Multimedia Tools and Applications
https://doi.org/10.1007/s11042-020-09314-6

Fingerprint classification using deep learning approach

Beanbonyka Rim[1] **· Junseob Kim**[1] **· Min Hong**[2]

Received: 20 March 2020 / Revised: 8 June 2020 / Accepted: 9 July 2020
Published online: 14 July 2020

Abstract
The AFIS (Automatic Fingerprint Identification System) which generally processes two steps: feature extraction and matching, has challenges with a large database of fingerprint images for the real-time application due to the huge number of comparisons required. Therefore, the additional step of classifying detailed information of fingerprint can speed up the process of distinguishing for individual identification in the AFIS. In this paper, we presented a classification method to identify a detailed fingerprint information using deep learning approach. The proposed method aimed to distinguish the specific fingerprint information such as left-right hand classification, sweat-pore classification, scratch classification and fingers classification. Due to high personalization and security issue, we privately constructed our own dataset of fingerprint images. Five state-of-the-art deep learning models such as classic CNN, Alexnet, VGG-16, Yolo-v2 and Resnet-50 were adapted to be trained from scratch for those four categories. In our experimental tests, we received the results as follows. The Yolo-v2 model provided the highest accuracy of **90.98%**, **78.68%** and **66.55%** for the left-right hand, scratch and fingers classification, respectively. For sweat-pore classification, the Resnet-50 model provided the highest accuracy of **91.29%**. It is also worth noted that both Yolo-v2 and Resnet-50 took at most **250.37 ms** per image.

Keywords Fingerprint classification · Deep learning · Alexnet · VGG · Yolo · Resnet

✉ Min Hong
mhong@sch.ac.kr

Beanbonyka Rim
rim.beanbonyka@sch.ac.kr

Junseob Kim
openfolder0220@gmail.com

[1] Department of Computer Science, Soonchunhyang University, Asan 31538, South Korea

[2] Department of Computer Software Engineering, Soonchunhyang University, Asan 31538, South Korea

1 Introduction

Human biometrics have played an essential role for recognizing a person in various applications such as personal identification, personal authentication, video surveillance, security investigation, border control and more. The common properties of biometrics which have been used regard to the physical characteristics of an individual person such as face, iris and fingerprint [1]. Since the fingerprint attributes are unique and rarely changed, the AFIS uses the fingerprint recognition for the above applications [2].

The goal of the fingerprint recognition system is to strive for robust and fast detection. The AFIS performs feature exaction and matching using fingerprint information to identify the personal and it requires tremendous number of comparison processes with generally huge size of fingerprint database. Therefore, AFIS is usually not able to return the result in real-time. Fundamentally, the AFIS categories fingerprint images into three levels: pattern (a flow and orientation of ridge), points (friction ridge terminations) and shape (edge, width and scar) [3]. Thus, the high quality and discrimination of those three-levels' characteristics are put into consideration to receive the best features.

DL (Deep learning) which is a successful approach in distinguishing 2D imaging analysis, composes of multiple hidden layers to learn features from raw data and outputs the result through loss function. Regardless to the characteristics of pattern, points or shape, DL defines a learning algorithm which analyses the 2D images via color distribution. The 2D images are flattened into 1D array and computed for a score of each category. Then, the loss function which is differentiated by result of score function and actual result, is used to define which category the image is belong to. Additionally, the huge contribution of DL approach is that DL may consume much time on training step depending on capacity of computer, number of training data and number of training parameters, but it is fast in detection step enabling DL to be the state-of-the-art method for the real-time classification applications [4, 5].

Therefore, this paper proposes a detailed finger-information classification in fingerprint images using DL approach. We aim to classify fingerprint into four categories: (1) left or right hand, (2) sweat pore or normal state, (3) scratch or normal state and (4) type of fingers such as thumb, index, middle, ring and little. Our main contributions are (1) to construct a dataset of detailed information of fingers; (2) to classify a detailed information of fingers from fingerprint images; and (3) to use color distribution of fingerprint image as a feature data for classification by exploiting the DL technique.

2 Related works

The classification of detailed finger information in fingerprint images has not been potentially conducted yet due to high personalization and security issue. The National Institution of Standards and Technology (NIST) [6] of U.S. department of commerce which publicly provides fingerprint dataset for research propose, still has a limitation for this detailed finger information classification. Therefore, in order to conduct a research on this classification, the data is privately collected by an individual institution.

By our observation, fingerprint classification has been oriented for three tasks. One is minutiae structure-based classification. It categorizes the fingerprint-images into finger patterns such as arch, loop and whorl group [7]. Gan et al. [8] proposed a deep learning method, particularly SqueezeNet model, for fingerprint classification. They used 2000 pairs fingerprints of NIST-DB4 with resolution of 512 × 512. Firstly, the region of interest (ROI) was extracted using gradient method. Then, SqueezeNet model extracted the features and classified into five groups: arch, tented arch, left loop, right loop and whorl. Their best accuracy using the transfer learning technique is 95.73%. Michelsanti et al. [9] employed two CNN models on NIST SD4 database for fast fingerprint classification into four classes: arch, left loop, right loop, and whorl. The pre-trained weight of VGG19-F model (for fast processing) and VGG19-S model (for slow processing) which were used as transfer learning with data augmentation (flipped and rotated transformation), were trained on NVIDIA GTX 950 M GPU with CuDNN library for 140 epochs. The training processes took around 9 h for VGG19-F and 30 h for VGG19-S and received an accuracy of 94.4% and 95.05%, respectively.

Second is finger information-based classification. It categorizes the fingerprint-images into finger types (such as left-hand, right-hand, thumb, index, middle, ring and little group) and finger states (such as wetness, scar and fragment). Shehu et al. [10] exploited the transfer learning technique of CNN to classify hand and fingers. They collected 3000 images for left-right hand classification and 600 images for each of the 10 fingers for fingers classification. After collected, those datasets are put publicly at Sokoto Coventry Fingerprint Dataset [11]. The Resnet model of DL was adapted to train for the hand and fingers classification with an accuracy of 93.5% and 76.72%, respectively. Kim et al. [12] conducted a left-right hand classification from fingerprint images using five deep neural networks. The four DL models such as CNN, Alexnet, VGG and Resnet were adapted to train using python with tensorflow framework and YOLO v3 was adapted to train using C++ with darknet framework. The models were trained on 9080 fingerprint images and validated on 1000 images. Among five models, Resnet model provided the highest accuracy of 96.80%. Kim et al. [13] conducted a wet fingerprint classification using DL approach. Five DL models were adapted to train the same as the work in [12] on 6858 images and validate on 1716 images. Among five models, Resnet model provided the highest accuracy of 96.17%.

