Computer-Aided Diagnosis Systems in the Classification of Neuroblastoma Histological Images

A Thesis Submitted for the Degree of

Doctor of Philosophy

By

Soheila Gheisari

in

School of Computer Science

UNIVERSITY OF TECHNOLOGY SYDNEY

AUSTRALIA

2019

UNIVERSITY OF TECHNOLOGY SYDNEY

SCHOOL OF COMPUTER SCIENCE

The undersigned hereby certify that they have read this thesis entitled "**Computer-Aided Diagnosis Systems in the Classification of Neuroblastoma Histological Images**" by **Soheila Gheisari** and that in their opinions it is fully adequate, in scope and in quality, as a thesis for the degree of **Doctor of Philosophy**.

Dated: December 2019

Principal Supervisor: Prof. Paul J. Kennedy

CERTIFICATE OF ORIGINAL AUTHORSHIP

I, Soheila Gheisari declare that this thesis, is submitted in fulfilment of the requirements for the award of Doctor of Philosophy, in the *school of Computer Science*, *faculty of Engineering and Information Technology* at the University of Technology Sydney. This thesis is wholly my own work unless otherwise reference or acknowledged. In addition, I certify that all information sources and literature used are indicated in the thesis. This document has not been submitted for qualifications at any other academic institution. This research is supported by the Australian Government research Training Program.

Production Note:
Signature removed prior to publication.

Signature of Author

Acknowledgements

I am greatly indebted to my supervisor, Prof. Paul J. Kennedy for all the efforts, advice, help, and patient guidance that he has provided during my PhD. He is such a nice, generous, helpful and kindhearted person. I feel really happy, comfortable and unconstrained with him during my PhD study. I owe my research achievements to his experienced supervision. Many thanks are also due to my co-supervisor A/P Daniel R. Catchpoole from the Tumour Bank at the Children's Hospital for his support, advice and numerous conversation with him. I would like to express grateful acknowledgement to my co-supervisor Dr. Amanda Charlton from department of histopathology in Auckland City Hospital. She supported me in collecting the dataset of neuroblastoma. She provided the project with necessary clinical resources. I appreciate her attempts in providing invaluable advice and necessary clinical resources. I would also like to thank Zsombor Melegh and Elise Gradhand from department of pathology in Southmead Hospital in Bristol. They supported me in providing neuroblastoma images during this project.

I would like to thank my parents that I could never have done this without their faith, support, and constant prayers. I also thank my dear husband Farshid, love of my life, for his support, patience and managing the life. I also want to thank my five years old daughter Shayna, who has been the light of my life. She has given me the extra motivation to get things done. You will be my greatest achievement.

Last but not least, the financial assistance of an Australian Postgraduate Award scholarship is gratefully appreciated.

To my father and my mother, who supported me emotionally. I always knew that you believed in me and wanted the best for me. To my husband Farshid, that his understanding, assistance and company during my study enabled the hours of research for me. To my little daughter Shayna you will always be my blessing from God. This thesis could not have been completed without your supports. I love you all.

Table of Contents

List of Tables

List of Figures

Abstract

Neuroblastoma is the most common extracranial solid malignancy in early childhood. Optimal management of neuroblastoma depends on many factors, including histopathological classification. Although histopathological classification by a human histopathologist is considered the gold standard, computers can help to extract many more features, some of which may not be recognisable by the human eye. Neuroblastoma histological images have a complex texture with complicated features which are different from appearance-based features. Computer-aided diagnosis (CAD) systems facilitate the analysis and classification of neuroblastoma histological images which are non-trivial tasks due to the differences in staining, intensity, and instrumentation. This motivates the thesis to work on the classification of neuroblastoma histological images.

In the past, a small number of methods were proposed by previous studies for the classification of neuroblastoma histological images. These methods are based on the geometry and appearance of the different cells. However, there is a high intra-class variation of intensity and size of the neuroblast cells within the same classification group. Therefore, these methods are not applicable to neuroblastoma histological images. This research proposes a solution based on traditional machine learning approaches and deep learning approaches to extract non-appearance-based

features in small regions. This thesis will investigate two research areas of feature extraction: low-level feature extraction and high-level feature extraction. Low-level features are minor details of the image such as lines, curves and edges. However, high-level features are on top of the low-level features to detect object and larger shape in the image. Feature extraction is aggregated with the classifier in this research to classify neuroblastoma histological images into five categories.

