Deep Learning Based Gender Determination Using Fingerprints

Roy Amante Salvador and Mary Grace Villanueva
Department of Computer Science
College of Engineering, University of the Philippines Diliman
Quezon City Philippines, 1101
(632) 434 3877

***Abstract*—Sex determination from fingerprints is an important task in the field of forensics. There are many features which have been studied in fingerprints in relation to gender. In this paper, we've approached the problem with Deep Learning and used Convolutional Neural Networks (CNN) to determine gender from raw fingerprint images. Four architectures are used and compared - LeNet, AlexNet and two of its modified versions utilizing Maxout and Exponential Linear Unit activation functions. With Dropout and Data Augmentation, we've achieved a more balanced performance between Male and Female classes and an overall performance around 90% while using an imbalanced dataset. In our result, we also see the thumb as the best finger to discriminate a male fingerprint, and the little finger for female fingerprints.**

Keywords*—*Gender determination, Fingerprint Analysis, Deep Learning, Convolutional Neural Networks,

I. Introduction

Fingerprints have been reliably used for identification because of its uniqueness. No two fingers are found to be identical even from those coming from the same gene pool. This is why fingerprints are commonly used for biometric systems. Successful identification however requires maintenance of a fingerprint database for feature matching. In the event that a person is not yet in the database, fingerprints still give some hints on the person's identity. This helps narrow down the list of suspects. Gender determination is significant in forensic investigations to identify a criminal.

Some studies have shown that there are visual features in fingerprints which can provide gender information [9]. In this project, we aim on validating this claim using Deep Learning. Some of the popular features used for gender determination are finger size, ridge count and density. With Deep Learning, there is no need to manually engineer and calculate for these features as this is automatically learned by the network. Convolutional Neural Networks (CNN) in particular have been the state-of-the-art classifier for the past years when it comes to image classification.

The outline of this paper is as follows: we briefly describe some of the methods previously used for gender determination using fingerprints in section II, followed by description of the dataset and the CNN architectures we've considered in Section III, ended by experimental results and some analysis in Sections IV and V.

II. Related Work

Previous work on determining sex from fingerprints, mostly revolved around relying on calculating visual features. Ridge density seems to be highly used for feature extraction in gender classification. The universal trend is that females have a higher ridge density compared to males. This is found consistent in Indian [12], Malay [1], American Indian [6], Chinese [2] and Spanish Caucasian [5] races. Wang et al. [9] experimented with the number of ridges and its density as well as finger size. By using a Multi Layer Perceptron (MLP), they've concluded that the combination of these features produced higher classification results compared to individual features.

Gnanasivam et al. [14] came up with a method by combining analysis of its frequency domain and spatial domain. Feature vector consisted of Discrete wavelet transform (DWT) and singular value decomposition (SVD) values and classified by K-Nearest Neighbor. Similar work was also done by Gornale et al. [8] but using Support Vector Machine (SVM) as classifier while employing 10 fold cross validation technique. They've achieved an overall classification rate of 88.28%. Ganesh et al. [4] focused on Frequency domain calculations. They simply applied thresholding on Fast Fourier Transform (FFT), Discrete Cosine Transform (DCT) and Power Spectral Density (PSD) values to discriminate between Male and Female classes and still able to get 78% and 88% accuracy for respective classes.

Deep Learning has been applied with Fingerprint classification. Han et al. [15] have adopted a deep network structure. A sparse auto encoder neural network containing three hidden layers have been used to learn the characteristics of each fingerprint class by unsupervised learning. The feature output of the network is a reconstructed orientation field of the fingerprint and it is fed to a Softmax regression model for classification. They achieved an overall accuracy between 92 to 93%.

As of this moment, the authors have not found published work on sex determination from fingerprints using Deep Learning approach.

III. Methodology

A. Dataset

The National Institute of Standards and Technology (NIST) Special Database 4 is used as dataset for this project. It contains 4000 fingerprint images belonging to 2000 individuals. Although it's intended for fingerprint type classification, each

Fig. 1. NIST Special Database 4 Fingerprint Sample

TABLE I. NIST SPECIAL DATABASE 4 CLASS COUNTS

	Male	Female	Total
Right Thumb	314	70	384
Right Index	360	82	442
Right Middle	332	72	404
Right Ring	328	68	396
Right Little	292	60	352
Left Thumb	328	68	396
Left Index	352	86	438
Left Middle	344	88	432
Left Ring	322	74	396
Left Little	284	72	356
Total	3250	750	4000

one is tagged with gender information. Fingerprint images are 8-bit grayscale and are of size 512 x 512 pixels. Figure 1 shows one of the fingerprint images from the dataset.

