creative commons
COMMONS DEED

저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:

저작자표시. 귀하는 원저작자를 표시하여야 합니다.

비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.

변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 이용허락규약(Legal Code)을 이해하기 쉽게 요약한 것입니다.

Disclaimer

dCollection

[UCI]I804:31001-200000108362

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

Master's Thesis

Deep Learning based Classification with Bio-medical Data

Thanh Quoc Phan

Department of Electrical Engineering

Graduate School of UNIST

2018

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

Deep Learning based Classification with Biomedical Data

Thanh Quoc Phan

Department of Electrical Engineering

Graduate School of UNIST

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

Deep Learning based Classification with Bio-medical Data

A thesis/dissertation
submitted to the Graduate School of UNIST
in partial fulfillment of the
requirements for the degree of
Master of Science

Thanh Quoc Phan

06 / 14 / 2018
Approved by

_____

Advisor
Se Young Chun

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

Thesis/Dissertation Title

Thanh Quoc Phan

This certifies that the thesis/dissertation of Thanh Quoc Phan is approved.

14 June 2018

_____

Advisor: Se Young Chun

_____

Jae-Young Sim: Thesis Committee Member #1

_____

Sung-Phil Kim: Thesis Committee Member #2

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

Acknowledgements

This thesis is the achievement for a long period of my studying and learning efforts. This result is absolutely the allegory of love and encouragement of my surrounding people; without them, I am nothing and have nothing. First of all, I would like to thank my parents, my friends, and my labmates. They are people always help me and support when I fail and in difficult circumstances.

I would also like to express my acknowledgment to my thesis committee: Professor Jae Young Sim and Professor Sung-Phil Kim for their kindness as well as attention. The appearances of Professor Jae and Professor Kim are really meaningful to me and help my thesis to become more valuable.

Finally, I greatly appreciate to my advisor Professor Se Young Chun for his motivation, patience, and patronage. His guidance is the skillful knowledge for my foreseeing life and career.

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

Abstract

Deep learning (DL) is one of the state-of-the art techniques in practical and theoretical perspective. As a significant method to extract features, deep learning is applied to many fields like computer vision, data analysis, and others. Convolutional Neural Network (CNN), which is the popular DL architectures, has been successfully providing excellent features for some computer vision tasks. Based on a large volume and diversity of database, deep learning figure out key features and solve problems on classification, recognition, or reconstruction.

Moreover, designing deep learning architectures as well as optimizers are popular and principal tendencies for enhancing the efficiency and performance of deep learning systems. In this project, we aim to solve one of the problematic issues of deep learning - the imbalance database. Usually, large medical databases with high quality and quantity are difficult to collect and be publicly available. Hence, some techniques were investigated as alternative methods to strengthen deep learning applications which are restricted and have a low performance when training database is not ideal such as using non-medical database for chest pathology detection training [5],visual features map [26], neuroimaging data model [35], and others. Thus, in this project, we proposed two methods to analyze two specific well-liked kinds of medical database: Skin cancer database and electrocardiogram signal.

To understand what good features are for skin cancer classification, we investigated image features of melanoma and benign images using VGG-19 that is a CNN model trained with the general ImageNet [50] database. Skin cancer data with labels (900 images) were obtained from 2016 IEEE ISBI challenge about Skin Lesion. These images were fed into VGG-19 [26] and each image yielded a vector with 1000 elements that are corresponding to classes in. These feature vectors were clustered into 8 categories using k-means algorithm with 3 different distance metrics: Euclidean distance, correlation, and cityblock distance (l1 norm). K-means algorithm with Euclidean distance yielded concentrated clusters that do not have any differentiation power. K-means with other metrics yielded histograms of clusters that are well-spread. Among them, k-means with cityblock yielded the best consistent images

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

for each category (e.g. images with circular patch or pale orange background). Most of these clusters contain both melanoma and benign images, but one category contains mostly benign images (about 100 images). This implies that noncancer related background information may affect CNN for melanoma detection when this database is used without care.

Additionally, we also proposed a specific method to solve the imbalance issue for signal classification. Biometrics due to the electrocardiogram(ECG) is one of the most promising techniques to substitute for classical methods such as iris, fingerprint, hand geometry, voice, or face. Though, the variations of ECG affect a lot on the performance of user identification. These variations derive from many multiple states (e.g exercises, stress, tension) as well as physiology changes of the heart. Thus, the proposed method is an augmentation method for QT interval correction (QTc augmentation). Aim to evaluate our investigated technique, we apply the model to publicly available ECG-ID database. The ECD-ID database with QTc augmentation is identified by several machine learning methods such as SVM with different kernels parameters, deep learning architectures (e.g RNN, CNN). From that, we evaluated the different performances of classifier techniques with and without QTc augmentation.

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

Contents

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

List of Figures

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

CHAPTER I

Introduction

Nowadays, deep learning has become a powerful technique when it surpassed the human level to solve many complex problems for numerous fields . It consists of many perceptron layers to form the deep neural network. In data processing, especially medical image and signal processing, this type of architecture can learn to recognize the features of the image hierarchically. The effectiveness of deep neural network, however, depends on the model architecture and the data to train it. Having a significant role for the performance of deep learning, the quality and quantity of database always affect on the generalization and accuracy of deep learning practicability. One of most challenging problem classes is the biased data problems. These problems that are very common on actual database.For examples, are when the number of data instances is distributed unequally on data labels or classes, data contains meaningless features, or ineffective database. In classification by deep learning, as a consequence, the performance of the class that has an abundance of database samples usually is biased while the accuracy of classifier on minority classes is very low. The volume of minority samples has the more important role than majority classes on the performance of deep learning classifier. To solve out this problem, many methods have been proposed such as low-level features feeding for CNN [56], random oversampling, adaptive sampling, and class-sensitive Accuracy [44]. In this work, we concentrated on solving the biased data of two kinds of medical databases: skin cancer ISIC database and electrocardiogram.

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

Skin cancer early detection is an effective method assisting to help the treatment more effectively. Due to a plenty of methods, people researchers are studying to explore a most specific method with high accuracy on skin cancer analysis. By using spectroscopic reflectometry as well as based on the characteristics of cell-nuclear size, oxyhemoglobin, and deoxyhemoglobin, Garcia-Uribe [19] could classify cancer cases from many kinds of skin cancer lesions. The experiments from this article achieved high accuracy with a collection of sophisticated databases. Have achieved 91.6 percent for skin cancer classification challenge in ISBI 2016, Gutman [22] ultilized simple linear iterative clustering algorithm [2] in order to figure out skin cancer features. Besides that, clinical database, which is considered as a simple and inexpensive database from digital camera, have been served for skin cancer classification by deep learning. This investigation was published by Nasr-Esfahani [29].With the CNN model, in this research, we utilize the VGG-19, which achieved the high rankings in the ILSVRC-2014 challenge, as the main model. In this paper, the VGG-19 that was pre-trained by ImageNet is applied to classify dermatologic images from the ISBI 2016 challenge for melanoma detection. This model was built up to analyze the tons of features for 1000 categories from 1.2 million images and applied to solve the classification and detection such as Texture synthesis [20], Facial Landmark Detection [34], etc.

Melanoma detection is one of the most interesting fields of cancer detection when the occurrence of cutaneous melanoma has risen every year, and the early detection demand is required. At this time, there were many methods and models to solve this problem such as artificial neural networks [36], decision tree [10], or Nestor Development System (NDS) [16]. On this paper, we used the VGG-19 to detect the skin cancer lesions. The skin cancer training data is a part of database of International Skin Imaging Collaboration (ISIC). It contains 900 dermoscopic skin images with their labels, which have been vetted by recognized melanoma experts. However, on these images, there are some unnecessary objects such as: circular patch or pale orange background. Although, in general, CNN, particularly VGG-19, can automatically analyze and extract features of data by their specific layers, the quality of dataset still has the significant role to improve the generalized function for output model. Because of the reason, we investigated the model to detect the biased features as well as invalid images of training dataset by k-means clustering.

Electrocardiogram has been investigated as a potential candidate for biometrics[29]. Each ECG pulse, which records the electricity signal of heart activity, contains a P wave, a QRS complex , and T wave. They are corresponding to three heart activity steps in one circle: atrial depolarization, ventricular depolarization, and ventricular repolarization. Due to specific

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

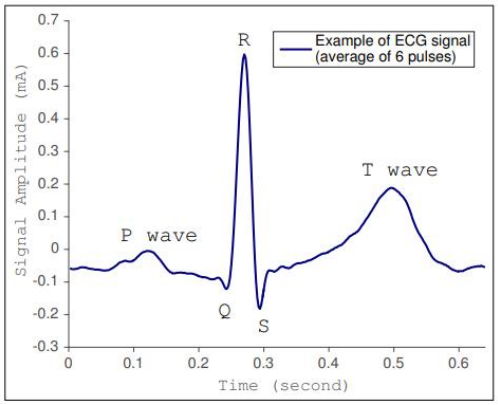

Figure 1.1: An ECG signal structure [13].

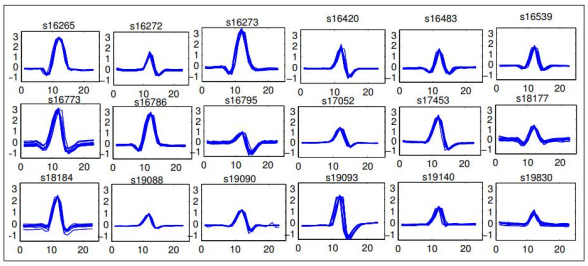

Figure 1.2: Specific QRS complexes of subjects [40].

characteristics of heart structure and circulatory system, ECG pulse of each person is different to of other people and even on various circumstances such as: alternative ECG device, distinctive recording days, or surrounding environment, ECG pulse is also changed.

To employ ECG signals as a powerful target on biometrics,excepting traditional features, Safie et al [51] investigated a new framework to extract a new feature which was called pulse active ratio. This method yielded to a greater performance than ECG processing by other features. Vu et al [40] utilized QRS complexes, which is evaluated as the most robust specific feature as Figure 1.2, to classify the user's identifications.

