adopted to build the different SVM and CNN classifier models for nuclei detection in biomedical images.

6.3.3 SVM and CNN Training and Testing Protocols

In this section, we describe the training and testing protocols we follow to build the various SVM models investigated for high-quality nuclei proposals generation and the three CNN classifiers developed for fine-grained nuclear object detection.

SVM Training and Testing Protocols

As indicated in the SVM models introduced in Section 6.2.2, we investigated four types of feature vectors such as HOG, LBP, GLCM, and PC and built SVMs models upon them. Based on various combinations of these four features, we built 14 different linear SVM fusion models at the feature level and one fusion model at score level using the four basic features. Training and testing routines of the different models were performed by 10-cross validation over splits (70% for training and 30% for testing) of our 250-image-labelled datasets.

In details, to train SVM models based only on HOG, we extracted from nuclei and background sample patches, 8100 vector lengths of HOG features and normalized them. The normalization was done by maximum ratio i.e. computing the ratio of the feature values extracted from image patches by the maximum values extracted from those patches. In the same way, to train SVMs models based only on LBP, we computed LBP images of the ground truth dataset and extracted 59-bins histograms from their corresponding nuclei and background sample patches

to fit the models. To build SVMs models based only on GLCM, we computed GLCM images of the NRVM ground truth dataset and extracted and normalized 236-vector-length features from nuclei and background samples patches to fit the models. The 236-vector length GLCM parameters consist of the concatenation of four texture statistics such as contrast, correlation, energy, and homogeneity. Finally, to build SVM models based on PC, we computed PC images of NRVM ground truth dataset and extracted 59-bins histograms from nuclei and background sample patches to fit the models.

For every round of 10 cross-validations, we reshuffled the training and testing sets. All the feature vectors extracted get max-normalized. We measured the average performance of our trained models through the computation of F_1 score, the level of sensitivity, specificity, and precision of the predictions. In such a way, we were able to identify the best SVM kernel that works well for our dataset, and feature types, i.e. a linear kernel of multi-class SVM investigated with LIBSVM library described in [17] and MATLAB proprietary functions. The result section presents a detailed performance of the different SVM models based on concatenation of HOG, LBP, GLCM and PC features for filtering nuclei proposal boxes generated initially by local adaptive thresholding.

CNN Training and Testing Protocols

To train the three CNN based image patch classifiers, we fine-tuned the pre-trained networks (AlexNet [51], GoogLeNet [83] and Inception V3 [84]) by transfer learning using a MATLAB GPU implementation over splits of our datasets made of 2734 nuclei and 2734 background patches(70% for training versus 30% for testing). By doing so, we kept the middle layers of the pre-trained networks and replaced the final layers with learnable weights by new final layers with higher learning rate and two output categories: nuclei and background. Thus, the fine-tuned networks are used for both feature extraction and image window classification. Our training used asynchronous stochastic gradient descent (SGD) with 0.9 momentum, a mini-batch size of 16, an initial learning rate of 0.0003, and a maximum number of six epochs per model. We minimized the SGD cost function and achieved training convergence for InceptionV3 model after five epochs, and AlexNet and GoogLeNet after six epochs.

To prevent overfitting, while increasing the accuracy of the network, and complying with the input size requirement, we performed image resizing, and data augmentation by random trans-

lation, reflection, and scaling in two directions. Theoretically, the augmentation techniques enabled us to multiply our working datasets size by a factor of six. The translation data augmentation technique involves moving the image along the X and Y direction. It is useful as it forces CNN to look everywhere for the labelled ROIs that have been translated in the augmented datasets. This learning ability is interesting in biomedical images where ROIs could be found anywhere in images i.e. nuclei to be pick up and background elements to be ignored. The reflection data augmentation technique involves reflecting the ROI along the X or Y direction. It is useful as it may help CNN recognize nuclei or cells in a cluster environment. The scaling technique involves onward or inward zooming of ROIs or images. It is helpful as it may also force the network to learn ROIs with multiple sizes. The model evaluations were performed using a running average of the network parameters computed over time.

To test the deep learning models based on AlexNet, GoogLeNet and Inception V3 for image patch classification as described in Section 6.2.3, we performed two types of experiments. The first one used a known set of nuclei and background patches to test the different CNN classifiers while the second relied on the segmentation results of adaptive thresholding and SVM filtering to do the testing. In that way, the effect of the guided search on the system overall performance could be easily assessed. In Section 6.4. all the findings are reported.

6.4 Result and Discussion

In this section, we explore the performance of our customized R-CNN solution for automated nuclei detection based on adaptive proposal generation against our ground truth datasets made of 250 labelled images, 2734 nuclei patches and 2734 background patches. Since our automated routine consists of three main steps which are initial nuclei proposal generation by local adaptive thresholding, proposals filtering by SVM classifiers and image patch classification by CNN, the results of the findings have been presented and discussed accordingly.

6.4.1 Initial Nuclei Detection by Local Adaptive Thresholding

The experiments over 250 NRVM nuclei images showed that the method for initial nuclei proposal boxes generation by local adaptive thresholding achieved a good performance of 99.40% GTCR. This ground truth coverage ratio means that our technique based on thresholding enable

to cover almost all potential nuclei locations in NVRVM images. However, the mean-average precision recorded is low ($mAP = 0.24604$) with an average of 60.844 nuclei proposal boxes being generated per image as the sampling rate.

Compared with the state-of-the-art object detection techniques investigated in Chapter 5, the small sampling rate (61 candidate locations per image) output by our proposed technique makes it more realistic to implement. Indeed, in Chapter 5, we reported that the investigated edge-based strategy, the technique with the highest average precision (30.75%) and great ground truth coverage (95.43%) over NRVM nuclei images, generates in average 8364 proposal boxes per image which is 137 times bigger than the sampling rate of our proposed method.

Moreover, despite this huge sampling rate, the edge-based strategy, does not outperform our proposed technique in GTCR (95.43% vs 99.40%). Figure 6.15 reports our proposed technique performance at generating initial nuclei locations from adaptive thresholding over NRVM nuclei images.

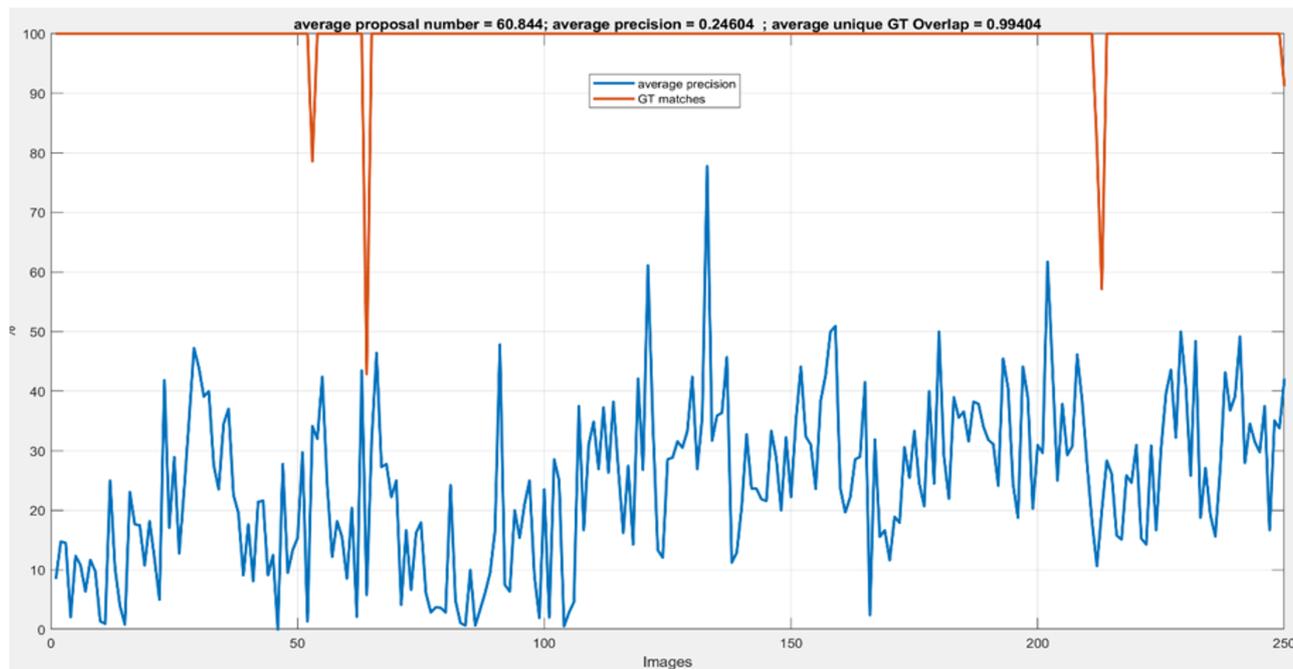


Figure 6.15: Performance of initial nuclei proposal box generation from adaptive thresholding mask. Red line: average ground truth coverage ratio per image. Blue line: average precision of proposal boxes per image. X-axis: NRVM image index or our dataset. Y-axis: values in percentage

In Figure, 6.15, red line shows the ground truth coverage per image. The blue shows the precision of proposal box generation per image. X-axis represents NRVM image index or our

dataset. Y-axis values in percentage. The red line in Figure 6.15 shows that most NRVM images truth nuclei locations are covered by the threshold-based candidates except few of them (5 in the current illustration). When analysing the outputs of our proposed routine, we noticed 2 fundamental limitations: 1) some overlapping or touching cells are detected as single cellular entities, and 2) some unique cell's locations are covered by multiple overlapping boxes. The detection issues related to overlapping or touching cells, are since local thresholding tends to generate single connected components rather than multiple as per cellular units. The sampling issues related to multiple overlapping boxes in the same cellular regions can be explained by the fact that our system fails to consider that several broken binary elements could belong to the same cells. Thus, a kind of region merging strategy may be needed.

To solve the cellular promiscuity problems, watershed transform-based techniques are well-known solutions that rely on image content segmentation into ridges and valleys [56], [85]. However, these techniques are sensitive to image illumination and thus can easily lead to over-segmentation of cellular regions. Here, we applied a simpler technique that separates two overlapping cells by using regional maxima segmentation. Regional maxima correspond to the consecutive pixels with constant intensity values and surrounded by pixels with lower values. We derived regional maxima of overlapping cells from morphological reconstruction by opening. To solve the problem associated to multiple bounding boxes around same cellular elements, a region merging technique could be the path to follow, but this is beyond the scope of this research. 6.16 illustrates the limitations of adaptive thresholding technique for identifying proposal boxes in touching or overlapping cells cases. Figure 6.17 illustrates our method for separating two overlapping cells in the same image patch.

Figure 6.16(A) shows 34 ground truth nuclei highlighted in white bounding boxes over an input NRVM image displayed in pseudo-colour. Blue shades correspond to smaller pixel intensities, whereas warm or red shades represent higher pixel values. Figure 6.16(B) shows local adaptive thresholding mask of the input NRVM image in A. In the middle and at the bottom right corner of B overlapping cells are represented as unique connected components (next to the red stars labels). Figure 6.16(C) shows 76 initial proposal boxes output from the threshold mask in B. Figure 6.16(D) represents the comparison of the proposal boxes with the ground Truth in A. Visually, it can be noticed that only 30 ground truth nuclei (white boxes) are covered by 32 true-positive proposal boxes (red bounding boxes). The missing four ground truth nuclei locations correspond to three image patches containing five overlapping cells identified as two

separate entities. The first group of overlapping cells is at the top centre of Figure 6.16(D) and are represented by a single proposal box (red labels) instead of two (white labels). The two remaining sets of overlapping cells are at the bottom right of Figure 6.16(D) and are represented by two separate proposal boxes (red labels) instead of five (white labels).

First column of Figure 6.17 shows greyscale patches of overlapping cells. Visually it can be noticed that there are two overlapping cells labelled (A), two touching cells labelled (E) and three touching cells labelled (I). Second column represents outputs of foreground elements marking after morphological opening-reconstruction. The foreground marking aims to increase the image contrast by brightening cellular regions. It uses a 15-size diamond structuring element. Visually it can be noticed that in cells labelled (B) and (F), cellular regions corresponding roughly to the centre of every nuclei unit are emphasized except in (J) where only 2 cells are marked over 3. Third column of Figure 6.17 shows global Otsu segmentation result of the reconstructed regions of the second column. Visually it can also be noticed that for each cellular region emphasized or marked, separated connected components have been generated (C, G). In Figure 6.17(K), interestingly only two regional maxima have been segmented as only two cells were correctly marked instead of three in total.

Last column of Figure 6.17 illustrates the theoretical border set to separate overlapping cells. The separation border between two overlapping cells is theoretically set in the middle of the centre of the connected components of the maxima regions obtained by segmentation in (C), (G) and (K). To estimate the borderline between two connected components, we computed the Euclidean distance between the two regions' centres and the slope of the line that is perpendicular to the line joining both centres.

As illustrated in the Figure 6.17 red-labelled (D), (H), and (L), the border between the cells is materialized in the red line, and each cell centre is displayed in a green dot. In this illustration, it can be noticed that the procedure for separating touching or overlapping cells works relatively well for two cells at a time but not for three touching cells at a time (Figure 6.17(L)). In total, we recorded a performance of 60.87% of precision (i.e. 14 overlapping cells correctly detected over 23 touching cells observed in NRVM nuclei patches we analysed). The following section explores the performance of SVM filtering of nuclei proposal boxes generated by local adaptive thresholding into high-quality candidate locations based on a fusion of multiple image features such as HOG, LBP, GLCM, and PC.

6.4.2 Filtering Initial Nuclei Proposals by SVM Models

When comparing the different SVMs models investigated, we focused on the TPR because our aim was to minimize the number of initial nuclei locations that would be wrongly discarded by the system rather than filtering out background patches to be filtered. In other terms, at this stage, we wanted to make sure that, after the initial generation of nuclei potential locations by local adaptive thresholding, all the good potential nuclei locations are effectively picked up by our trained SVMs models, and all the evident background patches are ignored, no matter the false-positive ratio, i.e. the percentage of background patches wrongly predicted as nuclei by the system.

To do so, we performed two types of experiments: 1) feed various SVM models with known nuclei and background patches for classification to identify the best models and 2), feed SVM models with initial nuclei locations generated by adaptive threshold segmentation to assess the true performance of our proposed technique. According to the two types of feature fusion analysed, the SVM models investigated were divided into two categories: 1) the models based on single or multiple feature concatenation of HOG LBP, GLCM and PC at feature level; and 2) those models based on a fusion of HOG, LBP, GLCM and PC features at score level. Figure 6.18 summarizes performances recorded for all SVMs models built upon a combination of HOG, LBP, GLCM and PC and using category-known images patches as testing samples. For all those experiments, the datasets were divided into 70% for training and 30% for testing through a 10-cross validation routine.

SVM Fusion Models at Feature Level to Increase Sensitivity

Figure 6.18 illustrates that without segmentation, linear SVM models based on concatenation of GLCM features achieved the highest sensitivity (0.84) and F1 score (0.83). In overall, these recorded performances of the SVMs models based on feature level combination show that an important proportion of TP nuclei locations are still ignored by the system (in average 16%). Since the highest sensitivity level recorded is 84%, 16% of potential nuclei candidate may be ignored by SVMs models based on features vectors concatenation. But a closer look at the patches predicted by individual SVM models using single feature type on challenging images give us varying insights of their relative sensibility.

-
- GLCM-based SVM classifiers are the most precise at detecting nuclei patches but may fail at recognising vague-contour cells. Figure 6.19 illustrates GLCM feature performance over an example NRVM nuclei image.
 - PC-based SVM mobility shows great ability at capturing background elements despite varying illumination artefacts with high confidence. Figure 6.20 illustrates PC feature performance over an example NRVM image.
 - HOG-based SVM models outperform LBP models in well-shaped nuclei images but underperform in the presence of heterogeneous background elements. Figure 6.21 illustrates HOG-based SVM models in such cases.
 - LBP-based SVM models, in contrast, show great ability at assigning the correct category to cells exhibiting intense texture stains. Figure 6.22 illustrates such a comparison between LBP and HOG models over two NRVM image examples.

Figure 6.19 shows the performance of GLCM based SVM model for NRVM patches classification without segmentation. The first column of Figure 6.19 represents input nuclei (Figure 6.19(A), yellow boxes) and background patches (Figure 6.19(B), black boxes) that are fed into the SVM model. The greyscale NRVM image displayed here is shown in pseudo-colour for visual purpose. Blue shades correspond to smaller pixel intensities, whereas warm or red shades represent higher pixel values. The remaining columns of Figure 6.19(C, D, E, and F) represent the GLCM based SVM model classification results. Figure 6.19(C) shows that the GLCM model was able to correctly predict 7 nuclei regions (TP) out of 8 with maximum confidence of 100% for all predictions. With regards to the background patches fed to the model, Figure 6.19(D) shows that the model was able to identify all of them (TN) with confidence greater than 90%. Figure 6.19(E) illustrates that there were no background regions misclassified as nuclei by the system. Finally, Figure 6.19(F) shows that GLCM based SVM model wrongly predicted a single nuclei patch as a background region with relatively low confidence of 73%. These results confirm our assumption derived from a visual representation of GLCM feature (in Figure 2.4): GLCM may not be enough for capturing all nuclei information inside NRVM images especially in the case of weak boundary cells like here.