Third is gender classification such as female and male group. Shehu et al. [10] exploited the transfer learning technique of CNN to classify gender. They collected 1230 images of male and 1230 images of female for gender classification. After collected, those datasets are put publicly at Sokoto Coventry Fingerprint Dataset [11]. The Resnet model of DL was adapted to train for the gender classification with an accuracy of 75.2%. Liu et al. [14] proposed deep learning method to predict female or male from 3D fingerprint-images. The 3D fingerprint-images were captured by Optical Coherence Tomography (OCT) from 25 Asian females and 34 Asian males. The ResNet-17 model was exploited for this classification with the best accuracy of 80.7%.

3 DL network models

3.1 Classic CNN model

The simple and classic CNN model is employed and Fig. 1 illustrates the CNN architecture. It consists of 1 convolutional layer, 1 maxpooling and 2 fully-connected layers. The first convolutional layer (C) filters the input image of size h by w and 3 color channels (RGB) with 16 kernels of size 3 × 3 and a stride of 2 × 2, followed by a maxpooling (P) with kernel of 2 × 2 and a stride 2 × 2. Then, it is dense by a fully-connected layer (FC) with 16 neurons. For the categories of left-right hand, sweat-pore and scratch, the last FC is dense with softmax of 2-

output units, as shown in Fig. 1a. For the fingers category, the last FC is dense with softmax of 5-output units, as shown in Fig. 1b.

3.2 Alexnet model

Alexnet model [15] which is the state-of-the-art CNN model training on ImageNet ILSVRC-2010 for classifying 1000 different classes, consisted of 5 convolutional layers, some were followed by maxpooling and 3 fully-connected layers with softmax of 1000-output units. We adapt the Alexnet [15] to train from scratch on our own dataset. We add 1 fully-connected layer as shown in Fig. 2.

The input image of size h by w and 3 color channels (RGB) firstly is filtered by a convolutional layer (C) with 96 kernels of 11×11 and a stride of 4×4, followed by a maxpooling (P) with kernel of 2×2 and a stride of 2×2. Next, a C with 256 kernels of 11×11 and a stride of 1×1, followed by a P with kernel of 2×2 and a stride of 2×2. Then, two Cs with 384 kernels of 3×3 and a stride of 1×1, and a C with 256 kernels of 3×3 and a stride of 1×1, followed by a P with kernel of 2×2 and a stride of 2×2. Then, it is dense by two fully-connected layers (FC) with 4096 neurons and a FC with 1000 neurons. For the categories of left-right hand, sweat-pore and scratch, the new added FC is dense with softmax of 2-output units, as shown in Fig. 2a. For the fingers category, the new added FC is dense with softmax of 5-output units, as shown in Fig. 2b.

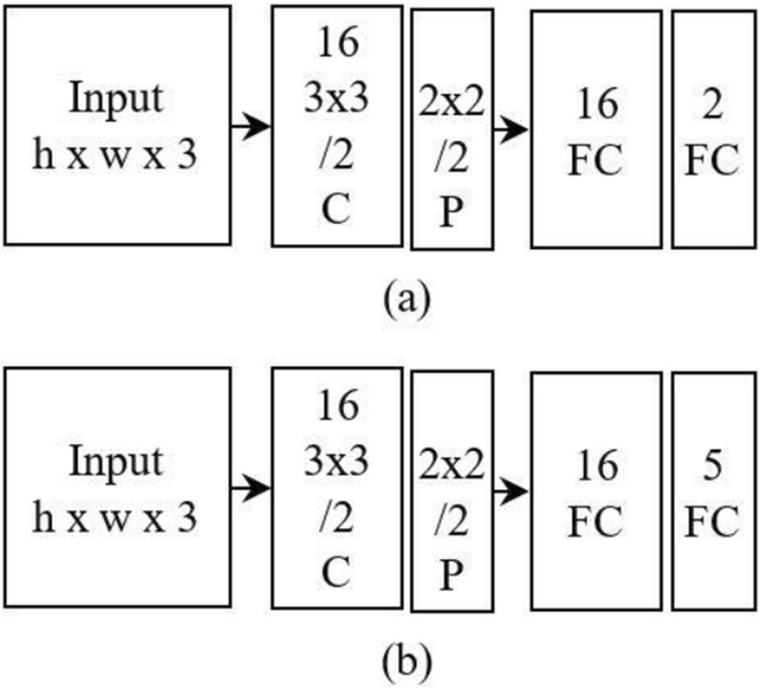

Fig. 1 Architecture of classic CNN model to classify a category consisting of (**a**) 2 classes and (**b**) 5 classes

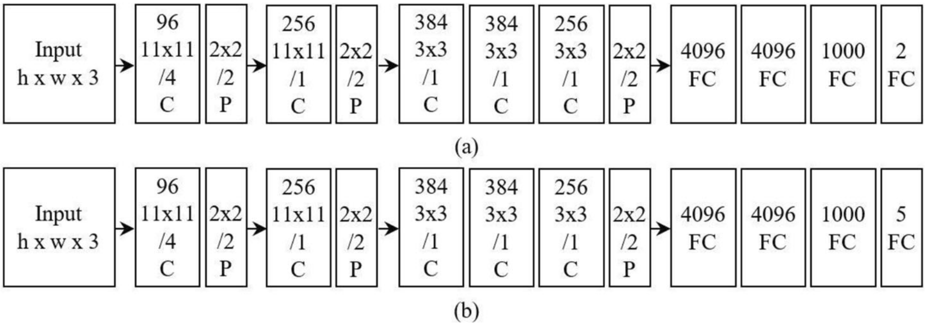

Fig. 2 Architecture of Alexnet model [15] to classify a category consisting of (**a**) 2 classes and (**b**) 5 classes

3.3 VGG-16 model

VGG model [16] which is the state-of-the-art CNN model training on ImageNet ILSVRC-2014 for classifying 1000 different classes, defined of 16-weight-layers model (VGG16) and 19-weight-layers model (VGG19). The VGG16 consists of 13 convolutional layers, some of which followed by maxpooling and 3 fully-connected layers with softmax of 1000-output units. The VGG19 consists of 16 convolutional layers, some of which followed by maxpooling and 3 fully-connected layers with softmax of 1000-output units. We adapt the VGG16 to train from scratch on our own dataset. We add 1 fully-connected layer as shown in Fig. 3.

The input image of size h by w and 3 color channels (RGB) firstly is filtered by two convolutional layers (C) with 64 kernels of 3×3 and a stride of 1×1, followed by a maxpooling (P) with kernel of 2×2 and a stride of 2×2. Next, two Cs with 128 kernels of 3×3 and a stride of 1×1, followed by a P with kernel of 2×2 and stride of 2×2. Next, three Cs with 256 kernels of 3×3 and a stride of 1×1, followed by a P with kernel of 2×2 and a stride of 2×2. Next, three Cs with 512 kernels of 3×3 and a stride of 1×1, followed by a P with kernel of 2×2 and a stride of 2×2. Next, three Cs with 512 kernels of 3×3 and a stride of 1×1, followed by a P with kernel of 2×2 and a stride of 2×2. Then, it is dense by two fully-connected layers (FC) with 4096 neurons and a FC with 1000 neurons. For the categories of left-right hand, sweat-pore and scratch, the new added FC is dense with softmax of 2-output units, as shown in Fig. 3a. For the fingers category, the new added FC is dense with softmax of 5-output units, as shown in Fig. 3b.

