

**AUTOMATIC TRANSLUCENCY DETECTION OF BASAL CELL CARCINOMA (BCC)
VIA DEEP LEARNING METHODS**

by

He Huang

B.CS. (Hons), Computer Science, Dalhousie University, 2016

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

MASTER OF APPLIED SCIENCE

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL STUDIES
(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA

(Vancouver)

August 2018

© He Huang, 2018

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, a thesis/dissertation entitled:

submitted by He Huang in partial fulfillment of the requirements for

the degree of Master of Applied Science

in Electrical and Computer Engineering

Examining Committee:

Z. Jane Wang

Co-supervisor

Tim Lee

Co-supervisor

Supervisory Committee Member

Rabab Ward

Additional Examiner

Additional Supervisory Committee Members:

Supervisory Committee Member

Supervisory Committee Member

Abstract

Translucency, defined as a jelly-like appearance, is a common clinical feature of basal cell carcinoma (BCC), the most common skin cancer. This feature plays an important role in diagnosing basal cell carcinoma at an early stage because the translucency can be observed readily in clinical examinations with a high specificity. Therefore, translucency detection is a critical component of computer aided systems which aim at early detection of basal cell carcinoma. In this thesis, we proposed two deep learning methods to automatically detect translucency.

First, we develop a convolutional neural network based framework to detect translucency of basal cell carcinoma. Furthermore, a sparse auto-encoder based framework is proposed for translucency detection on BCC images. Since currently two types of skin images are mainly used for diagnosis of basal cell carcinoma by doctors, which are dermoscopy images and clinical images, we evaluate two proposed methods on both types of skin images.

Our results showed that the two proposed methods yield similar detection performances. For detecting translucency in dermoscopy images, both proposed methods achieve comparable accuracy results, though the accuracy is not as good as we expected. For detecting translucency in clinical images, both methods achieve good performances. Compared the performances in both types of images, the proposed deep learning based methods seems more suitable for translucency detection in clinical images than in dermoscopy images.

Lay Summary

Basal cell carcinoma(BCC) is the most common type of skin cancer. Despite that BCC rarely causes mortality, the malignancy will destroy extensively the surrounding tissues and damage the skin structure aggressively at advanced stages, causing the high cost of treatment and increasing the suffering of patients. Therefore, early detection of BCC is of great importance for disease management. Translucency is one of the most important characteristic features of BCC. It can be detected in tiny lesions with a high specificity. Therefore, detecting translucency is a key function of computer aids systems for automatic diagnosis of BCC at an early stage. In this thesis, we develop two deep learning frameworks to detect translucency in BCC automatically, which can be used further to help a BCC diagnosis.

Preface

The research was jointly initiated by Dr. Z. Jane Wang, Dr. Tim K. Lee and the thesis author, and the majority work of the research was conducted by the author of this thesis, with valuable suggestions from Dr. Z Jane Wang and Dr. Tim K. Lee. Dr. Pegah Kharazmi also helped on the method part in Chapter 3. Dr. David I. McLean helped greatly on cancer images labeling in Chapters 4 and 5.

Table of Contents

Abstract.....	iii
Lay Summary	iv
Preface.....	v
Table of Contents	vi
List of Tables	ix
List of Figures.....	x
List of Abbreviations	xii
Acknowledgements	xiii
Chapter 1: Introduction	1
1.1 Images used for BCC diagnosis.....	2
1.1.1 Dermoscopy images.....	2
1.1.2 clinical images	4
1.2 Translucency	4
1.3 Related work	5
1.4 Motivation.....	6
1.5 Thesis organization	7
Chapter 2: Deep learning background	8
2.1 Artificial neural network.....	8
2.1.1 A single neuron.....	8
2.1.2 Architecture of artificial neural network.....	9
2.1.3 Backpropagation algorithm.....	11

2.2	Convolution neural networks	12
2.2.1	Convolutional layer	13
2.2.2	Pooling layer	14
2.2.3	Fully-connected layer	15
2.3	Autoencoder	15
Chapter 3: Proposed Methods		17
3.1	Patch strategy	17
3.2	Method one: A CNN-based framework	18
3.2.1	Overview of the proposed method	18
3.2.2	Architecture of CNN	20
3.3	Method two: A SSAE-based framework	21
3.3.1	Overview of the proposed method	22
3.3.2	Sparse autoencoder	23
3.3.3	High-level features learning through SSAE	24
3.3.4	Translucency detection through softmax classifier	25
3.4	Conclusion	25
Chapter 4: Translucency detection of BCC in dermoscopy images		27
4.1	Dataset	27
4.2	Patching and Labeling	28
4.3	Traning and testing set preparation	29
4.4	Study one: Automatic detection of translucency via a Convolutional neural network	30
4.4.1	Experimental setting	30
4.4.2	Experiment result	31
4.5	Study two: Automatic detection of translucency via a SSAE-based framework	33

4.5.1	Experimental setting	33
4.5.2	Experiment result	34
4.6	Discussion	35
4.7	Conclusion	43
Chapter 5: Translucency detection of BCC in clinical images		44
5.1	Dataset and pre-processing	44
5.1.1	Generating of ROIs	45
5.1.2	Patching and Labeling	45
5.2	Study one: Automatic detection of translucency via a Convolutional neural network	47
5.2.1	Experimental setting	47
5.2.2	Experiment result	48
5.3	Study two: Automatic detection of translucency via a SSAE-based method	51
5.3.1	Experimental setting	51
5.3.2	Experiment result	51
5.3.2	Discussion	52
5.4	Comparison of two proposed methods	54
5.4.1	The quantitative performance comparison	54
5.4.2	The localization performance comparison	55
5.4.3	Feature Visualization	57
5.5	Conclusion	58
Chapter 6: Conclusion and Future work		59
6.1	Conclusion and Contribution	59
6.2	Future work	60
Bibliography		62

List of Tables

Table 4.1 The quantitative result of translucency detection in dermoscopy images using CNN-based method	32
Table 4.2 The quantitative result of translucency detection in dermoscopy images using SSAE-based method	34
Table 5.1 Results of translucency detection in clinical images using CNN-based method.....	49
Table 5.2 Results of translucency detection in clinical images using SSAE-based method	52
Table 5.3 The quantitative performance comparison of three different combinations of nodes numbers.....	53
Table 5.4 The quantitative performance comparison of three different sizes of patches	54
Table 5.5 The quantitative performance comparison of proposed methods in clinical images....	55

List of Figures

Figure 1.1 A typical device of dermoscopy	2
Figure 1.2 Examples of dermoscopy images of Basal cell carcinoma	3
Figure 1.3 Examples of clinical images of Basal cell carcinoma	4
Figure 1.4 Dermoscopy images of BCC with translucency.....	5
Figure 2.1 A single neuron.....	8
Figure 2.2 A simple neural network	10
Figure 2.3 Architecture of Alexnet	12
Figure 2.4 An example of how convolution works.....	13
Figure 2.5 An example of max pooling	15
Figure 2.6 The architecture of a simple autoencoder.....	16
Figure 3.1 Diagram of CNN-based framework	19
Figure 3.2 Architecture of the designed CNN	20
Figure 3.3 Diagram of SSAE-based framework	22
Figure 3.4 The proposed framework of the stacked sparse autoencoder and softmax classifier for detecting translucency.....	24
Figure 4.1 Examples of dermoscopy basal cell carcinoma images	28
Figure 4.2 Examples of translucency segmentation in dermoscopy images by doctor	29
Figure 4.3 Examples of translucent(left) and non-translucent patches(right) of dermoscopy images	30
Figure 4.4 The model loss of CNN method.....	31

Figure 4.5 The ROC curve for translucency detection in dermoscopy images using CNN-based method.....	32
Figure 4.6 The ROC curve for translucency detection in dermoscopy images using SSAE-based method.....	34
Figure 4.7 Examples of translucency detection in pigmented BCC	36
Figure 4.8 Features of translucency extracted from pigmented BCC(middle), non-pigmented BCC(left) and both types of BCC(right).....	37
Figure 4.9 Examples of translucency detection in dermoscopy images	38
Figure 4.10 Examples of translucent patches with and without vessels	39
Figure 4.11 Examples of translucency detection in dermoscopy images	40
Figure 4.12 Examples of translucency detection in dermoscopy images with artifacts such as bubbles and hairs.....	42
Figure 5.1 Examples of clinical basal cell carcinoma images	44
Figure 5.2 Examples of ROIs generating.....	45
Figure 5.3 Examples of translucency segmentation in clinical images by doctor	46
Figure 5.4 Examples of translucent patches and non-translucent patches of clinical images	47
Figure 5.5 The model loss of CNN method.....	48
Figure 5.6 Translucency localization results of Basal cell carcinoma in clinical images using CNN-based method.....	50
Figure 5.7 The ROC curve for translucency detection of SSAE-based method.....	52
Figure 5.8 Localization results comparison between two proposed methods in clinical images ..	56
Figure 5.9 Examples of features learned from proposed methods.....	58

List of Abbreviations

ANN	Artificial Neural Network
AUC	Area Under Curve
BCC	Basal Cell Carcinoma
CNN	Convolutional neural network
ROC	Receiver Operating Characteristic Curve
SSAE	Stacked Sparse Autoencoder

Acknowledgements

I would like to express my great appreciation to my supervisors, Dr. Z. Jane Wang and Dr. Tim K. Lee for their help, support, encouragement and guidance for the research throughout my master study. I would also like to thank Dr. Pegah Kharazmi for her great technique support. Many thanks go to Dr. David I. McLean, who helped me label all skin medical images.

I would like to thank my friends and lab mates for their help and encouragement. Special thanks to Shuo Yang for his great support during my difficult period.

Last but not least, I own my deepest gratitude to my parents for their love, support and understanding.

Chapter 1: Introduction

Basal cell carcinoma (BCC) is the most common type of skin cancer among the white populations in the world [1]. In the United States, more than 4 million patients are diagnosed with skin cancer annually and 80% of them are BCC [2, 3]. In UK, over 3,0000 new cases of BCC are estimated per year [4]. In German, around 100,000 cases of BCC were reported in 2009 [5]. However, there are still over 2 million cases that are under-reported each year estimated by the World Health Organization(WHO) [5]. Even though so many people are diagnosed with BCC, the incidence of BCC is increasing every year [6-8]. The increasing rate is about 5% in Europe and it is about 2% in the United States [7].

Despite the fact that BCC rarely causes mortality, the malignancy will destroy extensively the surrounding tissues and damage the skin structure aggressively at advanced stages [9-11]. Therefore, early detection of BCC is important for disease management. However, the diagnosis of BCC is complicated. Early recognition of skin cancer heavily relies on visual examination by physicians, following by a confirmation diagnosis based on biopsy and a histological examination [12]. Due to such a large number of patients, the burden of BCC diagnosis is extremely heavy. Therefore, many research studies have been focusing on the development of computer-aided systems for detecting skin cancer automatically in order to relieve the pressure caused by the increasing rate of skin cancer and limited medical resources.

The majority of computer-aided systems are image-based systems which include image preprocessing, feature extraction, and image classification. These computerized systems have great potential for skin cancer detection and may lead to a higher diagnostic accuracy at early stages as well.

1.1 Images used for BCC diagnosis

Clinical images and dermoscopy images are the two most common imaging modalities for skin cancer diagnosis today. Both types of images are digital images taken from the skin surface.

1.1.1 Dermoscopy images

Dermoscopy, a non-invasive tool for skin cancer detection, enables the visualization of subsurface structures and patterns which are unable to be seen by naked eyes [13]. It is a reliable method for early detection of skin cancer [14].

There are two main types of dermoscopy. The first one is fluid immersion which includes a magnifier and non-polarized light system [15]. Usually, fluid immersion dermoscopy can achieve 10 – 20 times magnification [16]. This kind of dermoscopy directly contacts the skin which requires a liquid, such as alcohol or mineral oil, in the interface between the skin and the instrument [15]. The liquid between the interface is used to reduce the refractive index mismatch between air and skin tissues [17]. Therefore, doctors can observe the morphologic structures of skin cancers which are not seen by naked eyes. The other type of dermoscopy is polarized dermoscopy which has no need to contact the skin. This type of dermoscopy utilizes the polarized light which helps eliminate the skin light reflections to observe the subsurface structure [17]. A typical dermoscopy device is shown in Figure 1.1.



Figure 1.1 A typical device of dermoscopy

Dermoscopy can achieve two advantages. First, it can improve the accuracy of clinical diagnosis of skin cancer. It has been reported that it could result in a 35% increase in accuracy of clinical diagnosis [15]. Also, studies have shown that both the sensitivity and the specificity of diagnosis increased as well [18]. Second, dermoscopy can significantly reduce the number of patients who would be further referred to do a biopsy because of the increasing confidence of diagnosis [19]. However, there is also one limitation of dermoscopy: Generally, long time training and experience are required before a dermatologist can make professional diagnosis using dermoscopy. A study shows that a dermatologist who can achieve higher detection rates needs a formal training and at least 3 years of experience [19]. Figure 1.2 demonstrated some examples of Basal cell carcinoma in dermoscopy images.

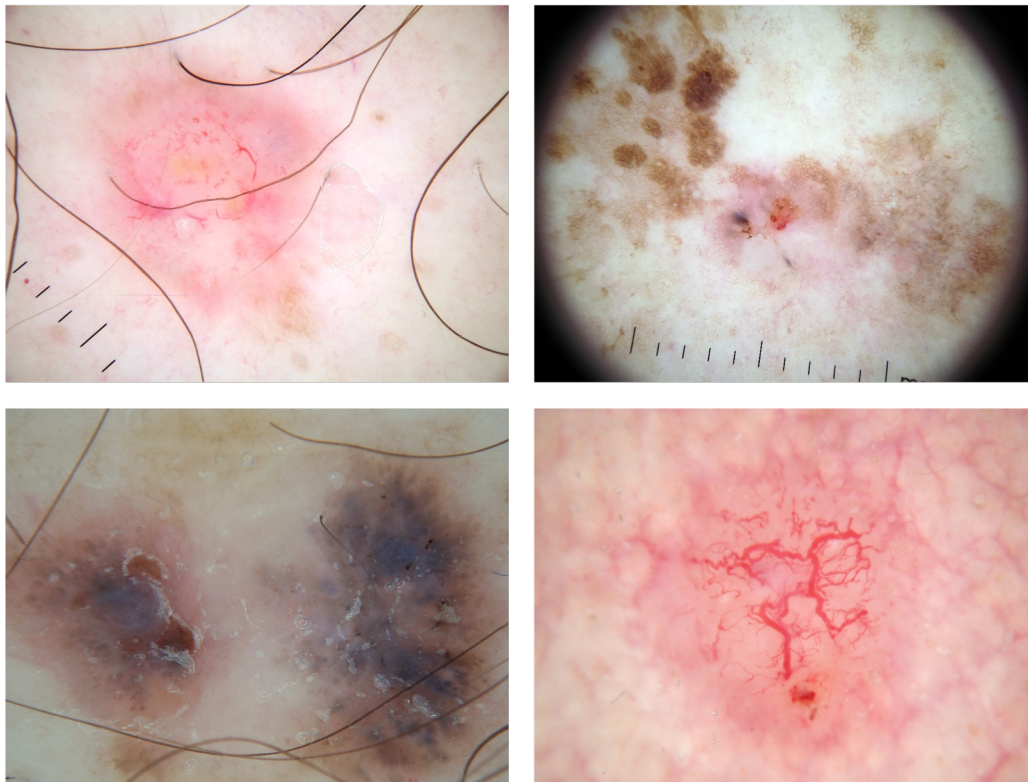


Figure 1.2 Examples of dermoscopy images of Basal Cell Carcinoma

1.1.2 Clinical images

However, a study shows that only 23% of dermatologists use dermoscopy for skin cancer diagnosis in United State [20]. Clinical images are still the most commonly used modality for skin imaging. Clinical images, taken by color digital cameras, are the most convenient method for capturing diagnostic features. Doctors can easily record images of a skin cancer surface from any angle and any direction without using other devices. Figure1.3 illustrates clinical images of skin cancers.



Figure 1.3 Examples of clinical images of Basal Cell Carcinoma.

1.2 Translucency

Translucency, defined as a jelly-like appearance, is an important characteristic feature of BCC. It is an optical phenomenon generated by the cancerous tumor [21, 22]. Translucency plays an important role in the diagnosis of BCC because this feature can be readily observed in clinical examinations with a high specificity of 93% [24]. In addition, it can be detected in very tiny BCCs where other features have not yet shown up [21, 22]. Therefore, detecting translucency is a key function of

computer aided systems aimed at discriminating BCCs from benign skin conditions and other types of skin cancers at an early stage. Figure 1.4 illustrates several examples of BCC with translucency in dermoscopy images.

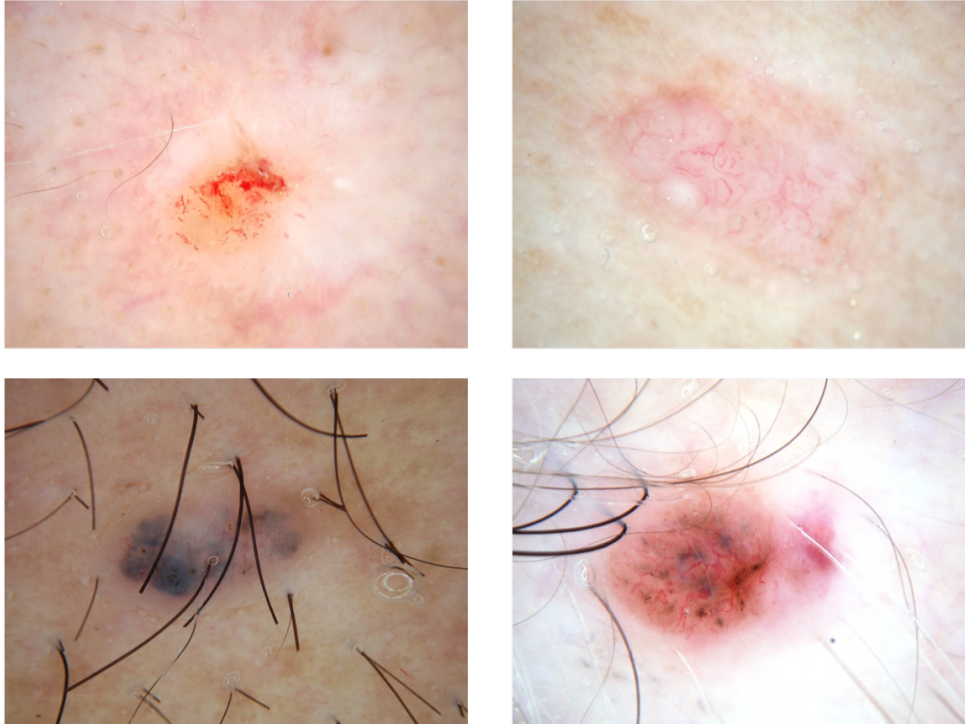


Figure 1.4 Example dermoscopy images of BCC with translucency.