One thing to note in this dataset is while the number of samples per finger class are relatively near with each other, it is highly imbalanced on the number of male and female volunteers. The number of male fingerprint samples are more than four times the number of female fingerprints. Counts per each class can be found on Table I.

The training set is composed of 50% of the samples for each finger class in each gender. The remaining 50% comprises the test set. This partitioning prevents further problems with the distribution as opposed to just randomly picking samples. Furthermore, it also provides insight on what finger best discriminate fingerprints between the genders.

B. CNN Architectures

Convolutional Neural Networks (CNNs) are multi-layer neural networks designed to recognize visual patterns directly from pixels of images. It is said to be robust to distortion and variabilities such as basic geometric transformations. Given this, images do not require a lot of preprocessing and no longer need a manual feature extraction phase.

The CAFFE [10] Deep Learning Framework is used for training various CNN Architectures. In CAFFE, they are defined in prototxt files and many of the popular models are available due to its active community. We have used and compared performance of the following architectures:

1) LeNet: One of the earliest CNN architectures is the LeNet developed and popularized by LeCun et al [13]. It was successfully used for hand and machine written character recognition. We have reason to believe it can be applicable too in an Image-based Fingerprint recognition system. The LeNet comprises of an input layer, layers of alternating convolution and pooling layers then closed by a fully connected output layer. The size of the network used is given by Table II.

2) AlexNet: In 2012, Krizhevsky et. al [11] won the Imagenet Large Scale Visual Recognition Challenge (ILSVRC) by a huge margin. Their model which is now known as AlexNet is a large, deep convolutional neural network trained on raw RGB pixel values. The neural network has 60 million parameters and 650,000 neurons. It consists of five convolutional layers, some of which are followed by max-pooling layers, and three globally-connected layers with a final 1000-way softmax for the object classes. They trained it in parallel using two NVIDIA Graphical Processing Units for about a week. To regularize the network, the fully connected layers were employed with hidden-unit "dropout" which was a recently developed method back then to reduce overfitting. Its architecture can be found in Table III.

3) Maxout Network: The Maxout activation function was introduced by Goodfellow and Bengio [7]. It was shown that this activation function compared to Rectified Linear Units (ReLU) yields better performance when the network is trained with Dropout. They've tested this in public datasets such as MNIST, CIFAR-10, CIFAR-100 and Street View House Numbers (SVHN) and had achieved state-of-art performance in 2013. The Maxout network used in this project is similar to Krizhevsky's [11] AlexNet model with ReLUs replaced by Maxout Activation Function as seen on Table IV.

4) Exponential Linear Unit (ELU) Network: Clevert et al. recently introduced the Exponential Linear Unit (ELU) activation function [3]. They've used an ELU network in the Imagenet Large Scale Visual Recognition Challenge (ILSVRC) 2015 and found that ELU networks considerably speed up learning compared to Rectifier networks with similar classification performance. It also yielded the best published result on CIFAR-100, without the use of multi-view evaluation or model averaging. Compared to Rectified Linear Units (ReLUs), ELUs have negative values which allows them to push mean unit activations closer to zero. Zero mean is said to speed up learning because they bring the gradient closer to the unit natural gradient. For this architecture in this project, our ELU network is basically the Maxout Architecture previously mentioned with Maxout activation functions replaced by ELUs in the convolutional layers. Table V shows this change.

IV. EXPERIMENT RESULTS

Models are trained and tested using the CAFFE [10] Deep Learning Framework in a laptop computer equipped with NVIDIA GeForce GTX 970M graphics card. All images are resized to 256 x 256 pixels for the networks' consumption.

A. Selection of CNN Architecture

Accuracy and Training Speed are the two criteria for choosing the CNN Architecture.

TABLE II. LENET ARCHITECTURE

Layer	1	2	3	4	5	Output
Stage	Conv	Max Pool	Conv	Max Pool	Full	Full
Spatial Input	256x256	252x252	126x126	122x122	61x61	1x1
# of Channels	20	20	50	50	500	2
Filter Size	5x5	2x2	5x5	2x2		
Stride	1	2	1	2		
Activation Function					ReLU	Softmax