Figure 6.20 shows the PC feature ability to capture background elements despite varying illumination artefacts. The first column of Figure 6.20 represents input nuclei (Figure 6.20(A), yellow boxes) and background patches (Figure 6.20(B), black boxes) that are fed into the SVM

model. The greyscale NRVM image displayed here is shown in pseudo-colour for visual purpose. The remaining columns of Figure 6.20(C, D, E, and F) represent the PC based SVM model classification results. Figure 6.20(C) shows that the PC SVM model was able to only predict 8 nuclei locations (TP) over 11 ground truth with confidences varying from 53% to 99%. With regards to the background patches fed to the model, Figure 6.20(D) shows that the model was able to identify almost all of them (10/11 TN) with confidence greater than 80% despite bright spots and varying background illumination all over the input image. The only background patch that was wrongly predicted as nuclei (1 FP) is highlighted in Figure 6.20(E) with long confidence of 52%. At this stage, FP ratio is not a big deal because we are aiming at detecting all candidate nuclei locations, i.e. maximizing the TP rate. However, Figure 6.20 (F) shows that PC based model could also fail at recognizing nuclei regions with weak boundaries (3 FN) with varying low confidences (less than 77%). These results confirm our former assumptions that PC may be good at recognizing background elements, hence increasing detector precision but contradict the fact they may outperform GLCM at detecting nuclei with weak boundaries.

Figure 6.21 compares HOG and LBP based SVM models over the same NRVM cells. It illustrates HOG feature ability against LBP to capture nuclei objects with diverse shapes despite their low illumination and weak boundaries. In details, Figure 6.21(A) shows that over 14 input nuclei patches, LBP based model was able to predict 9 TP. The 5 (FN) misclassified nuclei patches are displayed in Figure 6.21(F). Visually, they are effectively characterised by low illumination and week boundary artefacts.

In contrast, over the same input cells, Figure 6.21(C) shows that HOG-based SVM classifier achieved better performances. 14/14 nuclei sample patches were correctly predicted with 100% of confidence. With regards to background prediction, Figure 6.21(D) also confirms HOG-based model outstanding performance (11 TN) over LBP-based model (10 TN) whose performance is reported in 6.21(B). All in one, these results confirm our hypothesis in 2.2 about HOG features' ability at encoding nuclei shapes despite illumination artefacts.

Figure 6.22 describes LBP feature ability against HOG to capture nuclei objects with high texture intensity inside image patches. These challenging image patches are due to abnormal accumulation of Green Fluorescence Protein in cells. The performance of LBP-based models is verified in Figure 6.22(A) where all nuclei locations are correctly predicted by LBP against 1 missed by HOG as emphasized in Figure 6.22(C) and Figure 6.22(F).

With regards to the performance of both models at predicting background patches, Figure 6.22(B) and Figure 6.22(D) demonstrates that LBP still outperformed HOG with a 2 TN margin despite the disparity in confidence values. Finally, although LBP-based models relatively perform better in the presence of GFP accumulation, Figure 6.22(E) shows that further filtering will be needed to deal with FP background patches, i.e. the background patches wrongly predicted as nuclei. Moreover, the disparity of all SVM models based on unique feature vectors confirmed our hypothesis to build a fusion model. Next Section 6.4.2 discuss our findings about SVM models based on fusion at score level to filter initial proposal boxes.

SVM Fusion Model at Score Level to Increase Sensitivity

As stated earlier in this Chapter, we introduced initial proposal filtering by SVM fusion models to make sure that all potential nuclei locations are kept by the system while ignoring as much as possible evident background patches i.e., to make a trade-off between sensitivity and specificity of the system performances as sensitivity depends on the true positive rate (TPR) and specificity on the opposite (True Negative Rate, TNR). To do so, we built a SVM score fusion model based on HOG, GLCM, LBP and PC features, then recorded and compared their average sensitivity and accuracy at different fusion threshold values as described in Section 6.2.2. Figure 6.23 demonstrates this ambivalence between sensitivity and specificity of the system at the stage of initial proposal filtering. Indeed, when pushing down the scoring threshold, system sensitivity increases, while its specificity decreases (Figure 6.23(B)).

Figure 6.24 summarizes the results obtained by the score-fusion model we built over HOG, GLCM, LBP and PC single SVM models and tested over user-input nuclei and background patches extracted from our entire labelled dataset. Experiments were done in 10-cross validation manner where 70% of the datasets were used for training the models and 30% for testing. As reported in the graph, the sensitivity of the score fusion model is close to 100% (0.98). However, the average specificity (0.41) and accuracy (0.7) reported tells that many FP need further processing, i.e., the proportion of background patches assigned to the wrong category. That is why, we introduced an ultimate classifier of the image patches by CNN as CNN can generalize well across image content variability and can provide a rich and autonomous feature extraction [31], [40].

Figure 6.25 reports performances output by the same score fusion SVM model tested but this

time over nuclei locations initially generated by thresholding. The filtering by SVM fusion model at score level enables to increase the accuracy (52.21%) and sensitivity (97.23%) of the initial nuclei locations derived from local thresholding. But many misclassified background patches need further filtering.

Figure 6.26 illustrates the performance of the score fusion SVMs at filtering initial nuclei locations generated by local adaptive thresholding over an NRVM image. Figure 6.26(A) represents an input NRVM image with 15 ground truth nuclei locations highlighted by yellow boxes. Figure 6.26(B) shows the first 56 initial nuclei proposal boxes generated by local adaptive thresholding. The Proposal boxes are highlighted by white boxes for visual purpose and overlaid on the input image A. Figure 6.26(C) represents the true positive predictions output by the system based on SVM fusion at score level (18 TP highlighted in red boxes and red scores). A comparison of the filtered proposal boxes with the nuclei ground truth shows that all right nuclei have been covered. However, there is a great concentration of boxes around inter-cellular spaces and few background areas that are away from nuclei regions as potential cellular units.

Figure 6.26(D) shows that our proposed SVM fusion model was able to correctly predict 9 background patches (highlighted in black boxes and scores) over 38 potential ones. Figure 6.26(E) represents the FP predictions output by the fusion model, i.e. the background patches that were misclassified by the model. Visually it can be noticed that most of the FP predictions, i.e. the background patches assigned to the wrong category correspond to boxes containing parts of nuclei elements that are less than 50% of the total cell area.

Finally, Figure 6.26(F) shows that there is no FN recorded, i.e. no nuclei element has been wrongly predicted as background. Which means that the SVM filtering does not introduce further errors into the initial proposal box generation procedure using thresholding in terms of potential nuclei locations being discarded. In contrast, the SVM classification enables the filtering of proposals by ignoring a proportion of candidate locations that are truly part of the background of NRVM images. In the current example, the total number of initial proposals boxes get reduced from 56 to 47 after 9 TN boxes got filtered. However, many FP predictions need further assessment specially in inter-cellular regions.

To sum up, by analysing single SVM feature model performances, we found that HOG, LBP, GLCM and PC features performed unequally in critical situations such as heterogeneous background illuminations, bright spot artefacts, high-intensity texture cells, fade nuclei texture,

weak nuclei boundaries or vague contour cells. Based on the example images that we analysed and the confidence ratio output by individual classifiers, it can be argued that GLCM features and HOG are roughly good at encoding nuclei characteristics whereas PC and LBP may be best at detecting background elements. GLCM-based models output the highest sensitivity equalling to 0.84, and F1score equalling to 0.83; HOG-based models reach 0.82 of sensitivity and 0.72 of F1score; while models based on a concatenation of LBP and PC achieved the highest true-negative rate or sensitivity equalling to 0.80 (according to Figure 6.18 Performances of SVM models based on fusion at feature level for nuclei detection).

However, the contrasting sensibilities exhibited by the different SVM models made us take a fusion approach at score level of the four features. The fusion model at score level enabled us to filter the initial nuclei proposals generated by thresholding into better candidate locations by increasing the sensitivity of the model and reducing the number of the sampling rate. Experiences proved that for a threshold score set at -0.275, we could reach an acceptable compromise in terms of sensitivity (97.23%), accuracy (52.21%) and specificity (29.5%) of nuclei proposal boxes as demonstrated in Figure 6.23. Thus, we concluded that further filtering of the proposal boxes is necessary to increase the overall system accuracy. Using CNN for final classification of the proposal boxes initially generated by local adaptive thresholding and filtered by score fusion SVM model is the object of the next experiment report.

6.4.3 Findings on Nuclei Bounding Boxes Classification by CNN

The filtering of initial nuclei locations by SVM fusion models enables to capture roughly 97.23% of all potential nuclei regions inside NRVM images, with an overall accuracy of 52.21% against 24.60% for the initial proposals generated by local adaptive thresholding. This accuracy level of the fusion model means that despite the filtering of higher-quality proposals boxes, nearly half of them should have been discarded by the system as FP. To overcome such problem by increasing the overall accuracy while maintaining the sensitivity level, we fine-tuned three pre-trained CNNs (AlexNet, GoogLeNet and Inception V3) and analysed their performances in a 10 cross-validation routine as explained in Section 6.3.

With a nuclei patch detector based on AlexNet, we achieved a validation accuracy of 94.51%. For GoogLeNet, we achieved a validation accuracy of 95.24%, and for InceptionV3, the validation accuracy increases to 96.27%. These results confirmed the general assumption that the

deeper go CNNs, the more accurate they become until reaching a certain plateau [35] [84]. They also validated our former hypothesis made in the literature review at Chapter 2 that deep learning strategies can outperform classification models based on hand-crafted features vectors and generalize well over variable image conditions in terms of cell illumination, and ROI characteristics.

Figure 6.27 illustrates the power of CNN learner at identifying ambiguous nuclei and background patches with strong confidence compared to SVMs classifiers based on hand-crafted feature engineering. Figure 6.28 illustrates the ability of CNN detector at outperforming SVMs classifiers in extreme low-contrast images. In both figures, the cell images are displayed in true greyscale colour to emphasize on the cell illumination disparities.

Figure 6.28(A) shows a low contrasted NRVM image that contains 14 nuclei and 14 background used to compare our proposed SVM and CNN models trained for nuclei patches classification. Figure 6.28(B) and (F) show that CNN outperforms SVM in detecting an extra fade nucleus with weak boundary located in the middle right of the image. In total, over 14 ground truth (GT) nuclei, SVM fusion enables to detect 11, while CNN enables to find 12. Figure 6.28(C) and Figure 6.28(G) show these differences in terms of FN predictions (3 FN for SVM and 2 FN for CNN). Figure 6.28(D) and (H) illustrates that CNN still outperforms SVM at predicting more background patches (13 for CNN and 7 for SVM for a total of 14 GT background). These differences are also illustrated in Figure 6.28(E) and Figure 6.28(I) where the corresponding FP predictions made by both classifiers are displayed.

Figure 6.29 and Figure 6.30 summarize the average performances recorded in a 10-cross validation routine over splits of our dataset (70%, 30%) without or with segmentation. Both experiments respectively exhibit a standard deviation of 0.16% and 0.34% between the rounds and for all the performance measurements.

6.5 Summary on Nuclei Detection by Customized R-CNN

In summary, our practical investigation about automated nuclei detection by customized R-CNN showed the following. First, a data-driven based technique for generating nuclei proposals is efficient as it helps reduce the sampling rate (less than 61 boxes per image) and improve the

GTC ratio of nuclei proposal boxes (97.23%) prior to fine-grained classification by CNN object detector described in Section 6.2.3. Second, Hand-crafted feature engineering based on HOG, LBP, GLCM, and PC is tedious and may not be sufficient to encode all nuclei features in challenging situations such as heterogeneous image backgrounds, vague-contour cells, broken-edges cells, fading and high-intensity texture cells. Third, a deep image patch classifier based on Inception V3 CNN increases the accuracy of the whole system with a fine margin of accuracy = 0.75.

In details, the region proposal generation technique we developed consists of 2 steps: 1) local adaptive thresholding for computing initial nuclei locations; 2) a score-fusion model using four linear SVM Classifiers built upon HOG, LBP, GLCM, and PC features for filtering the initial proposal boxes. Combining local adaptive thresholding and SVM classification enables us to perform a kind of selective based on image content rather than a blind search involving window sliding over entire input images. Compared to the most promising object detection technique we investigated in Chapter 5, the edge-box based strategy, which generates more than 8 000 bounding boxes per NRVM image with an overall accuracy of 30.61%, our solution is more realistic and applicable to FM datasets, as it generates fewer candidate locations and more accurate ones. Indeed, the SVM filtering procedure, we applied to the initial proposal boxes generated by thresholding, enabled to output candidate nuclei locations that cover 97.23% of the nuclei ground truth with 52.21% of accuracy.

The ultimate object detection routine we applied after SVM filtering relies on the power of CNN, such as Inception V3. The rationale behind the usage of CNN for final nuclei patches detection was to reduce the number of FP predictions of SVM models, i.e. to filter out background patches wrongly classified as nuclei. Results discussed in Section 6.4.3 proved that CNN features could effectively generalize well and outperform SVM models based on hand-crafted feature vectors with better accuracy and confidence scores.

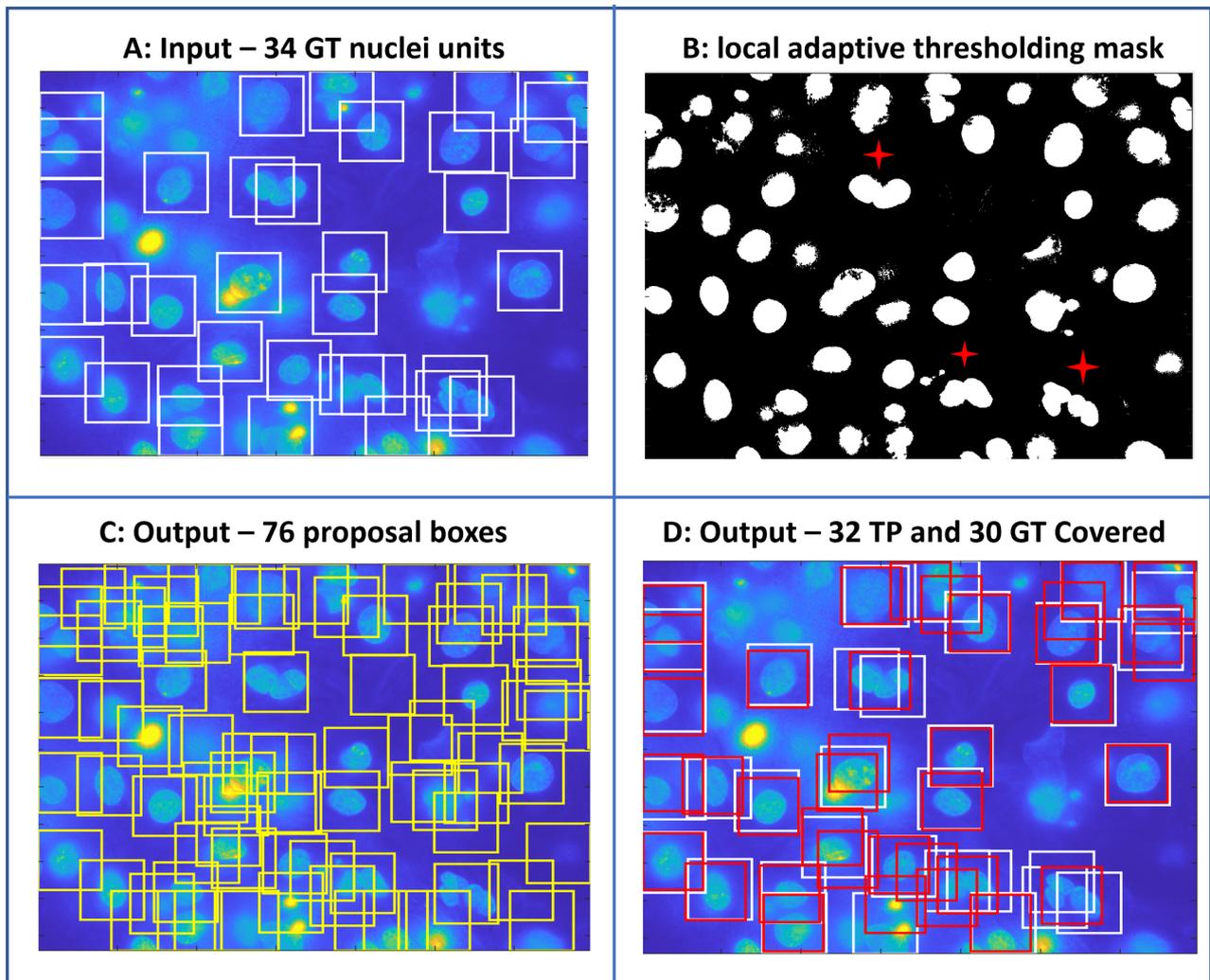


Figure 6.16: Limitations of initial nuclei proposal generation in handling cell overlaps. 3 cellular regions showing overlapping candidates are labelled with red stars in B. A: original reference image with 34 ground truth nuclei labels. B: binary mask after applying local adaptive thresholding. C: results of the initial proposal box generation. D: comparison of the initial proposals with the reference image.