With the astonishing development trend, deep learning is also a candidate technique to solve the ECG biometric topic. A combination of Deep Belief Networks and Restricted Boltzman Machines showed an impressive performance when this process achieved a high accuracy

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

96.1 percent [30]. Comparing to other techniques, Neural Network has shown an outstanding ability on figuring data features adaptively and efficiently [58]. Furthermore, one of the most strength aspects of deep learning is analyzing a huge volume of data. Via a deep neural network with a big size of database, Pranav et al was successful in mapping ECG sequences to arrhythmia signals.

In this study, we have experimented with many kinds of machine learning techniques as well as efficient deep learning models to solve the ECG identification problem. In this time, though, there are several works on ECG authentication as well as classification. Many of them have achieved very good performances [45, 46, 53] . However, in some mentioned challenging situations, it is really difficult to solve out. Thus, not only applying deep learning to ECG classification, but in this article, we also apply a new investigation to increase the ECG database which usually is constrained. This is the augmentation method that analyses active QT interval correction to improve number of effective features of ECG database at different hearth rates.
In this thesis, chapter II describes the background of two powerful DL architectures. Next, chapter III reviews some previous works on skin cancer classification and discusses the proposed framework of our method as well as results. Then, chapter IV represents a literature review of ECG processing for identification, proposes the Qtc augmentation for classification. Finally, chapter V concludes this study with a summary and a general evaluation.

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

CHAPTER II

Background

In this study, we concentrated on how to improve performances of different kinds of deep learning architectures. The first one is convolutional neural network, which is very popular and effective machine learning technique as well as deep learning method. In general, CNN is familiar to the neural network; however, it requires more sophisticated unit activations than the regular neural network. The second one is the recurrent neural network(RNN) that is utilized commonly with sequential data with three specific kinds of RNN cells: traditional RNN cell, gated recurrent units (GRUs) cell, and long short-term memory(LSTM) cell. These kinds of RNN model were implemented to experiment on the ECG identification task.

2.1 Convolutional Neural Network

Convolutional neural network is created of neurons that have learnable kernels and biases. These neurons analyze their inputs, perform a scalar product and optionally with a non-linearity. There are three main kinds of layers to construct a ConvNet architectures that are sequences of kernels: Convolutional Layer, Pooling Layer, and Fully-Connected Layer(F). Besides that, CNN model was evaluated and strengthened by optimizing a loss function on the last layer which is normally a fully connected layer as an ouput layer. Generally, the Convnet's function is transforming the image volume or the signal (1D input) into an output of class scores.

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

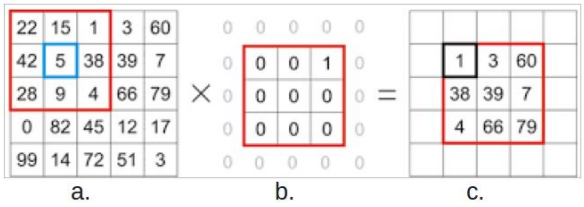

Figure 2.1: Convolutional layer. a. Input matrix, b. Kernel filter, c. Ouput matrix.

2.1.1 Convolutional Layer

Convolutional layers is the primary component of ConvNet to extract features of input and be a mainly computational element of a model. In practical, a convolutional layer performs as a set of 2D filter in certain conditions and extends through the volume of the input data. The weighted outputs of pixels are calculated as a linear function and a new neuron is obtained.

$$f = W^T X + B \tag{II.1}$$

By convolving input signal x with a band of K filters f and biases b, we could obtain an output signal y:

$$y_{i''j''d''} = b_{d''} + \sum_{i'=1}^{H'} \sum_{j'=1}^{W'} \sum_{d'=1}^{D'} + f_{i'j'd'} \times x_{i''+i'-1,j''+j'-1,d',d''} \tag{II.2}$$

where H : height, W: the width, D: the depth dimension. The characteristics of each convolutional layer are depended on the three hyperparameters which determine the the output volume: the depth, padding numbers, and stride.

- The depth hyperparameters is the number of filters as the dimension of the convolution layer. This parameter is also the depth of the output volume. (Figure 2)
- Sometimes, it is necessary to add several zeros around the border to pad the input data. The numbers of this zero-padding are advantageous for setting the spatial final volume of convolutional layer.
- When we slide kernel filters on input matrix, the stride is the step of the convolutional layer. This will produce smaller output volumes. (Figure 3)

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

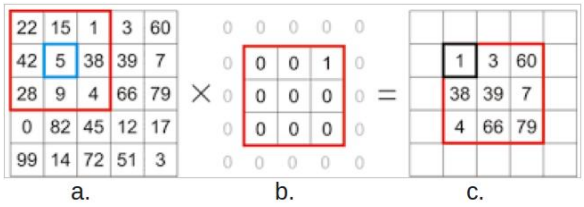

Figure 2.2: Ouput image depth [14]

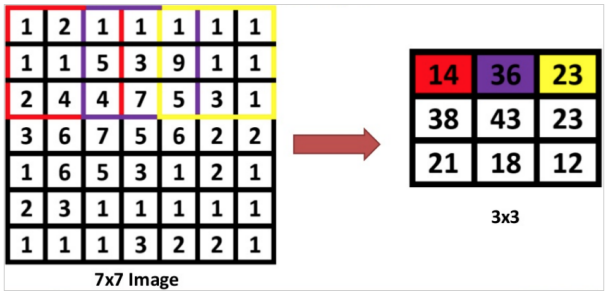

Figure 2.3: A 3x3 convolutional layer with stride = 2

In particular, the output volume of neuron is measured by the equation:

$$Ouput = \frac{I - F + 2Z}{T} + 1 \tag{II.3}$$

with the I is the input size, the F is size of convolutional layer, Z is the zero-padding number, and T is tride option. However, the output of convolutional layer is a result of a linear function, we need an activation function to avoid this phenomenon. Hence, Retified Linear activation function (ReLU) is utilized as an important factor for ConvNet. Without ReLU, the network would aim to be a linear classifier. The predominantly used ReLU activation function in almost CNN models with convolution layers defined as the position elements of its argument.

$$f(x) = max(0, x) \tag{II.4}$$

where x is the ouput of kernel size apply to ReLU .

2.1.2 Max Pooling Layer

The results, after convolving by the bands of filters, are the significant features which have been extracted from the input neuron. Though, the deep Convnet with many filters accumulates a large amount of information, this problem affects on the speed of the networks optimizer. Because of this reason, the Max Pooling layer becomes a powerful tool to reduce the feature

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

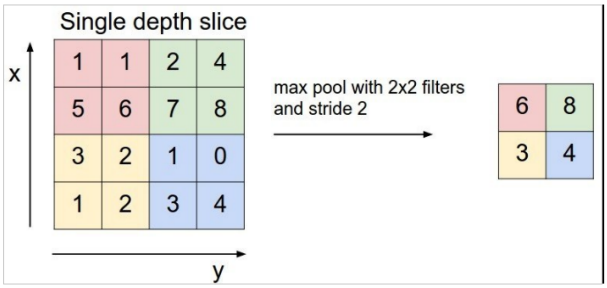

Figure 2.4: Max Pooling layer with 2x2 filter. Source: Stanford's CS231n course

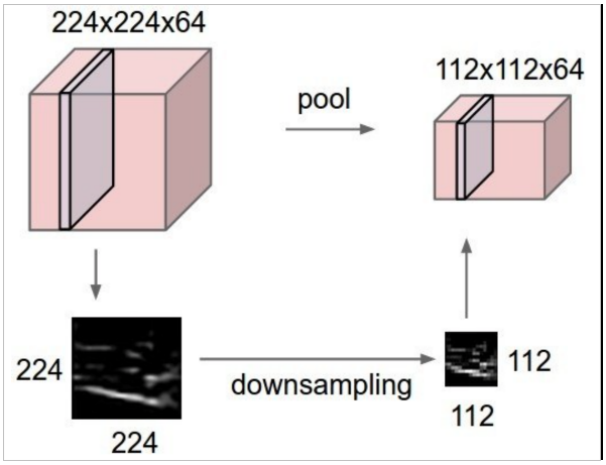

Figure 2.5: The max pooling layer extracts the significant feature. Source: Stanford's CS231n course.

space from convolution layers. For a layer with the size WxH, we have the formulation [26].

$$y_{i''j''d''} = max_{1 \leqslant i' \leqslant H', 1 \leqslant j' \leqslant W'} \chi_{i''+i'-1, j''+j'-1, d} \tag{II.5}$$

For max pooling layer, there are two main hyperparameters: filter size and stride. Figure 4 is a max pooling layer 2 by 2 filter size and stride 2. In the real case, this kind of layer is a filter that is utilized to deflate size of the output of convolutional neuron and just extract the remarkable information as Figure 2.5.

2.1.3 Fully Connected Layer

In particular, the fully connected layer, which is final kernel, is a layer of regular neural network. Each activation of FC layer connected all activations of the last layer. For classification problems, after using the fully connected layers, we need to squash the output to between 0 and 1 which are equivalent to a probability distribution. This represents these probabilities as the

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

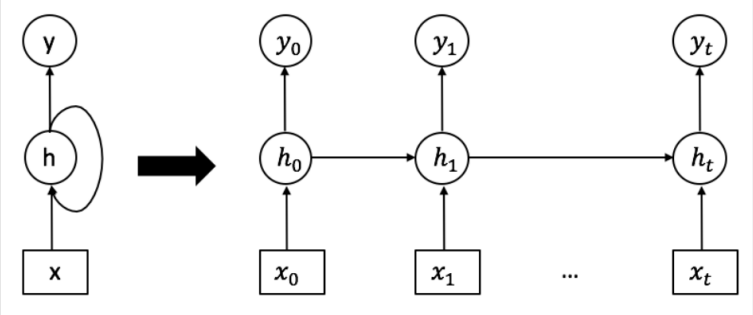

Figure 2.6: Recurrent Neural Network

classified label outcomes. The softmax equation is shown [35].

$$\sigma(z)_j = \frac{e^{z_j}}{\sum_{k=1}^{K} e^{z_k}} \tag{II.6}$$

for $j = 1,...,K$, with K is the size of input volume from the final FC layer.