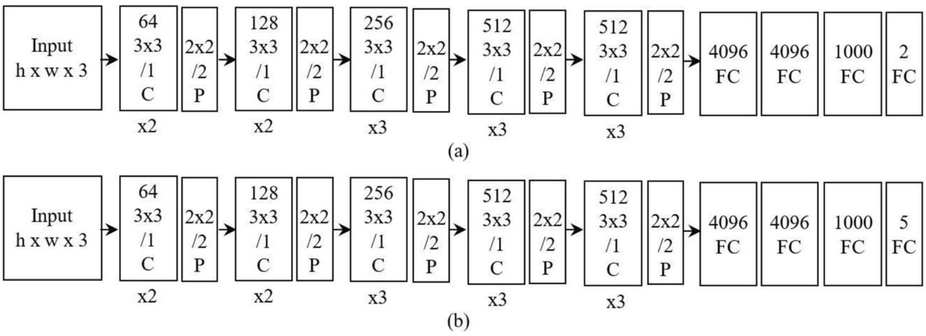

Fig. 3 Architecture of VGG16 model [16] to classify a category consisting of (**a**) 2 classes and (**b**) 5 classes

3.4 Yolo-v2 model

You Only Look Once (Yolo) model [17] which is an excellent fast real-time detection system, defined a classification model named Yolo-v2. The Yolo-v2 consists of 19 convolutional layers, some were followed by 5 maxpooling layers and 1 fully-connected layer with average pooling and softmax of 1000-output units. We adapt the Yolo-v2 to train from scratch on our own dataset. We add we add 1 fully-connected layer as shown in Fig. 4.

The input image of size h by w and 3 color channels (RGB) firstly is filtered by a convolutional layers (C) with 32 kernels of 3 × 3 and a stride of 1 × 1, followed by a maxpooling (P) with kernel of 2 × 2 and a stride of 2 × 2. Next, a C with 64 kernels of 3 × 3 and a stride of 1 × 1, followed by a P with kernel of 2 × 2 and a stride of 2 × 2. Next, a C with 128 kernels of 3 × 3 and a stride of 1 × 1, a C with 64 kernels of 1 × 1 and a stride of 1 × 1, a C with 128 kernels of 3 × 3 and a stride of 1 × 1, followed by a P with kernel of 2 × 2 and a stride of 2 × 2. Next, a C with 512 kernels of 3 × 3 and a stride of 1 × 1, a C with 256 kernels of 1 × 1 and a stride of 1 × 1, a C with 512 kernels of 3 × 3 and a stride of 1 × 1, a C with 256 kernels of 1 × 1 and a stride of 1 × 1, a C with 512 kernels of 3 × 3 and a stride of 1 × 1, followed by a P with kernel of 2 × 2 and a stride of 2 × 2. Next, a C with 1024 kernels of 3 × 3 and a stride of 1 × 1, a C with 512 kernels of 1 × 1 and a stride of 1 × 1, a C with 1024 kernels of 3 × 3 and a stride of 1 × 1, a C with 512 kernels of 1 × 1 and a stride of 1 × 1, a C with 1024 kernels of 3 × 3 and a stride of 1 × 1, followed by a P with kernel of 2 × 2 and stride of 2 × 2. Then, it is dense by a fully-connected layer (FC) with 1000 neurons. For the categories of left-right hand, sweat-pore and scratch, the new added FC is dense with softmax of 2-output units, as shown in Fig. 4a. For the fingers category, the new added FC is dense with softmax of 5-output units, as shown in Fig. 4b.

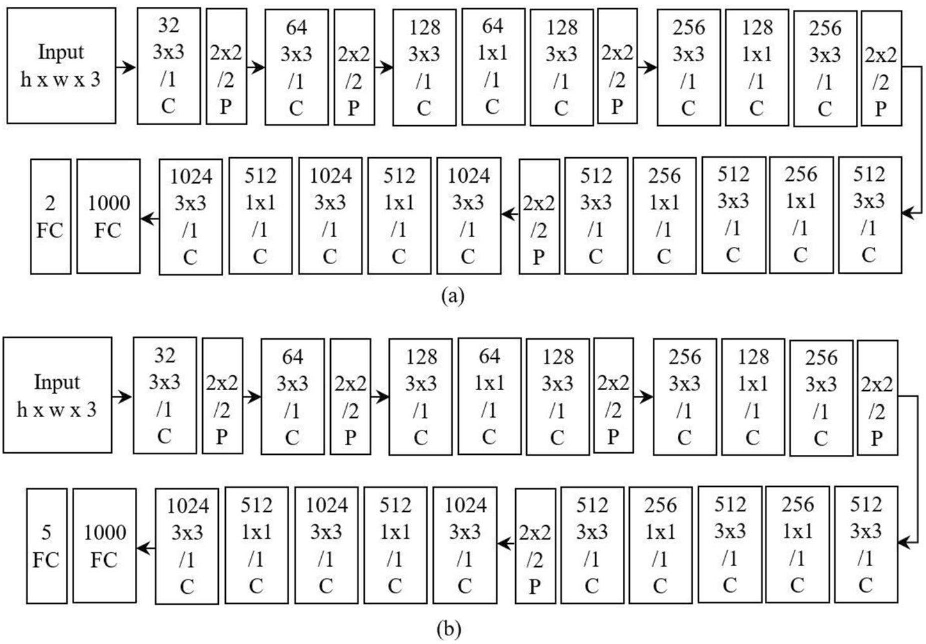

Fig. 4 Architecture of Yolo-v2 model [17] to classify a category consisting of (**a**) 2 classes and (**b**) 5 classes

3.5 Resnet-50 model

Resnet model [18] which is the outstanding deep residual learning for image recognition on ImageNet ILSVRC-2015 for classification task and COCO-2015 for detection and segmentation task, defined bottle-neck building blocks architecture for Resnet-50, Resnet-101 and Resnet-152. The Resnet-50 consists of 1 convolutional layer, 16 bottle-neck blocks and a fully-connected layer with softmax of 1000-output units. Each block consists of 3 convolutional layers. In total, the Resnet-50 consists of 50 convolutional layers. We adapt the Resnet-50 to train from scratch on our own dataset. We add we add 1 fully-connected layer as shown in Fig. 5.