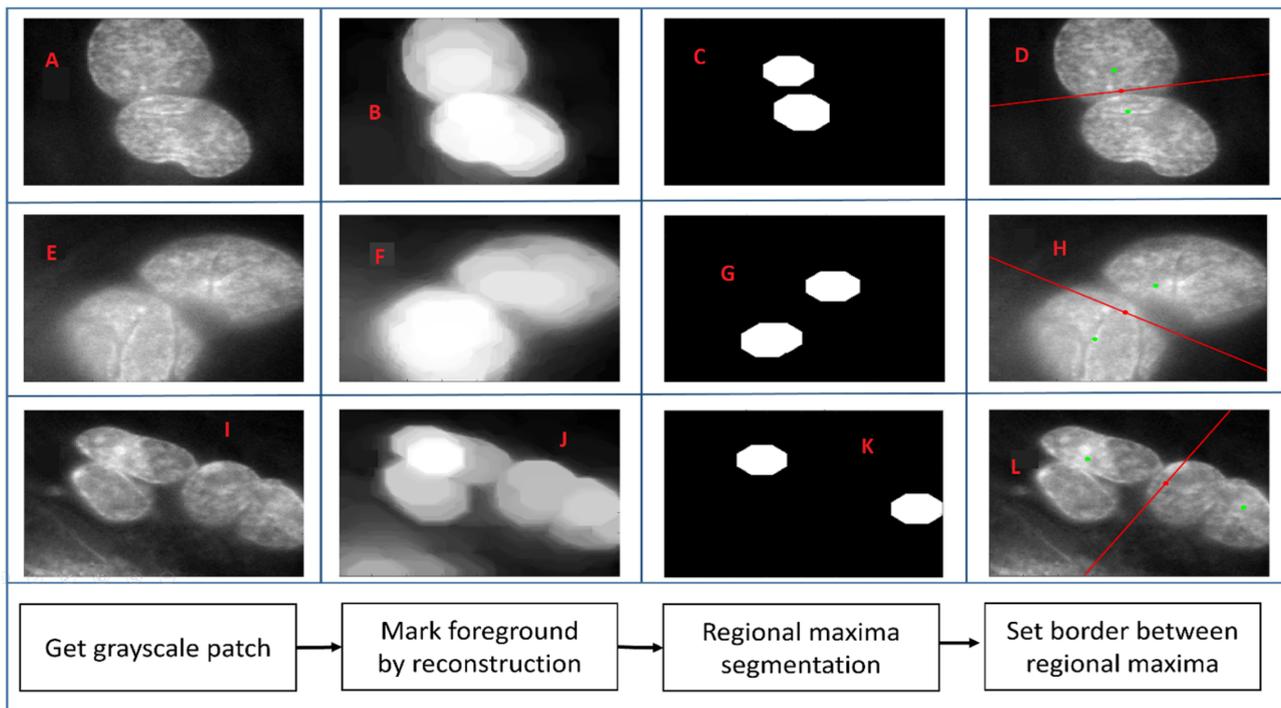


Figure 6.17: Finding locations of overlapping or touching cells from regional maxima centres. A, E, I: 3 original grayscale nuclei patches with overlapping or touching elements. B, F, J: foreground element marking by morphological opening-reconstruction. Regional maxima segmentation mask. Separation edges set between touching or overlapping cells.

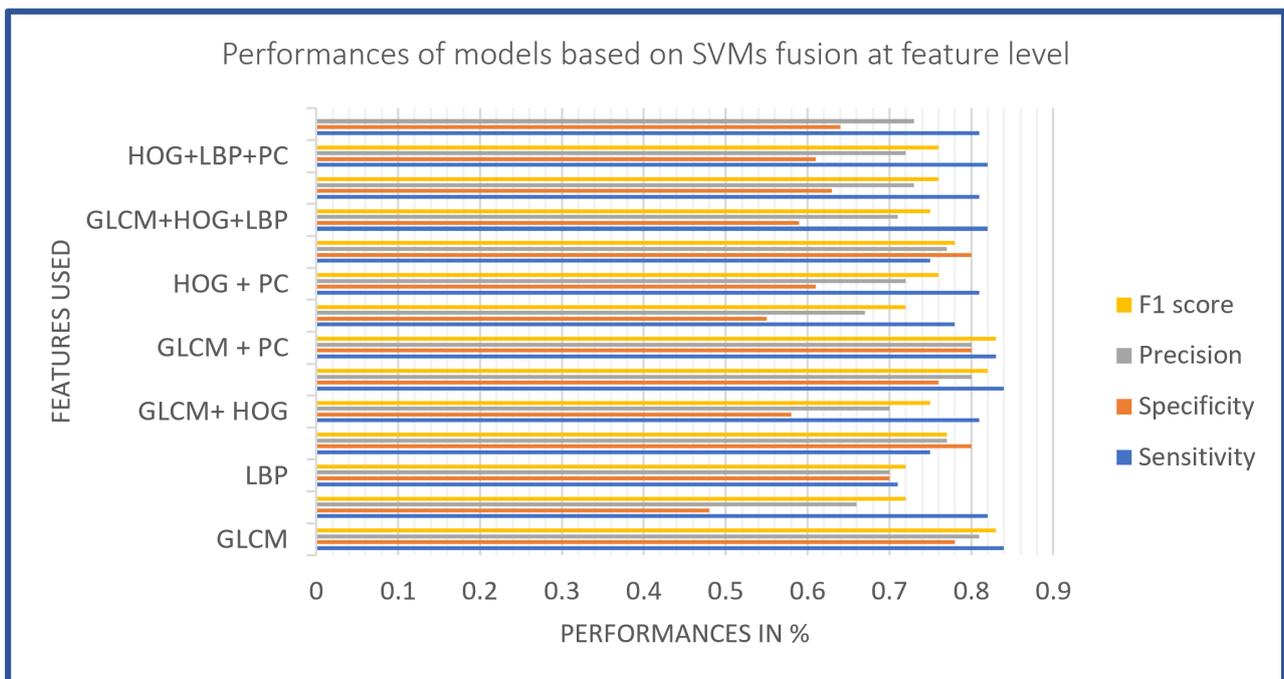


Figure 6.18: Average performances of SVM models based on fusion at feature level for nuclei detection.

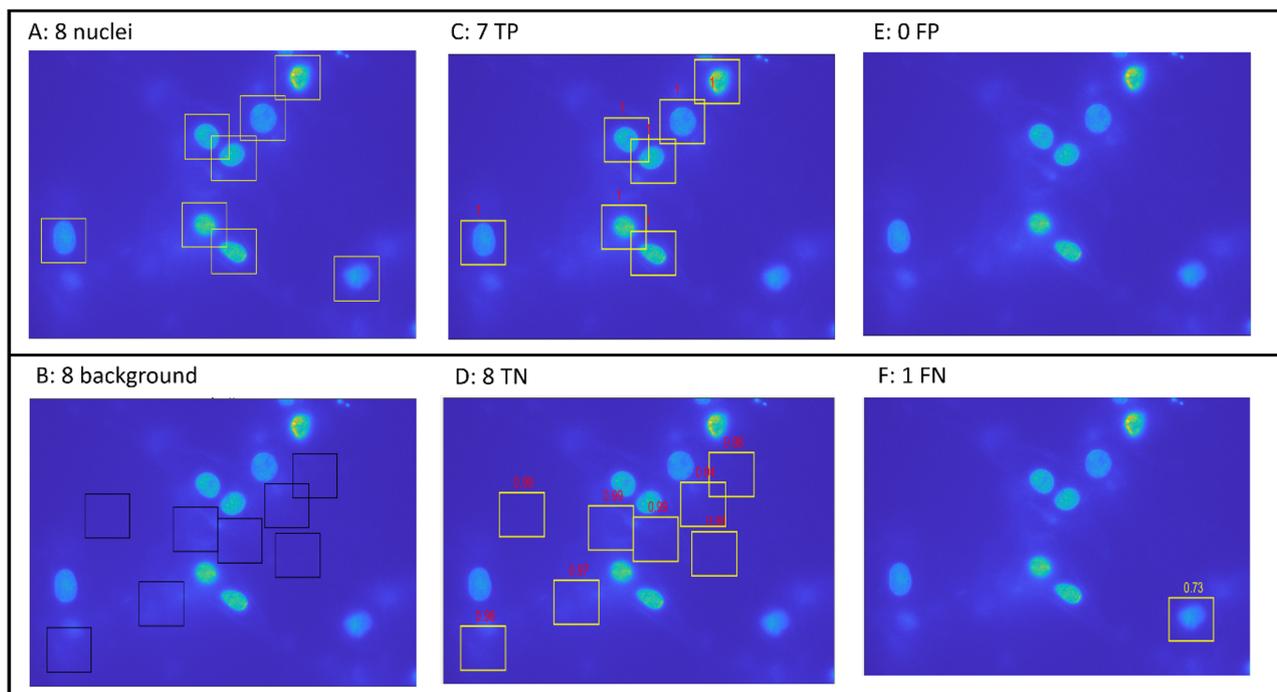


Figure 6.19: GLCM SVM classifiers outperform in presence of vague contour cells but underperform in heterogeneous image backgrounds. A: reference image with 8 ground truth nuclei locations. B: reference image with 8 ground truth sample background patches. C: 7 TP nuclei locations predicted by the system built on GLCM features. D: 8 TN background patches predicted by the system. E: no FP predicted by the system. F: only 1 FN predicted by the system. Decimal numbers represent the level of confidence of the classifier.

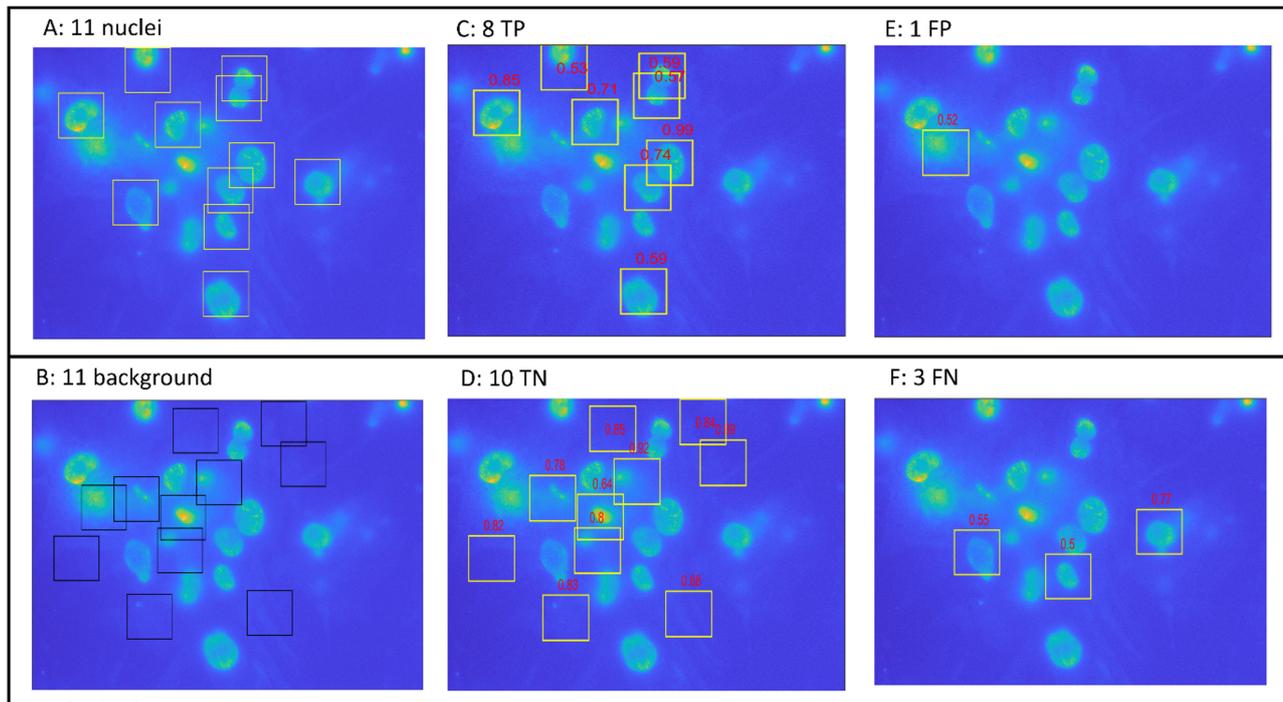


Figure 6.20: PC SVM classifiers outperform in presence of heterogeneous background. A: reference image with 11 ground truth nuclei locations. B: reference image with 11 ground truth sample background patches. C: 8 TP nuclei locations predicted by the system built on PC features. D: 10 TN background patches predicted by the system. E: only 1 FP predicted by the system. F: 3 FN predicted by the system. Decimal numbers represent the level of confidence of the classifier.

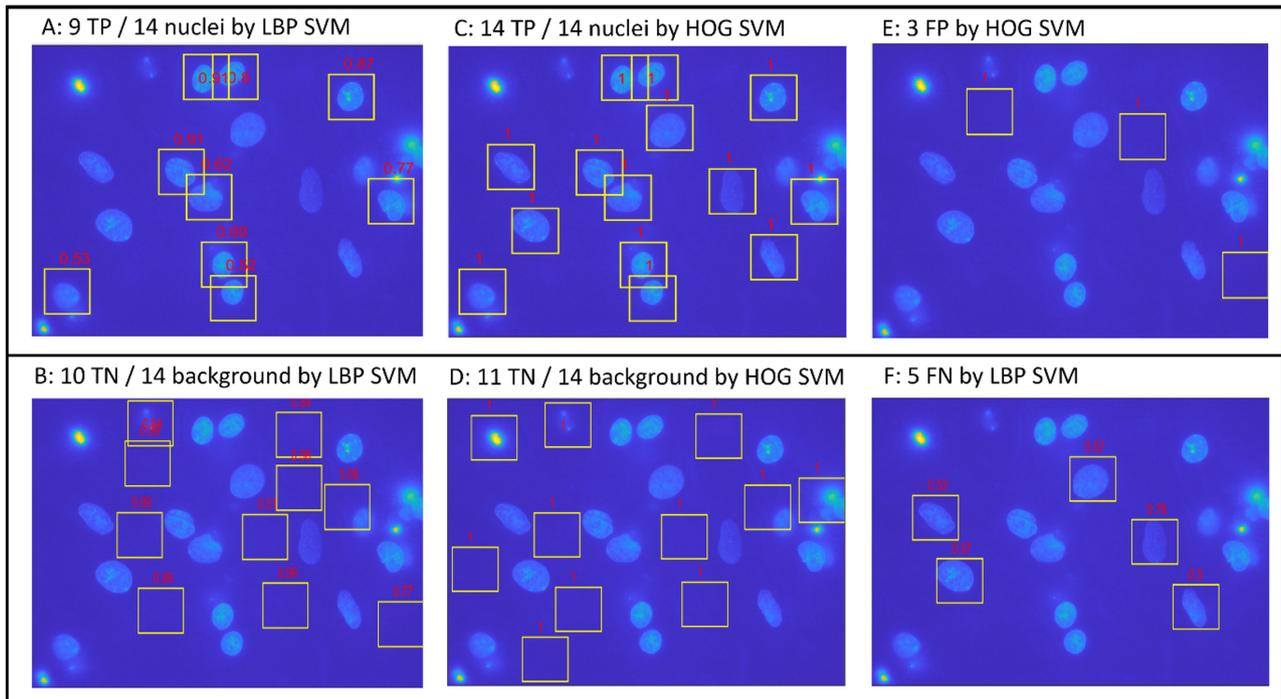


Figure 6.21: HOG SVM may outperform LBP SVM in well-shaped nuclei images but underperform in presence of fading texture. A: reference image with 14 ground truth nuclei locations. B: reference image with 14 ground truth sample background patches. C: 14 TP nuclei locations predicted by the system built on HOG features. D: 11 TN background patches predicted by the system built on HOG. E: 3 FP predicted by the HOG based system. F: 5 FN predicted by the LBP based system. Decimal numbers represent the level of confidence of the classifiers