1.3 Related work

In recent years, W. V. Stoecker's group paid close research attention to analyze the translucency of BCC. In [22], they analyzed RGB and chromaticity color features from manually marked translucent areas of a lesion and measured the corresponding texture properties by a set of intensity histograms. Totally, they obtained 6 color features and 6 texture features. Through statistical analyses, they concluded that texture features are more important than color features and the most significant feature of translucency is smoothness. In [23], a segmentation-based method was proposed to detect translucent areas in BCC. Firstly, the smoothness of each image block is

analyzed to find the candidates of translucent areas and then the uncorrected candidates are eliminated through analyzing the brightness.

1.4 Motivation

As discussed previously, translucency is an important biomarker in BCC diagnosis at an early stage. Therefore, translucency detection is considered as a key function for computer aided systems towards early diagnosis of BCC. A few papers published the image properties of translucency using conventional image analysis methods; however, to the best of my knowledge, there is no fully automated computer system on detection of translucency of BCC so far. Therefore, efficient methods which can detect translucency automatically are still needed.

In recent years, deep learning has been reported to achieve excellent performances on numerous image analysis tasks [44-46]. Unlike the conventional methods, deep learning methods are data-driven methods which can learn high-level features directly from the input data [50]. Therefore, there has been an increasing interest in the community in employing deep learning methods to solve medical image problems [51].

Inspired by the success of deep learning methods in numerous medical image analysis tasks, the objective of this thesis is to develop deep learning methods for translucency detection of BCC automatically. In this thesis, we propose two deep learning frameworks, one is based on the convolutional neural network (CNN) and the other is based on the stacked sparse autoencoder (SSAE), to achieve the research goal. Both CNN and SSAE are typical deep learning models.

1.5 Thesis organization

The organization of this thesis is as follows:

Chapter2: Deep learning background. In this chapter, we introduce some basic concepts of deep learning methods which will be explored for our work in later chapters. Specifically, we review the concepts of ANN, CNN and Autoencoder.

Chapter3: Proposed methods. In this chapter, two deep learning methods are proposed for translucency detection of basal cell carcinoma. One is the CNN-based method and the other is the SSAE-based method. The framework and architecture of the proposed methods are described in detail in Chapter 3.

Chapter4: Translucency detection of Basal cell carcinoma in dermoscopy images. In this chapter, we detect translucency of basal cell carcinoma in dermoscopy images using the proposed methods. We report the results and discuss the limitations of the proposed methods when applied to this specific application.

Chapter5: Translucency detection of Basal cell carcinoma in clinical images. In this chapter, two proposed methods are applied to detect translucency of basal cell carcinoma in clinical images. The experimental setting and results are reported and the performances of the proposed methods are compared.

Chapter6: Conclusion and future work. In this chapter, we summarize the contributions of this thesis and discuss the future direction of this research.

Chapter 2: Deep learning background

The objective of this thesis is to take advantages of deep learning methods to achieve translucency detection of basal cell carcinoma automatically. In this chapter, some basic concepts of deep learning methods, which will be used in this thesis research, are introduced. In section 2.1, we overview the artificial neural network. In section 2.2 and section 2.3, two deep learning models, convolution neural network and autoencoder, are introduced respectively.

2.1 Artificial neural network

Artificial Neural Network (ANN) is a mathematics model which simulates the biological neural networks [26]. It is motivated by simulating the human brain to process the information. In the current literature, ANN is the foundation of almost all complex neural networks such as convolutional neural networks and autoencoders which we will introduce later.

2.1.1 A single neuron

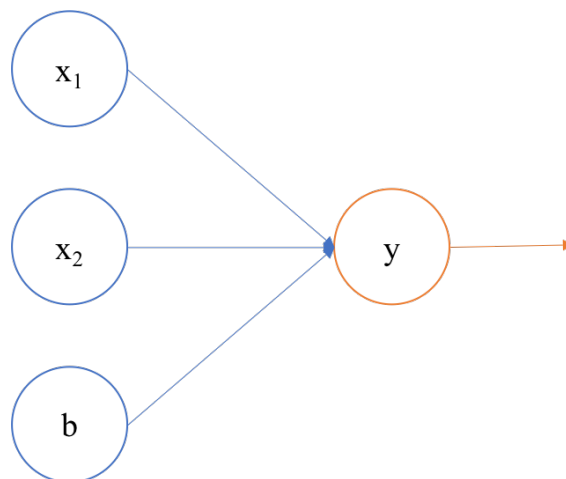


Figure 2.1 A single neuron.

Artificial neural networks are based on neurons. Figure 2.1 illustrates a simple neuron.

The mathematical model of the neuron can be expressed as:

$$h_{w,b}(x) = f(\sum W_i x_i + b)$$

where x_i means the i^{th} input, W_i is the weight of x_i and $\sum W_i x_i$ is the weighted summation of the inputs. b is the bias and $f(.)$ means the activation function. The activation function transfers the input to the output in ANN [27]. In the literature, there are three commonly used activation functions [28]:

Sigmoid: It has the output value between 0 and 1

$$f(z) = \frac{1}{1 + \exp(-z)}$$

Tanh: It has the output value between -1 and 1

$$f(z) = \frac{2}{1 + e^{-2z}} - 1$$

ReLU: It has the output value above 0

$$f(z) = \max(0, z)$$

2.1.2 Architecture of artificial neural network

The structure of an artificial neural network simulates biological neural networks [29]. There are generally three types of layers in an artificial neural network: the input layer, the hidden layer and the output layer. All such layers are composed of neurons. Nodes in two adjunct layers are connected by edges. Each edge has its corresponding weight value. A simple artificial neural network is illustrated in Figure 2.2.

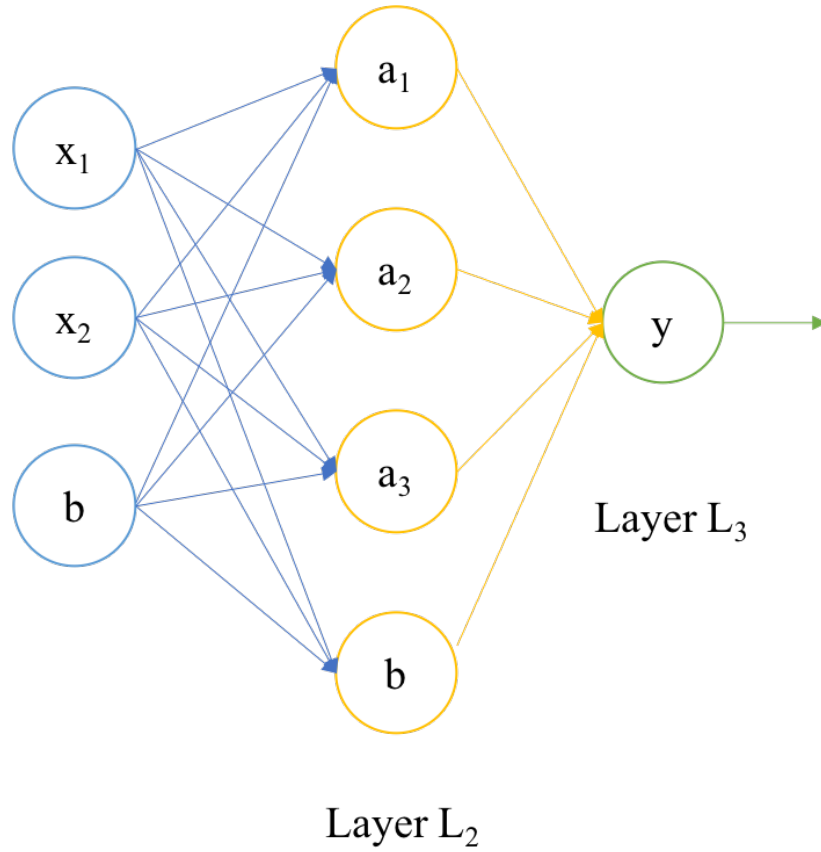


Figure 2.2 A simple neural network

The nodes at the input layer are used to receive the input. Also, nodes at hidden layers perform computations of the information received from the input layer and transfer the computational results to the output layer. Finally, the output nodes give a final result of the network.

The computational procedure of the network illustrated in Figure 2.2 can be represented by:

$$a_i = f(W_{ij}x_i + b)$$

$$y = f(W_{ij}a_i + b)$$

where W_{ij} is the weight of the connection between the i^{th} node at L_{n-1} layer and the j^{th} node at L_n layer, a_i represents the output value of the i^{th} node at the hidden layer, x_i represents the i^{th} node at the input layer, b is the bias and y represents the output value of this neural network.

From the above equation we can note that each a_i is determined by the values of all nodes and their corresponding weights at the previous layer. This kind of network is called the feedforward neural network in which all information is transferred layer by layer and there is no feedback between the interconnected layers.

2.1.3 Backpropagation algorithm

The backpropagation algorithm is probably the most commonly used algorithm in the current deep learning literature. It is widely used to train artificial neural networks [30]. The idea of the backpropagation algorithm can be summarized as follows:

Step-1: The training data are input into the artificial neural network and forward each layer of the network. Then the output is calculated. This is the feedforward process of an artificial neural network.

Step-2: Definitely, there is discrepancy between the real output and the estimated output, and thus the error is calculated and propagated back to previous layers until the input layer. In the processing of backpropagation, the parameters of each node are adjusted according to the error. The steps above are iterated until the error convergence.

The main purpose of the backpropagation algorithm is to minimize the output error. We need a cost function to define the error [25]:

$$J(W, b) = \frac{1}{2N} \sum \|h_{W,b}(x) - y\|^2$$

where $h_{W,b}(x)$ is the output of the neural network and y is the desired output.

2.2 Convolutional neural networks

Convolutional neural network (CNN) is one of the most popular deep neural network models in the current literature. It has been shown to perform excellently in numerous image analysis tasks such as face recognition and object detection [31, 32]. Besides, CNN became a powerful tool for medical image analysis in past few years, e.g., for cancer detection [33]. Therefore, there are increasing research focuses on using CNN to solve image analysis problems.

The structure of a convolutional neural network is similar to that of an original artificial neural network. It consists of one input layer, one output layer and several hidden layers. Also, each layer is made up of neurons which learn weights and bias through training. However, the neurons in CNN are different in the sense that they are organized in three dimensions: height, width, depth [28]. In addition, most fully-connected layers in CNN are replaced by convolutional layers. Figure 2.3 illustrates the architecture of Alexnet which is a typical CNN model. There are three main types of layers in the Alexnet model: convolutional layers, pooling layers, and fully-connected layers.

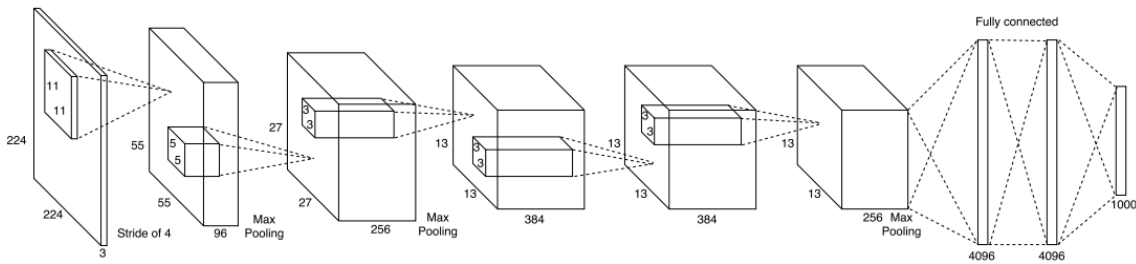


Figure 2.3 The architecture of Alexnet [35].

2.2.1 Convolutional layer

Convolutional layer is the core layer of a convolutional neural network. The purpose of a convolutional layer is to extract specific features from the input image which are used to make the final classification. Feature maps, the output of the convolutional layer, are generated from the input image which is convolved with filters. Figure 2.4 illustrates a simple example to show how convolution works.

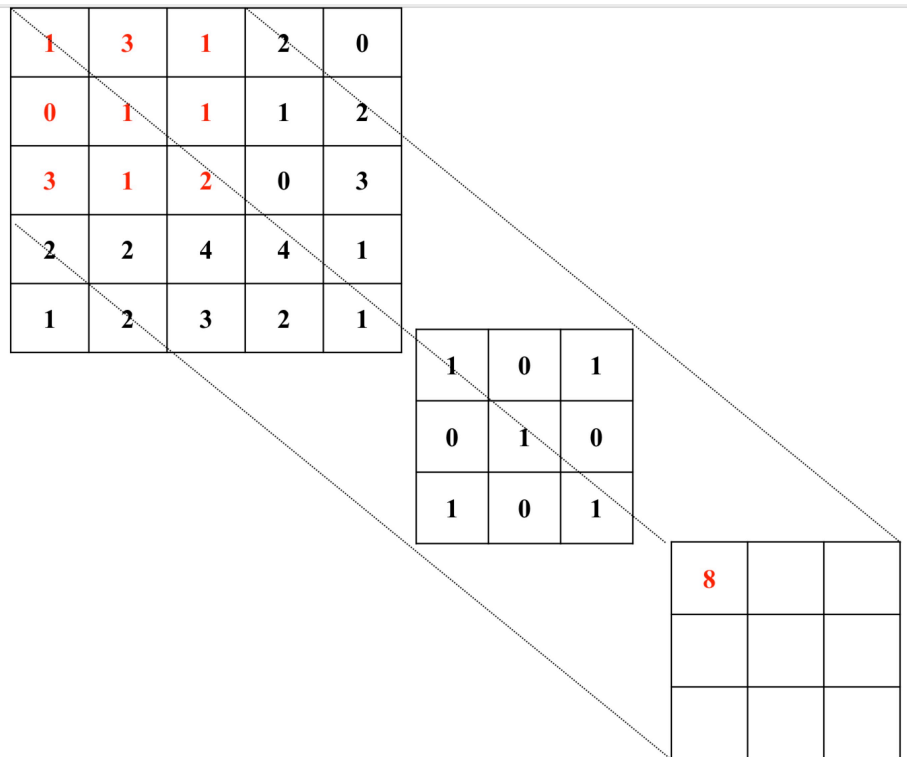


Figure 2.4 An example of how convolution works

At a convolutional layer, filters convolve with the receptive field of the input image and then slide over the input image until all receptive fields are reached and the output computes followed the formula:

$$h_i = f(W_k * I_i + b)$$

where h_i is the i^{th} feature map generated from convolutional layer, f is the activation function, I_i is the i^{th} region in input image, W_k is the weight of filter and b is the bias. The number of feature maps and the size of the filter are predetermined at network building stage and the optimal value is independent for each case.

There are three parameters that can decide the size of feature maps [28].

- Depth: decide the number of the outputs generated from the convolutional layer.
- Stride: decide the size of sliding step of the filter.
- Zero-padding: add a number of zeros on the border which is convenient to slip the end position from the initial position with the step length.

2.2.2 Pooling layer

Pooling layers are inserted into a CNN periodically followed after convolutional layers. Pooling layers downsample the spatial size of feature maps to reduce the number of parameters and amount of calculation, which enable network to learn features more effective and control overfitting better [28]. The most effective and commonly used pooling method is max pooling, which simply takes the max value of feature maps over a pooling window. Figure 2.5 illustrate an example of max pool with size of 2×2 and stride 2.

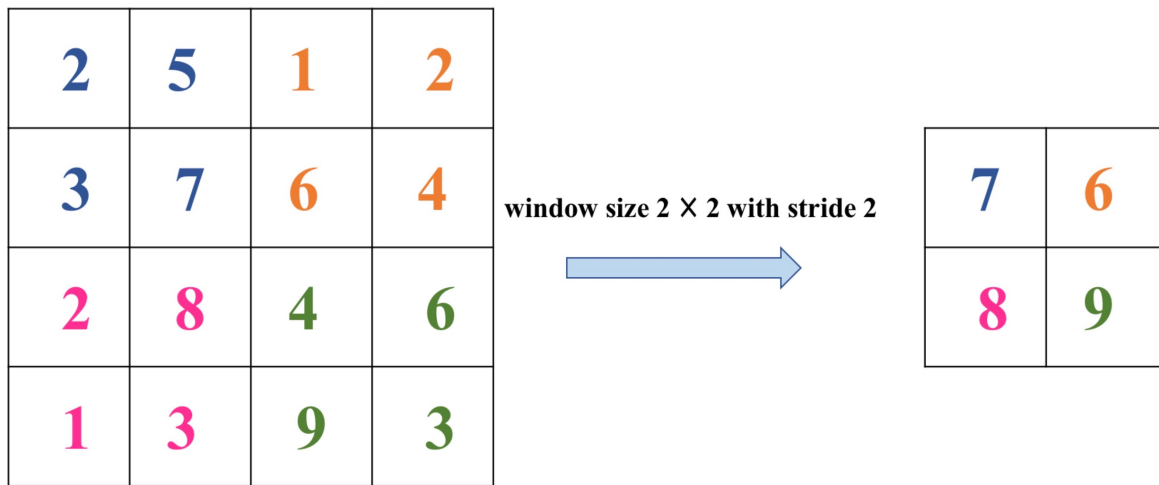


Figure 2.5 An example of max pooling

2.2.3 Fully-connected layer

Fully-connected layer provides a non-linear combination of features so that output layer can use these features to classify the input into corresponding class. Neurons in fully connected layers have a full connection between two adjacent layers.

2.3 Autoencoder

Autoencoder is one of the most frequently used unsupervised learning method. It is a neural network that aims to reconstruct the input at the output layer [34]. Therefore, an autoencoder attempts to find the function $h_{W,b}(x) \approx x$ which will reconstruct the input x as \tilde{x} , where W is the weight of each hidden neuron and b is the bias [49]. Autoencoder is mainly composed of an encoder and a decoder. Figure 2.6 illustrates the architecture of a simple autoencoder.