The input image of size h by w and 3 color channels (RGB) firstly is filtered by a convolutional layers (C) with 64 kernels of 7 × 7 and a stride of 2 × 2, followed by a maxpooling (P) with kernel of 3 × 3 and a stride of 2 × 2. Next, three bottle-neck blocks of the C with 64 kernels of 1 × 1 and a stride of 1 × 1, a C with 64 kernels of 3 × 3 and a stride of 1 × 1, and a C with 256 kernels of 1 × 1 and a stride of 1 × 1. Next, four bottle-neck blocks of the C with 128 kernels of 1 × 1 and a stride of 2 × 2, a C with 128 kernels of 3 × 3 and a stride of 1 × 1, and a C with 512 kernels of 1 × 1 and a stride of 1 × 1. Next, six bottle-neck blocks of the C with 256 kernels of 1 × 1 and a stride of 2 × 2, a C with 256 kernels of 3 × 3 and a stride of 1 × 1, and a C with 1024 kernels of 1 × 1 and a stride of 1 × 1. Next, three bottle-neck blocks of the C with 512 kernels of 1 × 1 and a stride of 2 × 2, a C with 512 kernels of 3 × 3 and a stride of 1 × 1, and a C with 1024 kernels of 1 × 1 and a stride of 1 × 1. Then, it is dense by a fully-connected layer (FC) with 1000 neurons. For the categories of left-right hand, sweat-pore and scratch, the new added FC is dense with softmax of 2-output units, as shown in Fig. 5a. For the fingers category, the new added FC is dense with softmax of 5-output units, as shown in Fig. 5b.

4 Fingerprint dataset

Since there is not any publicly available dataset including NIST [6] about detailed information of fingers, we have collaborated with Soonchunhyang University of South Korea and Royal University of Phnom Penh of Cambodia to collect the fingerprint images. The 1069 fingerprints were collected from 1008 Cambodian people and 61 Korean people, who volunteering

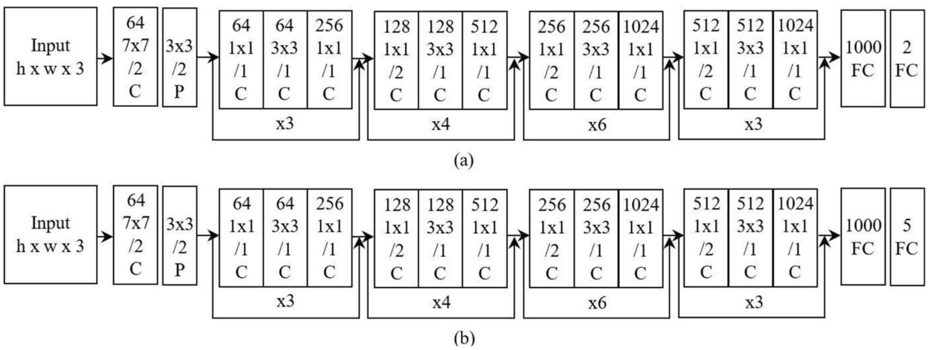

Fig. 5 Architecture of Resnet-50 model [18] to classify a category consisting of (**a**) 2 classes and (**b**) 5 classes

to provide their fingerprints under research propose only. Most majority, they age between 18 to 25 years old.

4.1 Data acquisition

The Integrated Biometrics (IB) Sherlock scanner [19] is used to capture the fingerprint. The scanner scans the image with resolution of 800 × 750 pixels, 24 bits and BMP image format. 10 fingers were scanned per person along with recorded of fingers' types. Thus, for 1069 subjects, we received 10,690 fingerprint images. While scanning, the fingers of all subjects were in various orientation, state and strength pressure. Those diversities of scanning behavior result broadly different pattern and state such as too light, too dark, blur, moisture or normal. However, we put all of them into classification method because those patterns are occurred by everyday life.

4.2 Data enhancement

Figure 6 illustrates the data enhancement. Since the DL approach uses color distribution to extract the feature, we strived for only the pattern of fingerprint. Therefore, the image is cropped to remove the whitespace, as illustrated in Fig. 6a. Then, the image is rotated to be in portrait position, as shown in Fig. 6b. Additionally, all the DL models that we proposed use the input image in fixed size, except Yolo-v2 model. Therefore, to be consistent principle of image size for inputting into all models, the image is scaled to fixed resolution of 224 × 224 pixels, as shown in Fig. 6c.

4.3 Data Split

The data is split into training set, validation set and test set. Since the dataset from 2 countries is not enough for equally distributed, we decided to perform the training and validation of network model with Cambodian fingerprint images and perform a test with Korean fingerprint images. Thus, 10,080 images from Cambodian people are used as training set and validation set and 610 images from Korean people are used as testing set as listed in Table 1.

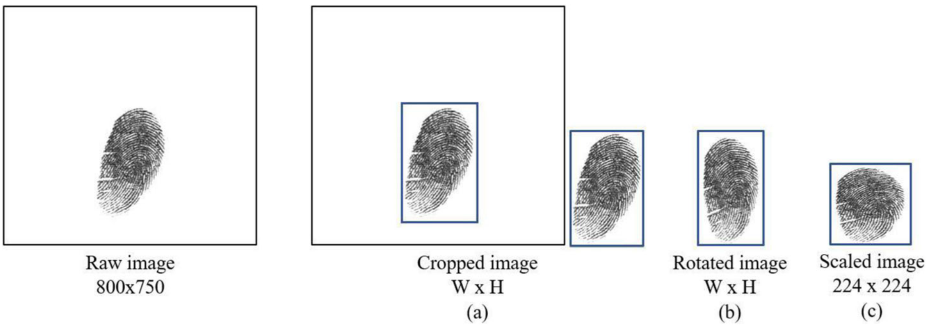

Fig. 6 Image enhancement by (**a**) cropping, (**b**) rotating and (**c**) scaling

Table 1 Data split for training, validation and testing set

Category	Class	Training dataset	Validation dataset	Testing dataset
Left-Right hand	Left	4540	500	305
	Right	4540	500	305
	Total	**9080**	**1000**	**610**
Sweat-pore	Sweat-pore	1225	307	125
	Normal	5633	1409	485
	Total	**6858**	**1716**	**610**
Scratch	Scratch	3072	769	218
	Normal	3864	967	392
	Total	**6936**	**1736**	**610**
Fingers	Thumb	1612	404	122
	Index	1612	404	122
	Middle	1612	404	122
	Ring	1612	404	122
	Little	1612	404	122
	Total	**8060**	**2020**	**610**

5 Experimental tests and results

5.1 Training

Our experimental test was conducted on a system equipped with a GPU of NVIDIA GeForce RTX 2070 SUPER and CuDNN library for accelerating the training. The code was implemented in Python language and Tensorflow framework with Keras API. The training was carried out by optimizing using stochastic gradient descent (SGD) with momentum to 0.9, weight decay to 0.0005, learning rate to 0.0001 and batch size to 16. To avoid an overfitting of training, the training was regularized by dropout for all fully-connected layers to ratio of 0.5, except the last dense layer. The networks were trained with random initialization. We trained using 3 types of epoch with size of 250, 500 and 1000 in order to be sure that the model has been trained long enough and there is not any room for improvement on the validation and test data. For the training outputs of 3 trained-weights, we named as w-250, w-500 and w-1000.