2.2 Recurrent Neural Network

Originally, recurrent neural network is an artificial neural network method to solve the sequential data prediction problem. This is a kind of network that is created as a guided graph along a sequence by links among nodes. A simple recurrent neural network consists of a input layer x which is a variable length sequence $x = (x_1, ..., x_T)$,a hidden layer h, and a output layer y. The input data of RNN is a vector of information for a time sequence, at each time t, the input to the network is $x(t)$, the state of the network is $h(t)$, the output is defined as $y(t)$.

The input layers, hidden layers, and output layers are formulated [42].

$$x(t) = w(t) + h(t-1) \tag{II.7}$$

$$h(t) = f(\sum_i x_i(t) u_{ij}) \tag{II.8}$$

with $f(s) = \frac{1}{1+e^{-z}}$ is a sigmoid activation function.

$$y_k(t) = g(\sum_j s_j(t) v_{kj}) \tag{II.9}$$

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

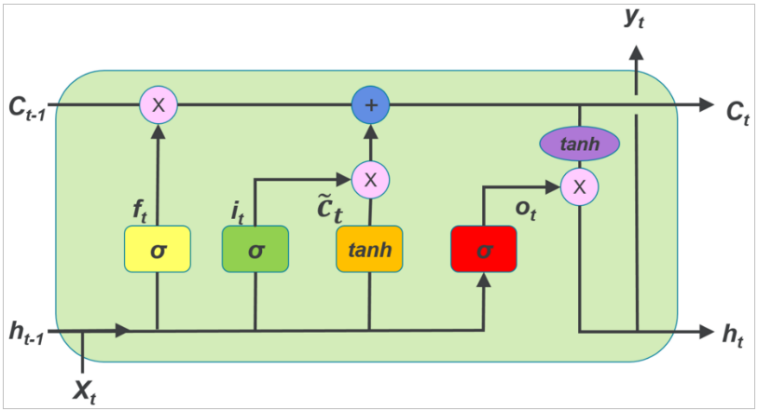

Figure 2.7: Long Short-term Memory Structure [60]

with $g(z)$ is a softmax function as II.6. The output layer is a softmax function [12].

$$p(x_{t,j}|x_{t-1}, ..., x_1) = \frac{e^{w_j h_t}}{\sum_{j=1}^{K} e^{w_j h_t}} \tag{II.10}$$

for $j = 1, , K$ and w_j are the rows of weight matrix.

2.2.1 Long short-term memory

In RNN, the hidden layer determines the characteristics of the network. With different kinds of hidden layer, RNN shows an alternative specific performance. Besides that, RNN against a problem with its memory to understand a sequence which contains too much irrelevant information. The reason of this problem is of Vanishing Gradient. One of the most popular hidden unit for RNN layer to solve this kind of problem is long short-term memory (LSTM) which was proposed by Sepp Hochreiter [25]in 1997. The RNN model using LSTM as the main unit for hidden layers is called as an LSTM model.

In general, a LSTM unit is included an input gate, an output gate, and a forget gate [60]. First of all, the LSTM unit ultilize the forget gate $f(t)$, which is defined as a sigmoid function, to filter input information $x(t)$ and output of previous LSTM unit . The outcome of forget gate is a probability value between 0 and 1 for each number in the cell state $C(t-1)$

$$f(t) = \sigma(W_f \cdot [h(t-1), x(t)] + b_f) \tag{II.11}$$

$$i(t) = \sigma(W_i \cdot [h(t-1), x(t)] + b_i) \tag{II.12}$$

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

$$\widetilde{c}(t) = tanh(W_i \cdot [h(t-1), x(t)] + b_f) \tag{II.13}$$

Next, to determine what new information should be stored in the cell C. There are two steps. The first step is activating a sigmoid input gate II.12 to consider what information to update. Subsequently, on the second step, a tanh function II.13 is applied to create a vector of potential information for the cell C updating.

Then, the new cell state $c(t)$ of this LSTM unit will be updated by the cell state $c(t-1)$ of the last LSTM cell based on $f(t)$, $i(t)$, and $\widetilde{c}(t)$. It follows the formulation below.

$$c(t) = f(t) * c(t-1) + i(t) * \widetilde{c}(t) \tag{II.14}$$

Finally, we need two more parts to decide the value for the output gate. Firstly, a output gates is used to analyse the input signals as first step. After that, another tanh function is applied to cell state c(t) and multiplies by o(t) to output the appropriate content.

$$i(t) = \sigma(W_o \cdot [h(t-1), x(t)] + b_o) \tag{II.15}$$

$$h(t) = o(t) * tanh(c(t)) \tag{II.16}$$

In each gate, there are the parameterss weigth W and bias b, correspondingly. In general, the LSTM based on these gates could evaluate and consider about the meaningful information to store. This assists to the RNN model to avoid the vanishing gradient. From that, it is improving the performance by higher accuracy and faster optimization.

2.2.2 Gated Recurrent Unit

In 2014, Kyunghyun Cho [12] proposed a new method to overcome the vanishing gradient problem which is a drawback of the traditional RNN. In fact, the GRU cell has a similar design with LSTM cell unless it has fewer parameters as well as the absence of an output gate. However, this kind of hidden layer shown a more positive performance on small dataset. The GRU cell consists of two main parts: update gate and reset gate. Both of these gates provide for RNN model with an ability to analyze what information should be kept in reserve. Comparing to LSTM cell, this hidden layer merges the cell state and hidden state. Additionally, while LSTM activates a forget gate and an input gate, GRUs combines them to be an update gate which is also a sigmoid function.

$$z(t) = \sigma(W_z \cdot [h(t-1), x(t)]) \tag{II.17}$$

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

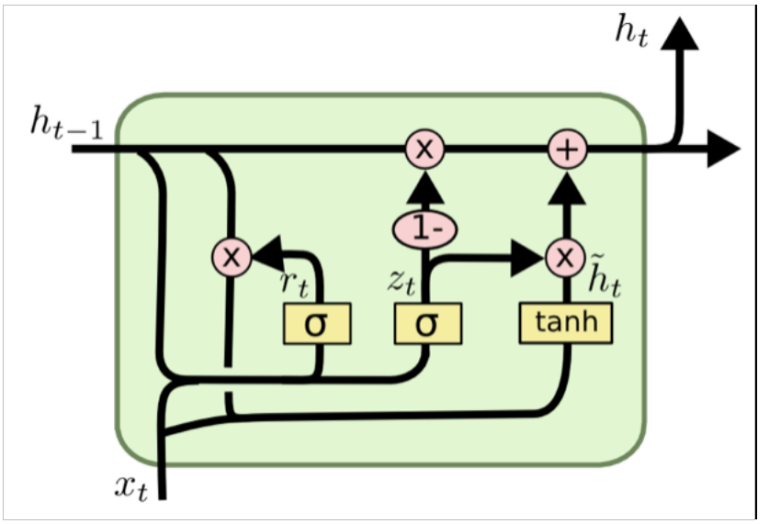

Figure 2.8: GRU Structure. Source: Wikipedia

The function of LSTMs forget gate is now performed on the reset gate of GRU.

$$r(t) = \sigma(W_z \cdot [h(t-1), x(t)]) \tag{II.18}$$

After using a reset gate to figure out the relevant information from the input and the last hidden layer, a tanh function is called to store these information.

$$\widetilde{h}(t) = tanh(W_i \cdot [h(t-1), x(t)]) \tag{II.19}$$

The final step is combining the result from the reset gate and the update gate to output them to hidden state and cell outcomes.

$$h(t) = (1 - z(t)) * h(t-1) + z(t) * \widetilde{h}(t) \tag{II.20}$$

Using gating mechanism as LSTM, GRUs are able to filter irrelevant information by using update and re set gates. This shows that GRU cell is a great way to eliminate the vanishing gradient problem for RNN.

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

CHAPTER III

Towards Good Features for Skin Cancer Classification

3.1 Related work

A several years ago, malignant melanoma has became one of the most appealing fields of cancer detection. It could be due to the rapidly increasing of skin cancer around the world, especially United State. According to the statistics of Siegel RL et al. [54], the estimated new cases and estimated deaths of skin cancer in US mutatively rise to 76,380 and 10,130 in 2016. Hence, early detection is required as an advantageous way to assist doctors on diagnosis. In efforts to solve out this problem, A.Estela et al [17] evaluated the performance of a DL architecture that was the state-of-the-art architecture at that time - the GoogleNet Inception v3 CNN architecture and compare to the performance of dermatologists. Transfer learning with ImageNet was applied to improve the performance of model when their clinical dataset is 129,450 images that is significantly smaller than ImageNet. Moreover, Ulzii-Orshikh Dorj et al [15] incorporated the convolutional neural network Alexnet with ECOC SVM. Using CNN features from Alexnet which is suitable for small size dataset images, ECOC SVM classified skin cancer images into several groups.

On the other hand, many authors have used CNN segmentation to extract the skin lesions before classifying them. We could believe that this is the way to avoid the meaningless informa-

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

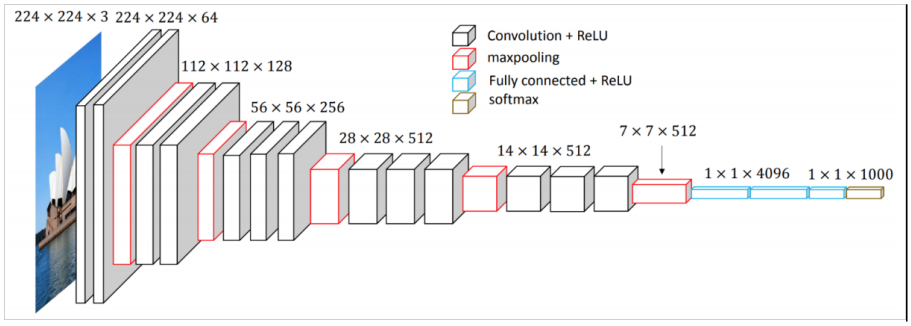

Figure 3.1: VGG-19 Architecture [3]

tion. To be comparable to classify without segmentation tasks, the performance showed more efficient as [9]. For skin lesion segmentation, there are many deep-learning-based techniques that have achieved significant performances such as [29], [23], and [57]. Besides that, there have been recent works to figure out effective features from skin cancer database. Bhuiyan et al [7] has proposed that Otsus method, Gradient Vector Flow, Color Based Image Segmentation Using K-mean Clustering are the beneficial methods for features extraction approaching to skin cancer analysis. Simple Adaptive Thresholding Algorithm [31] was shown an incredible performance on skin cancer texture detection.