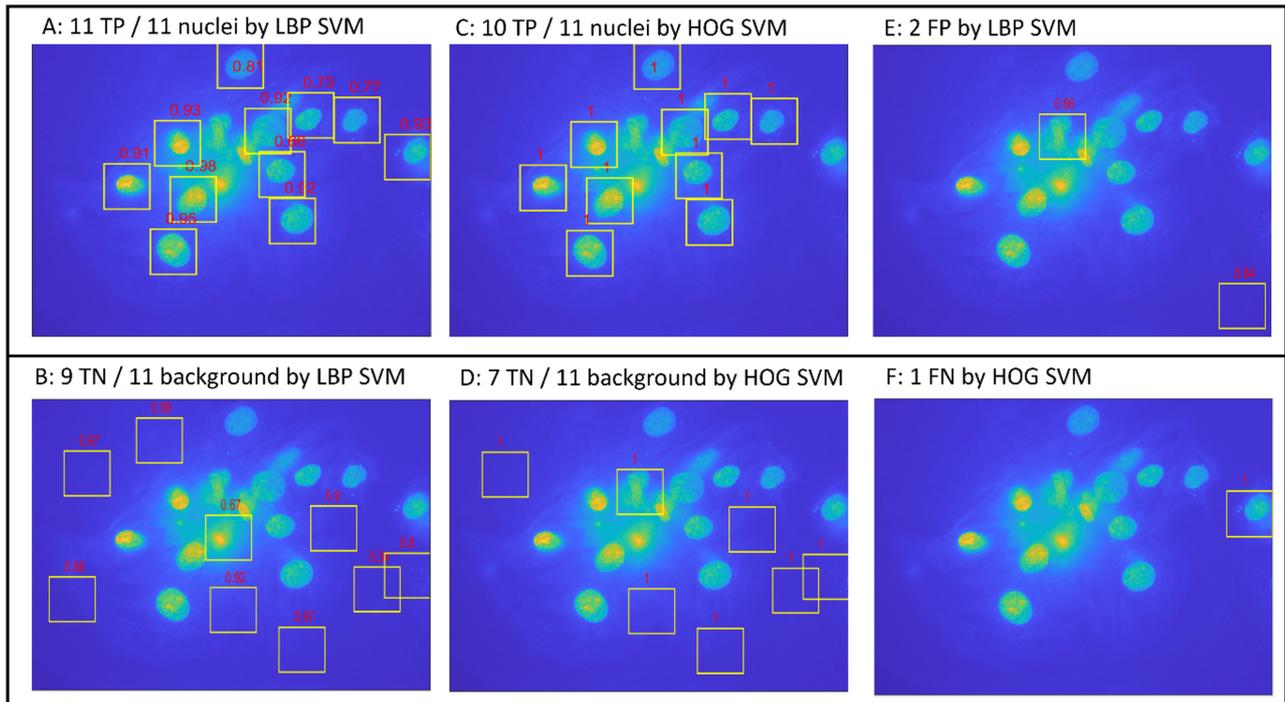


Figure 6.22: LBP SVM may outperform HOG SVM in intense-texture nuclei images. A: All TP nuclei detected by LBP based model. B: 9/11 TN predicted by LBP based model. C: 10 TP/11 nuclei locations predicted by the HOG based system. D: 7 TN/11 background patches predicted by the HOG based system. E: 2 FP nuclei predicted by the LBP based system. F: only 1 FN predicted by the HOG based system. Decimal numbers represent the level of confidence of the classifiers.

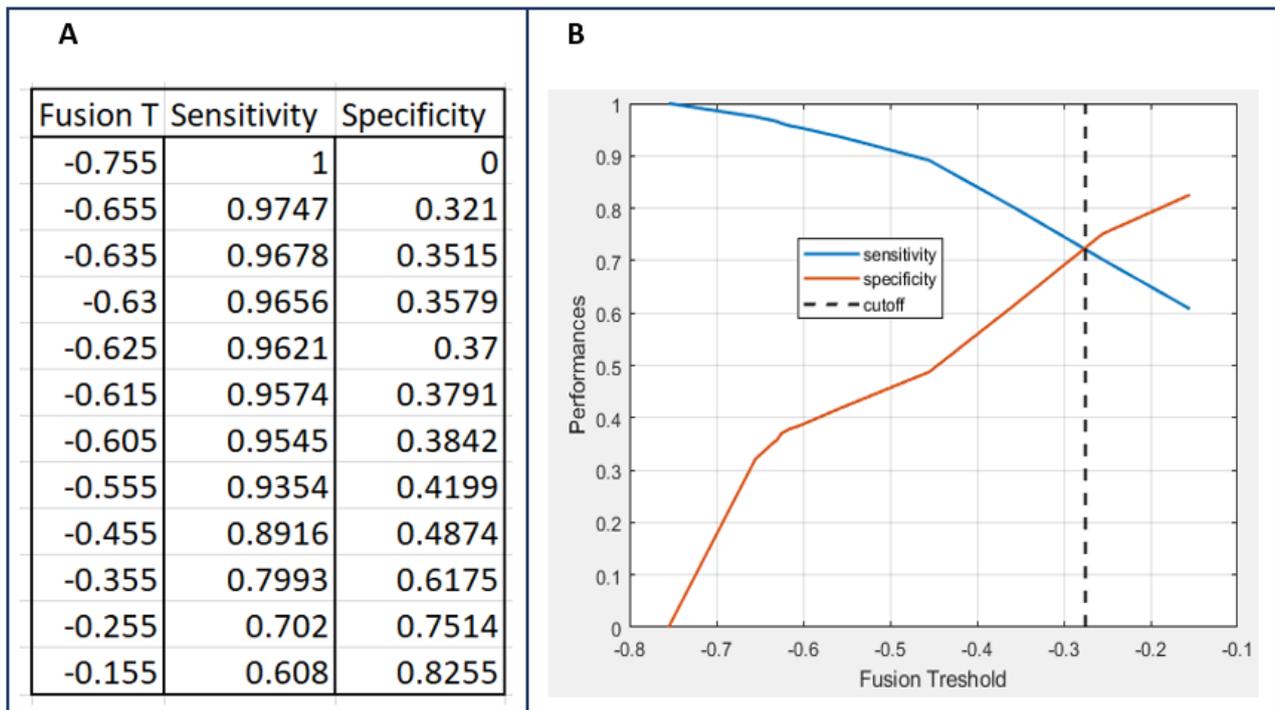


Figure 6.23: Optimum score fusion Threshold: trade-off between sensitivity and specificity. A: Table of sensitivity and accuracy performances recorded for filtering initial proposals by SVM fusion at score level with different thresholds applied over half of the testing dataset. B: plot of sensitivity (blue line) and specificity (orange line) performances over fusion threshold values. Optimum threshold is the point of junction of both curves i.e., the cut-off dashed line with $x = -0.275$

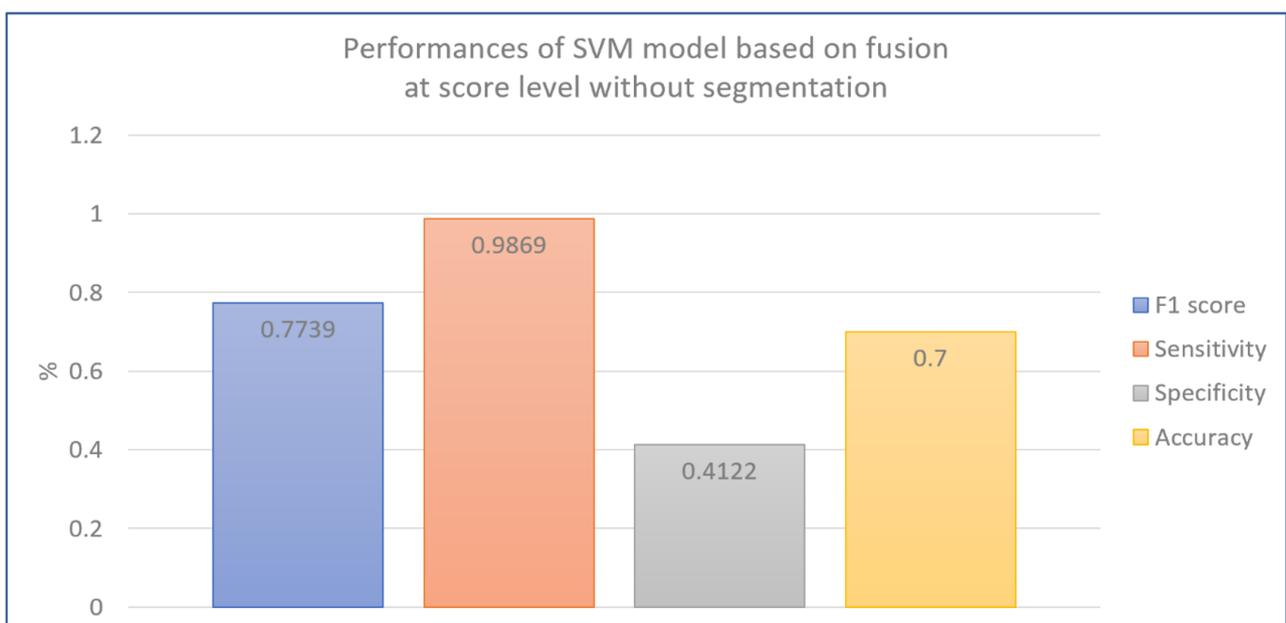


Figure 6.24: Average performances recorded after 10 Cross-validation of SVM model based on Fusion at score level for image patch classification without segmentation.

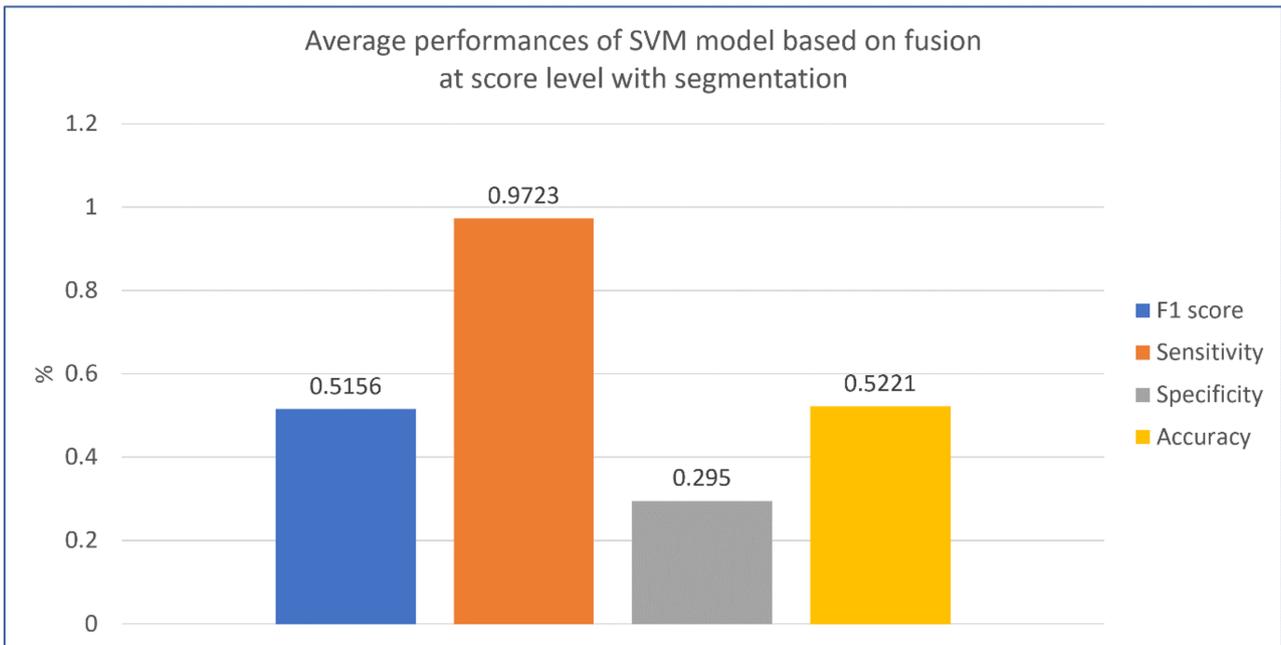


Figure 6.25: Average performances recorded after 10 Cross validation of SVM model based on Fusion at score level for image patch classification with segmentation.

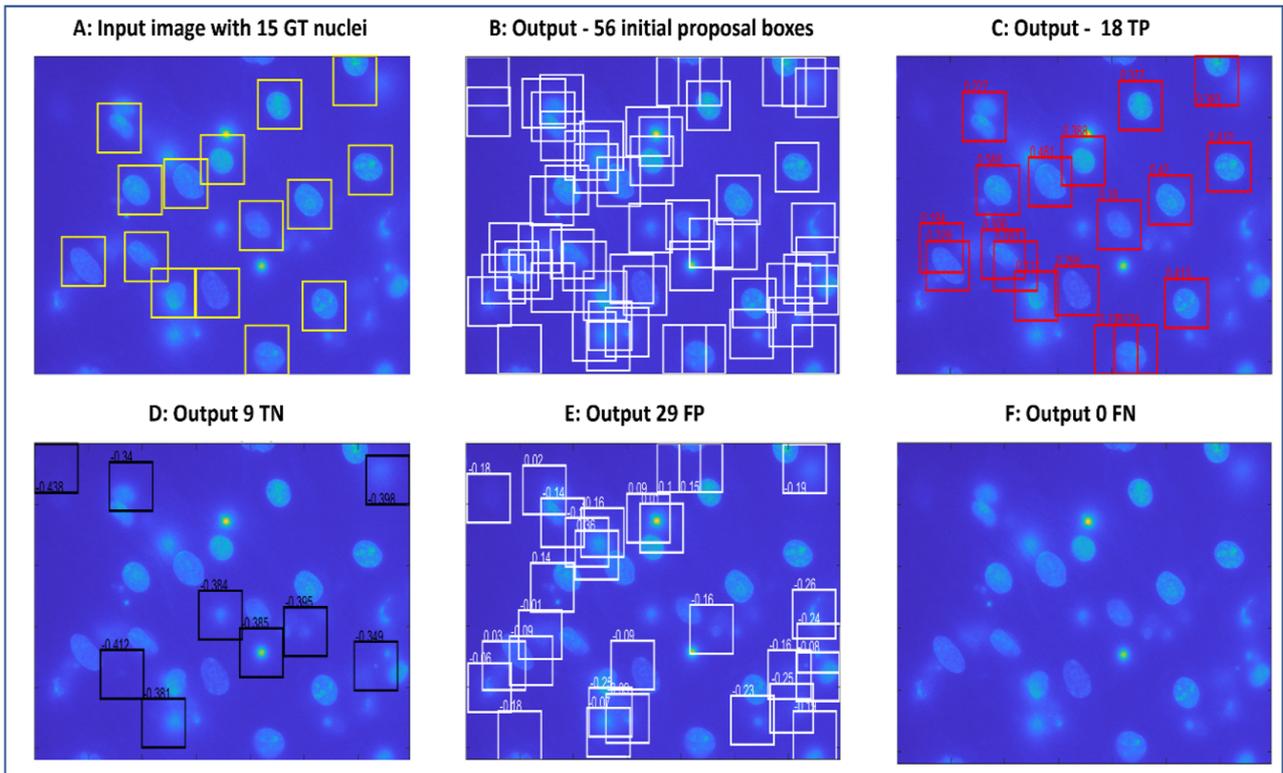


Figure 6.26: Initial threshold nuclei location filtering by SVMs models using fusion at score level. A: original input NRVM image with 15 ground truth labelled nuclei. B: 56 initial locations generated by local adaptive thresholding. C: 18 high-quality boxes filtered by SVMs using score fusion. Confidence scores plotted in red numbers. D: 9 TN boxes filtered out by SVMs using score fusion. Confidences scores plotted in black numbers. E: 29 FP boxes output by SVMs. Confidence scores plotted in white numbers. F: SVMs model ability at filtering all background samples from the initial proposals.

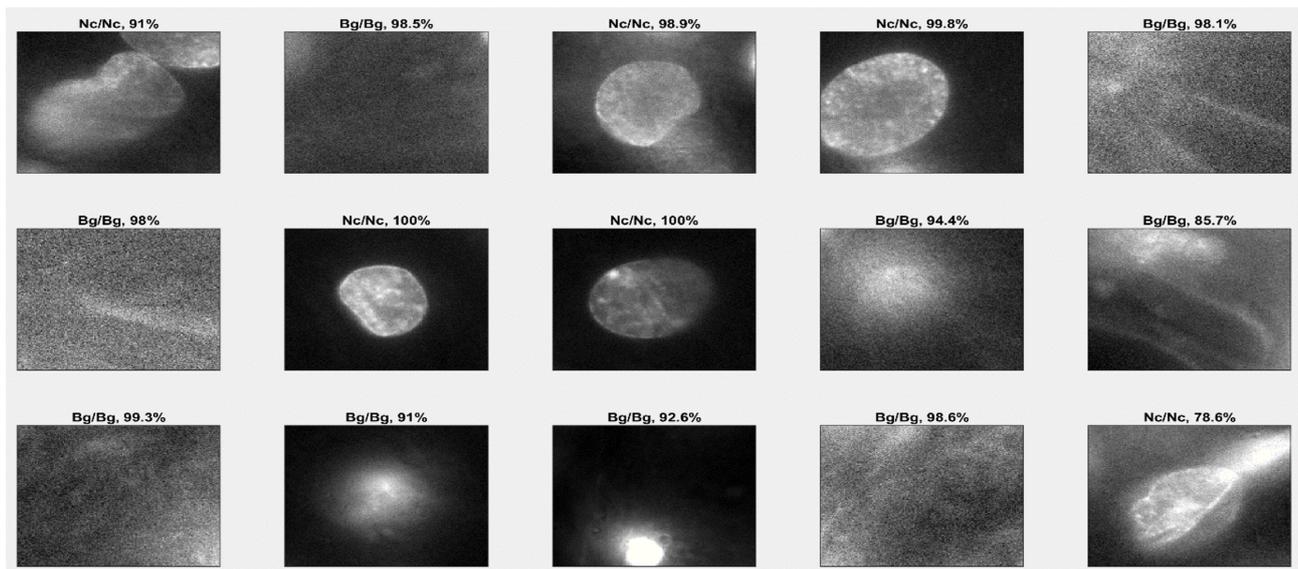


Figure 6.27: Inception V3 based R-CNN detector is highly confident at detecting challenging nuclei and background. The label [Nc/Nc, 100%] means the classifier has predicted a nuclei class at a location labelled as nuclei patch with 100% confidence. In the same way, the label [Bg/Bg, 98.6%] means the classifier has predicted a background class at a location labelled as background patch with 98.6% of confidence.