5.2 Validating

To avoid an overfitting of validating, the validation and testing set are in different data, as described in Section 4.3. The validation dataset is evaluated by the 3 trained-weights (w-250, w-500 and w-1000) across five DL models (classic CNN, Alexnet, VGG-16, Yolo-v2 and Resnet-50). The model that corresponds to the highest performance value in the validation set was selected as the final model. The final model was applied to the testing set for the prediction.

Table 2 shows the performance of the left-right hand classification using five DL models. Among those models, we can see that the Yolo-v2 model of trained-weight (w-500) which takes 70s 8 ms for training and 3 s 3 ms for validation, provides the loss score of 0.1412 to the ground truth with the highest accuracy of 95.80% comparing to other models.

Table 3 shows the performance of the sweat-pore classification using five DL models. Among those models, we can see that the Resnet-50 model of trained-weight (w-1000) which

Table 2 The performance of left-right hand classification

Model	Weight	Training (time/epoch)	Training lose score	Training accuracy (%)	Validation (time/epoch)	Validation lose score	Validation accuracy (%)
Classic CNN	w-250	9 s	0.4629	70.07	1 s	0.3778	89.50
	w-500		0.4269	76.06		0.3246	90.10
	w-1000		0.3951	79.47		**0.2925**	**90.50**
AlexNet	w-250	20s 2 ms	0.0258	99.25	1 s 1 ms	0.1774	94.00
	w-500		0.0221	99.32		0.1974	94.10
	w-1000		0.0118	99.71		**0.1646**	**94.80**
VGG-16	w-250	105 s 12 ms	0.5112	87.25	5 s 5 ms	**0.4673**	**86.60**
	w-500		0.6755	84.26		0.6718	84.80
	w-1000		0.6860	78.66		0.6846	81.60
Yolo-v2	w-250	70s 8 ms	0.0052	99.91	3 s 3 ms	**0.1406**	95.10
	w-500		0.0041	99.96		0.1412	**95.80**
	w-1000		0.0025	99.97		0.1667	95.80
Resnet-50	w-250	52 s 6 ms	0.1307	95.14	3 s 3 ms	0.2357	90.20
	w-500		0.1003	96.76		0.2385	90.40
	w-1000		0.0614	98.48		**0.2160**	**91.90**

takes 39 s 6 ms for training and 4 s 3 ms for validation, provides the loss score of 0.1511 to the ground truth with the highest accuracy of 93.99% comparing to other models.

Table 4 shows the performance of the scratch classification using five DL models. Among those models, we can see that the Yolo-v2 model of trained-weight (w-1000) which takes 53 s 8 ms for training and 5 s 3 ms for validation, provides the loss score of 0.8353 to the ground truth with the highest accuracy of 80.18% comparing to other models.

Table 5 shows the performance of the fingers classification using five DL models. Among those models, we can see that the Yolo-v2 model of trained-weight (w-250) which takes 62 s 8 ms for training and 6 s 3 ms for validation, provides the loss score of 0.5700 to the ground truth with the highest accuracy of 81.29% comparing to other models. Our performance result

Table 3 The performance of sweat-pore classification

Model	Weight	Training (time/epoch)	Training lose score	Training accuracy (%)	Validation (time/epoch)	Validation lose score	Validation accuracy (%)
Classic CNN	w-250	7 s 1 ms	0.1951	90.77	2 s	0.1813	92.54
	w-500		0.1904	91.18		0.1767	92.42
	w-1000		0.1843	91.72		**0.1740**	**92.77**
AlexNet	w-250	15 s 2 ms	0.0754	97.96	2 s 1 ms	0.2516	90.90
	w-500		0.0379	99.26		**0.2332**	**92.48**
	w-1000		0.0253	99.66		0.2680	92.48
VGG-16	w-250	80s 12 ms	0.1478	93.92	9 s 5 ms	0.1804	92.89
	w-500		0.1484	93.86		0.1821	93.24
	w-1000		0.1239	94.74		0.1648	**93.47**
Yolo-v2	w-250	53 s 8 ms	0.0154	99.84	5 s 3 ms	**0.2030**	93.76
	w-500		0.0130	99.94		0.2265	93.53
	w-1000		0.0083	99.99		0.2454	**93.93**
Resnet-50	w-250	39 s 6 ms	0.0905	96.05	4 s 3 ms	**0.1382**	93.58
	w-500		0.0672	97.62		0.1585	93.47
	w-1000		0.0587	98.10		0.1511	**93.99**

Table 4 The performance of scratch classification

Model	Weight	Training (time/epoch)	Training lose score	Training accuracy (%)	Validation (time/epoch)	Validation lose score	Validation accuracy (%)
Classic CNN	w-250	7 s 1 ms	0.6573	55.71	2 s	0.6534	55.70
	w-500		0.6551	59.50		0.6500	63.07
	w-1000		0.6467	60.97		**0.6445**	**63.94**
AlexNet	w-250	15 s 2 ms	0.0459	98.92	2 s 1 ms	**0.8256**	71.42
	w-500		0.0241	99.48		0.9286	73.38
	w-1000		0.0187	99.51		0.9191	**73.90**
VGG-16	w-250	80s 12 ms	0.6864	55.71	8 s 4 ms	0.6863	55.70
	w-500		0.6861	55.71		0.6862	55.70
	w-1000		0.6854	55.71		**0.6853**	**55.70**
Yolo-v2	w-250	53 s 8 ms	0.0074	99.90	5 s 3 ms	**0.7830**	78.68
	w-500		0.0061	99.90		0.9081	76.44
	w-1000		0.0040	99.94		0.8353	**80.18**
Resnet-50	w-250	39 s 6 ms	0.2388	91.13	4 s 2 ms	**0.4534**	**79.78**
	w-500		0.0657	98.69		0.6381	75.05
	w-1000		0.0529	98.85		0.6605	76.72

provides somewhat higher validation accuracy of 81.29% comparing to Shehu et al. [10] which was 76.72%.

5.3 Testing

In this paper, in order to confirm whether fingerprints are affected by regional environments, five DL models were compared with fingerprint information of Cambodian people, and then, we conducted to discriminate fingerprints of Korean people based on the best model. For the left-right hand classification, the Yolo-v2 model of w-500 takes 4 s 6 ms and provides the loss score of 0.3037 with an accuracy of 90.98%, as shown in Table 6.