3.2 Methodology

3.2.1 The proposed approach

On this paper, the convolutional network (ConvNet) is the deep neural network which is up to 19 layers and feasible to use the small convolutional layers, VGG-19.

With the one of the best performances on the challenge of ImageNet, we applied a pre-trained VGG-19 model with ImageNet to our experiments. Following the procedure:

First of all, the ISIC skin cancer training database was applied through the VGG-19 model which was trained by ImageNet to test our skin lesion images. The output of this step is the 900x1000 matrix that indicates the correspondence between skin lesions features and 1000 ImageNet objects features. Then, the matrix was partitioned into categories by k-mean clustering. On our experiments, k-means clustering algorithm with different distance measures presents clusters the corresponding matrix of training database and ImageNet to eight categories. Based on the centroids locations, the variant category of cityblock histogram was interpreted to check their corresponding ImageNet objects and contains almost cicular patches and pale orange background images of ISCI skin cancer database. The obstacles as the colourful circular patches in Fig 3.4 do not have the specific features of skin lesions. Their existence provides the noise

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

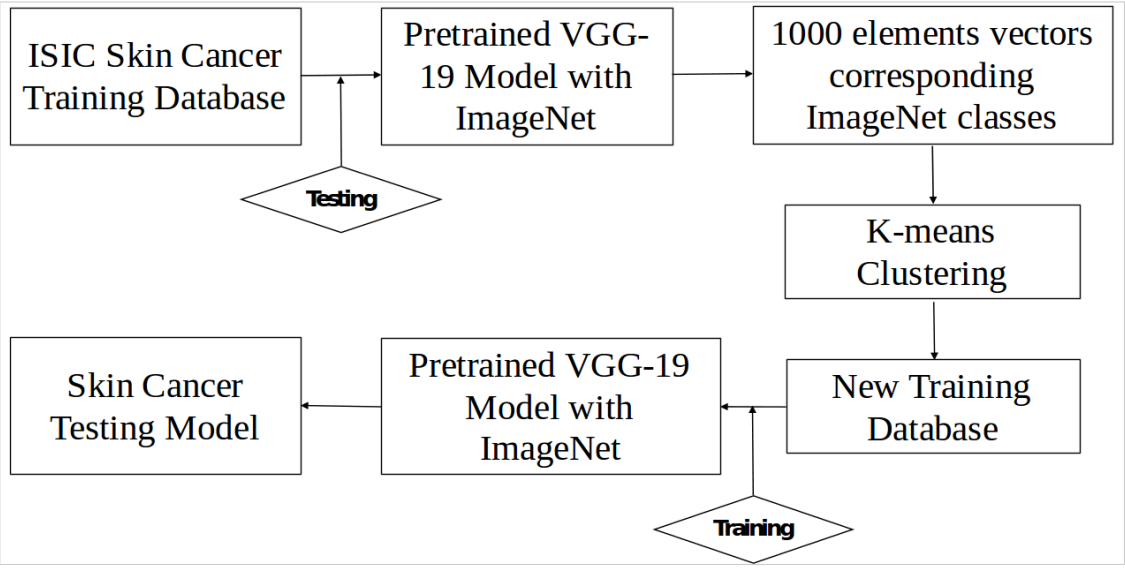

Figure 3.2: Experiment Procedure

features for our deep learning models. Because of the reason, we need to detect and remove the images of the variant category.

3.2.2 K-means clustering algorithm

Our proposed approach is based on k-mean clustering technique that is a principle algorithm of machine learning on unsupervised learning. This method follows a simple way to classify unlabel data to separate groups or clusters. The preeminent idea of this algorithm is to detect the collection of points that is assigned to a center of cluster(centroid). Each centroid is updated based on the assigned points via optimizing the loss function L.

$$L(x, v) = \sum_{i=1}^{c} \sum_{c_i}^{j=1} (\|x_i - v_j\|)^2 \tag{III.1}$$

where j is the volume database, i is the number of centroids or labels, c_i is quantity of data point in i_th centroid, x is data points, and v is the cluster center.

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

Algorithm 1: K-means Clustering

Input: Data $X = (x_1, x_2, ..., x_n)$ is the set of data

Output: Centroids $V = (v_1, v_2, ..., v_n)$ is the set of data

```
Step 1: Randomly choose $c$ centroids.
Step 2: Measure the distance between each data points $X$ and centroids.
Step 3: Assign the data point to the nearest centroids.
Step 4: Updating the centroids by the mean function.
Step 5: Recalculate the distance and updating the centroids.
Step 6: Until centroids do not change.
```

As function III.1 shown, the loss function utilizes the Euclidian function to calculate distance between data points and centroids. In this article, we evaluate the results of k-means clustering on several kinds of distance functions.

The cityblock $L_1 - norm$ distance :

$$d(x, c) = \sum_{i=1}^{n} (|x_j - c_j|) \tag{III.2}$$

The correlation distance:

$$d(x, c) = 1 - \frac{(x - \bar{x})(c - \bar{c})}{\sqrt{(x - \bar{x})(x - \bar{x})}\sqrt{(c - \bar{c})(c - \bar{c})}} \tag{III.3}$$

where

$\bar{x} = \frac{n}{1}(\sum_{i=1}^{n} x_i \bar{1_p})$

$\bar{c} = \frac{n}{1}(\sum_{i=1}^{n} c_i \bar{1_p})$

$\bar{1}_p$ is a row vector of p vector

3.2.3 Experimental Results

While other metrics distribute histograms of clusters of benign and melanoma uniformly, cityblock distance ($l1 - norm$) presents that there is the special category which yields almost benign images with circular patches or pale orange background on this category (Fig 3.4a). This category is category 5 in Figure 3.4.

Almost images that accumulate cicular patches and pale orange background are on the variant category 5 and they predicted by three main objects: butternut squash, toilet seat, and plectrum.