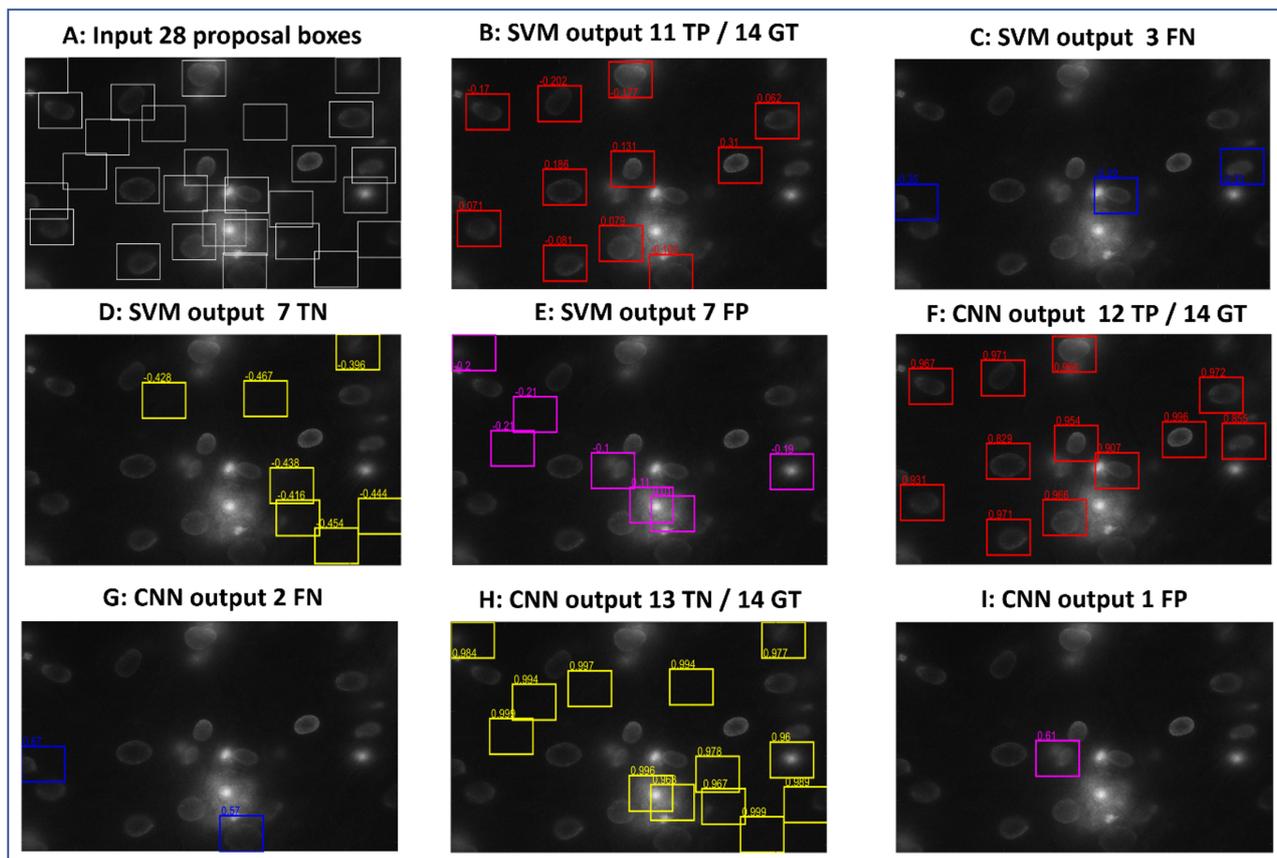


Figure 6.28: CNN classifier outperforms hand-crafted feature based SVM models for nuclei patch classification. White boxes in A: represent initial known proposal boxes fed to SVM fusion model and CNN for classification. Red boxes in B: and F: represent TP predictions made by both classifiers i.e., nuclei boxes correctly classified as nuclei. Blue boxes in C: and G: represent FN predictions made by the classifiers i.e., nuclei boxes wrongly classified as background. Yellow boxes in D: and H: represent TN predictions made by the classifiers i.e., background boxes correctly classified as background. Magenta boxes in E: and I: represent FP predictions made by the classifiers, i.e., background patches wrongly classified as nuclei

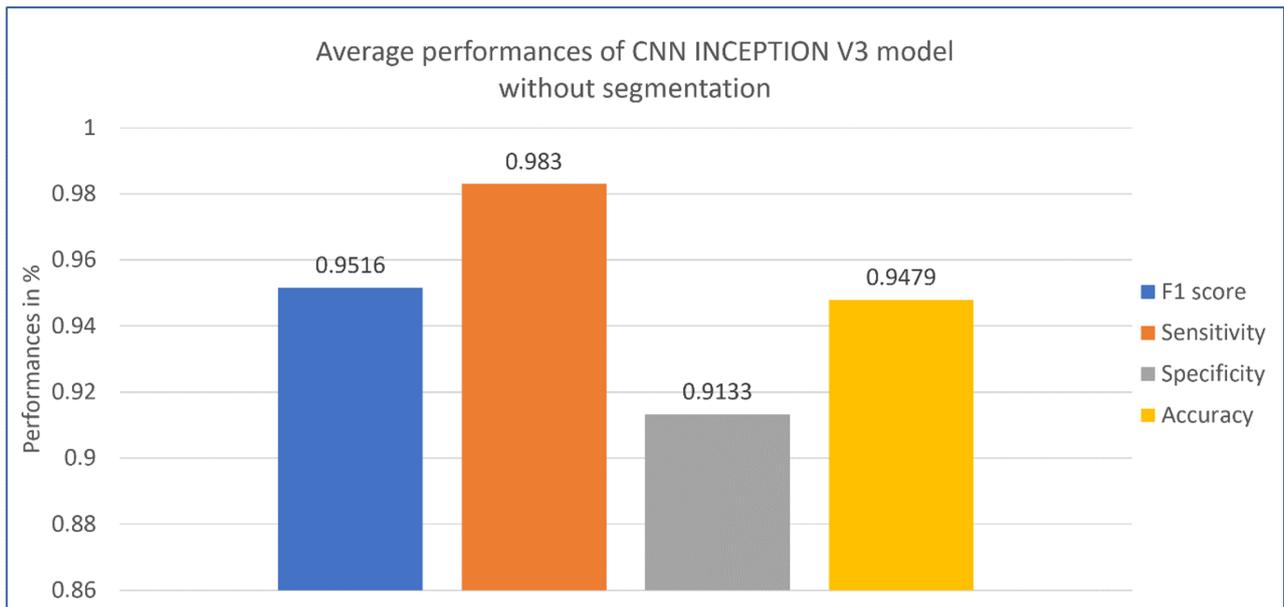


Figure 6.29: Performance of nuclei patch classification by CNN Inception V3 without segmentation.

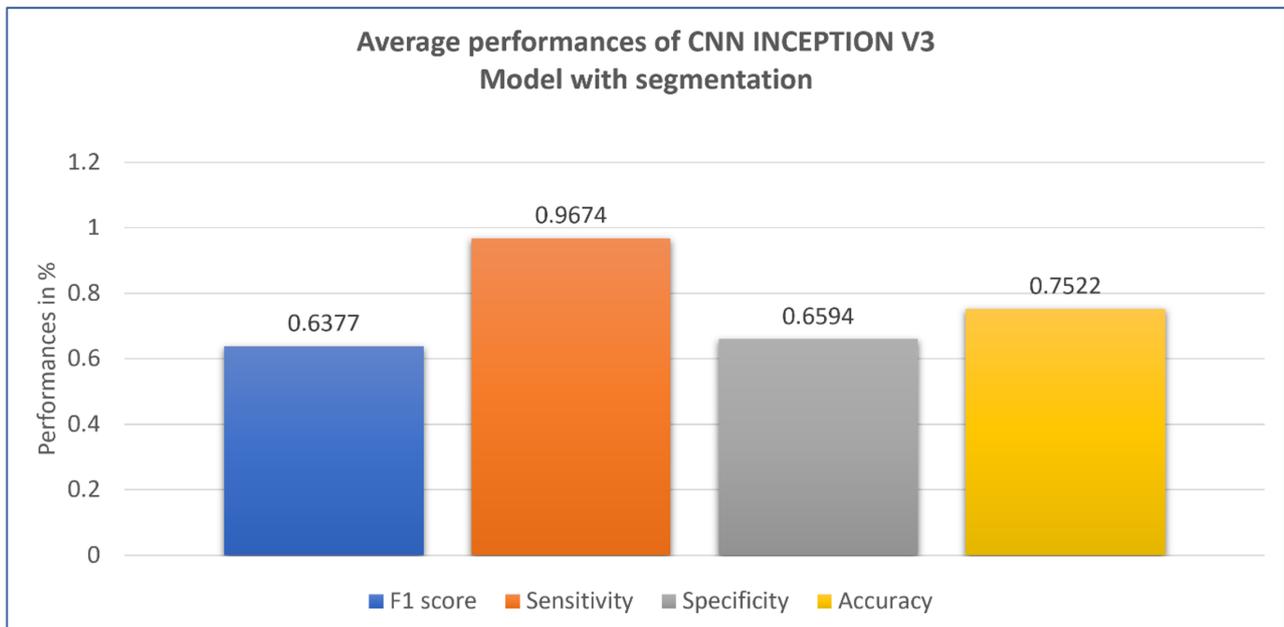


Figure 6.30: Performance of nuclei patch classification by CNN Inception V3 with segmentation.

Chapter 7

General Conclusion

This chapter presents our thesis conclusion in terms of achievements, limitations, and areas for future research in automated analysis of cellular parameters based on fluorescence microscopy (FM) video imaging. The following sections will report those ending notes about our research for automated calcium signalling analysis in cardiac myocytes, and nuclei detection in FM 2D images.

7.1 Achievements

To sum up, our investigation about automated analysis of intracellular signalling parameters through FM video imaging enabled us to reach to some achievements in both computing and cell biology under the following two topics: intracellular calcium signalling in cardiac myocytes and nuclear detection.

7.1.1 Calcium Signal Analysis for Understanding Arrhythmia

Regarding intracellular calcium analysis, we proposed automated solutions for calcium hotspots detection and measurement and calcium hotspot signal processing. The calcium hotspot detection technique proposed was based on a two-level segmentation strategy that combines calcium signal segmentation by multi-level thresholding and the hit-ratio computation of calcium signal locations (3). To the best of our knowledge, our proposed routine for calcium signal hotspot de-

tection in FM video is the first of its kind for automated analysis of best active cardiac myocytes' frequency of calcium signals without prior knowledge in terms of size of the ROIs. Compared to pioneering techniques developed by Cheng [21], our proposed technique outperformed with an average precision margin of 12.2%. The measurement of calcium signal monitored in hotspots sets the platform for our next investigation: calcium hotspots signal processing (4).

The fundamental contribution of our research about calcium signal processing is related to photobleaching detection and correction in calcium signals of cardiac myocytes through FM image sequences. Photobleaching is a degenerative photochemistry phenomenon that shows fluorescence molecules losing their ability to re-emit light at a certain expected wavelength with time [90]. In calcium imaging using FM, the advent of photobleaching causes loss of information about the real level of signal amplitudes. The photobleaching detection and correction we proposed and tested in cardiac myocytes enables scientists to accurately measure of intracellular signalling regardless of experimental conditions.

Chapter 4 demonstrates that photobleaching correction in cardiac myocytes loaded with calcium indicators can be modelled by mono-exponential curve fitting with an MSE varying between [0.09013, 6.41796]. The photobleaching signal recovery method we derived from the curve fitting model is based on signal baseline adjustment. The adjustment is made by adding to the original signals the differences between the expected flat baselines and the estimated baselines exhibiting photobleaching loss. When applied to 80 hotspots exhibiting photobleaching decays in recorded calcium traces, our method achieved a performance of 0.7875 of correction.

Several cardiac dysfunctions, like arrhythmia involving PVCs, one of the cell types used in this study, are caused by abnormal intracellular calcium signals within the cardiac myocytes. Our automated solutions allow a faster and more reliable analysis of these processes and enable experiment replicability. Our proposed routine enabled the automation of the data analysis extracted by biologists from the observation in the laboratory of cells undergoing arrhythmia and transient calcium signals. For the easiness of use, we developed a graphical user interface MATLAB toolbox (INCAS software) for such analysis of video of cardiac myocytes stained with calcium indicators. The INCAS stands for Intracellular Calcium Signal Software. Its description is available in the appendix A.

7.1.2 Nuclei Detection for Autophagy Evolution Analysis

Cell and/or nuclei detection is the starting point for multiple biological-image analysis analyses. Our investigations reported in Chapter 5, and Chapter 6 about deep neural network strategies for nuclei detection in NVRM images showed that data-driven proposal box generation combined to CNN could achieve great performances over ground truth expert labelled dataset.

Experiments demonstrated that standard object proposal techniques such as the sliding window search, the selective search, the objectness search, and the edge box strategy showed huge limitations in terms of the sampling rate (more than 8 000 proposals/image) and mAP (less than 0.3) when applied to NRVM nuclei images. The data-driven proposal box generation we proposed used local adaptive thresholding mask to generate initial nuclei locations, then SVM models to filter the initial locations into better candidates. With such a data-driven strategy, our proposed technique achieved a TPR of 0.9723, an accuracy of 0.57, and an average sampling rate of 61 proposal boxes per image. The final nuclei patch classification by CNN enabled to keep similar TPR (0.9830) and to increase the accuracy by roughly 0.4 (0.9479). All in one, the automated detection of NRVM nuclei undergoing autophagy through FM images enabled us to provide biologist with accurate measurement for autophagy evolution monitoring.

7.2 Limitations

As the contributions, the limitations associated to our research about automated analysis of cellular parameters through FM video imaging can be classified into two main groups: calcium signal analysis and cell segmentation.

7.2.1 Automated Solution for Calcium Signal Analysis

As stated in Section 3.1 cardiac myocytes can show multiple types of calcium signals according to their spatiotemporal characteristics: rising time, spatial spread, amplitude. These calcium signals are essentially calcium puffs, calcium sparks that are localized signals and calcium oscillations, and waves that can are high-intensity repetitive phenomenons and can spread through several neighbouring cells. Our proposed technique for hotspot detection and measurement does not provide such classification of calcium signals in particular, as there is no cell boundary

detection associated to the hotspot segmentation method we investigated.

However, since the automated procedure for intracellular active regions (hotspots) detection depends on initial segmentation of calcium signals events in individual frames no matter their spatiotemporal properties, someone can argue our technique's ability to generalize well onto multiple datasets involving different calcium signal types. Indeed, unlike techniques discussed in the literature that define a fixed size radius for calcium signal event ([34]), or user-input threshold value ([28]), our techniques default parameters for hotspot detection are automatically computed by the algorithm based on the initial calcium signal segmentation output.

Besides, for the validation of our solutions, it can be argued that the manually annotated ground truth dataset we used may not be enough as there were no other available ground truth nuclei images obtained by FM to test against. The validation of the photobleaching detection and correction techniques suffers from the same limitations associated to the lack of other ground truth calcium signals altered by photobleaching decay.

7.2.2 Automated Solution for Cell Detection

The proposed solution for nuclei detection by adaptive R-CNN based on data-driven proposal box generation shows essential limitations at two levels: the initial generation of candidate nuclei location by thresholding and the classification of best candidate locations by CNN. Although the locally adaptive thresholding mask enables to capture almost all potential nuclei objects in NRVM images, pairs of touching and overlapping cells are still output as single entities. Our proposed solution for separating touching cells is semi-automated as it supposes the touching cells patches identified. Besides, the proposed technique for separating overlapping cells depends on the local maxima of image regions. That may be misleading as experiences in Section 6.2 showed that background objects could be brighter than the cells in some extreme cases.

7.3 Areas for Future Research

Areas for future research could include calcium signal classification in FM cardiac myocytes images to give valuable insights to biologists about the cell's frequency of calcium signals.