This thesis makes four contributions. Contribution 1 is the construction of a dataset comprising neuroblastoma histological images which are labeled by an expert histopathologist. Contribution 2 is the proposal of a local feature extraction method which can extract local features which are robust to high intra-class variations of intensity. Contribution 3 is the extraction of discriminative features which are robust to high intra-class variation of scale of the neuroblast cells within the same class. Contribution 4 is the proposal of deep networks to extract high-level features which are difficult for the human eye to recognise. The performance of all the proposed methods in this research is evaluated on a dataset collected from The Children's Hospital at Westmead, Sydney, Australia. As there was no publicly available dataset in this field, the proposed algorithms were evaluated on the second dataset of neuroblastoma provided by the University of Bristol and the public breast cancer dataset. All the results are compared with state-of-the-art methods. The results indicate the effectiveness of the proposed algorithms.

This is the first time that neuroblastoma histological images have been classified into five subtypes using low-level and high-level features. However, there are limitations in this research. The specificity is not 100% compared with the gold standard.

Moreover, the proposed algorithms are confused in the distinction between poorly-differentiated and differentiating neuroblastoma, a distinction that human pathologists also find difficult in limited fields of view.

Publications

Following is the list of publications resulting from the work in this thesis:

1. Gheisari, S., Catchpoole, D.R., Charlton, A., Kennedy, P.J. (2017), 'Patched Completed Local Binary Pattern is an Effective Method for Neuroblastoma Histological Image Classification' Communications in Computer and Information Science (Springer), pp. 57-71.

2. Gheisari, S., Catchpoole, D.R., Charlton, A., Kennedy, P.J. (2018) 'Computer Aided Classication of Neuroblastoma Histological Images using Scale Invariant Feature Transform with Feature Encoding', Journal of Diagnostics, 8(3), 1-18; doi:10.3390/diagnostics8030056.

3. Gheisari, S., Catchpoole, D.R., Charlton, A., Melegh, Z., Gradhand, E., Kennedy, P.J. (2018) 'Convolutional Deep Belief Network with Feature Encoding for Classification of Neuroblastoma Histological Images', Journal of Pathology Informatics, Vol. 9, No. 1, pp. 1-17, 2018.

4. Gheisari, S., Charlton, A., Catchpoole, D.R., Kennedy, P.J. (2017) 'Computers can Classify Neuroblastic Tumours from Histopathological Images using Machine Learning', Pathology, No. 49, pp. 72-73, 2017.

Table of Symbols

Symbol	Description
CAD	Computer aided diagnosis
PCLBP	Patched completed local binary pattern
SIFT	Scale-invariant feature transform
CDBN	Convolutional deep belief network
CNN	Convolutional neural network
CT	Computed tomography
CLBP	Completed local binary pattern
SBP	Sign binary pattern
MBP	Magnitude binary pattern
DSBP	Decimal sign binary pattern
DMBP	Decimal magnitude binary pattern
HSBP	Histogram of sign binary patterns in image
HMBP	Histogram of magnitude binary patterns in image
LH	Local histogram
k-NN	k nearest neighbour
SVM	Support vector machine
RBF	Radial basis function

C_C	Contrast threshold
C_E	Edge threshold
UD	Undifferentiated neuroblastoma
PD	Poorly-differentiated neuroblastoma
D	Differentiating neuroblastoma
GN	Ganglioneuroma
GNB	Ganglioneuroblastoma
QDA	Quadratic Discriminant analysis
GLCM	Gray level co-occurrence matrices
LPQ	Local phase quantization
ORB	Oriented fast rotated brief
PFTAS	Parameter-free threshold adjacency statistics
CDBN	Convolutional deep belief network
CRBM	Convolutional restricted botzmann machine
H	Hidden layer
V	Visible layer
TMA	Tissue microarray
MKI	Mitosis karyorrhexis Index
H & E	Hematoxylin and eosin
G (x, y)	Gaussian function
L(x, y, σ)	Scale space of an image
x and y	Image coordinates
σ	Width of Gaussian function
DoG	Difference of gaussian

F_l	Feature maps of convolutional layer l
$y_{i,j}^{\phi,l}$	Output of neuron (i, j, ϕ, l) at position i, j
$w_{\Delta i,\Delta j}^{\phi,\phi',l}$	Weight of each connection between the neurons in convolutional layer (i, j, ϕ, l) and all the neurons in the previous subsampling layer ϕ'
b_j	Hidden unit biases for neuron j
c_i	Visible unit biases for neuron i