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

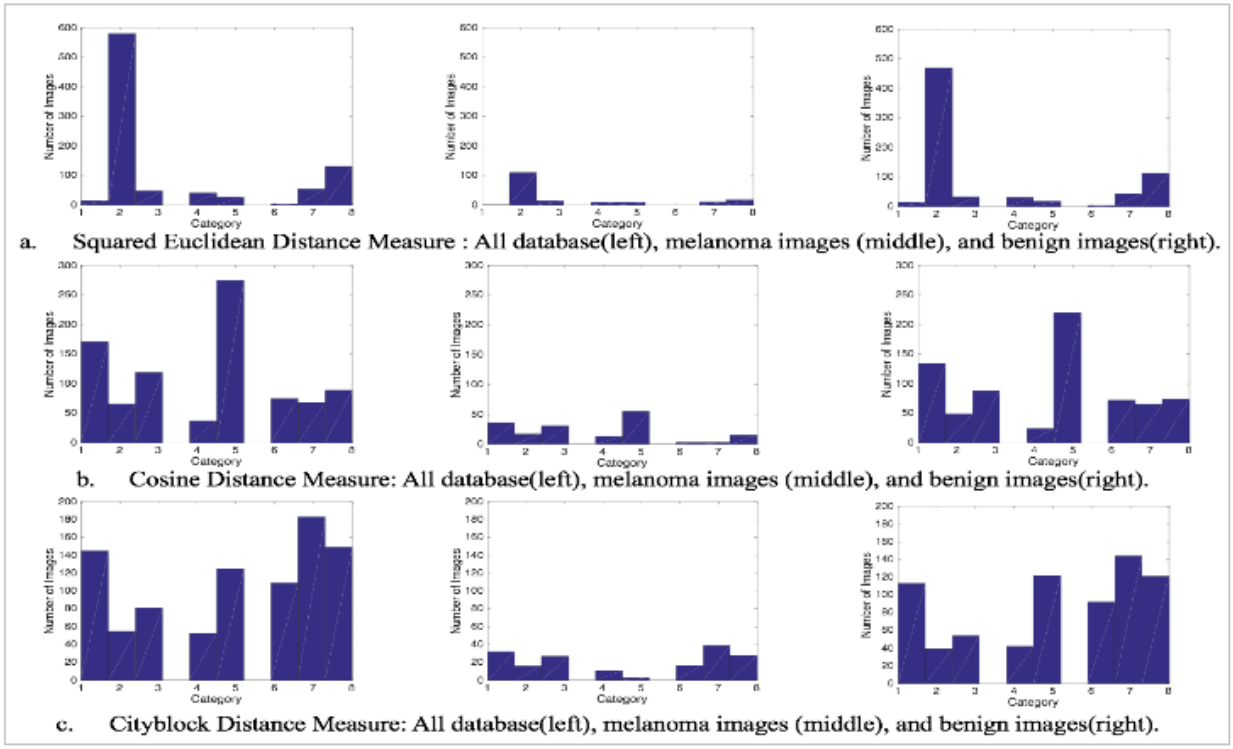

Figure 3.3: K-means clustering Histograms

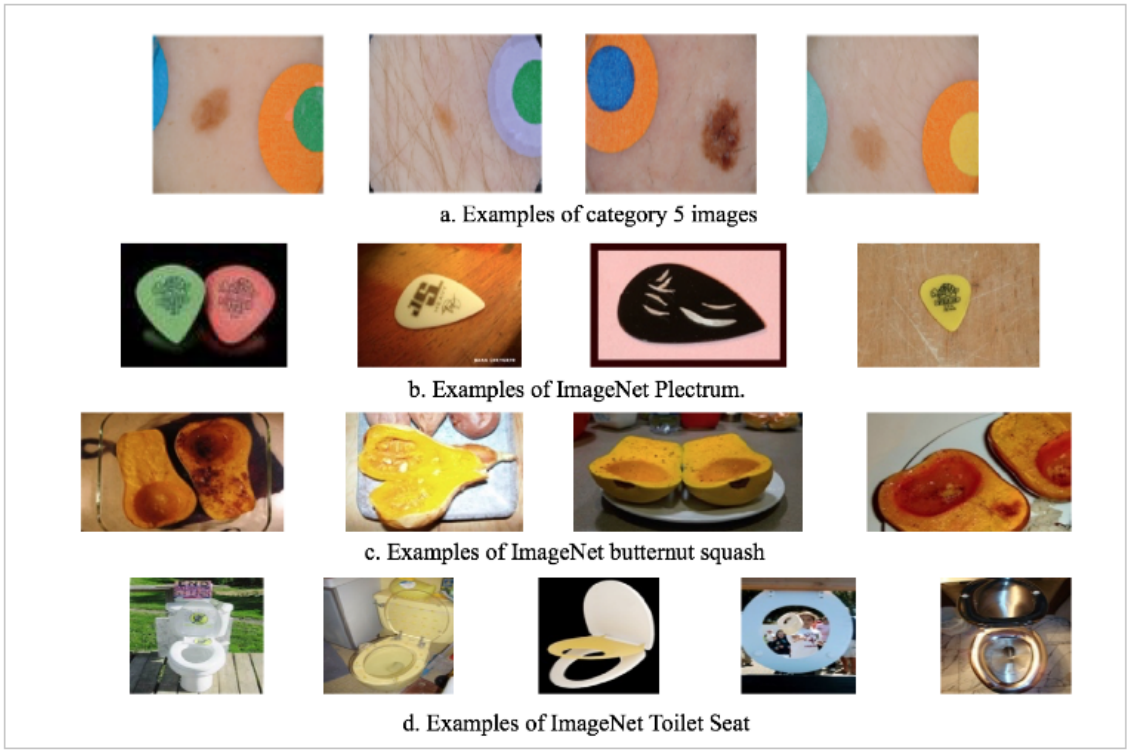

Figure 3.4: Representative images of category 5

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

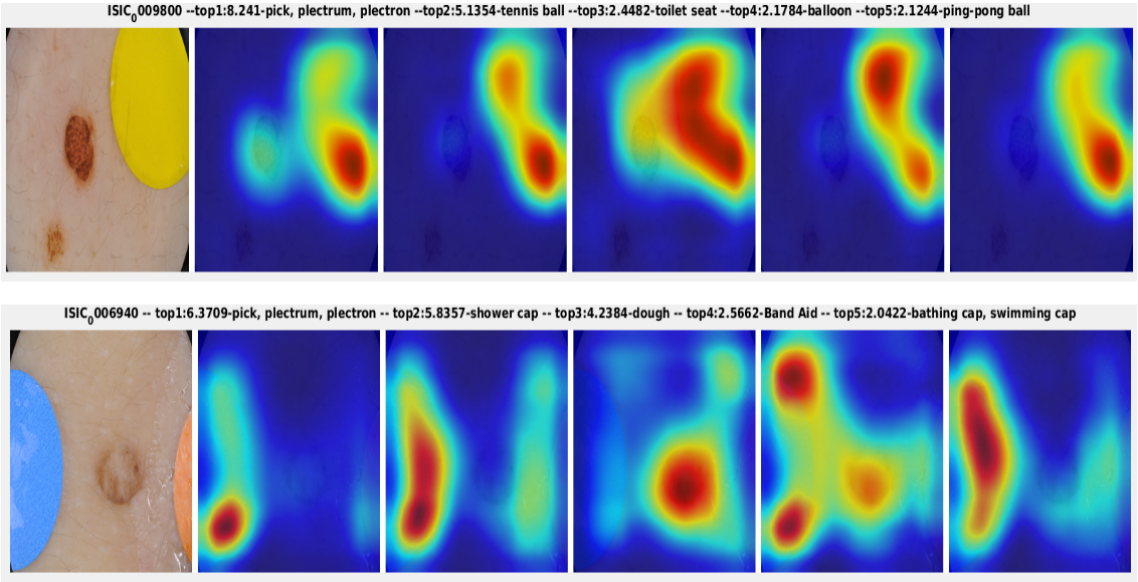

Figure 3.5: Examples for Heat Map of Skin Cancer Classification

Table III.1: The Performances of Testing Set with 379 images

Model	Accuracy	Sensitivity	Specificity	Average precision	Database Volume
Original Database	81.53%	97.36%	17.33%	47.8%	900 images
New Database	82.06%	98.02%	17.33%	51.06%	775 images

To figure out what are biased features of category 5, by using the discriminative localization method of Zhou et al [61], we recognize that the activity of the CNN model concentrates on the circular patches and pale background as Figure 3.5. The red areas are corresponding to unexpected objects. Specially, these areas do not detect skin lesions, their targets almost are the patches.

We evaluate the efficiency of our model by applying new training database to VGG-19 model. The database is trained by VGG-19 model with pre-trained ImageNet VGG-19 model. Compare to the original training database, the accuracy of validation test on training task is higher and the loss of training is more stable. The validation accuracy is 80 percent for the model of original training database, but increases to 84,4 percent for the model of new training database. The model of our new training database also shows the more positive performance than the model of primitive database as table III.1.

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

3.2.4 Discussion

In this section, we proposed the framework to analyze the training data for skin cancer classification. After training data with the powerful deep CNN model like pre-trained ImageNet VGG-19 and k-means clustering, we can detect and remove the invalid images of our database. Although, the new database has smaller volume, it improves the productivity for our training model.

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

CHAPTER IV

ECG Classification via Deep Learning

4.1 ECG signal models for biometrics.

Electrocardiogram has been investigated as a potential target for biometrics. However, the intra-personal varibility is the serious challenges when ECG is applied as a biometric objective. There are several works which have been shown that intra-personal variation has an negative impact on the verification and identification performance [27], [49], and [43]. Odinaka et al [43] presented the accuracy of verification with intra-personal ECG data is always lower than without intrapersonal one.

Recently, ECG biometric has been further investigated. Israel et al [28] by using fiducial features proposed an ECG-based recognition system for identification biometric via linear discriminant analysis(LDA). Biet et al [8] extracted 30 fiducial feature points from amplitude, duration, and deflection of QRS complex and achieved a 100 percent recognition rate for the test experiment, the soft independent modeling of a class analogy was applied for classification. Moreover, there are many further investigations using fiducial features such as beat duration, amplitudes, wave features for ECG biometric [55], [41], and [48] . Besides that, unlike fiducial approach-based studies, non-fiducial feature-based studies, which do not require precise information of ECG waveforms, also have been investigated for biometrics with the combination of Euclidian classifier and wavelet transform [11], linear predictive coding [37], linear discriminant analysis [4].

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

In this study, we proposed an evaluation survey for the performance of deep learning techniques on ECG identification when they against to the intra-personal problem of ECG. In this time, deep learning studies for ECG classification have been investigated, though, almost of these works are for the non-intrapersonal ECG dataset. Adam Page et al [45] proposed an embedded system using neural network classifier. The system could achieve a good performance with an EER (equal error rate) of 0.0582 percent. Palva et al [46] suggested a deep convolutional neural network with an ECG signal processing framework using R peaks normalization and distance measures. Another CNN model was presented by Zhang et al [59] as a core component of a multi-resolution identification system with segmented ECG signal by autocorrelation and wavelet transform.

In addition, we introduce an ECG augmentation method using active QT interval conrrection . In fact, QT interval correction have been studied as a normalization method and feature extraction [52], [6] ,and [18]. Lugovaya [38] proposed to use a QT interval correction to normalize ECG signal and concluded that T wave features could improve the performance of verification. This proposed method is able to be a promising way to augment ECG database and aim to increase the volume as well as the quantity of useful features.

4.2 Methodology

4.2.1 Deep Learning models for ECG biometrics

In this time, using deep learning for automatically signal detection is becoming a new tendency based on the outstanding characteristics of deep learning methods such as working on large volume of database, high accuracy, and creating new features. Although, in this project, our target is classifying the single pulse ECG database being collected by wearable devices, the attribute of ECG signal is sequential. Thus, in many kinds of deep learning methods, RNN is an expected technique which has the feedback connections that is suitable for sequential data processing. Salloun et al [53] showed that RNN model with few hidden layers could identify ECG-ID database effectively when they analyze three or nine ECG pulses simultaneously. The model is shown in Figure 4.1

The model was evaluated with different kinds of RNN hidden layers units: traditional, LSTM, and GRU. Besides that, their input sequence was selected as a group of continuous ECG pulses. This step is to ensure that the sequential features of signals are obtained. In fact, ECG data is a high noise signal that was mentioned above. So, single ECG pulse is more effortless to denoise than sequential 1-lead ECG signal or multiple lead signal even though amount of features is constrained. However, CNN model is still popular to handle the sequential data which is really

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

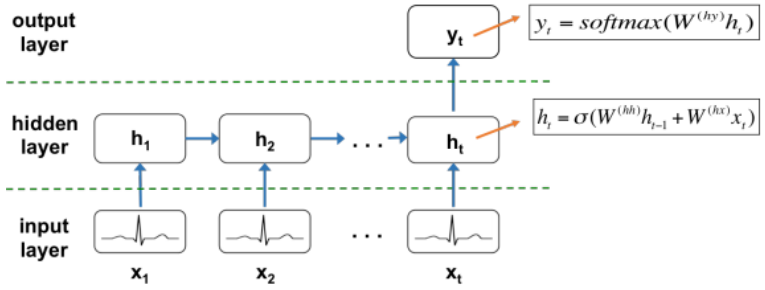

Figure 4.1: Block diagram of a traditional RNN by Salloun [53]