Integrating into the proposed routine for automated calcium signal detection and measure, cell boundary segmentation and thus ultimate calcium signal differentiation or classification would be highly beneficial to cell biologists.

Although, the proposed automated routine for calcium analysis was tested on cardiac myocytes image sequences, the hotspot detection strategy based on thresholding at frame and video level and using connected component filtering and the hit-ratio principle can be customized and applied to a range of variety of intensity-based measurements of cellular domains involving video recordings such other ions channels (sodium, potassium), or single molecules stains, or cellular domains tracking (marked cancer cells, virus).

Another topic for further research could be to investigate and develop an automated detection and separation of overlapping cells based on CNNs and/or regression algorithms rather than standard image processing techniques such as advanced morphological operations and watershed transforms.

Finally, an area for future study could be the development of biological ground truth datasets at the level of ImageNet made of millions of images classified into different cell domain categories or types (blood cells, neurons, cardiac myocytes, other muscle cells, epithelial cells, bone cells, teeth cells, cancer cells, healthy cells, lysosomes, vesicles, mitochondria, nuclei, etc.). Such huge publicly available datasets may enable the development of deep neural networks with the same performances as the state-of-the-art techniques developed for object detection, recognition, and image classification over natural images.

Bibliography

- [1] B. Alexe, T. Deselaers, and V. Ferrari. “Measuring the Objectness of Image Windows”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34.11 (Nov. 2012), pp. 2189–2202. DOI: 10.1109/tpami.2012.28. URL: <https://doi.org/10.1109/tpami.2012.28>.
- [2] Bogdan Alexe, Thomas Deselaers, and Vittorio Ferrari. “What is an object?” In: *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, June 2010. DOI: 10.1109/cvpr.2010.5540226. URL: <https://doi.org/10.1109/cvpr.2010.5540226>.
- [3] Pablo Arbelaez et al. “Multiscale Combinatorial Grouping”. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, June 2014. DOI: 10.1109/cvpr.2014.49. URL: <https://doi.org/10.1109/cvpr.2014.49>.
- [4] Maedeh Beheshti et al. “Fluorescence microscopy image segmentation based on graph and fuzzy methods: A comparison with ensemble method”. In: *Journal of Intelligent & Fuzzy Systems* 34.4 (Apr. 2018), pp. 2563–2578. DOI: 10.3233/jifs-17466. URL: <https://doi.org/10.3233/jifs-17466>.
- [5] Avi Ben-Cohen et al. “Anatomical data augmentation for CNN based pixel-wise classification”. In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE, Apr. 2018. DOI: 10.1109/isbi.2018.8363762. URL: <https://doi.org/10.1109/isbi.2018.8363762>.
- [6] T. Bernd, M. Kleutges, and A. Kroll. “Nonlinear Black Box Modelling – Fuzzy Networks versus Neural Networks”. In: *Neural Computing & Applications* 8.2 (May 1999), p. 151 162. DOI: 10.1007/s005210050017. URL: <https://doi.org/10.1007/s005210050017>.

-
- [7] Michael J. Berridge. “Calcium regulation of neural rhythms, memory and Alzheimer’s disease”. In: *The Journal of Physiology* 592.2 (July 2013), pp. 281–293. DOI: 10.1113/jphysiol.2013.257527. URL: <https://doi.org/10.1113/jphysiol.2013.257527>.
- [8] Michael J. Berridge, Peter Lipp, and Martin D. Bootman. “The versatility and universality of calcium signalling”. In: *Nature Reviews Molecular Cell Biology* 1.1 (Oct. 2000), pp. 11–21. DOI: 10.1038/35036035. URL: <https://doi.org/10.1038/35036035>.
- [9] Donald M. Bers. “Calcium Cycling and Signaling in Cardiac Myocytes”. In: *Annual Review of Physiology* 70.1 (Mar. 2008), pp. 23–49. DOI: 10.1146/annurev.physiol.70.113006.100455. URL: <https://doi.org/10.1146/annurev.physiol.70.113006.100455>.
- [10] Donald M. Bers. “Cardiac excitation–contraction coupling”. In: *Nature* 415.6868 (Jan. 2002), pp. 198–205. DOI: 10.1038/415198a. URL: <https://doi.org/10.1038/415198a>.
- [11] M. D. Bootman et al. “Loading Fluorescent Ca²⁺ Indicators into Living Cells”. In: *Cold Spring Harbor Protocols* 2013.2 (Feb. 2013), pdb.prot072801–pdb.prot072801. DOI: 10.1101/pdb.prot072801. URL: <https://doi.org/10.1101/pdb.prot072801>.
- [12] Martin D Bootman and Katja Rietdorf. “Atrial myocytes demonstrate the diversity of cardiac calcium signalling”. In: *Channels* 9.5 (Sept. 2015), pp. 219–220. DOI: 10.1080/15384101.2015.1086203. URL: <https://doi.org/10.1080/15384101.2015.1086203>.
- [13] Martin D. Bootman et al. “Ca²⁺-sensitive fluorescent dyes and intracellular Ca²⁺-imaging”. In: *Spring Harb* (2013). URL: <http://cshprotocols.cshlp.org/content/2013/2/pdb.top066050.full.pdfTopic>.
- [14] Martin D. Bootman et al. “Calcium Signalling - An Overview”. In: *Seminars in Cell and Developmental Biology* (2001). URL: <https://doi.org/10.1006/scdb.2000.0211>.
- [15] Martin D. Bootman et al. “The regulation of autophagy by calcium signals: Do we have a consensus?” In: *Cell Calcium* 70 (Mar. 2018), pp. 32–46. DOI: 10.1016/j.ceca.2017.08.005. URL: <https://doi.org/10.1016/j.ceca.2017.08.005>.
- [16] Joao Carreira and Cristian Sminchisescu. “Constrained parametric min-cuts for automatic object segmentation”. In: *2010 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*. IEEE, June 2010. DOI: 10.1109/cvpr.2010.5540063. URL: <https://doi.org/10.1109/cvpr.2010.5540063>.

-
- [17] Chih-Chung Chang and Chih-Jen Lin. “LIBSVM”. In: *ACM Transactions on Intelligent Systems and Technology* 2.3 (Apr. 2011), pp. 1–27. DOI: 10.1145/1961189.1961199. URL: <https://doi.org/10.1145/1961189.1961199>.
- [18] Ken Chatfield, Karen Simonyan, and Andrew Zisserman. “Efficient On-the-fly Category Retrieval Using ConvNets and GPUs”. In: *Computer Vision ACCV 2014*. Springer International Publishing, 2015, pp. 129–145. DOI: 10.1007/978-3-319-16865-4_9. URL: https://doi.org/10.1007/978-3-319-16865-4_9.
- [19] Tala Chehab. “The Role of Calcium Signalling in Autophagy Tala Chehab , BSc”. In: (2017).
- [20] Tsai-Wen Chen et al. “In Situ Background Estimation in Quantitative Fluorescence Imaging”. In: *Biophysical Journal* 90.7 (Apr. 2006), pp. 2534–2547. DOI: 10.1529/biophysj.105.070854. URL: <https://doi.org/10.1529/biophysj.105.070854>.
- [21] Heping Cheng et al. “Amplitude Distribution of Calcium Sparks in Confocal Images: Theory and Studies with an Automatic Detection Method”. In: *Biophysical Journal* 76.2 (Feb. 1999), pp. 606–617. DOI: 10.1016/s0006-3495(99)77229-2. URL: <https://doi.org/10.1016/s0006-3495%5C%2899%5C%2977229-2>.
- [22] N. Dalal and B. Triggs. “Histograms of Oriented Gradients for Human Detection”. In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR’05)*. IEEE. DOI: 10.1109/cvpr.2005.177. URL: <https://doi.org/10.1109/cvpr.2005.177>.
- [23] A DARSZON et al. “T-type Calcium channels in sperm function”. In: *Cell Calcium* 40.2 (Aug. 2006), pp. 241–252. DOI: 10.1016/j.ceca.2006.04.028. URL: <https://doi.org/10.1016/j.ceca.2006.04.028>.
- [24] Piotr Dollar and C. Lawrence Zitnick. “Fast Edge Detection Using Structured Forests”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 37.8 (Aug. 2015), pp. 1558–1570. DOI: 10.1109/tpami.2014.2377715. URL: <https://doi.org/10.1109/tpami.2014.2377715>.
- [25] Bo Dong et al. “Deep learning for automatic cell detection in wide-field microscopy zebrafish images”. In: *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*. IEEE, Apr. 2015. DOI: 10.1109/isbi.2015.7163986. URL: <https://doi.org/10.1109/isbi.2015.7163986>.

-
- [26] Niraj P. Doshi and Gerald Schaefer. “Texture classification using compact multi dimensional local binary pattern descriptors”. In: *2013 International Conference on Informatics, Electronics and Vision (ICIEV)*. IEEE, May 2013. DOI: 10.1109/iciev.2013.6572563. URL: <https://doi.org/10.1109/iciev.2013.6572563>.
- [27] Simon S. Du and Yining Wang. “How Many Samples are Needed to Estimate a Convolutional Neural Network?” In: *2018 Conference on Neural Information Processing Systems (NeurIPS)*. NIPS, July 2018. DOI: 10.1109/iciscae.2018.8666928. URL: <https://doi.org/10.1109/iciscae.2018.8666928>.
- [28] Kyle L. Ellefsen et al. “An algorithm for automated detection, localization and measurement of local calcium signals from camera-based imaging”. In: *Cell Calcium* 56.3 (2014), pp. 147–156. ISSN: 15321991. DOI: 10.1016/j.ceca.2014.06.003. URL: [/pmc/articles/PMC4162823/](https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4162823/)<https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4162823/?report=abstract><https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4162823/>.
- [29] Ian Endres and Derek Hoiem. “Category-Independent Object Proposals with Diverse Ranking”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 36.2 (Feb. 2014), pp. 222–234. DOI: 10.1109/tpami.2013.122. URL: <https://doi.org/10.1109/tpami.2013.122>.
- [30] Mark Everingham et al. “The Pascal Visual Object Classes (VOC) Challenge”. In: *International Journal of Computer Vision* 88.2 (Sept. 2009), pp. 303–338. DOI: 10.1007/s11263-009-0275-4. URL: <https://doi.org/10.1007/s11263-009-0275-4>.
- [31] C. J. Fearnley, H. L. Roderick, and M. D. Bootman. “Calcium Signaling in Cardiac Myocytes”. In: *Cold Spring Harbor Perspectives in Biology* 3.11 (Aug. 2011), a004242–a004242. DOI: 10.1101/cshperspect.a004242. URL: <https://doi.org/10.1101/cshperspect.a004242>.
- [32] Pedro F. Felzenszwalb and Daniel P. Huttenlocher. “Efficient Graph-Based Image Segmentation”. In: *International Journal of Computer Vision* 59.2 (Sept. 2004), pp. 167–181. DOI: 10.1023/b:visi.0000022288.19776.77. URL: <https://doi.org/10.1023/b:visi.0000022288.19776.77>.
- [33] David Forsyth. “Object Detection with Discriminatively Trained Part-Based Models”. In: *Computer* 47.2 (Feb. 2014), pp. 6–7. DOI: 10.1109/mc.2014.42. URL: <https://doi.org/10.1109/mc.2014.42>.

-
- [34] Michael Francis et al. “Automated region of interest analysis of dynamic Ca^2 signals in image sequences”. In: *American Journal of Physiology-Cell Physiology* 303.3 (Aug. 2012), pp. C236–C243. ISSN: 0363-6143. DOI: 10.1152/ajpcell.00016.2012. URL: <https://www.physiology.org/doi/10.1152/ajpcell.00016.2012>.
- [35] Ross Girshick et al. “Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation”. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition*. IEEE, June 2014. DOI: 10.1109/cvpr.2014.81. URL: <https://doi.org/10.1109/cvpr.2014.81>.
- [36] Rafael C. Gonzalez, Richard E. Woods, and Barry R. Masters. “Digital Image Processing, Third Edition”. In: *Journal of Biomedical Optics* 14.2 (2009), p. 029901. DOI: 10.1117/1.3115362. URL: <https://doi.org/10.1117/1.3115362>.
- [37] Igor Griva, Stephen G. Nash, and Ariela Sofer. *Linear and Nonlinear Optimization*. SIAM, Jan. 2009. DOI: 10.1137/1.9780898717730. URL: <https://doi.org/10.1137/1.9780898717730>.
- [38] Kaiming He et al. “Mask R-CNN”. In: *2017 IEEE International Conference on Computer Vision (ICCV)*. IEEE, Oct. 2017. DOI: 10.1109/iccv.2017.322. URL: <https://doi.org/10.1109/iccv.2017.322>.
- [39] Reka Hollandi et al. “A deep learning framework for nucleus segmentation using image style transfer”. In: (Mar. 2019). DOI: 10.1101/580605. URL: <https://doi.org/10.1101/580605>.
- [40] C.Y. Hong, B.N. Chiang, and P. Turner. “CALCIUM ION IS THE KEY REGULATOR OF HUMAN SPERM FUNCTION”. In: *The Lancet* 324.8417-8418 (Dec. 1984), pp. 1449–1451. DOI: 10.1016/s0140-6736(84)91634-9. URL: <https://doi.org/10.1016/s0140-6736%5C%2884%5C%2991634-9>.
- [41] Jan Hosang, Rodrigo Benenson, and Bernt Schiele. “How good are detection proposals, really?” In: *Proceedings of the British Machine Vision Conference 2014*. British Machine Vision Association, 2014. DOI: 10.5244/c.28.24. URL: <https://doi.org/10.5244/c.28.24>.
- [42] Humayun Irshad et al. “Methods for Nuclei Detection, Segmentation, and Classification in Digital Histopathology: A Review—Current Status and Future Potential”. In: *IEEE Reviews in Biomedical Engineering* 7 (2014), pp. 97–114. DOI: 10.1109/rbme.2013.2295804. URL: <https://doi.org/10.1109/rbme.2013.2295804>.

-
- [43] Andrey Kan. “Machine learning applications in cell image analysis”. In: *Immunology & Cell Biology* 95.6 (Apr. 2017), pp. 525–530. DOI: 10.1038/icb.2017.16. URL: <https://doi.org/10.1038/icb.2017.16>.
- [44] Chung-Chia Kang, Wen-June Wang, and Chung-Hao Kang. “Image segmentation with complicated background by using seeded region growing”. In: *AEU - International Journal of Electronics and Communications* 66.9 (Sept. 2012), pp. 767–771. DOI: 10.1016/j.aeue.2012.01.011. URL: <https://doi.org/10.1016/j.aeue.2012.01.011>.
- [45] Charles Kervrann et al. “A Guided Tour of Selected Image Processing and Analysis Methods for Fluorescence and Electron Microscopy”. In: *IEEE Journal of Selected Topics in Signal Processing* 10.1 (Feb. 2016), pp. 6–30. DOI: 10.1109/jstsp.2015.2505402. URL: <https://doi.org/10.1109/jstsp.2015.2505402>.
- [46] Kirill Kiselyov et al. “Aberrant Ca²⁺ handling in lysosomal storage disorders”. In: *Cell Calcium* 47.2 (Feb. 2010), pp. 103–111. DOI: 10.1016/j.ceca.2009.12.007. URL: <https://doi.org/10.1016/j.ceca.2009.12.007>.
- [47] Y. Al-Kofahi et al. “Improved Automatic Detection and Segmentation of Cell Nuclei in Histopathology Images”. In: *IEEE Transactions on Biomedical Engineering* 57.4 (Apr. 2010), pp. 841–852. DOI: 10.1109/tbme.2009.2035102. URL: <https://doi.org/10.1109/tbme.2009.2035102>.
- [48] Yousef Al-Kofahi et al. “A deep learning-based algorithm for 2-D cell segmentation in microscopy images”. In: *BMC Bioinformatics* 19.1 (Oct. 2018). DOI: 10.1186/s12859-018-2375-z. URL: <https://doi.org/10.1186/s12859-018-2375-z>.
- [49] Peter Kovesi. “Phase congruency detects corners and edges”. In: *Digital Image Computing: Techniques and Applications* (2003).
- [50] Peter Kovesi. “Phase congruency: A low-level image invariant”. In: *Psychological Research* 64.2 (Dec. 2000), pp. 136–148. DOI: 10.1007/s004260000024. URL: <https://doi.org/10.1007/s004260000024>.
- [51] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. “ImageNet classification with deep convolutional neural networks”. In: *ILSVRC-2012* 60.6 (May 2012), pp. 84–90. DOI: 10.1145/3065386. URL: <http://www.cs.toronto.edu/~hinton/absps/imagenet.pdf>.