TABLE III. ALEXNET ARCHITECTURE

Layer	1	2	3	4	5	6	7	8	9	10	Output
Stage	Conv	Max Pool	Conv	Max Pool	Conv	Conv	Conv	Max Pool	Full + Dropout	Full + Dropout	Full
Spatial Input	256x256	62x62	31x31	31x31	15x15	15x15	15x15	15x15	7x7	1x1	1x1
# of Channels	96	96	256	256	384	384	256	256	4096	4096	2
Filter Size	11x11	3x3	5x5	3x3	3x3	3x3	3x3	3x3			
Stride	4	2	2	2	1	1	1	2			
Padding Size			2		1	1	1				
Group			2			2	2				
Activation Function	ReLU		ReLU		ReLU	ReLU	ReLU		ReLU	ReLU	Softmax
Normalization	LRN		LRN								

TABLE IV. MAXOUT ARCHITECTURE

Layer	1	2	3	4	5	6	7	8	9	10	Output
Stage	Conv	Max Pool	Conv	Max Pool	Conv	Conv	Conv	Max Pool	Full + Dropout	Full + Dropout	Full
Spatial Input	256x256	62x62	31x31	31x31	15x15	15x15	15x15	15x15	7x7	1x1	1x1
# of Channels	96	96	256	256	384	384	256	256	4096	4096	2
Filter Size	11x11	3x3	5x5	3x3	3x3	3x3	3x3	3x3			
Stride	4	2	2	2	1	1	1	2			
Padding Size			2		1	1	1				
Group			2			2	2				
Activation Function	Maxout		Maxout		Maxout	Maxout	Maxout		Maxout	Maxout	Softmax

TABLE V. ELU ARCHITECTURE

Layer	1	2	3	4	5	6	7	8	9	10	Output
Stage	Conv	Max Pool	Conv	Max Pool	Conv	Conv	Conv	Max Pool	Full + Dropout	Full + Dropout	Full
Spatial Input	227x227	62x62	31x31	31x31	15x15	15x15	15x15	15x15	7x7	1x1	1x1
# of Channels	96	96	256	256	384	384	256	256	4096	4096	2
Filter Size	11x11	3x3	5x5	3x3	3x3	3x3	3x3	3x3			
Stride	4	2	2	2	1	1	1	2			
Padding Size			2		1	1	1				
Group			2			2	2				
Activation Function	ELU		ELU		ELU	ELU	ELU		Maxout	Maxout	Softmax

1) Accuracy: All CNN Architectures mentioned in Section III-B are trained and tested first with the raw imbalanced data set. This is to see how powerful the architectures are on their own. Three models are trained per architecture to reduce the effect of randomness in measuring performance. Each underwent 10,000 iterations in training phase. Table VI shows the Average Accuracy of the CNN architectures.

In classifying Male fingerprints which is the dominant class, all of them scored around 95%. On the other hand, Female fingerprints which is the undersampled class scored poorly for all the models. Overall accuracy for AlexNet, Maxout and ELU is at 88%, while LeNet gives lower score of 86%.

It can be seen that LeNet architecture wasn't able to beat pure guessing for Female class as it heavily favored on predicting Male class. While the accuracy of the other architectures AlexNet, Maxout and ELU Networks are still poor around 60% for the undersampled class, there is significant improvement seen from LeNet's. This is due to them implementing Dropout on the fully connected layers of the network. With Dropout, ensemble of models are generated for each training step and are averaged. This prevents overfitting thus somehow lowering the effect of the oversampled class. Moreover Lenet is an older and smaller model.

Maxout Network edged out AlexNet in Accuracy for both Male and Female classes. This is consistent with Goodfellow's claim [7] that Maxout networks are better partnered with Dropout over Rectifier networks such as AlexNet. Because of the nature of its activation function ReLU, rectifier networks exhibit diminished gradient flow on the deeper layers of the network. Perhaps if we have a bigger model, performance

TABLE VI. PERFORMANCE OF CNN MODELS WITH IMBALANCED TRAINING SET

	Average Accuracy (%)		
	Male	Female	Overall
LeNet	95.50	47.50	86.07
AlexNet	94.84	58.93	88.12
Maxout Network	94.98	59.64	88.45
ELU Network	94.91	61.06	88.48

TABLE VII. TRAINING TIME OF CNN MODELS FOR 10,000 ITERATIONS

	Average Training Time (minutes)
AlexNet	25.67
Maxout Network	31.00
ELU Network	31.00

between AlexNet and Maxout will be more apparent.

Both ELU and Maxout models employ Maxout activation function on the fully connected layers of the network. They differ in employing Exponential Linear Unit (ELU) activation function on the convolutional layers. ELU network has about the same Overall accuracy as Maxout but has generalized the Female class a little better.

2) Training Speed: At this point, AlexNet, Maxout and ELU Networks are the only ones considered based on results on Accuracy. Furthermore, LeNet doesn't have the same number of units with the other ones so it doesn't make sense to compare training time and validation loss with them.