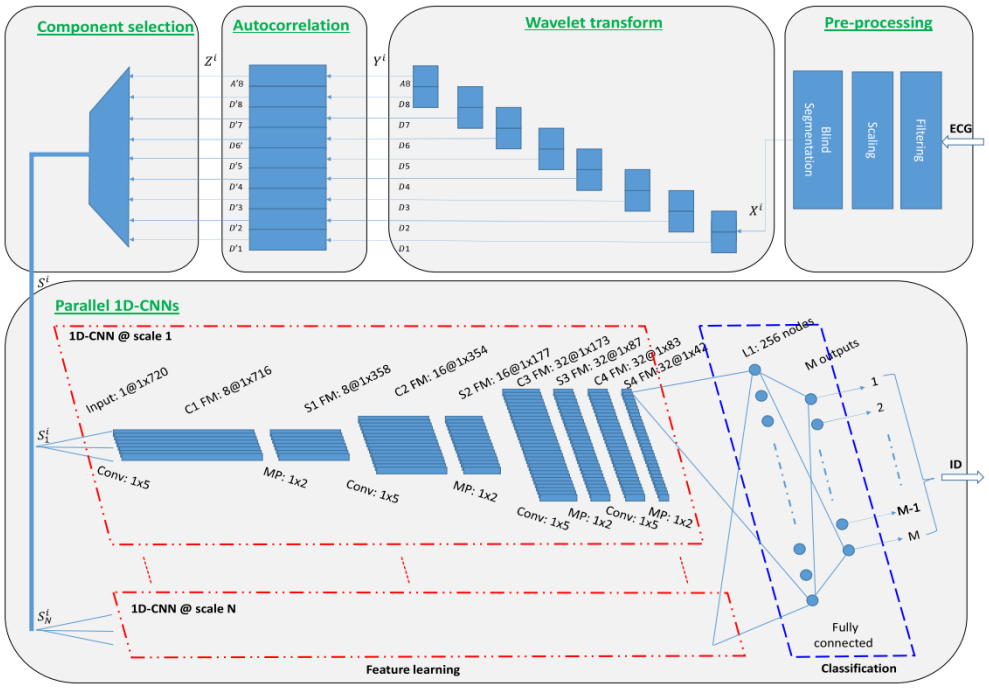

Figure 4.2: The system diagram of the HeartID CNN model of Zhang et al [59]

long with many features. These models are usually definitely deep and request a complicated pre-processing task. Of course, they could strengthen their models by cooperating to others machine learning techniques. HeartID [59] model is one of the specific examples. Published on 2017 by Zhang et al, HeartID is an ECG classifier for sequential database. It is not only a deep CNN model, but also a "wide" CNN model. Combining with wavelet transform and autocorrelation, their framework is able to improve the efficient of deep learning model when wavelet transform interprets the ECG signal on different time frequency patterns. Besides, auto-correlation has a function of filter and provides more data presentation.

In this study, aim to develop a DL model for only single ECG pulse which is fast on training and high accuracy, we proposed a simple CNN model on section IV.3.b and use the QTc augmentation to deal with ECG features at different heart rates. Furthermore, we also would

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

like to experiment the impact of the augmentation when apply QTc to each pulse of sequential data on classification by RNN.

4.2.2 QT correction model for ECG biometric

Heart rate variation has an important affect to the structure of the ECG signal. Israel et al [28] is one the first study interpreting the phenomenon of the invariance to variable heart rate. They experimented and proposed to normalize the features related to P and T waves. Besides that, they assumed that QRS complex is invariable under heart rate alteration. Similar to what Labati et al [33] has proposed, QRS complexes are effcient components for biometric. In their study, they extracted QRS complex as the prime feature for their Deep-ECG CNN model to handle the ECG identification. Although, a combination between a flexible ability of P and T waves and a stable QRS complex is a promising characteristic to advance the ECG biometric achievement.

The most prominent QT interval correction method is proposed by Bazett [6]:

$$QTcB = QT(RR)^{-0.5} \tag{IV.1}$$

That QT as a QT interval and RR as an interval between two R peaks. After many years, Framingham formulation, which is a linear bas model, was proposed [52]

$$QTcF = QT + 154(1 - RR) \tag{IV.2}$$

The Hodges's formula:

$$QTcF = QT + 105(\frac{1}{R} - 1) \tag{IV.3}$$

In 2004,S.Luo et al [39], after evaluating many of QTc formulas, has concluded that Hodges QT interval correction formula is the most appropriate method for a large scale ECG database which is essential and useful point to have an achieved deep learning classifier.

4.2.3 Guided for ECG biometric

Guided filter (GF) is the popular method in computer vision for many problems such as filtering, artifact analysis and data sampling [24]. It performs edge-preverving smoothing on a input signal , and using the information of another signal to conform the filter. Recently, Chun [13] proposed a method to utilize a 1D GF for ECG authentication. The user template guided filter employed the enrolled ECG guide signal t which is a template to filter a single

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

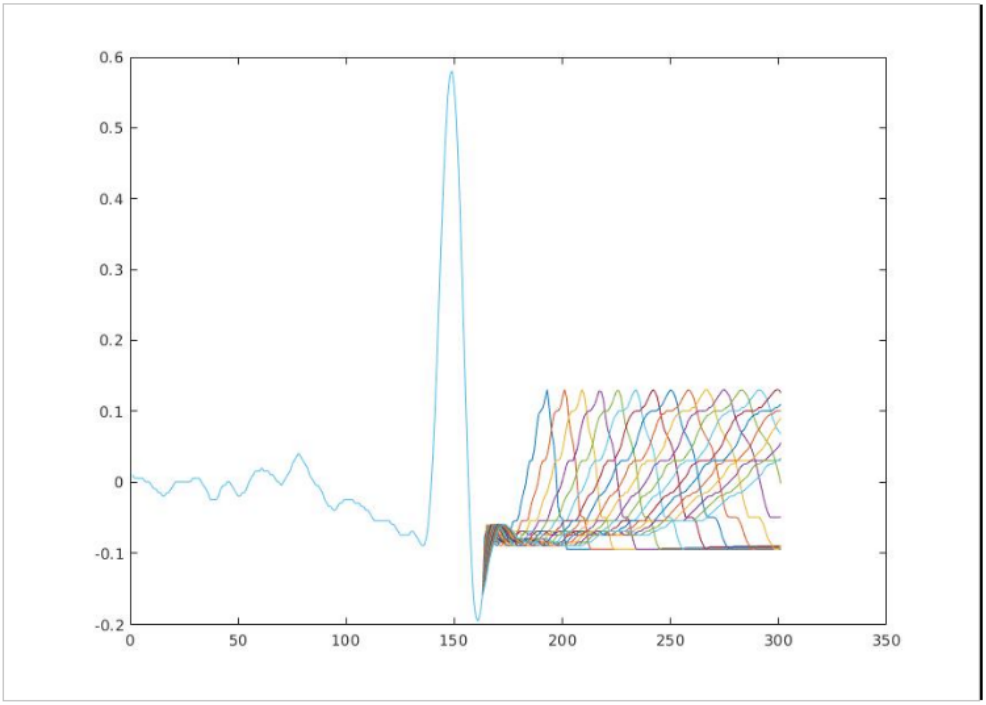

Figure 4.3: An example of single ECG pulse with QTc Augmentation.

ECG pulse. For a user template t as a guide signal with low noise and a noise signal g, we denote this GF as:

$$\hat{g} = GF(g; t) \tag{IV.4}$$

After guided filtering, the $\hat{g}$ is the new database for our experiments.

4.3 Results

4.3.1 Experiment Settings

The ECG recording used in this work is from the public Physionet dataset ECG-ID [21]. The database is a collection of ECG records from 89 healthy subjects using one-lead ECG sensor. Each raw ECG record was obtained for about 20 second with the sampling rate frequency 500Hz, 12-bit resolution.All of these ECG signals were pre-processed to denoise.

First of all, to segment ECG pulses separately from a whole record, the R peak, which is a noticeable feature of ECG waveform, was detected by the Pan-Tompkins method [47]. To generate a database with all of single pulses, from the R peak sample, a desirable number of samples before and after the R point were extracted. In this case, each ECG record were segmented to a plethora of single pulses with 320 samples corresponding to 0.64 second each

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

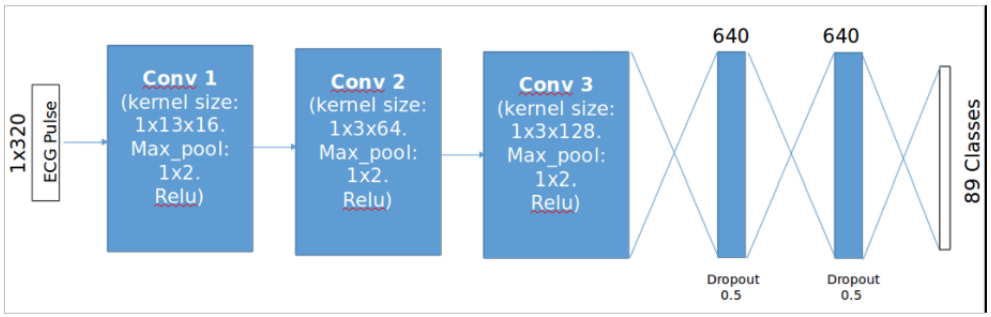

Figure 4.4: Our CNN Architecture

pulse which included 134 samples before R peak and 185 samples after R peak. Lastly, each record is composed of 1068 single ECG pulses with 12 pulses for each subject .Besides, based on the information of ECG-ID database, the records were divided into sets of database. Two records of each subject are called ECG Set S that were different sessions of acquisitions of a same day. ECG Set A, which is four records of ECG pulses, includes two records from ECG Set S and 2 other records with 25 subjects.

In our simulation studies, there are two main parts. On the first part, one record of ECG Set S is the training database in order to apply to classifier for learning. Another is the testing set to evaluate the efficiencies of classifiers. Then, the training set and test set are swapped together to cross-validate the dataset. This study aims to show the efficiency of classifiers when they work on the less intrapersonal data. On the second part, the experiments realized with two additional records with the similar process as the first one.

Moreover, in this study, there are three tasks for ECG classification. The first task is the original ECG database classification. In second, a user template guided filter is applied to denoise both sets of ECG signal before classifying. Finally, a Qtc augmentation with a user template guided filter is utilized to process these ECG sets. In these tasks, the training is from ECG Set S and testing database ECG Set A for the interpersonal experiments. For the non interpersonal cases, the training is one record of ECG Set S and another is testing set. Cross-validation is applied.

4.3.2 Biometric Classifier Models

In this section, we experiment for the behavior of a great deal of machine learning models as well as deep learning architectures to evaluate efficiency of them on classifying ECG signal with and without interpersonal database that corresponding to ECG Set S and ECG Set A . Besides that, CNN and RNN were chosen to classify out ECG data sets. The CNN model is defined as Fig 4.4. The RNN models are built up with three different hidden layer units:

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

Table IV.1: Performance summary of baseline, RNN models, and CNN model with original ECG sets.