Table 5 The performance of fingers classification

Model	Weight	Training (time/epoch)	Training lose score	Training accuracy (%)	Validation (time/epoch)	Validation lose score	Validation accuracy (%)
Classic CNN	w-250	8 s 1 ms	1.6094	20.00	2 s	**1.6094**	**19.99**
	w-500		1.6094	20.00		1.6094	19.99
	w-1000		1.6094	20.00		1.6094	19.99
AlexNet	w-250	18 s 2 ms	0.1641	95.72	3 s 1 ms	**0.6960**	75.23
	w-500		0.0852	97.92		0.7938	**73.84**
	w-1000		0.0429	99.18		0.8577	72.95
VGG-16	w-250	93 s 12 ms	1.6066	33.16	10s 5 ms	1.6059	38.41
	w-500		1.6013	42.77		**1.6002**	**48.83**
	w-1000		1.6021	40.40		1.6009	45.26
Yolo-v2	w-250	62 s 8 ms	0.0202	99.93	6 s 3 ms	**0.5700**	**81.29**
	w-500		0.0170	99.88		0.6226	79.95
	w-1000		0.0123	99.98		0.6098	80.74
Resnet-50	w-250	46 s 6 ms	0.5383	80.40	5 s 3 ms	0.6581	74.09
	w-500		0.4511	84.71		**0.6156**	**76.42**
	w-1000		0.4327	85.00		0.6218	75.13

Table 6 The performance of testing

Category	Model	Testing (time/epoch)	Testing lose score	Testing accuracy (%)
Left-Right hand	Yolo-v2 (w-500)	4 s 6 ms	0.3037	90.98
Sweat-pore	Resnet-50 (w-1000)	3 s 5 ms	0.2505	91.29
Scratch	Yolo-v2 (w-1000)	4 s 6 ms	0.9116	78.68
Fingers	Yolo-v2 (w-250)	4 s 6 ms	1.2317	66.55

When a fingerprint image is input, the Yolo-v2 model distinguishes whether the input image is the left or right hand. Figure 7 illustrates some examples which show the prediction results on testing set. The blue and red color presents correct and false prediction, respectively. There are 9 fingerprint images which are randomly selected for example result visualization. Among 9 examples, there are 8 examples which predicted correctly. The lowest accuracy of prediction is 55.00% and the highest accuracy is 73.00%. However, there is 1 example which predicted false. The ground truth is the fingerprint image of the right hand, but it is false to predict as the left hand.

For the sweat-pore classification, the Resnet-50 model of w-1000 takes 3 s 5 ms to process and provides the loss score of 0.2505 with an accuracy of 91.29%, as shown in Table 6. When a fingerprint image is input, the Resnet-50 model distinguishes whether the input image is clammy or normal hand. The fingerprint is categorized as a sweat-pore pattern as shown in Fig. 8. Figure 8a represents a normal pattern. The Fig. 8b, c and d represent a sweat-pore pattern when the finger is rubbed by wet towel, the wet finger is rubbed by dry towel and the wet finger is directly scanned without rubbed, respectively. Fig. 9 illustrates examples which show the prediction results on testing set. There are 9 fingerprint images which are randomly selected for example result visualization. Among 9 examples, there are 8 examples which predicted correctly. The lowest accuracy of prediction is 72.00% and the highest accuracy is 73.00%. However, there is 1 example which predicted false. The ground truth is the fingerprint image of the normal, but it is false to predict as the sweat-pore.

For the scratch classification, the Yolo-v2 model with trained-weight of w-1000 takes 4 s 6 ms to process and provides the loss score of 0.9116 with an accuracy of 78.68%, as shown in

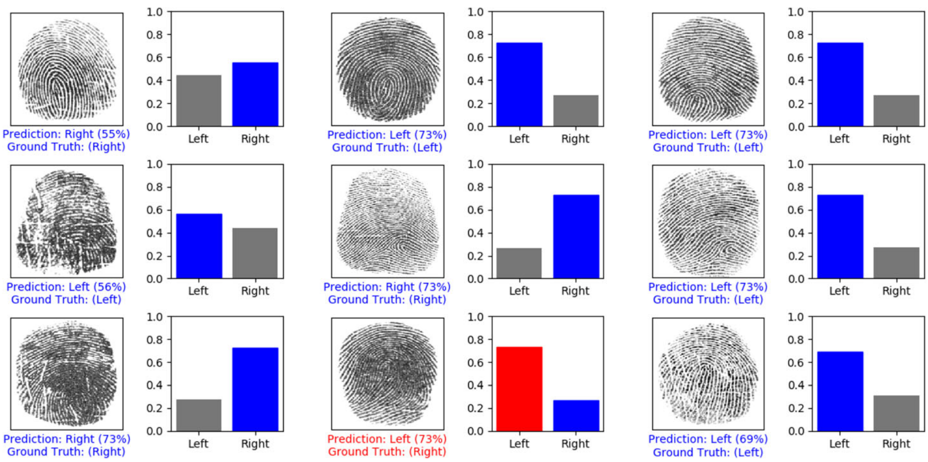

Fig. 7 The visualization examples of the left-right hand classification

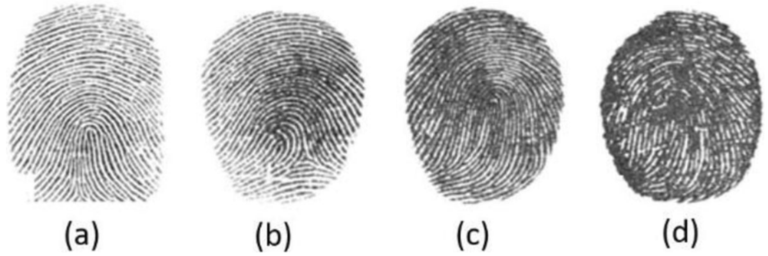

Fig. 8 Fingerprint in state of: (**a**) normal and (**b, c, d**) sweat-pore pattern

Table 6. When a fingerprint image is input, the Yolo-v2 model returns whether the input image is scratched or normal hand. The fingerprint is categorized as a scratch pattern when there are more than 3 scars occurred, otherwise it is considered as a normal pattern. Figure 10 illustrates examples which show the prediction results on testing set. There are 9 fingerprint images which are randomly selected for example visualization. Among 9 examples, there are 7 examples which predicted correctly. The lowest accuracy of prediction is 63.00% and the highest accuracy is 73.00%. However, there are 2 examples which predicted false. Case 1, the ground truth is the fingerprint image of the normal, but it is false to predict as the scratch. Since there are many whitespaces within the ridge of the pattern, that leads the classifier predicted as a scratch. Case 2, the ground truth is the fingerprint image of the scratch, but it is false to predict as the normal. Due to the strong clear ridge pattern and light scratch scars, the classifier confused predicting it as normal.