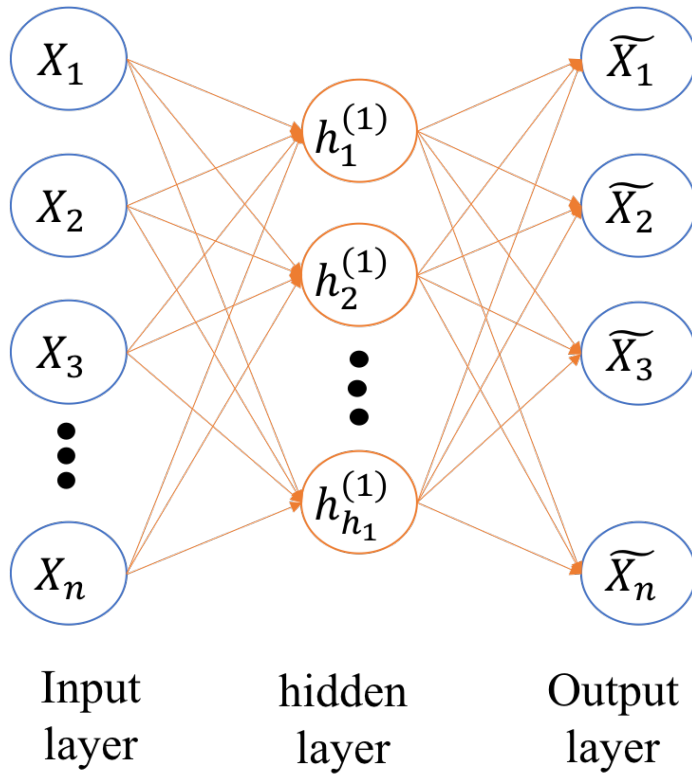


Figure 2.6 The architecture of a simple autoencoder.

There is an encoder at the input layer which encodes the input data. Then the encoded data are transfer to the hidden layer, and the hidden layer learns features of the input data. Then through the decoder at the output layer, autoencoder can reconstruct the input data.

By applying a backprobagation algorithm to train the autoencoder, the optimal (W,b) will be learned by minimizing the discrepancy between the input x and its reconstruction \tilde{x} . The cost function of training an autoencoder can be expressed as [25]:

$$J_{AE} = \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^H (x_{in} - \widetilde{x}_{in})^2$$

where N is the number of the input data and H means the number of hidden neurons.

Chapter 3: Method

Previous works [22,23] on translucency of BCC mainly employed conventional methods which analyzed features of translucency manually. These methods required comprehensive knowledge of translucency to select and analyze the features. In this chapter, we proposed two deep learning methods to achieve translucency detection of BCC automatically. One is based on the convolutional neural network (CNN) and the other is based on the stacked sparse autoencoder (SSAE). Both proposed methods are data-driven methods that can learn features directly from the input data.

3.1 Patch strategy

In this thesis research, we decide to apply patches to our proposed methods. There are three reasons for doing this:

- Patches can provide better presentations of characteristics of translucency and be more suitable for detection and localization translucency in BCC. As we mention earlier, translucency is a clinical feature and it usually accounts for a small portion of the images. Therefore, using the whole image to predict the presence or absence of translucency is unreasonable.
- Patches will increase the number of data points for the learning process. As we know, training A deep neural network needs a large amount of data. Lack of labeled data is the common challenge of the medical image analysis problems. Patching is a good way to increase the dataset.
- Patches will decrease the dimensionality of input data so that the network can be carried efficiently.

3.2 Method one: A CNN-based framework

In recent years, CNN becomes the most popular method for computer vision tasks and also achieves superior performances for these tasks such as face detection [44], image classification [46] and object recognition [45]. Therefore, there is an increasing trend of applying CNN in medical image analysis tasks [37, 47-48]. CNN is a data-driven method which can learn distinctive features directly from the raw data so that the requirement of domain knowledge can be reduced. In addition, according to its special structure, CNN can extract high-level and robust features in layer-by-layer manner with a low computational cost. Therefore, we decide to take an advantage of CNN and apply a CNN to our study.

CNN is different from a traditional neural network. CNN contains a special kind of layer, convolutional layer, which can extract features directly from the inputs by using a set of convolutional filters across over the inputs. These learned features make a contribution to the final classification. There are two properties of a convolutional layer: weight sharing and partial connectedness [36]. These properties greatly reduced the number of parameters. Another special layer of CNN is a pooling layer which down-samples the inputs. It decreases the dimension of the inputs, thereby reducing the number of parameters as well. These two kinds of layers make the learning of CNN extremely efficient. So, it allows network going deeper with a reasonable computational time. Finally, fully-connected layers, which are connected at the end of network, provide a combination of features for making the final detection. The details of the proposed CNN method will discuss in the following sections.

3.2.1 Overview of the proposed method

In this study, we proposed a CNN-based framework for detecting translucency automatically. The diagram of the proposed method is illustrated in Figure 3.1. From the diagram we can see that a

CNN is applied to labeled patches. Patches are labeled as translucent or non-translucent by expert dermatologist according to the presence or absence of translucency. Then patches are fed into the CNN. The CNN learns high-level features directly from the patches and makes predictions for translucent or non-translucent of patches with a softmax function.

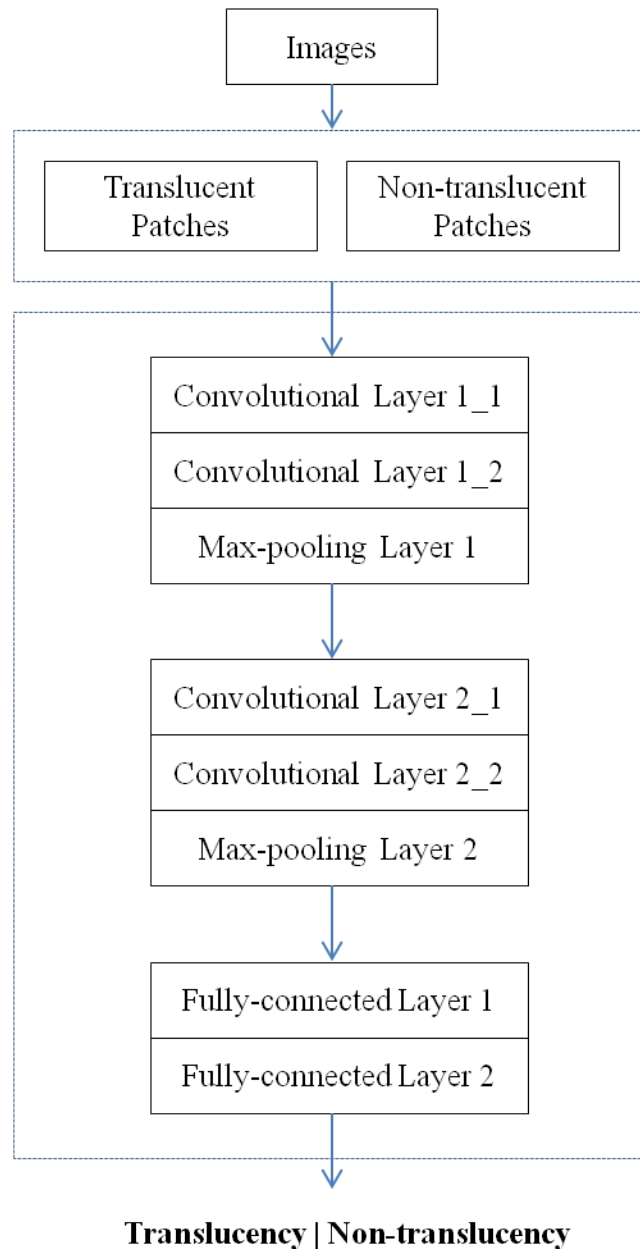


Figure 3.1 Diagram of CNN-based framework

3.2.2 Architecture of CNN

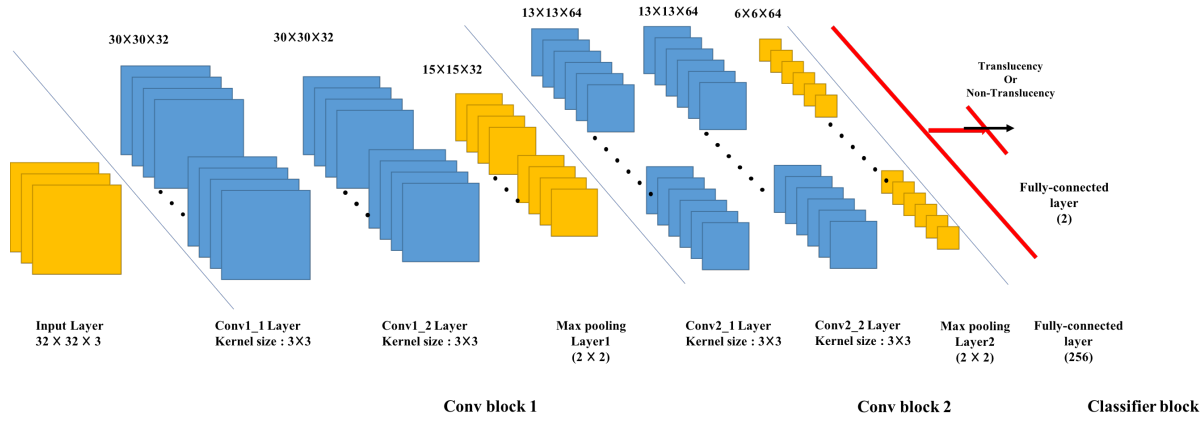


Figure 3.2 Architecture of the designed CNN

The architecture of CNN is shown in Figure 3.2. There are two convolution blocks and a classifier block contained in CNN. In each convolutional block, two 3×3 convolutional layers are stacked and a max pooling layer with size of 2×2 followed at the end of the block. There are 32 feature maps and 64 feature maps generated from first and second convolution block separately. Therefore, the output of first convolution block is 32 feature maps with size of 15×15 and 64 feature maps with size of 6×6 are generated from second convolution block. The classifier block consists of two fully-connected layers. The first fully-connected layer contains 256 neurons. The last fully-connected layer is the output layer which equipped with a softmax function as its activation function. The number of neurons in the last fully-connected layer is required to be same as the number of classes [28]. In our case, it is a binary classification problem, thereby the number of neurons is 2. With the help of the softmax function, the output layer gives the probabilities of two classes so that patches can be categorized to be translucent or non-translucent with the highest

probability. Totally, there are four convolutional layers, two max-pooling layers and two fully-connected layers in our CNN.

As we know, different tasks require different architectures of CNN. The choice of the architecture of CNN in our study depends on the number of images in our dataset and their sizes. Because we do not have a large amount of data point, we cannot choose an overly deep architecture of CNN which may cause the overfitting. Since the size of our images is only 32 x 32, we choose a small filter size. Moreover, the final decision of the architecture is based largely on the experiment results in our preliminary work.

3.3 Method two: A SSAE-based framework

In method one, we proposed a CNN-based method to detect the translucency of basal cell carcinoma. CNN is a supervised model which is training by a set of labeled data. Nowadays, the amount of labeled medical images in general is far from the needs of scientific research. Therefore, the unsupervised learning methods gradually attract people's attention. Stacked sparse autoencoder (SSAE) is one of the most common used unsupervised learning methods. It is a type of deep neural network which can learn high-level features through reconstruct the input at output layer [52]. SSAE can be trained by a large number of unlabeled data and learns features directly from the raw data. Thus, in method two, we proposed a framework based on SSAE in order to take its advantage for our study. Unlike CNN which is a partial connected network, SSAE is a fully-connected neural network where an encoder represents the original image at the input layer and a decoder reconstruct the original image at the output layer [38]. And high-level features are learned through the hidden layers.

In this study, the proposed method can learn the high-level features directly from a set of unlabeled images by SSAE and fed the learned features into a classifier to make a translucency detection. The details of the proposed SSAE-based method will discuss in the following sections.

3.3.1 Overview of the proposed method

In this study, we proposed a SSAE-based framework for automatic translucency detection. The diagram of proposed method is shown in Figure 3.3. From the diagram we can see that SSAE are applied by patches. SSAE learned high-level features directly in an unsupervised manner from the input patches. The features are then fed into a softmax classifier, which, however, run in a supervised manner. Softmax classifier provides the probability of each class where the patch belongs to and categorizes patches into one of the label classes with the highest probabilities. Then method predicts each patch is translucent or non-translucent in a framework combining both unsupervised and supervised learning.

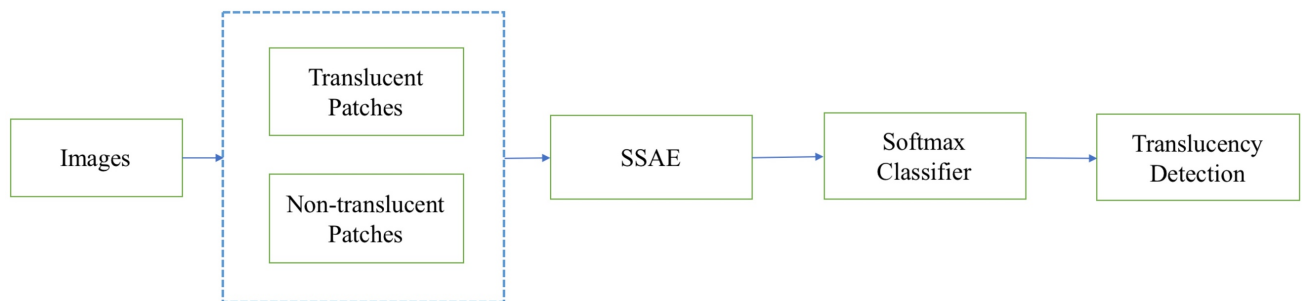


Figure 3.3 Diagram of SSAE-based framework

3.3.2 Sparse autoencoder

As describe in Chapter 2.3, autoencoder is a deep learning method which can learn high-level features in an unsupervised manner. It reconstructs the input data at the output layer in order to discovery a hidden feature representation of the input data [34]. Therefore, the autoencoder attempts to find the function $h_{W,b}(x) \approx x$ which will reconstruct the input x as \tilde{x} , where W is the weight of each hidden neuron and b is the bias [49]. A sparse autoencoder is a type of autoencoder with a sparsity constraint which will be discussed below. The same as the autoencoder, using a back propagation algorithm to train the sparse autoencoder, the optimal (W,b) will be learned by minimizing the discrepancy between the input x and its reconstruction \tilde{x} . The cost function of training a sparse autoencoder is [25]:

$$J_{SAE} = \frac{1}{N} \sum_{n=1}^N \sum_{i=1}^H (x_{in} - \widetilde{x}_{in})^2 + \eta ||W||^2 + \beta \sum_{j=1}^H KL(\rho || \tilde{\rho}_j),$$

where the first term is the mean square error term which describes the error between the input data and its reconstruction. N is the number of the input data and the H is the number of hidden neurons. The second term is a weight decay term which aims to decrease the magnitude of the overall neuron weight in order to avoid overfitting. η is the attenuation coefficient of the weigh decay. The third term is the sparsity constraint term which constrains the average activation value of each hidden neuron to be close to zero. ρ is the desired activation, a free sparsity parameter, which determines the proportion of neurons being active and ρ' is the average activation of j^{th} hidden neuron. The aim of sparse constrain is to minimize ρ_j using Kullback-Leibler (KL) divergence, $KL(\rho || \rho_j)$, between ρ_j and ρ . KL measures the difference of two distributions with the formulation: $KL(\rho || \rho_j) = \rho \log \frac{\rho}{\rho_j} + (1 - \rho) \log \frac{1-\rho}{1-\rho_j}$. β controls the weight of this penalty term.

3.3.3 High-level features learning through SSAE

A stacked sparse autoencoder is a neural network consisting of multiple layers of sparse autoencoders in which features are learned layer-by-layer [39]. The output of each layer is wired to the input of the successive layer [39]. For the purpose of detecting translucency from BCC we considered to use a two-layer sparse autoencoder. The architecture of the translucency detection framework is demonstrated in Figure 3.4.

As Figure 3.4 shows, the color input patches ($32 \times 32 \times 3$) are fed into the first layer that are transformed to the feature representations h_1 as the result of first layer training. Then the second layer is fed with the new feature representations to learn the high-level features h_2 . Finally, the high-level features learned from SSAE acts as the input to the softmax classifier for translucency detection.

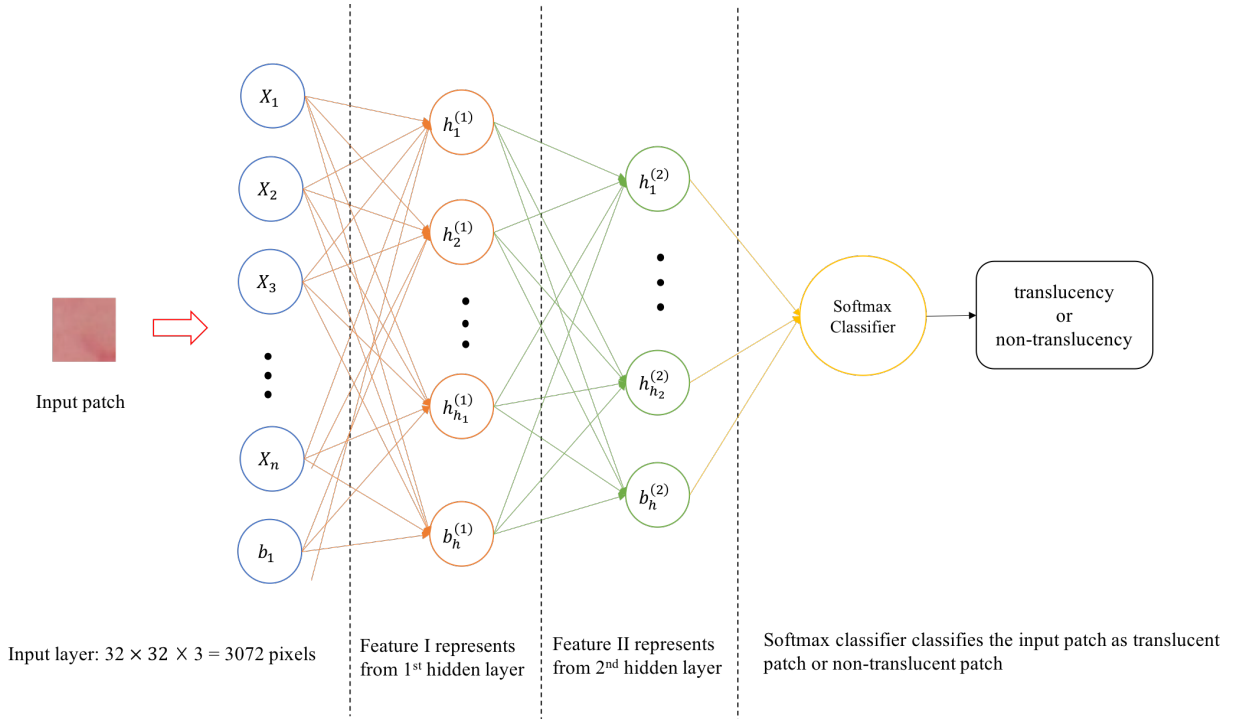


Figure 3.4 The proposed framework of the stacked sparse autoencoder and softmax classifier for detecting translucency

3.3.4 Translucency detection through softmax classifier

Softmax classifier is a supervised learning method which is a generalization of logistic regression [39]. Softmax categorizes the newly learned features into one of the label classes which has the highest probabilities. For instance, the classifier produces the probability of the presence of translucency in an input patch $t=1$ as follows:

$$P(t = 1|z) = \frac{1}{1 + e^{-z}}$$

where z is the learned high-level feature.