-
- [52] Angela M. Lee, Huw Colin-York, and Marco Fritzsche. “CalQuo 2: Automated Fourier-space, population-level quantification of global intracellular calcium responses”. In: *Scientific Reports* 7.1 (July 2017). DOI: 10.1038/s41598-017-05322-z. URL: <https://doi.org/10.1038/s41598-017-05322-z>.
- [53] Tsung Yi Lin and Piotr Dollar. “Microsoft COCO: Common Objects in Context”. In: *Computer Vision and Pattern Recognition* (Dec. 2014). DOI: arXiv:1405.0312. URL: <https://arxiv.org/abs/1405.0312>.
- [54] Matsilele Mabaso, Daniel Withey, and Bhekisipho Twala. “Spot Detection in Microscopy Images using Convolutional Neural Network with Sliding-Window Approach”. In: *Proceedings of the 11th International Joint Conference on Biomedical Engineering Systems and Technologies*. SCITEPRESS - Science and Technology Publications, 2018. DOI: 10.5220/0006724200670074. URL: <https://doi.org/10.5220/0006724200670074>.
- [55] Kate Maxwell and Giles N. Johnson. “Chlorophyll fluorescence—a practical guide”. In: *Journal of Experimental Botany* 51.345 (Apr. 2000), pp. 659–668. DOI: 10.1093/jexbot/51.345.659. URL: <https://doi.org/10.1093/jexbot%5C%2F51.345.659>.
- [56] Aymen Mouelhi et al. “Automatic image segmentation of nuclear stained breast tissue sections using color active contour model and an improved watershed method”. In: *Biomedical Signal Processing and Control* 8.5 (Sept. 2013), pp. 421–436. DOI: 10.1016/j.bspc.2013.04.003. URL: <https://doi.org/10.1016/j.bspc.2013.04.003>.
- [57] Aymen Mouelhi et al. “Fast unsupervised nuclear segmentation and classification scheme for automatic allred cancer scoring in immunohistochemical breast tissue images”. In: *Computer Methods and Programs in Biomedicine* 165 (Oct. 2018), pp. 37–51. DOI: 10.1016/j.cmpb.2018.08.005. URL: <https://doi.org/10.1016/j.cmpb.2018.08.005>.
- [58] Vidhya Nair, Eric C. Belanger, and John P. Veinot. “Lysosomal storage disorders affecting the heart: a review”. In: *Cardiovascular Pathology* 39 (Mar. 2019), pp. 12–24. DOI: 10.1016/j.carpath.2018.11.002. URL: <https://doi.org/10.1016/j.carpath.2018.11.002>.
- [59] Mark S. Nixon and Alberto S. Aguado. “Introduction”. In: *Feature Extraction & Image Processing for Computer Vision*. Elsevier, 2012, pp. 1–36. DOI: 10.1016/b978-0-12-396549-3.00001-x. URL: <https://doi.org/10.1016/b978-0-12-396549-3.00001-x>.

-
- [60] Thomas A. Nketia et al. “Analysis of live cell images: Methods, tools and opportunities”. In: *Methods* 115 (Feb. 2017), pp. 65–79. DOI: 10.1016/j.ymeth.2017.02.007. URL: <https://doi.org/10.1016/j.ymeth.2017.02.007>.
- [61] Rujuta O and Vyavahare AJ. “Review of Nuclei Detection, Segmentation in Microscopic Images”. In: *Journal of Bioengineering & Biomedical Science* 07.02 (2017). DOI: 10.4172/2155-9538.1000227. URL: <https://doi.org/10.4172/2155-9538.1000227>.
- [62] Nobuyuki Otsu. “A Threshold Selection Method from Gray-Level Histograms”. In: *IEEE Transactions on Systems, Man, and Cybernetics* 9.1 (Jan. 1979), pp. 62–66. DOI: 10.1109/tsmc.1979.4310076. URL: <https://doi.org/10.1109/tsmc.1979.4310076>.
- [63] Pawan Patidar et al. “Image De-noising by Various Filters for Different Noise”. In: *International Journal of Computer Applications* 9.4 (Nov. 2010), pp. 45–50. DOI: 10.5120/1370-1846. URL: <https://doi.org/10.5120/1370-1846>.
- [64] Thierry Pecot et al. “Background Fluorescence Estimation and Vesicle Segmentation in Live Cell Imaging With Conditional Random Fields”. In: *IEEE Transactions on Image Processing* 24.2 (Feb. 2015), pp. 667–680. DOI: 10.1109/tip.2014.2380178. URL: <https://doi.org/10.1109/tip.2014.2380178>.
- [65] Neerad Phansalkar et al. “Adaptive local thresholding for detection of nuclei in diversity stained cytology images”. In: *2011 International Conference on Communications and Signal Processing*. IEEE, Feb. 2011. DOI: 10.1109/iccsp.2011.5739305. URL: <https://doi.org/10.1109/iccsp.2011.5739305>.
- [66] Shanto Rahman et al. “An adaptive gamma correction for image enhancement”. In: *EURASIP Journal on Image and Video Processing* 2016.1 (Oct. 2016). DOI: 10.1186/s13640-016-0138-1. URL: <https://doi.org/10.1186/s13640-016-0138-1>.
- [67] Shaoqing Ren et al. “Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.6 (June 2017), pp. 1137–1149. DOI: 10.1109/tpami.2016.2577031. URL: <https://doi.org/10.1109/tpami.2016.2577031>.
- [68] Katja Rietdorf, Martin D. Bootman, and Michael J. Sanderson. “Spontaneous, Pro-Arrhythmic Calcium Signals Disrupt Electrical Pacing in Mouse Pulmonary Vein Sleeve Cells”. In: *PLoS ONE* 9.2 (Feb. 2014). Ed. by Mohamed Trebak, e88649. DOI: 10.1371/journal.pone.0088649. URL: <https://doi.org/10.1371/journal.pone.0088649>.

-
- [69] Katja Rietdorf et al. “Pulmonary vein sleeve cell excitation–contraction-coupling becomes desynchronized by spontaneous calcium transients”. In: *Biochemical Society Transactions* 43.3 (June 2015), pp. 410–416. DOI: 10.1042/bst20140299. URL: <https://doi.org/10.1042/bst20140299>.
- [70] Isabel Rodrigues and Joao Sanches. “Photoblinking/photobleaching differential equation model for intensity decay of fluorescence microscopy images”. In: *2010 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE, Apr. 2010. DOI: 10.1109/isbi.2010.5490226. URL: <https://doi.org/10.1109/isbi.2010.5490226>.
- [71] Isabel Rodrigues, Joao Sanches, and Jose Bioucas-Dias. “Denoising of medical images corrupted by Poisson noise”. In: *2008 15th IEEE International Conference on Image Processing*. IEEE, 2008. DOI: 10.1109/icip.2008.4712115. URL: <https://doi.org/10.1109/icip.2008.4712115>.
- [72] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. “U-Net: Convolutional Networks for Biomedical Image Segmentation”. In: *Lecture Notes in Computer Science*. Springer International Publishing, 2015, pp. 234–241. DOI: 10.1007/978-3-319-24574-4_28. URL: https://doi.org/10.1007/978-3-319-24574-4_28.
- [73] Lim Jae S. “Two-dimensional signal and image processing”. In: (1990). DOI: 10.4103/jpi.jpi_74_17. URL: <https://ui.adsabs.harvard.edu/abs/1990ph...book...L/abstract>.
- [74] D. Sage et al. “Automatic tracking of individual fluorescence particles: application to the study of chromosome dynamics”. In: *IEEE Transactions on Image Processing* 14.9 (Sept. 2005), pp. 1372–1383. DOI: 10.1109/tip.2005.852787. URL: <https://doi.org/10.1109/tip.2005.852787>.
- [75] Koen E. A. van de Sande et al. “Segmentation as selective search for object recognition”. In: *2011 International Conference on Computer Vision*. IEEE, Nov. 2011. DOI: 10.1109/iccv.2011.6126456. URL: <https://doi.org/10.1109/iccv.2011.6126456>.
- [76] Syed Islamuddin Shah et al. “CellSpecks: A Software for Automated Detection and Analysis of Calcium Channels in Live Cells”. In: *Biophysical Journal* 115.11 (Dec. 2018), pp. 2141–2151. ISSN: 15420086. DOI: 10.1016/j.bpj.2018.10.015. URL: <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6289074/>.

-
- [77] Evan Shelhamer, Jonathan Long, and Trevor Darrell. “Fully Convolutional Networks for Semantic Segmentation”. In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 39.4 (Apr. 2017), pp. 640–651. DOI: 10.1109/tpami.2016.2572683. URL: <https://doi.org/10.1109/tpami.2016.2572683>.
- [78] Osamu Shimomura, Martin Chalfie, and Roger Y. Tsien. “Discovery of Green Fluorescent Protein (GFP) (Nobel Lecture) / GFP: Lighting Up Life (Nobel Lecture) / Constructing and Exploiting the Fluorescent Protein Paintbox (Nobel Lecture) (Angew. Chem. Int. Ed. 31/2009)”. In: *Angewandte Chemie International Edition* 48.31 (July 2009), pp. 5555–5555. DOI: 10.1002/anie.200990159. URL: <https://doi.org/10.1002/anie.200990159>.
- [79] L. Song et al. “Photobleaching kinetics of fluorescein in quantitative fluorescence microscopy”. In: *Biophysical Journal* 68.6 (June 1995), pp. 2588–2600. DOI: 10.1016/s0006-3495(95)80442-x. URL: <https://doi.org/10.1016/s0006-3495%5C%2895%5C%2980442-x>.
- [80] RonaldJoe Stanley et al. “Deep Learning Nuclei Detection in Digitized Histology Images by Superpixels”. In: *Journal of Pathology Informatics* 9.1 (2018), p. 5. DOI: 10.4103/jpi.jpi_74_17. URL: https://doi.org/10.4103/jpi.jpi_74_17.
- [81] Tilo Strutz. “The Idea behind Least Squares”. In: *Data Fitting and Uncertainty*. Vieweg Teubner, 2011, pp. 146–166. DOI: 10.1007/978-3-8348-9813-5_6. URL: https://doi.org/10.1007/978-3-8348-9813-5_6.
- [82] Thattapon Surasak et al. “Histogram of oriented gradients for human detection in video”. In: *2018 5th International Conference on Business and Industrial Research (ICBIR)*. IEEE, May 2018. DOI: 10.1109/icbir.2018.8391187. URL: <https://doi.org/10.1109/icbir.2018.8391187>.
- [83] Christian Szegedy et al. “Going deeper with convolutions”. In: *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, June 2015. DOI: 10.1109/cvpr.2015.7298594. URL: <https://doi.org/10.1109/cvpr.2015.7298594>.
- [84] Christian Szegedy et al. “Rethinking the Inception Architecture for Computer Vision”. In: *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. IEEE, June 2016. DOI: 10.1109/cvpr.2016.308. URL: <https://doi.org/10.1109/cvpr.2016.308>.

-
- [85] Afaf Tareef et al. “Multi-Pass Fast Watershed for Accurate Segmentation of Overlapping Cervical Cells”. In: *IEEE Transactions on Medical Imaging* 37.9 (Sept. 2018), pp. 2044–2059. DOI: 10.1109/tmi.2018.2815013. URL: <https://doi.org/10.1109/tmi.2018.2815013>.
- [86] David Traore et al. “Automatic Hotspots Detection for Intracellular Calcium Analysis in Fluorescence Microscopic Videos”. In: *Communications in Computer and Information Science*. Springer International Publishing, 2017, pp. 862–873. DOI: 10.1007/978-3-319-60964-5_75. URL: https://doi.org/10.1007/978-3-319-60964-5_75.
- [87] J. R. R. Uijlings et al. “Selective Search for Object Recognition”. In: *International Journal of Computer Vision* 104.2 (Apr. 2013), pp. 154–171. DOI: 10.1007/s11263-013-0620-5. URL: <https://doi.org/10.1007/s11263-013-0620-5>.
- [88] Bino Sebastian V. “Grey Level Co-Occurrence Matrices: Generalisation and Some New Features”. In: *International Journal of Computer Science, Engineering and Information Technology* 2.2 (Apr. 2012), pp. 151–157. DOI: 10.5121/ijcseit.2012.2213. URL: <https://doi.org/10.5121/ijcseit.2012.2213>.
- [89] M. Veta et al. “Marker-controlled watershed segmentation of nuclei in H&E stained breast cancer biopsy images”. In: *2011 IEEE International Symposium on Biomedical Imaging: From Nano to Macro*. IEEE, Mar. 2011. DOI: 10.1109/isbi.2011.5872483. URL: <https://doi.org/10.1109/isbi.2011.5872483>.
- [90] Nathalie B Vicente et al. “Photobleaching correction in fluorescence microscopy images”. In: *Journal of Physics: Conference Series* 90 (Nov. 2007), p. 012068. DOI: 10.1088/1742-6596/90/1/012068. URL: <https://doi.org/10.1088/1742-6596/90/1/012068>.
- [91] Li Wang. “Tracking and detecting objects in image sequence”. PhD thesis. 2016. DOI: 10.32657/10356/65878. URL: <https://doi.org/10.32657/10356/65878>.
- [92] Jennifer C. Waters. “Accuracy and precision in quantitative fluorescence microscopy”. In: *The Journal of Cell Biology* 185.7 (June 2009), pp. 1135–1148. DOI: 10.1083/jcb.200903097. URL: <https://doi.org/10.1083/jcb.200903097>.
- [93] Fuyong Xing and Lin Yang. “Robust Nucleus/Cell Detection and Segmentation in Digital Pathology and Microscopy Images: A Comprehensive Review”. In: *IEEE Reviews in Biomedical Engineering* 9 (2016), pp. 234–263. DOI: 10.1109/rbme.2016.2515127. URL: <https://doi.org/10.1109/rbme.2016.2515127>.

-
- [94] Yao Xue et al. “A novel framework to integrate convolutional neural network with compressed sensing for cell detection”. In: *2017 IEEE International Conference on Image Processing (ICIP)*. IEEE, Sept. 2017. DOI: 10.1109/icip.2017.8296696. URL: <https://doi.org/10.1109/icip.2017.8296696>.
- [95] Huei-Fang Yang, Chu-Song Chen, and Xavier Descombes. “Detection and tracking of Golgi outposts in microscopy data”. In: *2015 IEEE 12th International Symposium on Biomedical Imaging (ISBI)*. IEEE, Apr. 2015. DOI: 10.1109/isbi.2015.7163991. URL: <https://doi.org/10.1109/isbi.2015.7163991>.
- [96] Su Yang et al. “Faster R-CNN based microscopic cell detection”. In: *2017 International Conference on Security, Pattern Analysis, and Cybernetics (SPAC)*. IEEE, Dec. 2017. DOI: 10.1109/spac.2017.8304302. URL: <https://doi.org/10.1109/spac.2017.8304302>.
- [97] C. Lawrence Zitnick and Piotr Dollár. “Edge Boxes: Locating Object Proposals from Edges”. In: *Computer Vision – ECCV 2014*. Springer International Publishing, 2014, pp. 391–405. DOI: 10.1007/978-3-319-10602-1_26. URL: https://doi.org/10.1007/978-3-319-10602-1_26.
- [98] Karel Zuiderveld. “Contrast Limited Adaptive Histogram Equalization”. In: *Graphics Gems*. Elsevier, 1994, pp. 474–485. DOI: 10.1016/b978-0-12-336156-1.50061-6. URL: <https://doi.org/10.1016/b978-0-12-336156-1.50061-6>.

Appendix A

Description of INCAS: Intracellular Calcium Analysis Software

This appendix presents the description of INCAS, the Intracellular Calcium Analysis Software developed in MATLAB through our research study to provide biologists with a Graphical User Interface (GUI) app integrating some of the automated solutions we presented in the thesis. INCAS can be defined as a data analysis toolbox for biomedical images processing and dedicated to calcium signal imaging. It works only on standard Windows Personal Computers set up with Windows 7 and the latest Windows Operating System. The software package can handle image stacks and videos of several formats: TIFF, MPEG, AVI, and MOV and process up to 1500 frames at a time. It output segmentation masks into image formats, and hotspots locations into MATLAB files. It also provides a set of measurements derived from changes of calcium concentration into text, excel sheets, and MATLAB files. The following sections give an overview of the software application in cell biology, its interface description, and some measurements output.

A.1 Overview

Monitoring intracellular calcium signalling in cardiac myocytes loaded with fluorescent calcium indicators often involves the detection of the most active cells, i.e. the calcium hotspots and the monitoring of the changes of calcium concentration inside them. To the best of our knowledge, there is no software solution that fully automates such process in a friendly manner. In contrast,

intracellular calcium hotspot analysis usually proceeds manually or semi-automatically through the combination of multiple proprietary software such as ImageJ and GraphPad Prism.

INCAS comes as an all-in-one solution that enables both image processing and statistical measurements for automated intracellular calcium signal analysis through calcium movies. The new software package we describe here is meant to fulfil biologists' need for a data analysis toolbox that allows the detection and measurement of intracellular calcium hotspots from 2-D fluorescent image stacks or video inputs of cardiac myocytes. The imaging data should involve cells stained with calcium indicators with a single wavelength. INCAS can automatically scan calcium imaging video experiments, detect the most active cells, and quantify signal measurements into curves, peaks, frequencies of peaks, amplitudes of maximum responses, inter-spikes intervals, rise times, and area under curves.