For the fingers classification, the Yolo-v2 model with trained-weight of w-250 takes 4 s 6 ms to process and provides the loss score of 1.2317 with an accuracy of 66.55%, as shown in Table 6. For the fingers classification, when a fingerprint image is input, the Yolo-v2 model distinguished which finger it is. Figure 11 illustrates examples which visualize the prediction results on testing set. There are 9 fingerprint images which are randomly selected for example result visualization. Among 9 examples, there are 5 examples which predicted correctly. The lowest accuracy of prediction is 34.00% and the highest accuracy is 40.00%. However, there are 4 examples which predicted false. Case 1, the ground truth is the fingerprint image of the index finger, but it is false to predict as the little finger. Case 2, the ground truth is the

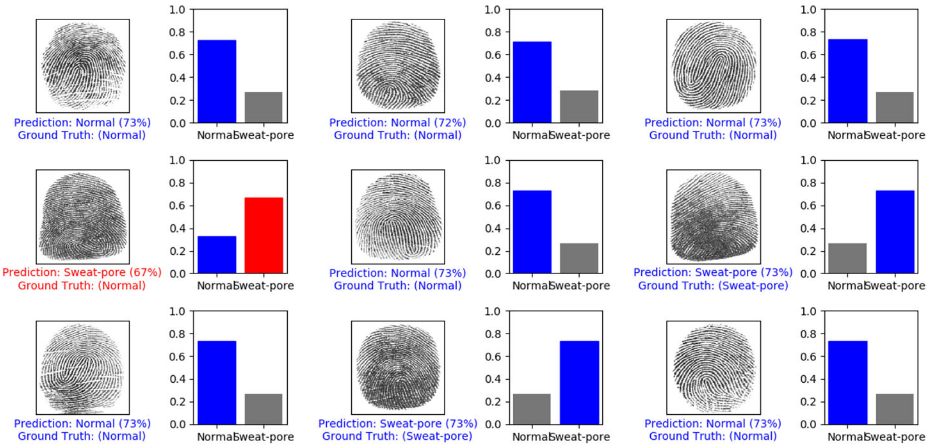

Fig. 9 The visualization examples of the sweat-pore classification

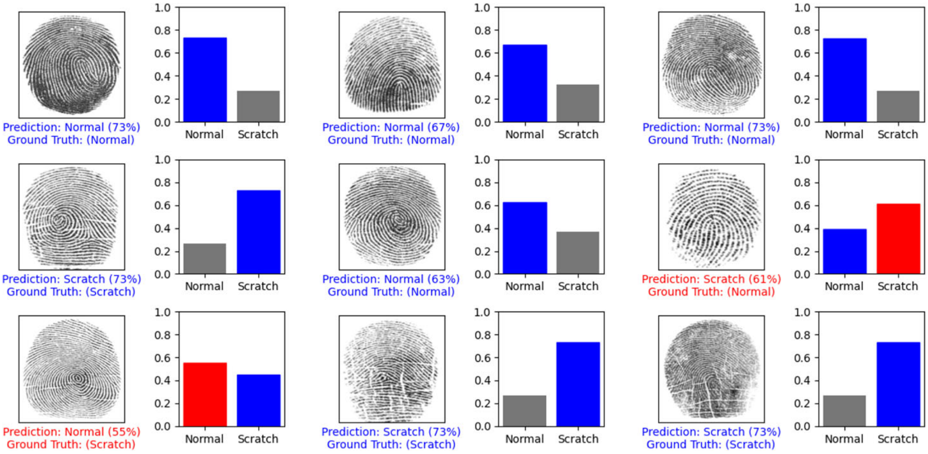

Fig. 10 The visualization examples of the scratch classification

fingerprint image of the middle finder, but it is false to predict as the ring finger. Case 3, the ground truth is the fingerprint image of the middle finger, but it is false to predict as the ring finger. Case 4, the ground truth is the fingerprint image of the ring finger, but it is false to predict as the little finger. The false predictions were occurred of ring and little finger.

The goal of setting the difference between validation and testing set is to strive for a generalization in detection. The more generalization it is, the more robust performance of our detection will be. The different characteristics pattern of these two sets are not only from an individual person to person, but also from two different regions (Cambodia and Korea). In our experimental tests, we can observe that the accuracy of testing is lower than validation, especially for fingers classification. The drop rates are 4.12%, 2.31%, 1.5% and 14.74% of the left-right hand, sweat-pore, scratch and fingers classification, respectively. We assume that the drop rates of these testing performances are due to the characteristic pattern of Cambodian fingerprints (in validation set) is rougher and tougher than characteristic pattern of Korean fingerprints (in testing set).

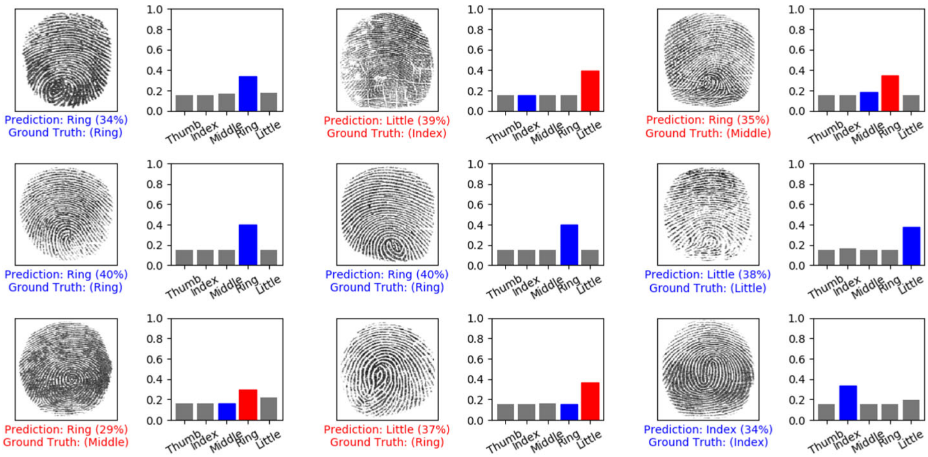

Fig. 11 The visualization examples of the fingers classification

6 Discussions

Even though our results are under state-of-the-art level, we have contributed three main points. One is a new dataset collection. We have collected 10,690 fingerprint-images from 1069 subjects.

Second is a propose of a baseline process adding up to the existing AFIS in order to significantly identify a fingerprint to a person. The high prediction accuracy of this baseline process beneficially minimizes the space for searching and maximizes the speed of matching a candidate fingerprint to a list of registered fingerprints in a large database. In practical, the fingerprint matching process of the existing AFIS is desirably conducted at least a two-stage search. The first stage generally matches the candidate fingerprint based on the global fingerprint characteristics. Then, the second stage specially matches based on the minutia features. Both stages are conducted within an entire database [20]. Therefore, if the candidate fingerprint is firstly recognized as a finger of left or right hand, the two above matching stages will be searched only within a left- or right-hand database, rather than within the entire database. Similarly, for the baseline process of finger-type classification, the matching stages will be conducted within a group of thumb, index, middle, ring or little, rather than within the entire database. By doing this, the matching process can be intuitively speed up by 2 and 5 times for adding up the baseline process of left-right hand and finger-type classification, respectively.