And the softmax classifier is trained by minimizing the cross-entropy between the estimated class q and the true class p :

$$H = \frac{1}{N} \sum_{j=1}^N \sum_{i=1}^C p_{ij} \log q_{ij} + (1 - p_{ij}) \log (1 - q_{ij}),$$

where N is the total number of inputs and C is the number of classes. In our study, it is a binary classification problem, thus, C is 2.

For training the softmax classifier, the high-level features learned from SSAE, which regard as the input, are fed into the classifier with the associate labels since softmax learns in supervised manner. Then softmax classifier will ready for detecting translucency according to the calculated probability of patches contain translucency.

3.4 Conclusion

In this chapter, we proposed two deep learning frameworks in order to achieve translucency detection of BCC automatically. The first one is CNN-based framework and the other one is SSAE-based framework. Both methods are applied to patches. For CNN-based framework, CNN learns the high-level features directly from the patches and with the help of the softmax function to make

predictions for presence or absence of translucency in patches. For SSAE-based framework, the high-level features learned from SSAE and are fed into a softmax classifier. Then the softmax classifier categorizes the newly learned features into translucency or non-translucency which has the highest probabilities.

Chapter 4: Translucency detection of Basal cell carcinoma in Dermoscopy images

Previous works [22, 23] related to translucency of BCC mainly examined image properties of translucent area in BCC that towards to discriminate BCC from other skin cancers. In [22], six color features and six texture features were analyzed from manually selected translucent areas and the features were used to classify BCC. In [23], a texture-based segmentation method is proposed which based on the manually selected features of translucency. Both works rely handcrafted features. In this chapter, two proposed frameworks which based on deep learning methods attempt to detect translucency of basal cell carcinoma in dermoscopy images automatically. Both methods are data-driven methods which can learn high-level features directly from input data.

4.1 Dataset

In this study, the dataset consists of 200 dermoscopy images of basal cell carcinoma from the University of Missouri. The size of images is 1024×728 pixels. The diagnoses of the lesions are included with the skin images. Figure 4.1 illustrates examples of dermoscopic BCC images.

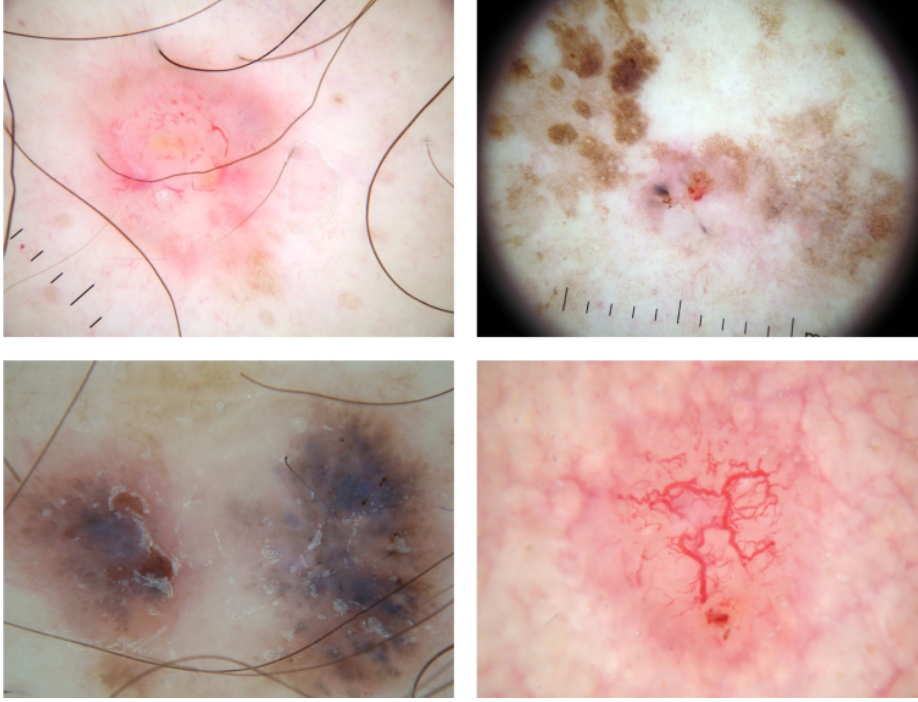


Figure 4.1 Examples of dermoscopy basal cell carcinoma images

4.2 Patching and Labeling

In this research, we decided to use a patching strategy to detect translucency of BCC automatically. For the purpose of training, all patches were labeled. However, labeling patches one by one is extremely hard and time consuming. Therefore, we first segmented translucency manually from images and then divided them into patches.

In our dataset, all images were manually segmented by an expert dermatologist. Examples of segmentation results are shown in Figure 4.2. The area within the red border is translucent and the area outside the red border is non-translucent. Then we divided the images into non-overlapped 32×32 patches from the top-left corner of an image. If the patch contained any translucent pixel, it was labeled as 1 which is a translucent patch. If the patch was entirely from the non-translucent area, it was labeled by 0 which is a non-translucent patch. The examples of patches are shown in Figure 4.3.

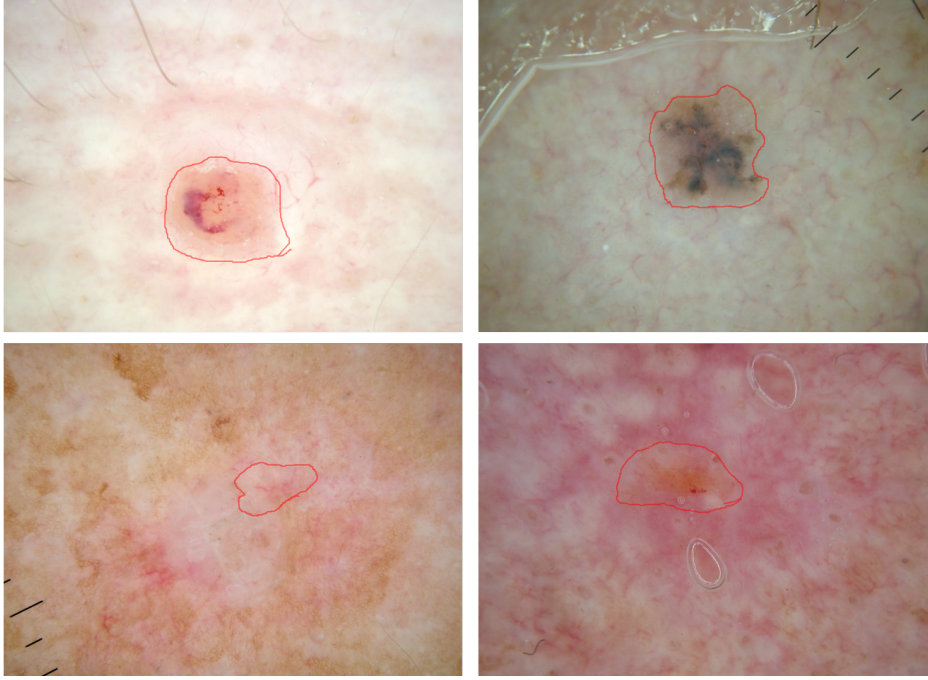


Figure 4.2 Examples of translucency segmentation in dermoscopy images by doctor

4.3 Training and testing set preparation

Totally we have 200 images. 160 images in the dataset were randomly selected as the training set which accounts for 80% of the images in the dataset. The rest 40 images were grouped into the testing set.

All images, both in the training set and testing set, were divided into non-overlapped patches of size 32×32 pixels which the size is large enough to present the features of translucency and small enough for localization later. The procedure of patch generation is followed Chapter 4.2. In the training set, the total number of patches was 25016 that 8141 patches were translucent and 16875 patches were non-translucent. In the testing set, 1740 patches were labeled as translucent and 4191 patches were labeled as non-translucent.

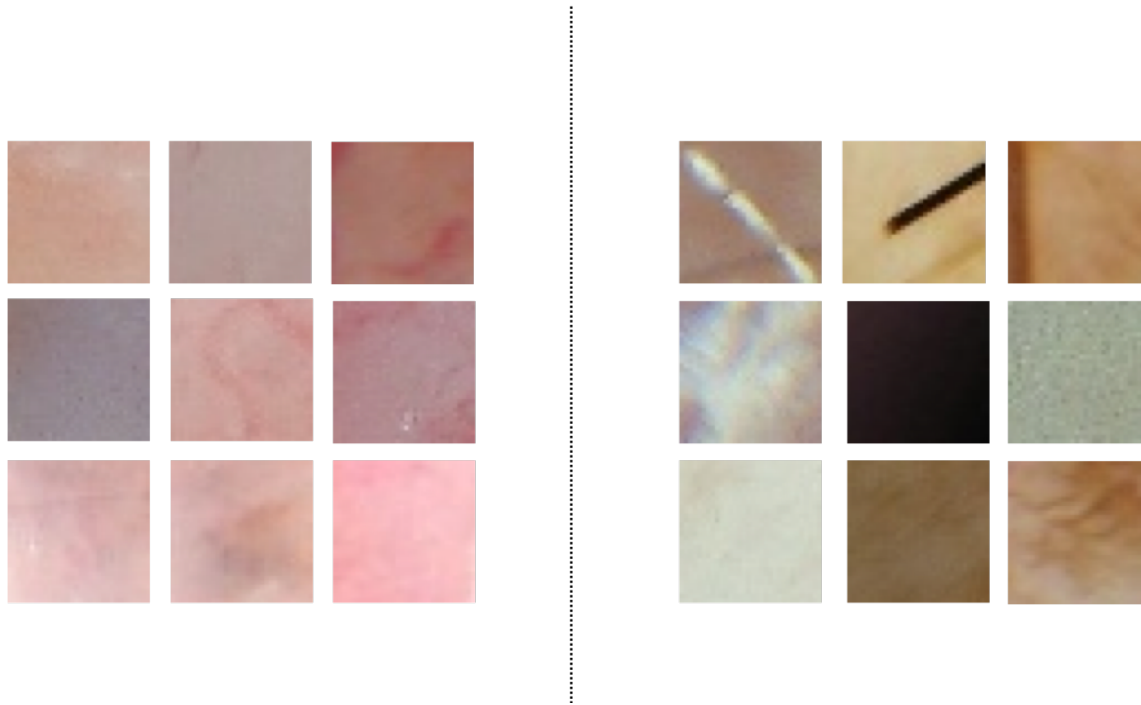


Figure 4.3 Examples of translucent(left) and non-translucent patches(right) of dermoscopy images

4.4 Study one: Automatic detection of translucency via a Convolutional neural network

In this study, the CNN-based framework which described in Chapter 3.2 was applied. First, dermoscopy images were divided into patches with associate labels. Then the patches were fed into the designed CNN. The CNN learned the features of translucency from the patches directly and made decisions whether the patches were translucent or not. The details of experiments are described as followed sections.

4.4.1 Experimental setting

For all experiments, the initial weights of the CNN were randomly generated with a uniform distribution. The dropout ratio of 0.25 was used for each convolution block and the value of 0.5 was used for the fully-connected layer. The model was trained for 80 epochs because the

convergence can be seen with 80 epochs. The model loss of CNN method is shown in Figure 4.4. The batch size was 16 and momentum was 0.9.

All experiments were implemented on Keras framework using a tensorflow backend and ran on a PC with an Intel core i7 processor, 16GB of RAM and a GeForce GTX NVIDIA Graphics Processor Unit.

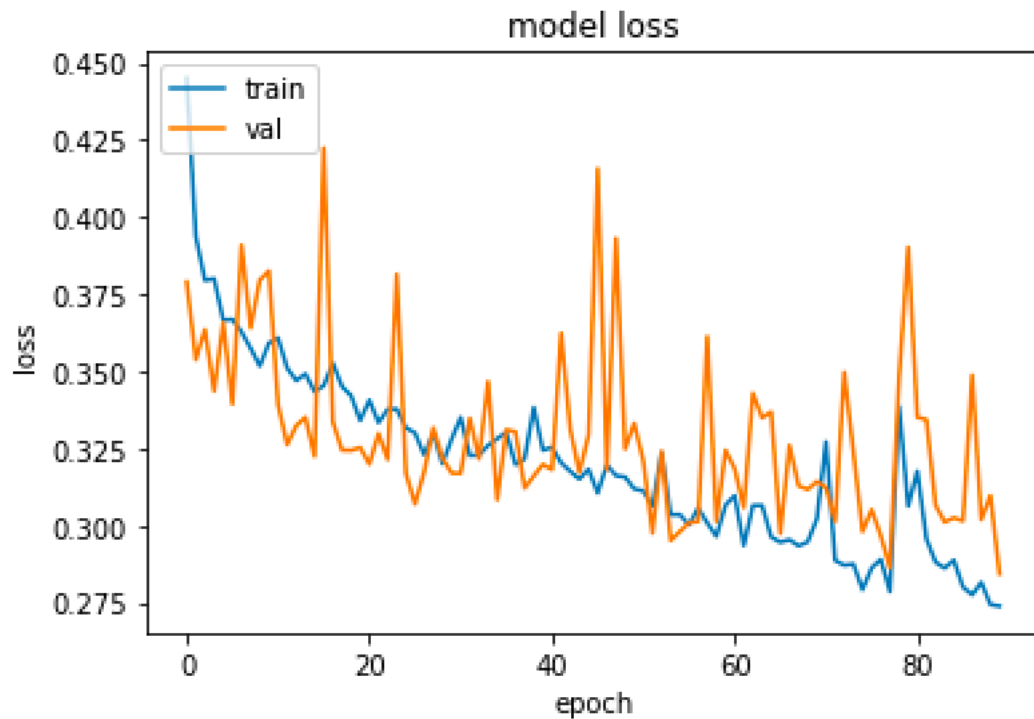


Figure 4.4 The model loss of CNN method

4.4.2 Experiment result

We used a metrics, which includes sensitivity, specificity, PPV, NPV, and accuracy, to evaluate the performance of the proposed method. The result is shown in Table 4.1 and the Receiver Operating Characteristic Curve (ROC) of the proposed method is shown in Figure 4.5. From Table 4.1 we can see that the proposed method achieved the accuracy of 77.8%. The sensitivity and PPV

were only 51.8% and 65.6%, respectively. From Figure 4.6, The Area Under the Curve (AUC) was 74.3%.

	Sensitivity	Specificity	PPV	NPV	Accuracy
CNN method	0.518	0.888	0.656	0.814	0.778

Table 4.1 The quantitative result of translucency detection in dermoscopy images using CNN-based method

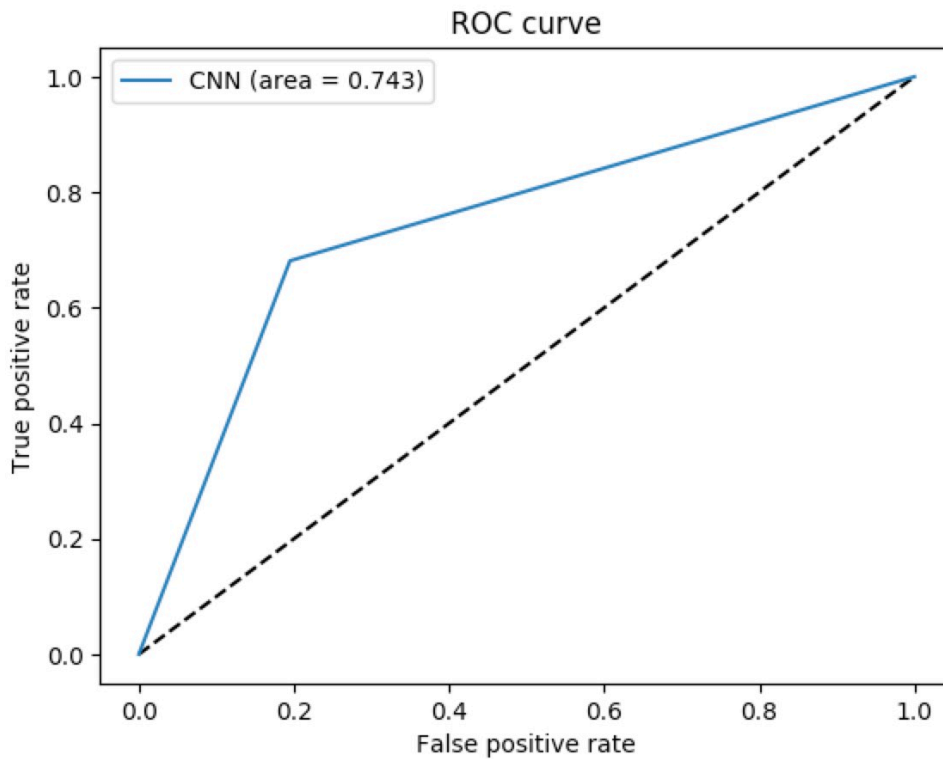


Figure 4.5 The ROC curve for translucency detection in dermoscopy images using CNN-based method

The experiment showed that CNN could be a fair translucency detector for BCC using dermoscopy images. Although the accuracy of the proposed method was almost 80%, the sensitivity and PPV were low. Sensitivity is the proportion of translucency detected among all true translucent patches. In this result, it means that a large number of translucent patches, about 48.2%,

was not detected. PPV is the proportion of the truly translucency detection among all detected results. In our case, only 65.6% of detected translucent patches were true positives. However, specificity and NPV were high at 88% and 82%, respectively.