Data Set	Method	Accuracy(%)
Data Set S	Baseline	89.1
	Traditional RNN	88
	GRU RNN	89.5
	LSTM RNN	90.5
	CNN	87.5
Data Set A	Baseline	64.8
	Traditional RNN	52
	GRU RNN	60
	LSTM RNN	54
	CNN	54.8

the traditional unit, the LSTM unit, and the GRU unit. Both of these deep learning models use Adam algorithm [32] to optimize their loss function. The implementations of them are performed on Tensorflow [1].

All of the results deep learning models are comparable to the baseline which is the classifier model using Euclidian algorithm.

4.3.3 Experimatal Results

Table IV.1 shows the performances of baseline and deep learning models. These ECG sets were preprocessed to denoise the base line drifting, power line interference, and other high frequency noise. In this task, the baseline is one of the highest accuracy method with 89.1 percent for ECG Set S and 64.8 percent for ECG Set A, while CNN model shows the low performance for both of ECG sets about 87.5 percent and 54.8 percent. Besides, the LSTM RNN with 90.5 percent achieved the best performance among RNN models as well as all of experiments on ECG Set S identification, but for interpersonal circumstances, the baseline with 64.8 percent is the most precise method for ECG biometric. In general, the performances of deep learning methods as well as baseline on ECG classification could not be improved by using the user template guided filter. There are no many considerable changes in Table IV.2.

In final experiment, the ECG data sets were augmented by QTcH algorithm. According to the Table IV.3, the deep learning models showed a significant improvement on their Data Set A classification. Although, the performances of RNN models on Data Set S experiments just slightly increase from 88 (Table IV.2) to 88.4 percent for traditional RNN, 89.5 (Table IV.2) to 90 percent for GRU CNN, on Data Set A experiments, RNN models jump rapidly from 53 to

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

Table IV.2: Performance summary of baseline, RNN models, and CNN model with the user template guided filter ECG sets

Data Set	Method	Accuracy(%)
Data Set S	Baseline	89.1
	Traditional RNN	88
	GRU RNN	89.5
	LSTM RNN	90.5
	CNN	88.5
Data Set A	Baseline	64.6
	Traditional RNN	53
	GRU RNN	59
	LSTM RNN	54
	CNN	54.7

Table IV.3: Performance summary of baseline, RNN models, and CNN model with the QTc correction ECG sets

Data Set	Method	Accuracy(%)
Data Set S	Baseline	90
	Traditional RNN	88.4
	GRU RNN	90
	LSTM RNN	90.4
	CNN	91.7
Data Set A	Baseline	64.9
	Traditional RNN	75.9
	GRU RNN	75
	LSTM RNN	77.7
	CNN	70.7

75.9 percent, 59 to 75 percent, and 54 to 77.7 percent for traditional, GRU, and LSTM RNN. Besides that, CNN model also shows a obvious accuracy growth for both experiment tasks. It is conspicuous that QTc augmentation is an effective method to improve performance of deep learning models on the intra-personal database classification.

If we compare the confusion matrices of RNN models which are shown as Figure 4.5, in all of three models, there are some subject databases that are too difficult to classify such as: subject 9, 11, 12, .etc.

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

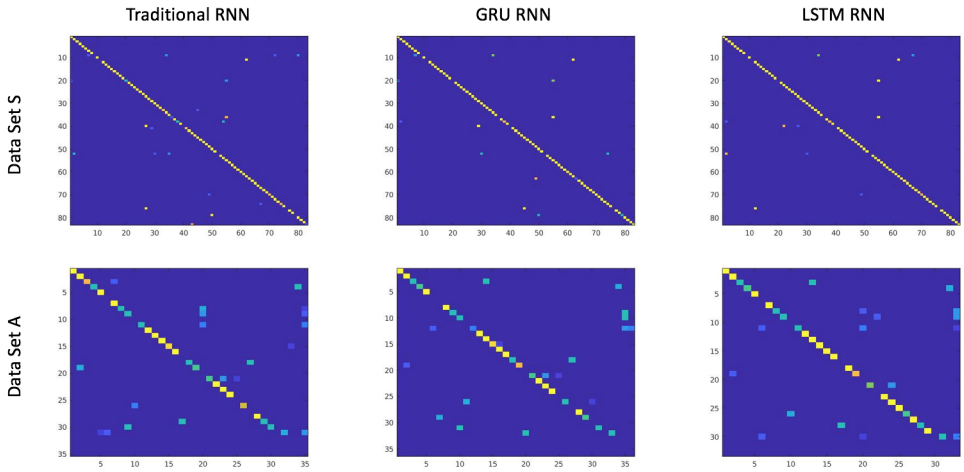

Figure 4.5: Confusion matrix comparison among RNN models of Experiment 3

4.3.4 Discussion

In this section, we proposed a new technique to improve the performance of deep learning models by using QTc augmentation. Specially, for intrapersonal data sets, while the deep learning models without QTc augmentation on Table IV.1 and IV.2 showed the lower-powered abilities to classify ECG Set A, with QTc augmentation, they becomes more powerful techniques to solve the identification problem.

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY

CHAPTER V

Conclusion

In this study, we investigated a framework using deep learning and k-means clustering to analyze and improve the quality of skin cancer dataset. From that, it significantly advances the performance of CNN model with smaller but higher quality database on skin cancer early detection. It does not only increase the accuracy of model, but also speed up the training task. Furthermore, this method is also a potential candidate to analyze many other databases.

Moreover, we propose an advantageous technique to improve the quantity of ECG database. This method has shown a good performance when it contributed to the obvious enhancement on the activity of deep learning models even on the specific challenging intrapersonal problem of ECG signal. The CNN model was yielded up from 88 percent to 91.7 percent for the non-intrapersonal dataset as well as 54.7 percent to 70.7 percent for the intrapersonal dataset. Besides, RNN models could increase their accuracy up to 24 percent for LSTM RNN case. This determines the efficiency of QTc augmentation on solving the intrapersonal problem.

All in one, this study aims to develop new methods to improve the performance of deep learning on biomedical data processing. Although, deep learning has achieved many state-of-the-art performances in numerous fields, on medical data processing, there are still many challenges such as variant database, noisy features, limitations on quality and quantity of database, etc.

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

References

[1] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, et al. Tensorflow: Large-scale machine learning on heterogeneous distributed systems. *arXiv preprint arXiv:1603.04467*, 2016. 26

[2] Radhakrishna Achanta, Appu Shaji, Kevin Smith, Aurelien Lucchi, Pascal Fua, and Sabine Süsstrunk. Slic superpixels. Technical report, 2010. 2

[3] Debaditya Acharya, Weilin Yan, and Kourosh Khoshelham. Real-time image-based parking occupancy detection using deep learning. vii, 14

[4] Foteini Agrafioti and Dimitrios Hatzinakos. Ecg biometric analysis in cardiac irregularity conditions. *Signal, Image and Video Processing*, 3(4):329, 2009. 20

[5] Yaniv Bar, Idit Diamant, Lior Wolf, and Hayit Greenspan. Deep learning with non-medical training used for chest pathology identification. In *Medical Imaging 2015: Computer-Aided Diagnosis*, volume 9414, page 94140V. International Society for Optics and Photonics, 2015. ii

[6] HC Bazett. An analysis of the time-relations of electrocardiograms. *Annals of Noninvasive Electrocardiology*, 2(2):177–194, 1997. 21, 23

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

[7] Md Amran Hossen Bhuiyan, Md Ibrahim Azad, and Kamal Uddin. Image processing for skin cancer features extraction. *International Journal of Scientific & Engineering Research*, 4(2):1–6, 2013. 14

[8] Lena Biel, Ola Pettersson, Lennart Philipson, and Peter Wide. Ecg analysis: a new approach in human identification. *IEEE Transactions on Instrumentation and Measurement*, 50(3):808–812, 2001. 20

[9] Jack Burdick, Oge Marques, Adrià Romero-Lopez, Xavier Giró Nieto, and Janet Weinthal. The impact of segmentation on the accuracy and sensitivity of a melanoma classifier based on skin lesion images. In *SIIM 2017 scientific program: Pittsburgh, PA, June 1-June 3, 2017, David L. Lawrence Convention Center*, pages 1–6, 2017. 14

[10] Chun-Lang Chang and Chih-Hao Chen. Applying decision tree and neural network to increase quality of dermatologic diagnosis. *Expert Systems with Applications*, 36(2):4035–4041, 2009. 2

[11] Chuang-Chien Chiu, Chou-Min Chuang, and Chih-Yu Hsu. A novel personal identity verification approach using a discrete wavelet transform of the ecg signal. In *Multimedia and Ubiquitous Engineering, 2008. MUE 2008. International Conference on*, pages 201–206. IEEE, 2008. 20

[12] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014. 10, 11

[13] Se Young Chun. Single pulse ecg-based small scale user authentication using guided filtering. In *Biometrics (ICB), 2016 International Conference on*, pages 1–7. IEEE, 2016. vii, 3, 23

[14] Ulzii-Orshikh Dorj, Keun-Kwang Lee, Jae-Young Choi, and Malrey Lee. The skin cancer classification using deep convolutional neural network. *Multimedia Tools and Applications*, pages 1–16, 2018. vii, 7

[15] Ulzii-Orshikh Dorj, Keun-Kwang Lee, Jae-Young Choi, and Malrey Lee. The skin cancer classification using deep convolutional neural network. *Multimedia Tools and Applications*, 77(8):9909–9924, Apr 2018. 13

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

[16] Ajaya Durg, William V Stoecker, JP Cookson, Scott E Umbaugh, and Randy Hays Moss. Identification of variegated coloring in skin tumors: neural network vs. rule-based induction methods. *IEEE Engineering in Medicine and Biology Magazine*, 12(3):71–74, 1993. 2

[17] Andre Esteva, Brett Kuprel, Roberto A Novoa, Justin Ko, Susan M Swetter, Helen M Blau, and Sebastian Thrun. Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542(7639):115, 2017. 13

[18] LS Fridericia. The duration of systole in an electrocardiogram in normal humans and in patients with heart disease. *Annals of Noninvasive Electrocardiology*, 8(4):343–351, 2003. 21

[19] Alejandro Garcia-Uribe, Nasser Kehtarnavaz, Guillermo Marquez, Victor Prieto, Madeleine Duvic, and Lihong V Wang. Skin cancer detection by spectroscopic oblique-incidence reflectometry: classification and physiological origins. *Applied Optics*, 43(13):2643–2650, 2004. 2