Third is a propose of applying deep learning method for classification. Our method proved that the deep learning method is more convenient comparing to the traditional machine learning, without human intervention such as hand-crafted feature extraction and pre-processing step. We do not need to be a professional in the knowledge of image processing in which it requires to enhance the input images in the pre-processing step. In our input-image enhancement, we tried to minimize the pre-processing as much as possible. We conducted only two simple processes of rotating and scaling. We kept the images as raw as possible, that is a reason our input image is in a format of three color-channels (i.e. R, G and B). Table 7 describes a performance comparison between our input images of three color-channels (RGB format) and the input images of one channel (gray-scale format) of the works in [12, 13]. Since the works in [12, 13] did not conduct a performance on Testing, we compare only on training and validation accuracy. For left-right hand classification, our accuracy is just 0.02% and 1.00% lower than the work of [12] for training and validation, respectively. For sweat-pore classification, our accuracy is just 1.89% and 2.18% lower than the work of [13] for training and validation, respectively. This proofs that the input image in RGB color format does not yield much differences comparing to the gray-scale format, while it beneficially and conveniently minimizes the pre-processing.

Table 7 Performance of input image in format of gray-scale and RBG

Category	Method	Input image format	Training accuracy (%)	Validation accuracy (%)
Left-Right	Kim et al. [12]	Gray-scale	99.98	96.80
	Our method	RBG	99.96	95.80
Sweat-pore	Kim et al. [13]	Gray-scale	99.99	96.17
	Our method	RBG	98.10	93.99

7 Conclusion

In this paper, we proposed a method to classify the detailed information of fingers from fingerprint image using deep learning approach. The fingerprint image is classified into four categories of the left-right hand, sweat-pore, scratch and finger type category. Five deep learning models (such as classic CNN, Alexnet, VGG-16, Yolo-v2 and Resnet-50) were applied to strive for high accuracy of validation. The trained-weight of highest accuracy is used to predict on the testing dataset. To be generalization of our experiment, the validation set is a data of Cambodian fingerprint, while the testing set is a data of Korean fingerprint. In our experimental tests, we received the results as follows: Yolo-v2 provided the highest accuracy of **90.98%**, **78.68%** and **66.55%** for the left-right hand, scratch and fingers classification, respectively. For sweat-pore classification, the Resnet-50 provided the highest accuracy of **91.29%**. It is also worth noted that both Yolo-v2 and Resnet-50 took at most 4 s 6 ms to classify 16 images which is around **250.37 ms** per image.

The fingerprint image is classified by 4 different trained-weights. In the future, we will extend our method by train a model which is able to classify a fingerprint image into 4 categories by a single trained-weight. The low accuracy of the finger classification due to lack of dataset. Hopefully, we will collect more data in order to improve the performance. Additionally, we will improve the performance of the classification by exploring a robust deep learning model network which yields for the state-of-art level of both processing and accuracy. We will also put the deployment of our experiments in a "real" fingerprint identification system into our interests for the future works.

Acknowledgements This work was supported by the Technology development Program(S2688148) funded by the Ministry of SMEs and Startups (MSS, Korea) and was supported by the Soonchunhyang University Research Fund.

References

1. Minaee S, Abdolrashidi A, Su H, Bennamoun M, Zhang D (2019) Biometric recognition using deep learning: a survey. arXiv preprint arXiv:1912.00271
2. Komarinski P (2005). Automated fingerprint identification systems (AFIS). Elsevier
3. Kim MJ, Kim WY, Paik J (2018) Optimum geometric transformation and bipartite graph-based approach to sweat pore matching for biometric identification. Symmetry 10(5):175. https://doi.org/10.3390/sym10050175
4. Johnson J, Yeung S, Li FF (2019). CS231n convolutional neural networks for visual recognition. Stanford University. http://cs231n.stanford.edu/2019/. Accessed 01 January 2020
5. Mu R, Zeng X (2019) A review of deep learning research. TIIS Trans Internet Inf Syst 13(4):1738–1764. https://doi.org/10.3837/tiis.2019.04.00
6. Fiumara G, Flanagan P, Grantham J, Bandini B, Ko K, Libert J (2019). NIST special database 300 uncompressed plain and rolled images from fingerprint cards. National Institute of Standards and Technology (NIST) https://doi.org/10.18434/T4/1502472. Accessed 01 January 2020
7. Win KN, Li K, Chen J, Viger PF, Li K (2019) Fingerprint classification and identification algorithms for criminal investigation: a survey. Futur Gener Comput Syst 110:758–771. https://doi.org/10.1016/j.future.2019.10.019
8. Gan J, Qi L, Bai Z, Xiang L (2019) Fingerprint classification based on lightweight neural networks. In: Sun Z, He R, Feng J, Shan S, Guo Z (eds) Biometric recognition. CCBR 2019, LNCS, vol 11818. Springer, Cham, pp 28–36. https://doi.org/10.1007/978-3-030-31456-9_4
9. Michelsanti D, Ene AD, Guichi Y, Stef R, Nasrollahi K, Moeslund TB (2017). Fast fingerprint classification with deep neural networks. In: Proc. of VISIGRAPP, pp. 202–209

10. Shehu YI, Ruiz-Garcia A, Palade V, James A (2018). Detailed Identification of Fingerprints Using Convolutional Neural Networks. In: Proc of 2018 17th IEEE international conference on machine learning and applications (ICMLA), pp 1161-1165. DOI: https://doi.org/10.1109/ICMLA.2018.00187
11. Shehu YI, Ruiz-Garcia A, Palade V, James A (2018). Sokoto Coventry fingerprint dataset. arXiv preprint arXiv:1807.10609
12. Kim J, Rim B, Sung NJ, Hong M (2020) Left or right hand classification from fingerprint images using a deep neural network. Computers, Materials & Continua 63(1):17–30. https://doi.org/10.32604/cmc.2020.09044
13. Kim J, Rim B, Sung NJ, Hong M (2019). Wet fingerprint classification using deep learning approach. In: Proc of KSII the 11th international conference on internet (ICONI), pp. 110-112
14. Liu H, Zhang W, Liu F, Qi Y (2019). 3D fingerprint gender classification using deep learning. In: Sun Z, He R, Feng J, Shan S, Guo Z (eds) Biometric Recognition. CCBR 2019, pp. 37–45. https://doi.org/10.1007/978-3-030-31456-9_5
15. Krizhevsky A, Sutskever I, Hinton GE (2012). Imagenet classification with deep convolutional neural networks. In Advances in neural information processing systems, pp 1097–1105
16. Simonyan K, Zisserman A (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556.
17. Redmon J, Farhadi A (2017). YOLO9000: better, faster, stronger. In: Proc. of the IEEE conference on computer vision and pattern recognition, pp. 7263–7271
18. He K, Zhang X, Ren S, Sun J (2016). Deep residual learning for image recognition. In: Proc. of the IEEE conference on computer vision and pattern recognition, pp. 770–778
19. Integrated Biometrics (IB) Sherlock scanner. https://integratedbiometrics.com/products/sherlock/. Accessed 01 January 2020
20. Automated fingerprint identification. https://en.wikipedia.org/wiki/Automated_fingerprint_identification. Accessed 01 June 2020