4.5 Study two: Automatic detection of translucency via a SSAE-based framework

In study one, we used the CNN-based framework to detect translucency. However, the result was less than our expectations. Therefore, in this study, the SSAE-based framework, which described in Chapter 3.3, was attempted to detect translucency in dermoscopy BCC images again. Also, the patches, generated from the dermoscopy images, were applied to the SSAE network. The SSAE learned high-level features through reconstructing the patches, from the input layer, at output layer. Then the learned features from SSAE were fed into a softmax classifier to make a final detection of translucency.

4.5.1 Experimental setting

In the experiment, patch size was chosen as 32×32 pixels. The input size for the SSAE was, thus, $32 \times 32 \times 3 = 3072$, because all the images are RGB color images. All three color channels are inputted to the network simultaneously. For the successive layers of SSAE, the number of hidden nodes in first and second layer were chosen as $h1 = 625$ and $h2 = 225$. For two control parameters, the sparsity parameter β was set to 4 and the weight decay parameter η was set to 0.001. ρ which is the desired activation parameter was set to 0.05. Initialization of bias and the weight of each neuron were random at the beginning of training. All experiments were performed on a PC with an Intel core i7 processor, 16GB of RAM and a Geforce GTX NVIDIA Graphics processor Unit.

4.5.2 Experiment result

The result is shown in Table 4.2 and the Receiver Operating Characteristic Curve (ROC) of the proposed method is shown in Figure 4.6. From Table 4.2 we can see that the proposed method just achieved the accuracy of 78.6%. The sensitivity and PPV were only 52.0% and 70.1% respectively.

	Sensitivity	Specificity	PPV	NPV	Accuracy
CNN method	0.520	0.902	0.701	0.810	0.786

Table 4.2 The quantitative result of translucency detection in dermoscopy images using SSAE-based method

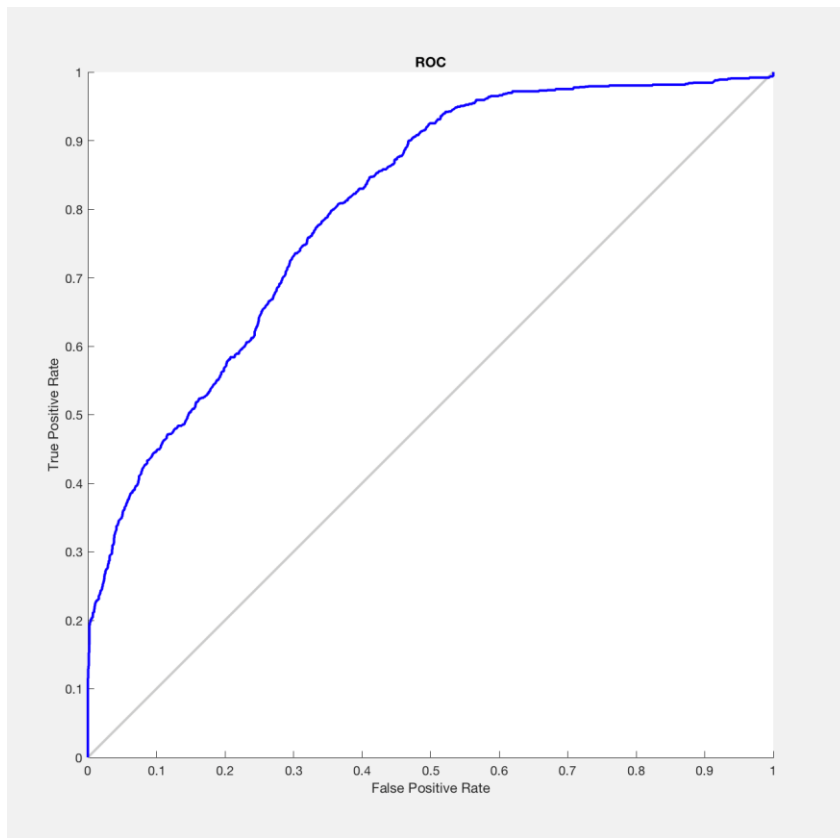


Figure 4.6 The ROC curve for translucency detection in dermoscopy images using SSAE-based method

From the results, the performance of the SSAE-based framework was also unexpected. The low value of sensitivity means only 52% translucent patches were detected. Also, only 70.1% of

detected translucent patches were correct detection according to the value of PPV. We cannot say the method achieved an excellent performance on translucency detection of BCC in dermoscopy images as well. However, the specificity and NPV were high.

4.6 Discussion

In this study, we use two proposed deep learning methods to detect translucency of BCC in dermoscopy images. However, both results were not satisfactory. In this section, we will discuss the possible reasons for unexpected results. From analyzing the failure cases, we could summarize four possible reasons.

Firstly, we find the translucency detection failed in pigmented BCCs. Figure 4.7 illustrates some of the failure cases.

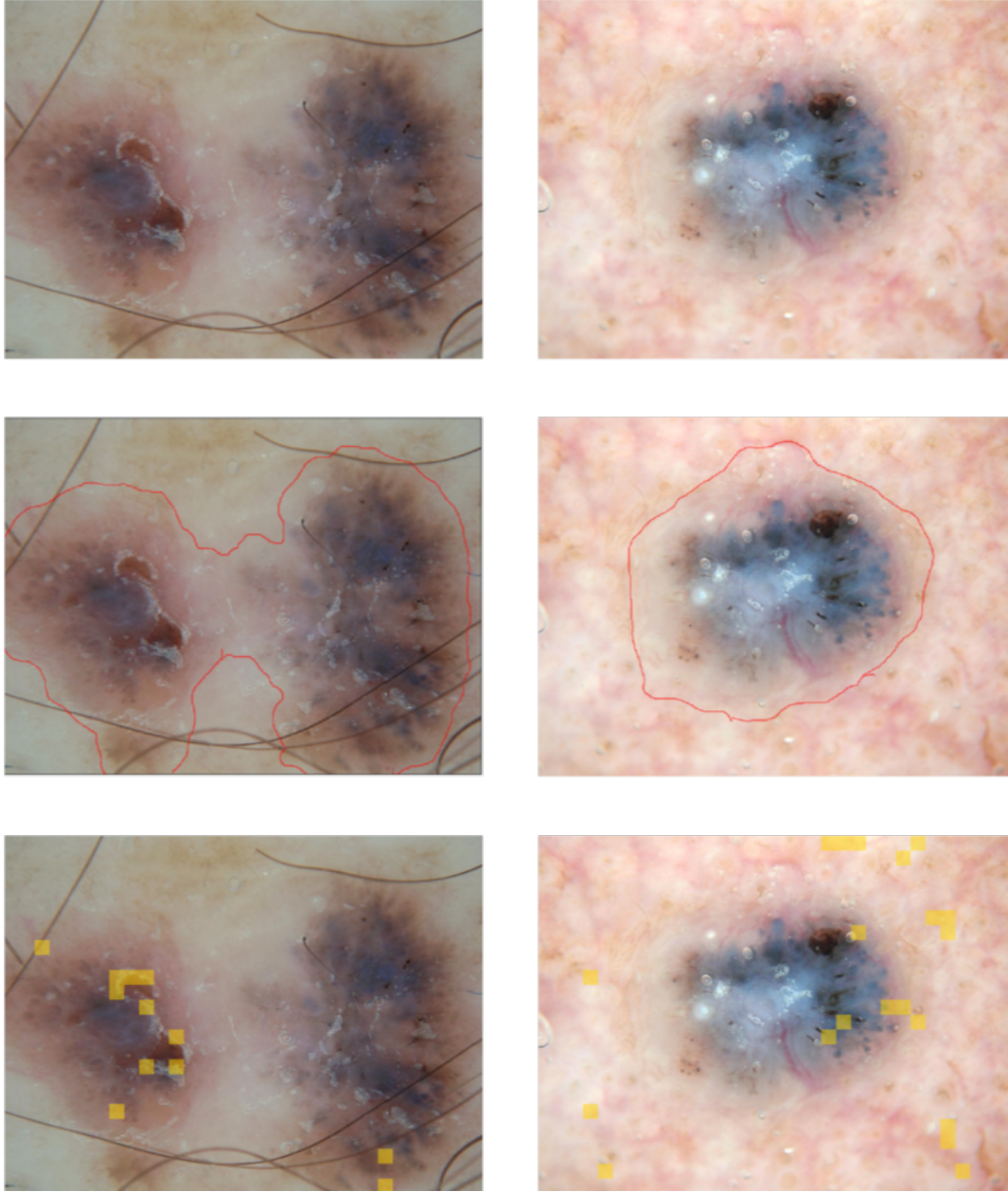


Figure 4.7 Examples of translucency detection in pigmented BCC.

Top: Original dermoscopy image. Middle: Manually segmented translucency in dermoscopy images. Bottom: Translucency detected results on dermoscopy images. The yellow blocks are the detected translucent patches by our proposed method.

Pigmented BCC accounts for around 8% of all BCCs [40]. In our dataset, there are 25 pigmented BCC with translucency. The pigmented BCC is greatly different from the other types of BCC like size, color, and structure of the lesion [41]. The translucency in pigmented BCCs is

highly pigmentation which is different from the translucency in non-pigmented BCCs. Figure 4.8 illustrates the features of translucency extracted from pigmented BCCs, non-pigmented BCCs and both kinds of BCCs by using SSAE. All the features are selected from the learned features in the first hidden layer of SSAE. From the Figure 4.8 we can see that the features of translucency are extremely different between pigmented BCCs and non-pigmented BCCs. Therefore, it is hard to find the common features which could detect translucency in both kinds of BCCs. From Figure 4.8 we can see that the common features of translucency which extract from both pigmented BCC and non-pigmented BCC are more similar to the translucent features extract from non-pigmented BCC rather than the features extract from pigmented BCC. Therefore, it is hard to detect translucency in pigmented BCC using common features of translucency.

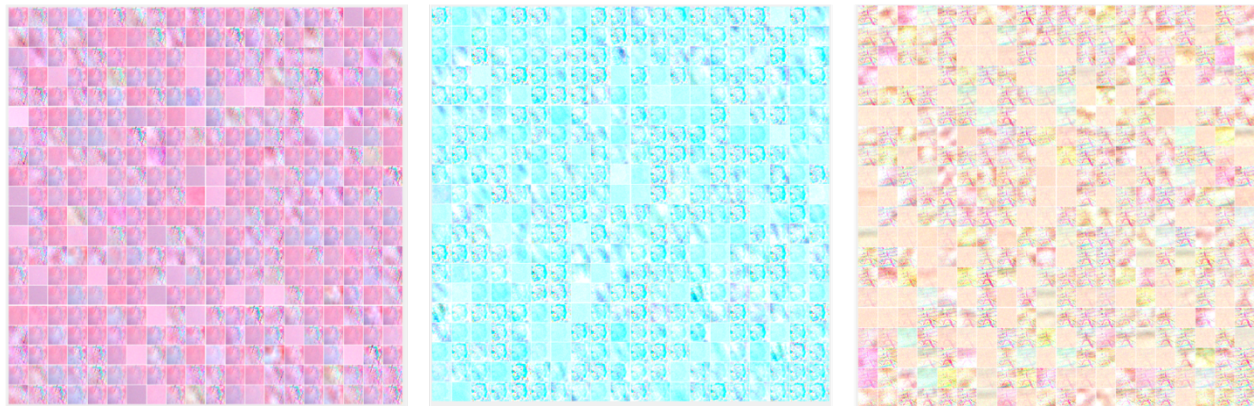


Figure 4.8 Features of translucency extracted from pigmented BCC(middle), non-pigmented BCC(left) and both types of BCC(right)

Secondly, we find translucency detection depends on the presence of vessels. Some examples of such results of translucency detection in dermoscopy images are shown in Figure 4.9.

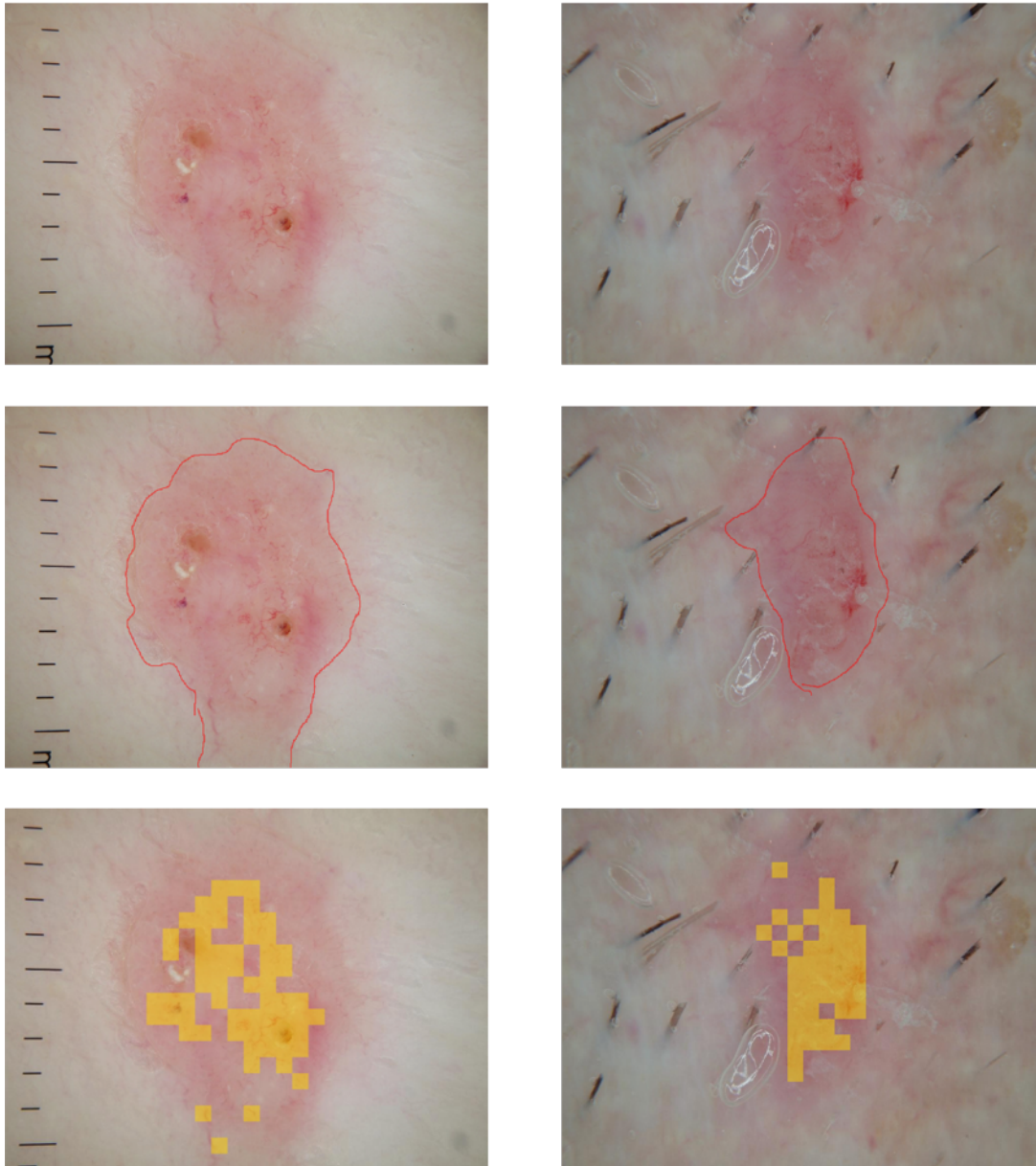


Figure 4.9 Examples of translucency detection in dermoscopy images

Top: Original dermoscopic images. Middle: Manually segmented translucency in dermoscopy images. Bottom: Translucency detected results. The yellow blocks are the detected translucent patches by our proposed method.

The detection results in Figure 4.9 are not bad. However, we find that the majority of detected translucent patches includes vessels but the missing patches did not include vessels. Blood vessels is one of the most important dermoscopy features of BCC [43]. Dermoscopy enables us to observe the structure of vessels which is unable to be seen by naked eyes, and translucency makes the skin become transparent so that vessels are easier to be observed. Near 80% translucent patches contain vessels. As a powerful biomarker, vessels may be seen as an important characteristic of translucency by proposed methods which select the translucency features automatically. Thereby, translucent patches are easier to detect when they contain vessels. However, not all translucent patches include vessels and the missed detection of these patches influences the performance. The examples of vessels present within and without translucent patches are shown in Figure 4.10

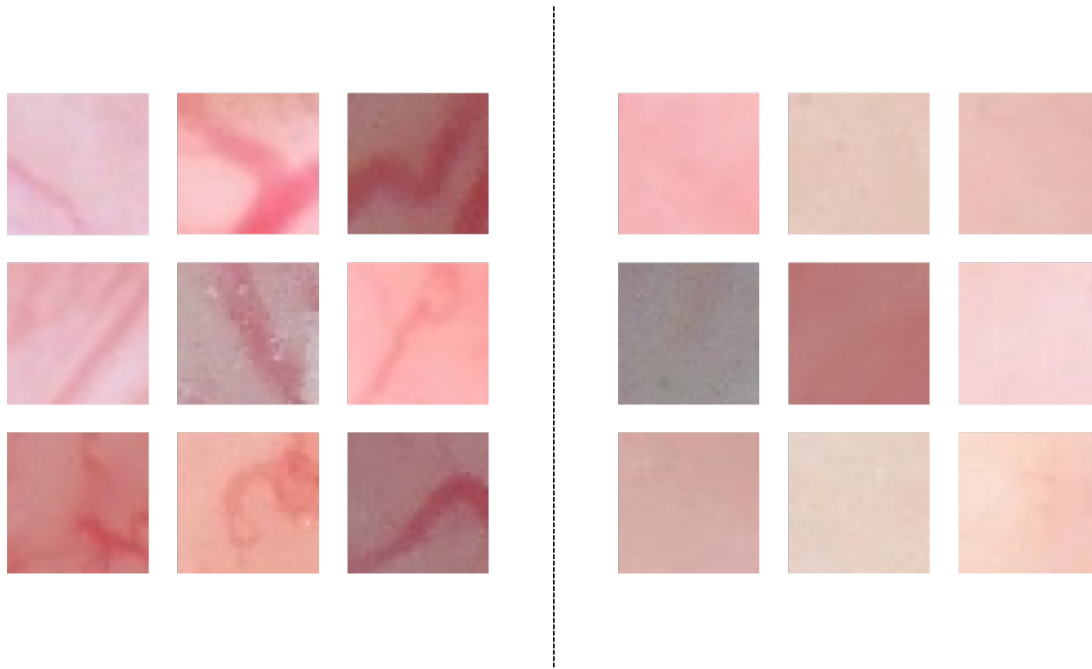


Figure 4.10 Examples of translucent patches with and without vessels
Left: Translucent patches contain vessels. Right: Translucent patches do not contain vessels.