[20] Leon Gatys, Alexander S Ecker, and Matthias Bethge. Texture synthesis using convolutional neural networks. In *Advances in Neural Information Processing Systems*, pages 262–270, 2015. 2

[21] Ary L. Goldberger, Luis A. N. Amaral, Leon Glass, Jeffrey M. Hausdorff, Plamen Ch. Ivanov, Roger G. Mark, Joseph E. Mietus, George B. Moody, Chung-Kang Peng, and H. Eugene Stanley. Physiobank, physiotoolkit, and physionet. *Circulation*, 101(23):e215–e220, 2000. 24

[22] David Gutman, Noel CF Codella, Emre Celebi, Brian Helba, Michael Marchetti, Nabin Mishra, and Allan Halpern. Skin lesion analysis toward melanoma detection: A challenge at the international symposium on biomedical imaging (isbi) 2016, hosted by the international skin imaging collaboration (isic). *arXiv preprint arXiv:1605.01397*, 2016. 2

[23] Mehdi Habibzadeh, Mahboobeh Jannesari, Zahra Rezaei, Hossein Baharvand, and Mehdi Totonchi. Automatic white blood cell classification using pre-trained deep learning models: Resnet and inception. In *Tenth International Conference on Machine Vision (ICMV 2017)*, volume 10696, page 1069612. International Society for Optics and Photonics, 2018. 14

[24] Kaiming He, Jian Sun, and Xiaoou Tang. Guided image filtering. In *European conference on computer vision*, pages 1–14. Springer, 2010. 23

[25] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997. 10

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

[26] Johannes Hofmanninger and Georg Langs. Mapping visual features to semantic profiles for retrieval in medical imaging. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 457–465, 2015. ii, 8

[27] John M Irvine and Steven A Israel. A sequential procedure for individual identity verification using ecg. *EURASIP Journal on Advances in Signal Processing*, 2009(1):243215, 2009. 20

[28] Steven A Israel, John M Irvine, Andrew Cheng, Mark D Wiederhold, and Brenda K Wiederhold. Ecg to identify individuals. *Pattern recognition*, 38(1):133–142, 2005. 20, 23

[29] Mohammad H Jafari, Nader Karimi, Ebrahim Nasr-Esfahani, Shadrokh Samavi, S Mohamad R Soroushmehr, K Ward, and Kayvan Najarian. Skin lesion segmentation in clinical images using deep learning. In *Pattern Recognition (ICPR), 2016 23rd International Conference on*, pages 337–342. IEEE, 2016. 2, 14

[30] V. Jindal, J. Birjandtalab, M. B. Pouyan, and M. Nourani. An adaptive deep learning approach for ppg-based identification. In *2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*, pages 6401–6404, Aug 2016. 4

[31] JC Kavitha and A Suruliandi. Texture and color feature extraction for classification of melanoma using svm. In *Computing Technologies and Intelligent Data Engineering (ICCTIDE), International Conference on*, pages 1–6. IEEE, 2016. 14

[32] Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014. 26

[33] Ruggero Donida Labati, Enrique Muoz, Vincenzo Piuri, Roberto Sassi, and Fabio Scotti. Deep-ecg: Convolutional neural networks for ecg biometric recognition. *Pattern Recognition Letters*, 2018. 23

[34] Hanjiang Lai, Shengtao Xiao, Yan Pan, Zhen Cui, Jiashi Feng, Chunyan Xu, Jian Yin, and Shuicheng Yan. Deep recurrent regression for facial landmark detection. *IEEE Transactions on Circuits and Systems for Video Technology*, 2016. 2

[35] Rongjian Li, Wenlu Zhang, Heung-Il Suk, Li Wang, Jiang Li, Dinggang Shen, and Shuiwang Ji. Deep learning based imaging data completion for improved brain disease diagnosis. In *International Conference on Medical Image Computing and Computer-Assisted Intervention*, pages 305–312. Springer, 2014. ii, 9

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

[36] Paulo J Lisboa and Azzam FG Taktak. The use of artificial neural networks in decision support in cancer: a systematic review. *Neural networks*, 19(4):408–415, 2006. 2

[37] Justin Leo Cheang Loong, Khazaimatol S Subari, Rosli Besar, and Muhammad Kamil Abdullah. A new approach to ecg biometric systems: a comparative study between lpc and wpd systems. *World Acad Sci Eng Technol*, 68:759–64, 2010. 20

[38] Tatiana S Lugovaya. Biometric human identification based on ecg. *PhysioNet*, 2005. 21

[39] Shen Luo, Kurt Michler, Paul Johnston, and Peter W Macfarlane. A comparison of commonly used qt correction formulae: the effect of heart rate on the qtc of normal ecgs. *Journal of electrocardiology*, 37:81–90, 2004. 23

[40] V. Mai, I. Khalil, and C. Meli. Ecg biometric using multilayer perceptron and radial basis function neural networks. In *2011 Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 2745–2748, Aug 2011. vii, 3

[41] Juan Pablo Martínez, Rute Almeida, Salvador Olmos, Ana Paula Rocha, and Pablo Laguna. A wavelet-based ecg delineator: evaluation on standard databases. *IEEE transactions on biomedical engineering*, 51(4):570–581, 2004. 20

[42] Tomáš Mikolov, Martin Karafiát, Lukáš Burget, Jan Černockỳ, and Sanjeev Khudanpur. Recurrent neural network based language model. In *Eleventh Annual Conference of the International Speech Communication Association*, 2010. 9

[43] Ikenna Odinaka, Po-Hsiang Lai, Alan D Kaplan, Joseph A O'Sullivan, Erik J Sirevaag, and John W Rohrbaugh. Ecg biometric recognition: A comparative analysis. *IEEE Transactions on Information Forensics and Security*, 7(6):1812–1824, 2012. 20

[44] Albert Orriols and Ester Bernadó-Mansilla. The class imbalance problem in learning classifier systems: a preliminary study. In *Proceedings of the 7th annual workshop on Genetic and evolutionary computation*, pages 74–78. ACM, 2005. 1

[45] Adam Page, Amey Kulkarni, and Tinoosh Mohsenin. Utilizing deep neural nets for an embedded ecg-based biometric authentication system. In *Biomedical Circuits and Systems Conference (BioCAS), 2015 IEEE*, pages 1–4. IEEE, 2015. 4, 21

[46] Joana S Paiva, Duarte Dias, and João PS Cunha. Beat-id: Towards a computationally low-cost single heartbeat biometric identity check system based on electrocardiogram wave morphology. *PloS one*, 12(7):e0180942, 2017. 4, 21

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

[47] Jiapu Pan and Willis J Tompkins. A real-time qrs detection algorithm. *IEEE transactions on biomedical engineering*, (3):230–236, 1985. 24

[48] Konstantinos N Plataniotis, Dimitrios Hatzinakos, and Jimmy KM Lee. Ecg biometric recognition without fiducial detection. In *Biometric Consortium Conference, 2006 Biometrics Symposium: Special Session on Research at the*, pages 1–6. IEEE, 2006. 20

[49] Fabienne Porée, Gaëlle Kervio, and Guy Carrault. Ecg biometric analysis in different physiological recording conditions. *Signal, image and video processing*, 10(2):267–276, 2016. 20

[50] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, et al. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision*, 115(3):211–252, 2015. ii

[51] S. I. Safie, J. J. Soraghan, and L. Petropoulakis. Electrocardiogram (ecg) biometric authentication using pulse active ratio (par). *IEEE Transactions on Information Forensics and Security*, 6(4):1315–1322, Dec 2011. 3

[52] Alex Sagie, Martin G Larson, Robert J Goldberg, James R Bengtson, and Daniel Levy. An improved method for adjusting the qt interval for heart rate (the framingham heart study). *American Journal of Cardiology*, 70(7):797–801, 1992. 21, 23

[53] Ronald Salloum and C-C Jay Kuo. Ecg-based biometrics using recurrent neural networks. In *Acoustics, Speech and Signal Processing (ICASSP), 2017 IEEE International Conference on*, pages 2062–2066. IEEE, 2017. vii, 4, 21, 22

[54] Rebecca L Siegel, Kimberly D Miller, and Ahmedin Jemal. Cancer statistics, 2016. *CA: a cancer journal for clinicians*, 66(1):7–30, 2016. 13

[55] Yongjin Wang, Foteini Agrafioti, Dimitrios Hatzinakos, and Konstantinos N Plataniotis. Analysis of human electrocardiogram for biometric recognition. *EURASIP journal on Advances in Signal Processing*, 2008(1):148658, 2007. 20

[56] Yilin Yan, Min Chen, Mei-Ling Shyu, and Shu-Ching Chen. Deep learning for imbalanced multimedia data classification. In *Multimedia (ISM), 2015 IEEE International Symposium on*, pages 483–488. IEEE, 2015. 1

ULSAN NATIONAL INSTITUTE OF SCIENCE AND TECHNOLOGY

[57] Yading Yuan, Ming Chao, and Yeh-Chi Lo. Automatic skin lesion segmentation using deep fully convolutional networks with jaccard distance. *IEEE transactions on medical imaging*, 36(9):1876–1886, 2017. 14

[58] Bo Zhang, Jiasheng Zhao, Xiao Chen, and Jianhuang Wu. Ecg data compression using a neural network model based on multi-objective optimization. *PloS one*, 12(10):e0182500, 2017. 4

[59] Qingxue Zhang, Dian Zhou, and Xuan Zeng. Heartid: a multiresolution convolutional neural network for ecg-based biometric human identification in smart health applications. *IEEE Access*, 5:11805–11816, 2017. vii, 21, 22

[60] Huiting Zheng, Jiabin Yuan, and Long Chen. Short-term load forecasting using emd-lstm neural networks with a xgboost algorithm for feature importance evaluation. *Energies*, 10(8):1168, 2017. vii, 10

[61] Bolei Zhou, Aditya Khosla, Agata Lapedriza, Aude Oliva, and Antonio Torralba. Learning deep features for discriminative localization. In *Computer Vision and Pattern Recognition (CVPR), 2016 IEEE Conference on*, pages 2921–2929. IEEE, 2016. 18

ULSAN NATIONAL INSTITUTE OF
SCIENCE AND TECHNOLOGY