A.2 Interfaces and Features Description

The software interface consists of two main Microsoft Window-like forms: a main interface for segmentation of calcium signal areas and hotspots (Figure A.1), and a second interface for plotting calcium signals and extracting measurements (Figure A.2).

As it can be noticed in Figure A.1, through the main interface, users can upload videos of calcium imaging experiments and perform thresholding and hotspot detection automatically or semi-automatically by setting up parameters through available user-friendly control buttons (red labels 1,3,4,7). The screenshot below shows that 43 hotspots have been detected from a video of 301 frames showing calcium imaging experiments involving PVC cells. By scanning all the frames, the software we're able to compute an optimum threshold of 0.64 to segment calcium signal areas in individual frames and to detect the best active regions highlighted in white (red label 9) based on default parameters automatically defined such as hotspot size, hit-ratio and separating distance (red label 4).

The whole main interface illustrated in Figure A.1 for essentially uploading calcium imaging videos, thresholding calcium signal area at two levels and detecting calcium hotspot can be described as the following:

- Red label 1 of Figure A.1 represents a set of functionalities that allow users to upload

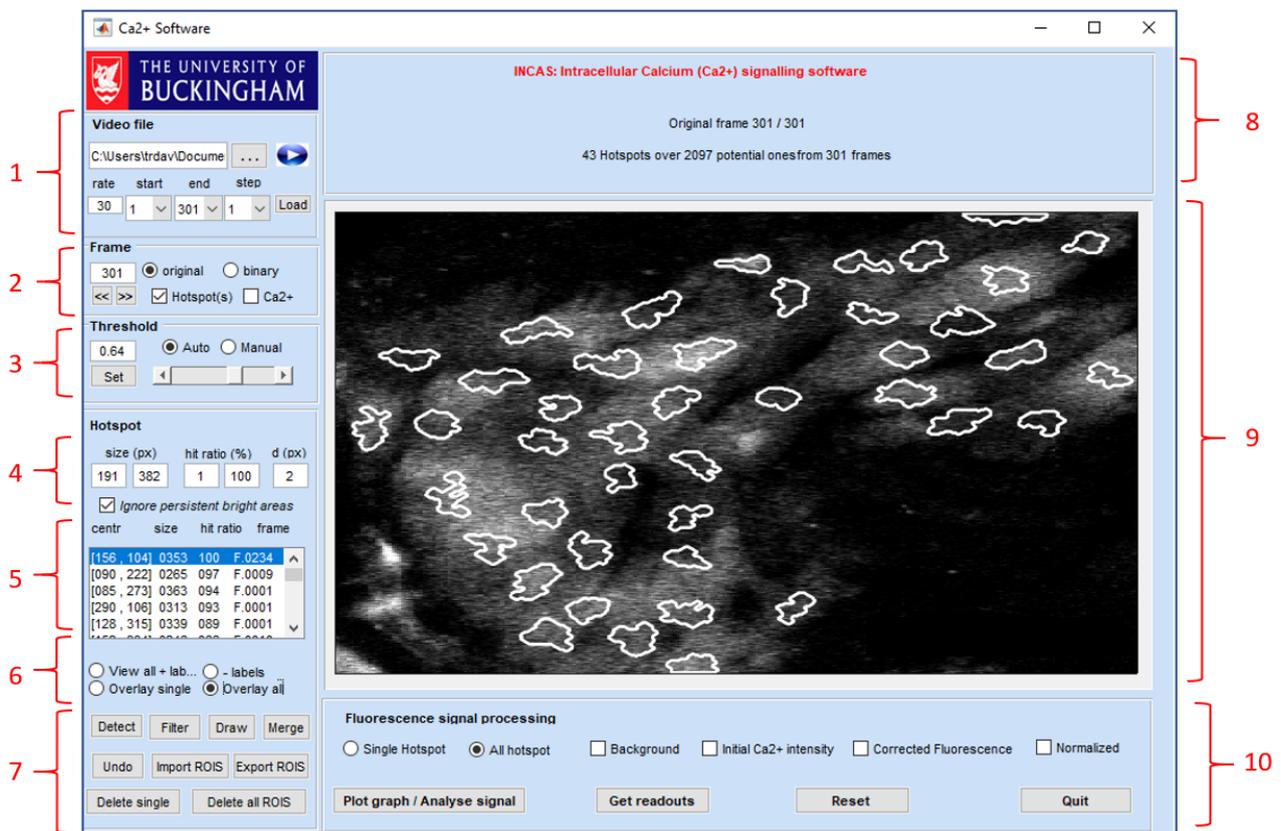


Figure A.1: INCAS main page - screenshot

fluorescent video into the app according to a range and step of frames, and to play the video once uploaded. During the loading process, an optimal threshold is computed by the system and enables the segmentation of calcium signal events in individual video frames according to the routine described in the thesis section 3.2. Here it can be noticed that a video recorded at a rate of 30 fps, has been uploaded from the D directory of a computer in a such way that all his 301 frames are accessible to the user and loaded into the working space of the app.

- Red label 2 shows a group of features that enable users to navigate through individual greyscale frames and binary frames obtained after optimum thresholding. Radio buttons “original” and “binary” indicate to users which frame to display. If hotspots have been detected, options button “Hotspot(s)” enables overlaying the hotspots locations on the frames to display as highlighted in white regions. The option “Calcium” allows users to overlay calcium signal locations on frames.
- Red label 3 shows a set of functionalities that enable users to manually set an optimum threshold to detect calcium signal events of individual frames. By default, the optimum threshold is computed at frame and video level according to the routine described in

Section 3.2.

- Red label 4 corresponds to a set of input boxes that allow users to set parameters for automated hotspots detection such minimum and maximum sizes, hit ratio minimum and maximum values, and the minimum distance between hotspots. By default, the size parameters are calculated based on the mean and the standard deviation of the size of calcium event segmented in frames. The default hit-ratio range is set between 1 to 100%, whereas the minimum distance is set to 2 pixels. When the option “ignore persistent bright areas” is selected, the system tries to get rid of noisy illumination artefacts that may be output as hotspots according to the principle described in the thesis Section 3.5.
- Red label 5 shows a list box that displays the list of hotspots detected by the system if they exist. The first column contains the centroid coordinates $[x, y]$ of hotspots, then follow their sizes, their hit-ratios, and the index of their frame of origin in the next columns.
- Red label 6 represents a set of radio buttons that enable users to view detected hotspots or overlay their locations on individual frames.
- Red label 7 and 10 represents a set of action buttons whose functions are suggested by their names.
- Red label 8 and 9 respectively correspond to a title section or an information panel and an output display area. Here, the system has output 43 hotspots and overlaid them over the last frame of a calcium video made of 301 frames.

The second interface in Figure A.2 illustrates the signal processing functionality of INCAS toolbox. It consists of two main features: - first, calcium signal adjustment by photobleaching detection and correction, - and second measurements of calcium signal parameters. The whole calcium signal processing interface that is illustrated in Figure A.2 can be described like the following:

- The output display panel of Figure A.2 red-labelled 6 shows plotting of curves of the 43 hotspots highlighted in Figure A.1. The y-axis represents concentration change units, and x-axis time in seconds that is calculated based on the video frame rate and length.
- The list box red-labelled 1 contains properties of the detected hotspots. The set of options buttons red-labelled 2 enable users to select the type of curve to plot i.e. an average

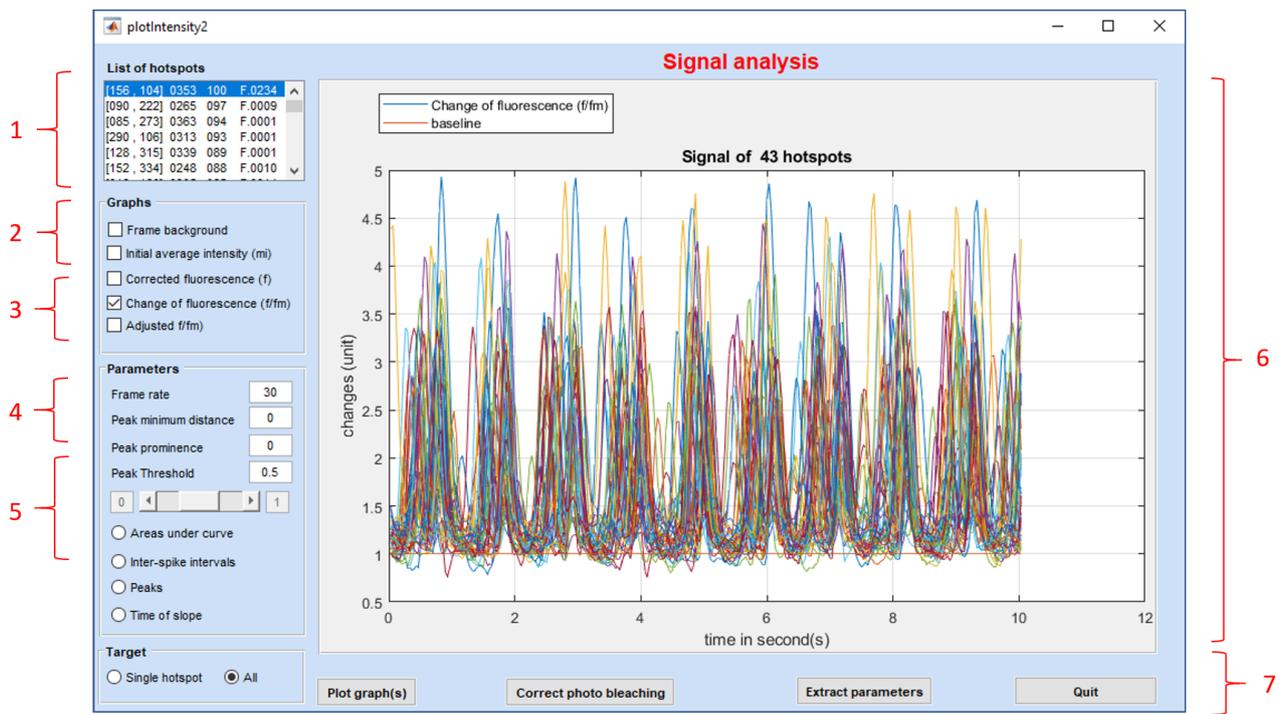


Figure A.2: Calcium signal processing - Screenshot

frame background illumination, initial average intensity of calcium traces, corrected fluorescence, the change of calcium concentration or the adjusted calcium measurement after photobleaching correction.

- The set of parameters red-labelled 3 and 4 enable users to select the types of statistical measurements to retrieve from the hotspots' calcium traces. Measurements can be the time of slopes, peaks frequencies, inter-spike intervals and areas under curves. Those measurements are made according to default parameters of frame rate, peak prominence, threshold, and minimum amplitude.
- The group of radio buttons red-labelled five allows to retrieve measurements from single or multiple hotspots.
- The group of actions buttons red-labelled 7 enable actions according to their respective names.

Figure A.3 illustrates INCAS ability to produce statistical measurements into dynamical tables that can be exported into MATLAB, text and excel sheet files. The measurements displayed here for illustration are from left to right, maximum amplitudes, Mean amplitudes, and peak frequencies of each hotspot (ROI) in the first form and all peak values per ROI in the second

Parameters: Peak amplitudes and frequencies			
	Max amplitude	Mean amplitude	Peak frequency
ROI - 10	1.9820	1.7500	1.4950
ROI - 11	2.1720	1.7410	2.0930
ROI - 12	2.2480	1.9190	1.9930
ROI - 13	3.1730	2.8270	1.4950
ROI - 14	2.7370	2.5440	1.8940
ROI - 15	3.0220	2.6790	1.3950
ROI - 16	2.8870	2.1250	2.3920
ROI - 17	1.7340	1.6140	1.8940
ROI - 18	2.6240	2.1580	1.9930
ROI - 19	2.2930	1.8300	1.7940
ROI - 20	3.0340	2.6830	1.6940
ROI - 21	3.3210	3.0250	1.8940
ROI - 22	2.8540	2.5780	1.5950
ROI - 23	2.6590	2.2630	1.8940
ROI - 24	2.7060	2.3630	1.5950
ROI - 25	3.5950	3.0660	1.6940
ROI - 26	3.5630	3.1750	1.6940
ROI - 27	2.8120	2.6980	1.7940
ROI - 28	2.3740	1.8310	1.5950
ROI - 29	2.6250	2.0210	1.8940
ROI - 30	3.1750	2.6890	1.7940

Parameters: Maximum amplitude per ROI											
	ROI - 1	ROI - 2	ROI - 3	ROI - 4	ROI - 5	ROI - 6	ROI - 7	ROI - 8	ROI - 9	ROI - 10	ROI - 11
Peak - 1	1.0258	1.3974	2.1503	1.6892	1.0617	1.0430	1.6594	1.2101	2.1019	1.7436	1.4919
Peak - 2	1.0314	1.2754	2.0675	1.8545	1.0692	1.0546	1.4298	1.3566	1.9672	1.5051	1.6797
Peak - 3	1.0660	1.1906	2.0933	1.4931	1.0609	1.0338	1.6776	1.3435	2.0499	1.7317	1.6507
Peak - 4	1.0627	1.4405	2.1105	1.6645	1.0971	1.0865	1.5473	1.4156	1.7651	1.6787	1.5685
Peak - 5	1.0583	1.1979	2.0353	1.2161	1.0332	1.0943	1.6196	1.1990	1.8038	1.7683	1.7835
Peak - 6	1.0342	1.1729	2.0889	1.6430	1.0831	1.0653	1.3816	1.4007	1.3488	1.8775	1.5774
Peak - 7	1.0295	1.3341	1.5275	1.5776	1.1352	1.0619	1.8713	1.2246	2.1830	1.6817	1.5581
Peak - 8	1.0367	1.1840	2.1817	1.4053	1.0579	1.0689	1.3590	1.2632	1.9534	1.7229	1.5637
Peak - 9	1.0246	1.2357	1.9655	1.7674	1.0404	1.0709	1.6612	1.6360	1.3175	1.8599	1.8516
Peak - 10	1.0341	1.5107	1.8588	1.5629	1.0584	1.1047	1.8252	1.4880	2.1727	1.7975	1.9715
Peak - 11	1.0247	1.2399	2.1693	1.6015	1.0751	1.0683	1.3846	1.4823	1.9534	1.4942	1.9531
Peak - 12	1.0567	1.3588	2.0002	1.5477	1.0722	1.0694	1.4547	1.2815	2.3251	1.9820	1.7622
Peak - 13	1.0467	1.5181	2.1199	1.7445	1.0393	1.0675	1.5904	1.6456		1.8611	1.7121
Peak - 14	1.0411	1.5288	2.2156	1.5084	1.0430	1.0570	1.4330	1.2539		1.7011	1.3684
Peak - 15	1.0443	1.2692	2.0968	1.7806	1.0760	1.0427	1.5345	1.7001		1.8426	1.6743
Peak - 16	1.0526	1.5458		1.5369	1.0384	1.0547	1.3839	1.3968			1.3056
Peak - 17	1.0241	1.3073		1.5266	1.0501	1.0430	1.5272	1.7474			2.1725
Peak - 18	1.0668	1.1648			1.0384	1.0757	1.8961				2.0660
Peak - 19	1.0586	1.4868			1.0786	1.0461					2.0360
Peak - 20	1.0259	1.1727			1.0660	1.0281					2.0431

Figure A.3: Statistical measurements - Screenshot

form. They are produced when users click on the button “Get readouts” after selecting options “All Hotspots” and “Normalized” of the main software interface. Those measurements can also be obtained, when users click on the button “Extract parameters” of the form “Signal analysis” displayed in Figure A.2.

A.3 Appendix Conclusion

All in one, INCAS is a first attempt to fully automate in one place intracellular calcium hotspots detection and measurement from a fluorescent video of cardiac myocytes. The software package has been tested over pulmonary veins sleeve cells and neonatal rat ventricular myocytes. We argue that the toolbox could be applied to any fluorescent video datasets showing dynamic cellular patterns that can be repetitive and quantified by illumination fluctuation.

All in one, INCAS is a first attempt to fully automate in one place intracellular calcium hotspots detection and measurement from a fluorescent video. INCAS is a standalone application that works only on Windows computers. It provides a graphical user interface to upload fluorescent video files, automatically detect areas and hotspots or set them up manually, and extract measurements from the calcium traces recorded over time. The main features of the toolbox consist of a set of functionalities such as a video frame visualization toolkit based on the functionality of MATLAB Image processing and Computer Vision toolbox. INCAS also integrated

a set of algorithms for hotspot detection and measurement based on thresholding and video fusion we presented in Chapter 3 of the thesis, and for photobleaching detection and correction using non-linear least square regression we introduced in Chapter 4. The INCAS software package has been tested over special cardiac cells such as PVCs and NRVMs. Still, we argue that the toolbox could be applied to any fluorescent video datasets showing dynamic cellular patterns that can be repetitive and quantified by illumination fluctuation. The code source of the INCAS toolbox is available at the School of Computing of the University of Buckingham (<https://www.buckingham.ac.uk/computing>).