Thirdly, we find that the translucent area which near the edge of a lesion are harder to be detected. Some examples are shown in Figure 4.11. From the Figure 4.11 we can see that the majority of detected translucency are near the center of the lesion and the most translucency which are not detected are in surrounding area of the lesion. Patches near the edge of the translucent area may contain both translucent and non-translucent features so that the detection of this kind of patches are difficult.

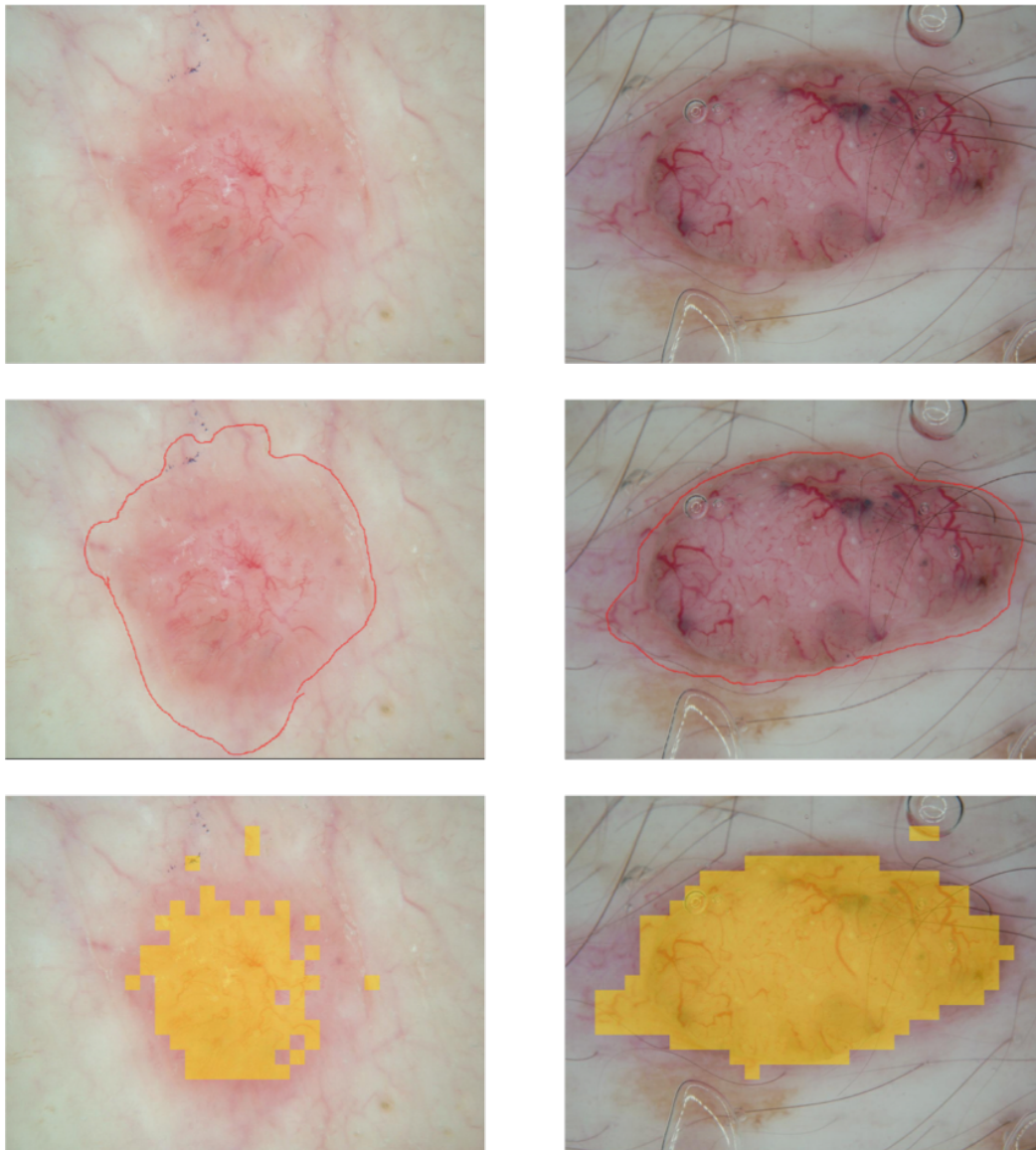


Figure 4.11 Examples of translucency detection in dermoscopy images

Top: Original dermoscopic images. Middle: Manually segmented translucency in dermoscopy images. Bottom: Translucency detected results. The yellow blocks are the detected translucent patches by our proposed method.

Finally, the presence of artifacts will affect the performance of translucency detection. The main artifacts of dermoscopy images are immersion oil bubbles and hairs [42]. The detection by our methods is based on the automatic learned features of translucency. Obviously, the presence of artifacts on the translucent area influences the features extraction and recognition on translucent area so that the detection accuracy will be affected. Figure 4.12 illustrates the poor results of translucency detection in two dermoscopy images where the immersion oil bubbles and hairs are present.

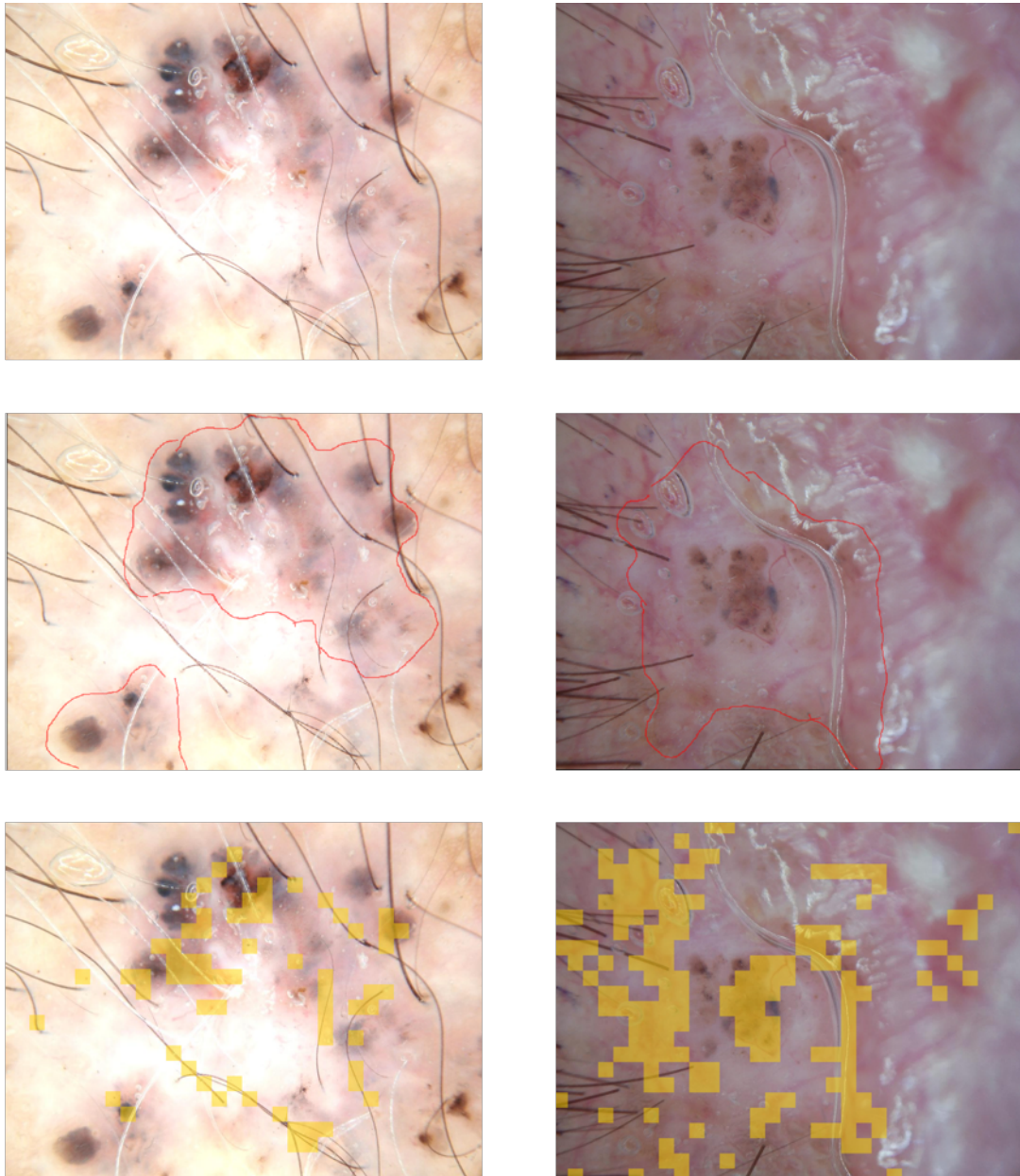


Figure 4.12 Examples of translucency detection in dermoscopy images with artifacts such as bubbles and hairs. Top: Original dermoscopy images. Middle: Manually segmented translucency in dermoscopy images. Bottom: Translucency detected results. The yellow blocks are the detected translucent patches by our proposed method.

4.7 Conclusion

In this chapter, two proposed methods were used to detect translucency of BCC in dermoscopy images. One was based on CNN and the other one was based on SSAE. However, both results were disappointing. Therefore, we analyzed the failure cases and found four possible reasons for fair performance. First, translucency detection failed in pigmented BCC. Second, translucency detection of BCC is affected by other features of BCC such as vessels. Third, the translucent patches near the edge of lesion is hard to be detected. Finally, the presence of artifacts in BCC influenced the accuracy of translucency detection.

Chapter 5: Translucency detection of Basal cell carcinoma in clinical images

Previous works in automatic detection of translucency in BCC have been done by using dermoscopy images. However, the results were not as good as our expectation. Clinical images, taken by color digital cameras, are another type of images used for diagnosis of BCC. Unlike dermoscopy, which is often pressed against a skin lesion and leads to distortion of the skin surface and color appearance, clinical images are captured free from skin contact, and, hence, free of distortion of the translucency feature. Thus, the performance of translucency detection in clinical images should potentially be better than the performance in dermoscopy images.

In this chapter, we apply the two proposed methods for detecting translucency of BCC in clinical images.

5.1 Dataset and pre-processing

The dataset we used consists of 32 clinical images of basal cell carcinoma collected from 32 patients in Vancouver Skin Care Centre. Figure 5.1 illustrates examples of clinical basal cell carcinoma images. The size of images is 3008×2000 pixels. The targeted lesion is near the center of the image and each lesion has a different magnification. All cases were confirmed by histopathological examinations.



Figure 5.1 Examples of clinical basal cell carcinoma images

5.1.1 Generating of ROIs

From the examples of clinical basal cell carcinoma images shown in Figure 5.1, we can see that the lesion accounts for a small portion of the image, which is surrounded by many irrelevant structures like hair, nose, eyes and so on. A large number of irrelevant data would affect the accuracy of the detection result. Therefore, we focused on regions of interest (ROIs), which were created by bounding boxes. The ROIs contain the target lesion with the surrounding skin. The size of each ROI depends on the size of the targeted lesion in the original clinical image. Three examples were shown in Figure 5.2. The size of the ROI image on the left is 704×800 pixels, the size of the ROI image in the middle is 1120×1216 pixels and the size of the ROI image on the right is 1120×1248 pixels.

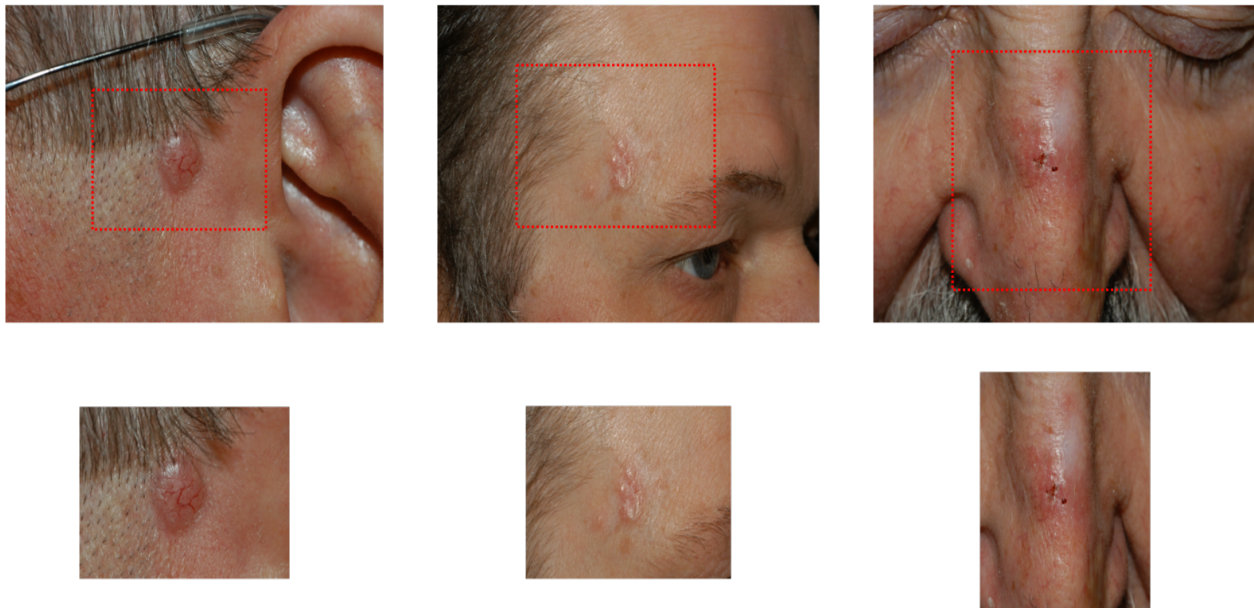


Figure 5.2 Examples of ROIs generating

Top: Images are original clinical basal cell carcinoma images with the bounding box. Bottom: ROI images

5.1.2 Patching and Labeling

As well, in clinical images, we still use the same patching strategy to detect translucency as described in Chapter 4. The translucent areas in clinical images were first segmented manually by

an expert dermatologist. Examples of translucency segmentation are shown in Figure 5.3. The area within the red border was translucent and the area outside the red border was non-translucent. Then we divided the images into non-overlapped patches starting from the top-left corner. If a patch contained any translucent pixel, it was labeled to 1, indicating a translucent patch. If a patch contained entirely non-translucent pixels, it was labeled as 0 which was a non-translucent patch.



Figure 5.3 Examples of translucency segmentation in clinical images by doctor

The total number of patches were 4401; there were 797 translucent patches and 3604 non-translucent patches. Figure 5.4 shows the examples of translucency and non-translucency patches generated from clinical images.

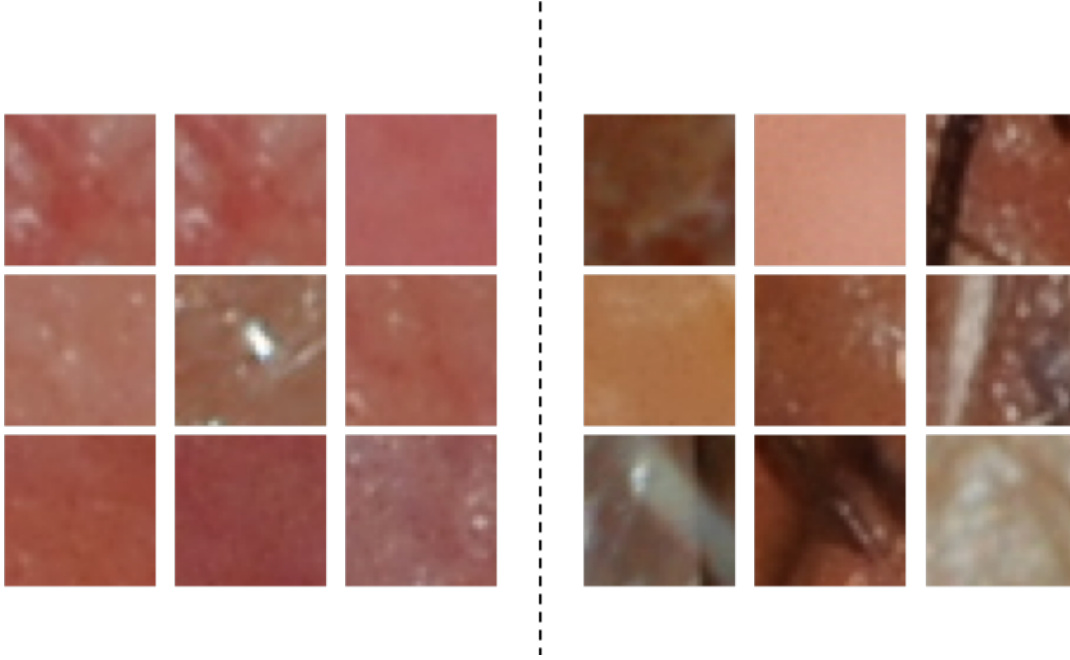


Figure 5.4 Examples of translucent patches and non-translucent patches of clinical images
Left: Translucent patches. Right: Non-translucent patches.

5.2 Study one: Automatic detection of translucency using a Convolutional neural network

In study one, the same proposed CNN was used for translucency detection in clinical images. The details of the method are described in Chapter 3.2. The patches generated from BCC clinical images were fed into the CNN and it learned high-level features from the input patches and made the final classification with the help of a softmax function. In following sections, we present the experiments and discuss the results.

5.2.1 Experimental setting

For all experiments, the initial weights of the CNN were randomly assigned by values with a uniform distribution. The dropout ratio of 0.25 was used for each convolution block and the value of 0.5 was used for fully-connected layer. The model was trained for 50 epochs since the error

convergence will be seen within 50 epochs. The model loss of CNN method is shown in Figure 5.5. The batch size was 32 and momentum was 0.9.

All experiments were implemented based on a Keras framework using a tensorflow backend and ran on a PC with an Intel core i7 processor, 16GB of RAM and a GeForce GTX NVIDIA Graphics Processor Unit.

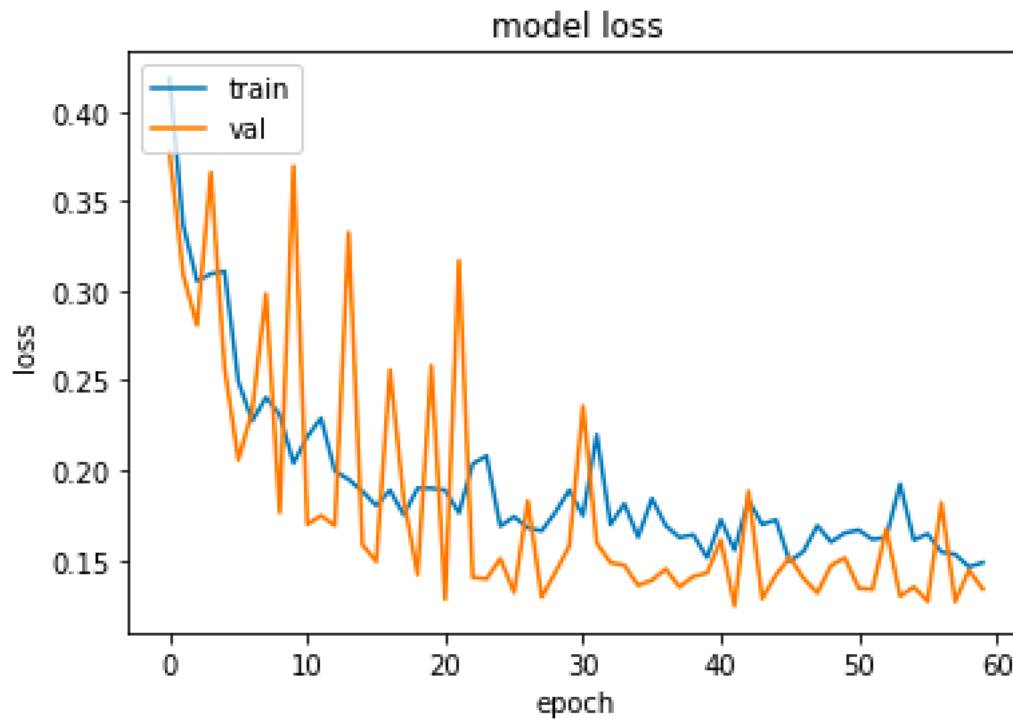


Figure 5.5 The model loss of CNN method

5.2.2 Experiment result

Applying the proposed convolutional neural network with a five-fold cross-validation to the patches, the result of translucency detection is illustrated in Table 5.1. From the Table1 we can see the method achieved an accuracy of 0.931, a sensitivity of 0.757, a specificity of 0.99, a PPV of 0.945 and a NPV of 0.937.

	Sensitivity	Specificity	PPV	NPV	Accuracy
CNN method	0.757	0.990	0.945	0.937	0.931

Table 5.1: Results of translucency detection in clinical images using CNN-based method

The results of translucency localization in clinical images are shown in Figure 5.6. The translucency localization is based on the detection results which used the patch strategy. Thus, the localization works for patches as well. First, we divided a test image into patches and predicted each patch was translucent or not. If the patch was translucent, we marked the patch as a yellow mask. Therefore, the part of the test image that is highlighted by a yellow block was the detected location of translucency. The lesions used in the localization analysis were randomly selected from different type of basal cell carcinoma.

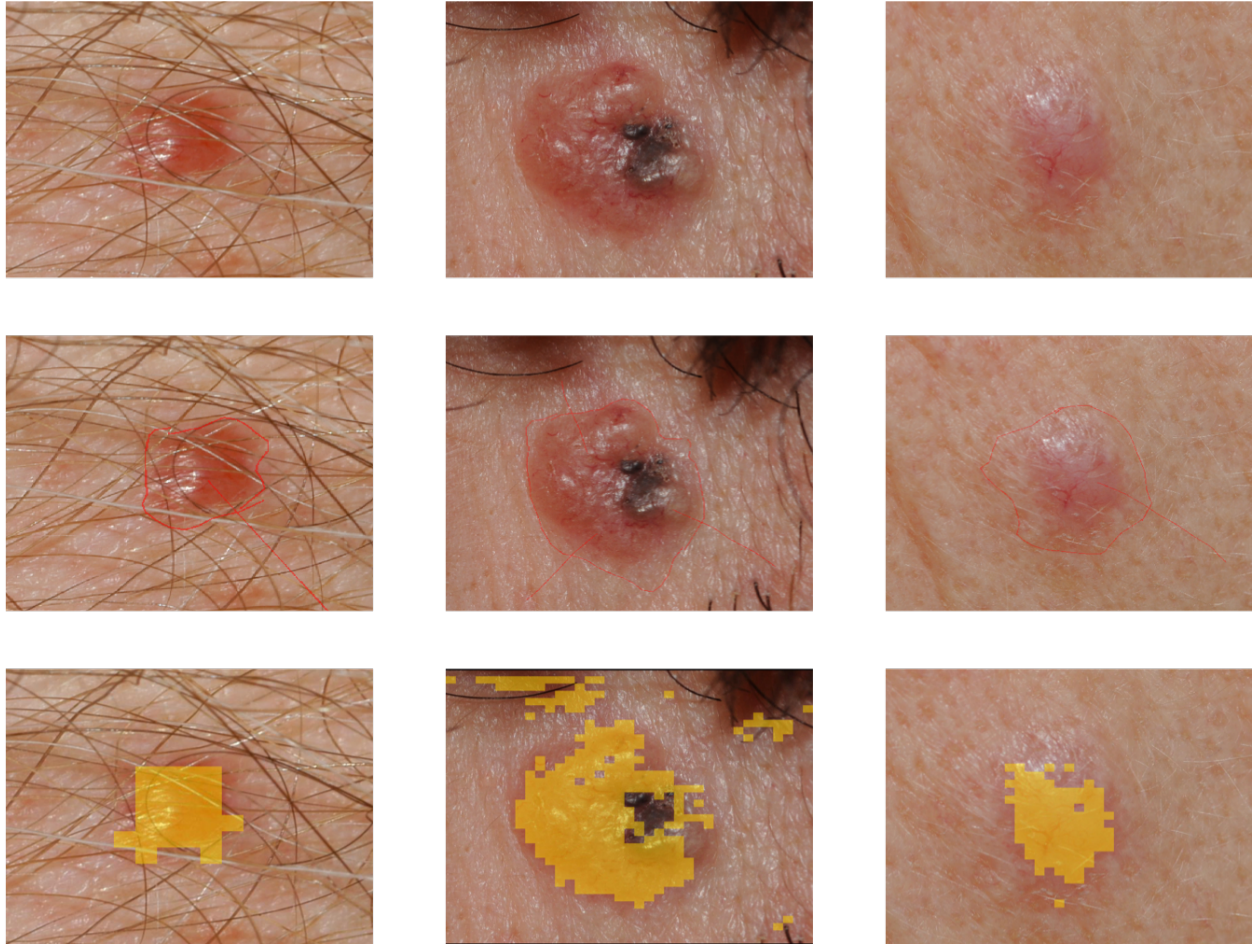


Figure 5.6 Translucency localization results of Basal cell carcinoma in clinical images using CNN-based method

Top: Original clinical basal cell carcinoma images. Middle: Manually segmented translucency in clinical images. Bottom: Translucent patches localization of basal cell carcinoma in original clinical image. The yellow blocks are the detected translucent patches by CNN method.

From the experiment results, we demonstrated that the convolutional neural network works well in detecting translucency of clinical images. Also, the experiment results demonstrated the proposed convolutional neural network enable to localize the translucency in clinical image based on the detection results.

5.3 Study two: Automatic detection of translucency using a SSAE-based method

In study one, the proposed CNN method achieved an outstanding performance. In this study, we applied the SSAE-based method, described in Chapter 3.3, to detect translucency of BCC in clinical images. First, the clinical images were divided into patches. The patches were applied to the SSAE. The SSAE method learned high-level features directly from pixel-level from the patches in an unsupervised manner; then the learned features were fed into a softmax classifier to make a final detection. In following section, we present the experiments and discuss the results.

5.3.1 Experimental setting

In the experiment, the input size of the deep network was $32 \times 32 \times 3 = 3072$. Because all the images were RGB, all three color channels were inputted to the network simultaneously. For successive layers of SSAE, the number of hidden nodes in first and second layer were chosen $h_1 = 225$ and $h_2 = 100$. For two control parameters, the sparsity parameter β was set to 4 and the weight decay parameter η was set to 0.001. ρ which is the desired activation parameter was set to 0.05. Initialization of bias and the weight of each neuron were randomly assigned at the beginning of training. All experiments were performed on a PC with an Intel core i7 processor, 16GB of RAM and a Geforce GTX NVIDIA Graphics processor Unit.

5.3.2 Experimental results

We applied the proposed SSAE-based method with a five-fold cross-validation to the patches generated from the dataset. The result of translucency detection is illustrated in Table 5.2. The SSAE-based method achieved an accuracy of 0.93, a sensitivity of 0.770, a specificity of 0.971, PPV of 0.873 and NPV of 0.942. The Receiver Operating Characteristic Curve (ROC) of the

proposed method is shown in Figure 5.7.

	Sensitivity	Specificity	PPV	NPV	Accuracy
SSAE+SMC	0.770	0.971	0.873	0.942	0.930

Table 5.2 Results of translucency detection in clinical images use SSAE-based method

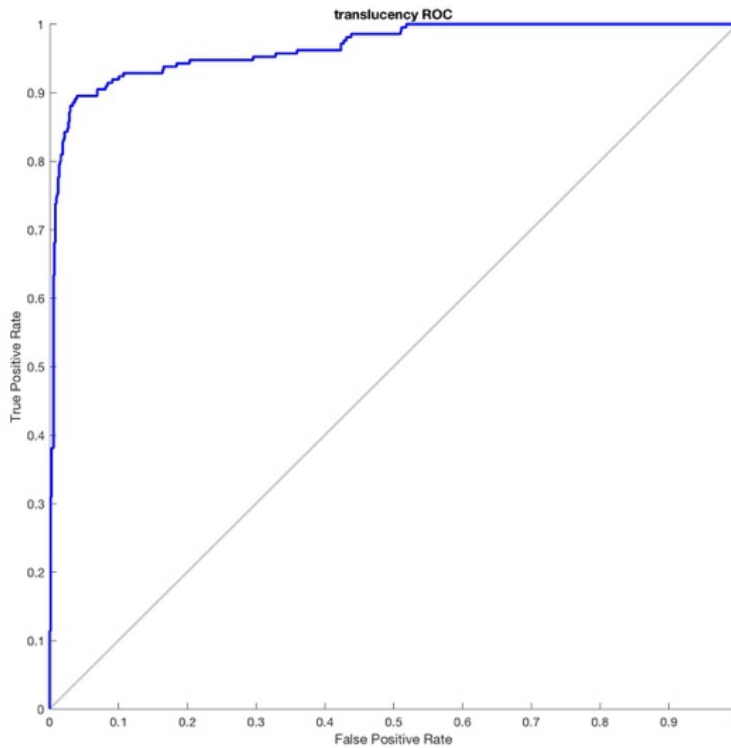


Figure 5.7: the ROC curve for translucency detection of SSAE-based method

5.3.3 Discussion

In this study, we proposed a SSAE-based framework to detect translucency in BCC. There were two key points affecting the performance of the proposed method. The first one was the number of

nodes in each hidden layer of SSAE and the other one was the size of patches. To achieve the best performance, we performed many experiments to evaluate different combinations.

The number of hidden nodes: In this study, we used two layers of SSAE so that we needed to decide the number of nodes for each hidden layer. To find the proper combination of the nodes number for two layers, we attempted three combinations: First, we set 625 nodes in first hidden layer and 225 nodes in second hidden layer. Second, we set 400 nodes in first hidden layer and 100 nodes in second hidden layer. Finally, we set 225 nodes in first hidden layer and 100 nodes in second hidden layer. The comparison of the quantitative performance of these three combinations of the number of hidden nodes is shown in Table 5.3. From Table 5.3 we can see that the performance of three combinations were almost the same. Only the combination that there are 225 nodes in the first hidden layer and 100 nodes in the second hidden layer achieved a slightly higher accuracy. Therefore, for our study, we set 225 nodes in first hidden layer and 100 nodes in second hidden layer.

	Sensitivity	Specificity	PPV	NPV	Accuracy
625-225	0.743	0.973	0.880	0.939	0.928
400-100	0.754	0.970	0.867	0.941	0.928
225-100	0.770	0.971	0.873	0.942	0.930

Table 5.3 The quantitative performance comparison of three different combinations of nodes numbers

The patch size: In this study, we used a patch strategy to detect translucency in BCC. Therefore, the size of patches was very an important factor. To tune the patch size, we tried three different sizes of patches: 8×8, 16×16, 32×32 pixels. The comparison of the performance of three

different sizes is shown in Table 5.4. From Table 5.4 we can see that the performance of the size 8×8 was inferior than the other two. For size 16×16 and size 32×32 , they achieved almost the same accuracy but the former had a higher specificity and PPV and the latter had a higher sensitivity and NPV. In my opinion, when accuracy is the same, the sensitivity is the most important evaluation indicator. Because our purpose is to detect translucency in BCC, sensitivity is the proportion of correctly detected translucent patches for all translucent patches; a higher sensitive indicates the ability to detect higher percentage of translucent patches [39]. Therefore, for our study, we select 32×32 as the size of patches which has the highest sensitivity value.

	Sensitivity	Specificity	PPV	NPV	Accuracy
8×8	0.687	0.964	0.856	0.920	0.928
16×16	0.739	0.980	0.91	0.937	0.928
32×32	0.770	0.971	0.873	0.942	0.930

Table 5.4 The quantitative performance comparison of three different sizes of patches

5.4 Comparison of two proposed methods

5.4.1 The quantitative performance comparison

The quantitative performance of the two proposed methods are shown in Table 5.5. It can be seen in Table 5.5, the accuracy of two proposed methods are almost the same. Therefore, to compare two methods, the other two statistical indicators are significant that are sensitivity and PPV. Sensitivity is the proportion of correct detection of translucent patches for all translucent patches. PPV is the ability of correctly detecting the translucency patch [39]. The greater the values of sensitivity, the more translucent patches are detected. In other word, there will be less missing cases.

Therefore, to address the detection problem which prefers a higher detection rate, and thus the higher sensitivity. PPV is the probability of truly translucent patches for detected translucent patches. PPV tells us the portion of detected translucent patches is true positives. The higher value of PPV means the better precision of translucency detection. For localization of translucency which is based on the results of translucency detection, the more accurate of detection, the better work can be done. Thus, for localization problem, performance is better when the value of PPV is higher. From Table 5.5 we can see that the SSAE-based method has a slightly higher value of sensitivity, which may imply that the technique would work a little better than the proposed CNN method for detecting the translucency of basal cell carcinoma in clinical images. On the other hand, for the localization of translucency of basal cell carcinoma in clinical images, the proposed CNN method may have a small edge than the SSAE-based method.

	Sensitivity	Specificity	PPV	NPV	Accuracy
CNN method	0.757	0.990	0.945	0.937	0.931
SSAE+SMC	0.770	0.971	0.873	0.942	0.930

Table 5.5 The quantitative performance comparison of proposed methods in clinical images

5.4.2 The localization performance comparison

The comparison of localization performance of two proposed methods are shown in Figure 5.7. From the Figure 5.8, we can obviously see that CNN-based method achieved better performance on translucency localization of BCC.

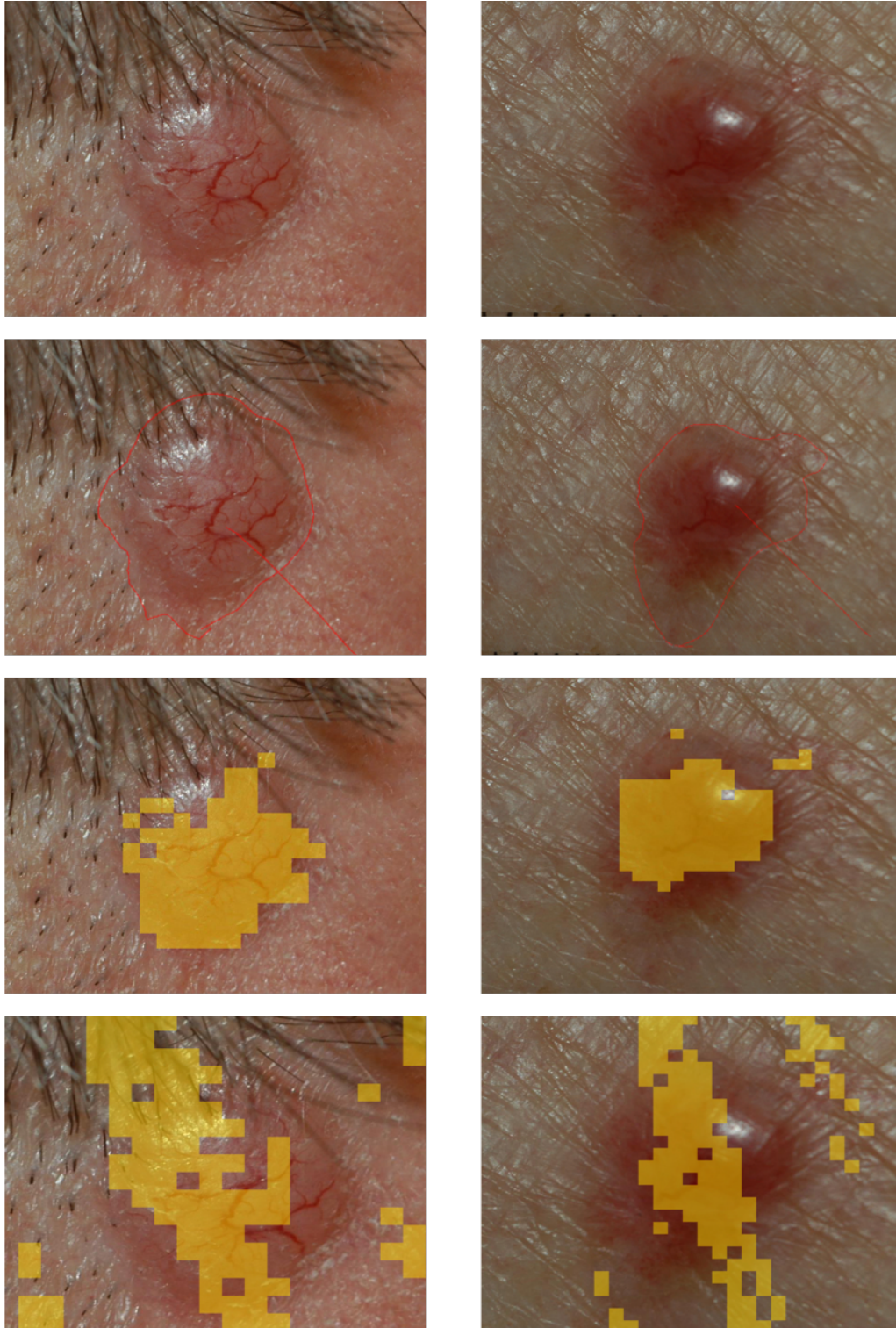


Figure 5.8 Localization results comparison between two proposed methods in clinical images.

Top one: Original clinical images. Top two: Manually segmented translucency in clinical images. Bottom two: Localization results using the CNN-based method. Bottom one: Localization results using the SSAE-based method.

5.4.3 Feature Visualization

Both proposed methods are feature learning networks which feed the learned features into a classifier to predict the presence or absence of translucency in a basal cell carcinoma. Therefore, all detections of translucency are based on the knowledge of learned features. Visualizing the features which are learned from the CNN and SSAE can help us understand what features methods learned and how methods understood the images. Figure 5.9 illustrated some examples of features learned from proposed methods. Left image in Figure 5.9 shows a set of features learned from the proposed CNN approach and the right one shows some examples of features learned from the proposed SSAE framework. The examples of the features, which were learned from CNN, are selected from the second convolution block that were the high-level features are fed into a softmax classifier to make a final decision. As well, the examples of features, which were learned from SSAE, are selected from the second hidden layer where the high-level features are generated. From Figure 5.9 we can see, features learned from the SSAE framework provide a good visual feeling of the translucency such as the color of the lesion and the brightness of the tumor. Whereas, the features learned through proposed CNN are abstract but the learned features still keep some local specification of translucency.

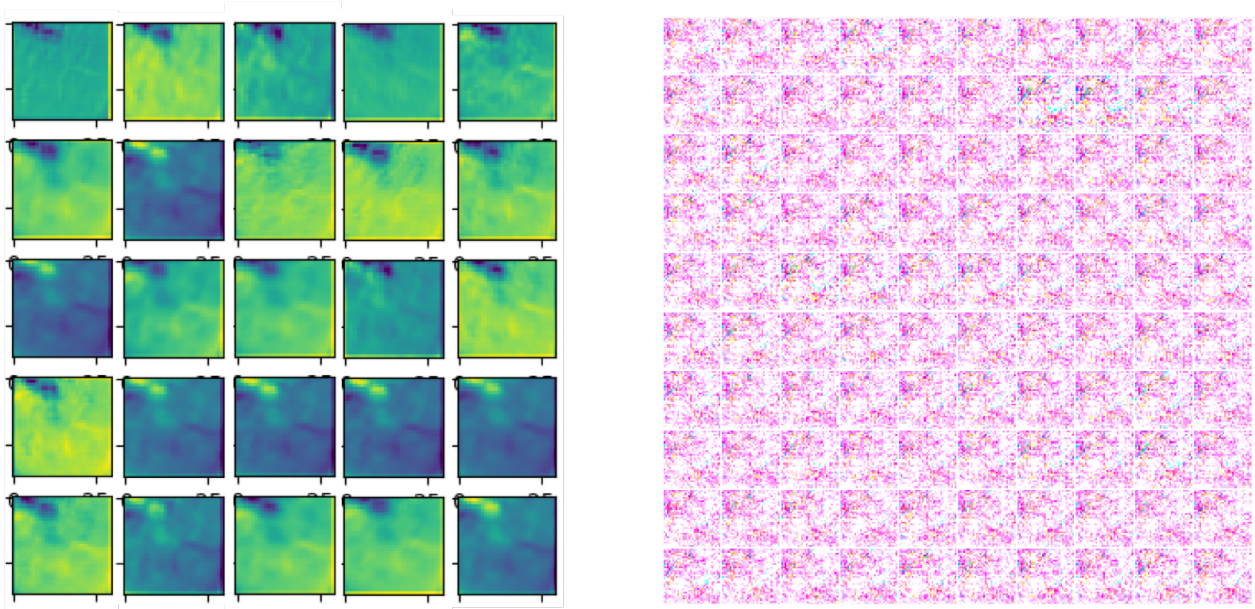


Figure 5.9 Examples of features learned from proposed methods.

Left: A set of features learned from proposed CNN. Right: A set of features learned from proposed SSAE framework.

5.5 Conclusion

In this chapter, two proposed methods were applied to detect translucency of BCC in clinical images. The performances of the two proposed methods are similar; both methods achieved superior accuracy where the CNN-based method achieved an accuracy of 93.1% and the SSAE-based method achieved an accuracy of 93%.

Chapter 6: Conclusion and Future work

6.1 Conclusion and Contribution

In this thesis, we proposed two deep learning based frameworks for detecting translucency of basal cell carcinoma (BCC) automatically. One framework is based on a designed convolutional neural network (CNN) and the other is based on a stacked sparse autoencoder (SSAE). Since the two common imaging modalities for skin cancer diagnosis are dermoscopy images and clinical images, we evaluated the proposed methods on both types of images. We note from our experimental results that the detection performances of the two proposed methods were similar. For detecting translucency in dermoscopy images, both proposed methods achieved comparable, but not so promising, accuracy results. For detecting translucency in clinical images, both proposed methods could achieve a promising performance.

In chapter 3, we developed two deep learning frameworks which aim to detect translucency of BCC automatically. The first one was based on a specially-designed CNN. The architecture of the CNN was composed of four convolution layers, two max-pooling layers and two fully-connected layers. The other deep network was based on the SSAE framework, which contained two stacked layers of sparse autoencoders.

In chapter 4, both proposed frameworks were applied to dermoscopy images. The performance results of the proposed frameworks were similar. The CNN-based framework achieved an accuracy of 77.8% and the SSAE-based framework achieved an accuracy of 78.6%. Both accuracy results were acceptable but did not meet our expectations. We discussed four possible reasons that might lead to the unexpected results. First, translucency detection seems failed in pigmented BCCs. Second, translucency detection of BCC is affected by other features of BCC

such as vessels. Third, the patches near the edge of a lesion is hard to be detected. Finally, the presence of artifacts in BCCs could influence the accuracy of translucency detection.

In chapter 5, we evaluated the proposed methods on clinical images. Unlike the performances in dermoscopy images, both proposed methods could achieve superior accuracy of translucency detection in clinical images. The CNN-based framework achieved an accuracy of 93.1% and the SSAE-based framework achieved an accuracy of 93.0%.

From our study, for translucency detection of BCC, I recommended to use clinical images rather than dermoscopy images. Dermoscopy enables the visualization of subsurface structures of BCC [40]. Especially some internal structures of BCC, which are not easily seen by naked eyes, could become clear in translucent areas under dermoscopy. These structures could complicate the translucency detection. However, in clinical images, subsurface structures of BCC are not visible so that translucency detection will be not affected. In addition, dermoscopy are often directly pressed against skin lesions. Such a contact action will likely cause distortion of skin surface and color appearance. It could affect the performance of translucency detection as well. For clinical images, which are taken by digital camera in a contact free manner, and, hence, free of distortion of translucency features. Moreover, translucency is the feature which can be seen by naked eyes. Therefore, there is no need to use another device such as a dermoscope to magnify the lesion along with its internal structures. To sum up, I think translucency detection of BCC is more suitable in clinical images than in dermoscopy images.

6.2 Future works

In recent years, transfer learning has been a powerful technique in image analysis tasks. It transfers prior knowledge of related tasks in the source domain to the new task in the target domain in order to achieve better performance. In this thesis, the proposed methods achieve superior performance

for translucency detection of BCC in clinical images but less promising performance in dermoscopy images. In the future, we can consider exploring transfer learning to transfer the knowledge of translucency learned in clinical images (the source domain) to dermoscopy images (the target domain) so that it may help achieve better performance in dermoscopy images.

In this thesis, we focus on translucency detection of BCC, since it is a key function of computer aids systems that aim to achieve accurate diagnosis of BCC at an early stage. By using clinical images, our proposed methods can achieve excellent performance. Therefore, in the future, we will incorporate our methods into the BCC diagnosis system in clinical images by integrating it with or without other feature detection functions.

Bibliography

- [1] Diepgen, T. L., and V. Mahler. "The epidemiology of skin cancer." *British Journal of Dermatology* 146.s61 (2002): 1-6.
- [2] Rogers, Howard W., et al. "Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population, 2012." *JAMA dermatology* 151.10 (2015): 1081-1086.
- [3] Eisemann, Nora, et al. "Non-melanoma skin cancer incidence and impact of skin cancer screening on incidence." *Journal of Investigative Dermatology* 134.1 (2014): 43-50.
- [4] Bath-Hextall, Fiona, et al. "Trends in incidence of skin basal cell carcinoma. Additional evidence from a UK primary care database study." *International Journal of Cancer* 121.9 (2007): 2105-2108.
- [5] Eisemann, Nora, et al. "Non-melanoma skin cancer incidence and impact of skin cancer screening on incidence." *Journal of Investigative Dermatology* 134.1 (2014): 43-50.
- [6] Arits, A. H. M. M., et al. "Trends in the incidence of basal cell carcinoma by histopathological subtype." *Journal of the European Academy of Dermatology and Venereology* 25.5 (2011): 565-569.
- [7] Verkouteren, J. A. C., et al. "Epidemiology of basal cell carcinoma: scholarly review." *British Journal of Dermatology* 177.2 (2017): 359-372.
- [8] Rubin, Adam I., Elbert H. Chen, and Désirée Ratner. "Basal-cell carcinoma." *New England Journal of Medicine* 353.21 (2005): 2262-2269.
- [9] Hakverdi, Sibel, et al. "Retrospective analysis of basal cell carcinoma." *Indian Journal of Dermatology, Venereology, and Leprology* 77.2 (2011): 251.
- [10] Zargarani, M., et al. "A clinicopathological survey of Basal cell carcinoma in an Iranian population." *Journal of Dentistry* 14.4 (2013): 170.
- [11] Rippey, J. J. "Why classify basal cell carcinomas." *Histopathology* 32.5 (1998): 393-398.
- [12] Esteva, Andre, et al. "Dermatologist-level classification of skin cancer with deep neural networks." *Nature* 542.7639 (2017): 115.
- [13] Argenziano, Giuseppe, and H. Peter Soyer. "Dermoscopy of pigmented skin lesions—a valuable tool for early." *The lancet oncology* 2.7 (2001): 443-449.
- [14] Thomas, Luc, and Susana Puig. "Dermoscopy, Digital Dermoscopy and Other Diagnostic Tools in the Early Detection of Melanoma and Follow-up of High-risk Skin Cancer Patients." *Acta dermato-venereologica* 97 (2017).

- [15] Oakley, A. Dermoscopy course. 2011; Available from: <https://www.dermnetnz.org/cme/dermoscopy-course/introduction-to-dermoscopy/>.
- [16] Kaliyadan, Feroze. "The scope of the dermoscope." *Indian dermatology online journal* 7.5 (2016): 359.
- [17] Bowling, Jonathan. "Introduction to dermoscopy." *Diagnostic Dermoscopy: The Illustrated Guide* (2012): 1-14.
- [18] Kittler, Harold, et al. "Diagnostic accuracy of dermoscopy." *The lancet oncology* 3.3 (2002): 159-165.
- [19] Herschorn, Andrea. "Dermoscopy for melanoma detection in family practice." *Canadian Family Physician* 58.7 (2012): 740-745.
- [20] Menzies, Scott W., and Iris Zalaudek. "Why perform dermoscopy?: The evidence for its role in the routine management of pigmented skin lesions." *Archives of dermatology* 142.9 (2006): 1211-1212.
- [21] Stoecker, William V., et al. "Semitranslucency in dermoscopic images of basal cell carcinoma." *Archives of dermatology* 145.2 (2009): 224-224.
- [22] Stoecker, William V., et al. "Detection of basal cell carcinoma using color and histogram measures of semitranslucent areas." *Skin Research and Technology* 15.3 (2009): 283-287.
- [23] Kefel, S., et al. "Adaptable texture-based segmentation by variance and intensity for automatic detection of semitranslucent and pink blush areas in basal cell carcinoma." *Skin Research and Technology* 22.4 (2016): 412-422.
- [24] Popadic, Mirjana. "Statistical evaluation of dermoscopic features in basal cell carcinomas." *Dermatologic Surgery* 40.7 (2014): 718-724.
- [25] Ng, Andrew, and Sparse Autoencoder. "CS294A Lecture notes." *Dosegljivo: https://web.stanford.edu/class/cs294a/sparseAutoencoder_2011new.pdf*. [Dostopano 20. 7. 2016] (2011).
- [26] Agatonovic-Kustrin, S., and R. Beresford. "Basic concepts of artificial neural network (ANN) modeling and its application in pharmaceutical research." *Journal of pharmaceutical and biomedical analysis* 22.5 (2000): 717-727.
- [27] Karlik, Bekir, and A. Vehbi Olgac. "Performance analysis of various activation functions in generalized MLP architectures of neural networks." *International Journal of Artificial Intelligence and Expert Systems* 1.4 (2011): 111-122.
- [28] Karpathy, Andrej. "Cs231n convolutional neural networks for visual recognition." *Neural networks* 1 (2016).

- [29] Basheer, Imad A., and M. Hajmeer. "Artificial neural networks: fundamentals, computing, design, and application." *Journal of microbiological methods* 43.1 (2000): 3-31
- [30] Cilimkovic, Mirza. "Neural networks and back propagation algorithm." *Institute of Technology Blanchardstown, Blanchardstown Road North Dublin* 15 (2015).
- [31] Parkhi, Omkar M., Andrea Vedaldi, and Andrew Zisserman. "Deep Face Recognition." *BMVC*. Vol. 1. No. 3. 2015.
- [32] Girshick, Ross, et al. "Rich feature hierarchies for accurate object detection and semantic segmentation." *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2014.
- [33] Nasr-Esfahani, Ebrahim, et al. "Melanoma detection by analysis of clinical images using convolutional neural network." *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the. IEEE*, 2016.
- [34] Goodfellow, Ian, et al. *Deep learning*. Vol. 1. Cambridge: MIT press, 2016.
- [35] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." *Advances in neural information processing systems*. 2012.
- [36] Zou, Liang, et al. "3D CNN based automatic diagnosis of attention deficit hyperactivity disorder using functional and structural MRI." *IEEE Access* 5 (2017): 23626-23636.
- [37] Nasr-Esfahani, Ebrahim, et al. "Melanoma detection by analysis of clinical images using convolutional neural network." *Engineering in Medicine and Biology Society (EMBC), 2016 IEEE 38th Annual International Conference of the. IEEE*, 2016.
- [38] Tao, Chao, et al. "Unsupervised spectral-spatial feature learning with stacked sparse autoencoder for hyperspectral imagery classification." *IEEE Geoscience and remote sensing letters* 12.12 (2015): 2438-2442
- [39] Ng, Andrew, et al. "UFLDL tutorial." 2012)[2014-08-12]. http://deeplearning.stanford.edu/wiki/index.php/UFLDL_Tutorial (2012)
- [40] Menzies, Scott W., et al. "Surface microscopy of pigmented basal cell carcinoma." *Archives of dermatology* 136.8 (2000): 1012-1016.
- [41] Aoyagi, Satoru, and Keyvan Nouri. "Difference between pigmented and nonpigmented basal cell carcinoma treated with Mohs micrographic surgery." *Dermatologic surgery* 32.11 (2006): 1375-1379.
- [42] Zhou, Howard, et al. "Feature-preserving artifact removal from dermoscopy images." *Medical Imaging 2008: Image Processing*. Vol. 6914. International Society for Optics and Photonics, 2008.

- [43] Kharazmi, Pegah, et al. "A Computer-Aided Decision Support System for Detection and Localization of Cutaneous Vasculature in Dermoscopy Images Via Deep Feature Learning." *Journal of medical systems* 42.2 (2018): 33.
- [44] Li, Haoxiang, et al. "A convolutional neural network cascade for face detection." Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2015.
- [45] Ren, Shaoqing, et al. "Faster r-cnn: Towards real-time object detection with region proposal networks." Advances in neural information processing systems. 2015.
- [46] Ciresan, Dan C., et al. "Flexible, high performance convolutional neural networks for image classification." IJCAI Proceedings-International Joint Conference on Artificial Intelligence. Vol. 22. No. 1. 2011.
- [47] Esteva, Andre, et al. "Dermatologist-level classification of skin cancer with deep neural networks." *Nature* 542.7639 (2017): 115.
- [48] Spanhol, Fabio Alexandre, et al. "Breast cancer histopathological image classification using convolutional neural networks." Neural Networks (IJCNN), 2016 International Joint Conference on. IEEE, 2016.
- [49] Germain, Mathieu, et al. "Made: Masked autoencoder for distribution estimation." International Conference on Machine Learning. 2015.
- [50] Vieira, Sandra, Walter HL Pinaya, and Andrea Mechelli. "Using deep learning to investigate the neuroimaging correlates of psychiatric and neurological disorders: Methods and applications." *Neuroscience & Biobehavioral Reviews* 74 (2017): 58-75.
- [51] Litjens, Geert, et al. "A survey on deep learning in medical image analysis." *Medical image analysis* 42 (2017): 60-88.
- [52] Xu, Jun, et al. "Stacked sparse autoencoder (SSAE) for nuclei detection on breast cancer histopathology images." *IEEE transactions on medical imaging* 35.1 (2016): 119-130.