For each remaining model, the training time is recorded for 10,000 iterations. The average training time of 10,000 iterations are shown in Table VII. It can be seen that the fastest is AlexNet. Maxout and ELU take almost the same time. Due to applying Maxout Activation function on their fully connected layer, their networks have more connections hence longer training time when comparing with the same number of n iterations.

Perhaps a more reliable way of comparing training of the networks is to see how fast they learn instead of actual training time for a set number of iterations. Figure 2 shows the test set loss as training progresses for the first 1000 iterations. Validation loss for ELU dips faster compared with the other two networks. ELUs are said to have improved learning characteristics compared to the units with other activation functions [3]. This suggests that ELU network will more likely need less number of iterations to learn the data.

B. Dealing with Imbalanced Dataset

The training of the best performing network Exponential Linear Unit (ELU) Network is improved by further addressing Imbalanced Data. Dropout which is an ensemble method has preliminarily increased performance of the Female fingerprint class. We've tried manipulating the training set by undersampling and oversampling.

In undersampling, we've randomly removed training instances from the Male class such that the two classes have the same number of instances. This led to reducing training

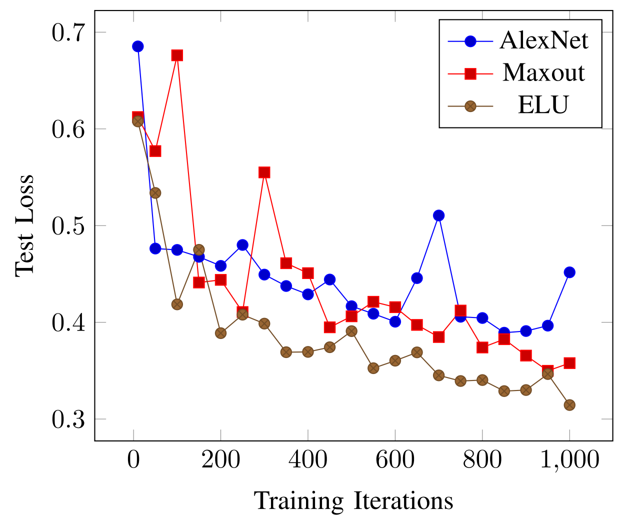

Fig. 2. Test Set Loss as Training Progresses

TABLE VIII. PERFORMANCE OF ELU NETWORK WITH DIFFERENT TRAINING SETS

	Accuracy (%)		
	Male	Female	Overall
Imbalanced Data	94.91	61.06	88.48
With Undersampling	58.56	84.40	63.04
With Data Augmentation	94.28	72.00	90.10

examples by about 4 times while the test set remains the same. Table VIII shows the effect of undersampling. Although the accuracy of the Female class went up, there's a steep decline of accuracy for the Male class thus dragging down Overall performance.

The easiest and most common method to reduce overfitting on image data is to artificially enlarge the dataset using label-preserving transformations [11]. For oversampling, we've used the out-of-the-box mirroring and cropping feature in the data layer in CAFFE. During the training phase random crops are taken and flipping is done for data augmentation. Patches of size 227 x 227 are randomly extracted from 256 x 256 training images. During test time, the network makes a prediction by extracting ten patches (the four corner patches and the center patch for the original and the horizontal mirror) and averaging the prediction made by the network for each patch.

The data augmentation procedure is one of Krizhevsky's methods used in his work on Imagenet Classification[11]. Table VIII supports the validity of this method. After data augmentation, classification performance for the Female class has risen significantly to 72% while still retaining high Accuracy for the Male class. Overall Accuracy also went up to more than 90%.

C. Effect of Finger Class to Gender Determination

The ELU network with Data Augmentation is the final classifier used to see which finger class best distinguishes a Male fingerprint from a female one. Table IX shows the results.

It can be seen that accuracy is highest overall with the middle fingers. Also, the undersampled Female class scored

TABLE IX. ACCURACY OF DATA AUGMENTED ELU NETWORK PER FINGER CLASS

	Accuracy (%)		
	Male	Female	Overall
Right Thumb	98.09	34.29	86.46
Right Index	92.22	65.86	87.33
Right Middle	99.40	80.56	96.04
Right Ring	95.03	79.49	92.00
Right Little	91.78	80.00	89.77
Left Thumb	96.95	47.06	88.38
Left Index	92.04	81.39	89.95
Left Middle	98.26	79.54	98.26
Left Ring	94.41	78.38	91.41
Left Little	83.10	88.89	84.27
Overall	94.28	72.00	90.10

above average relative to overall accuracy for Females in this finger class. This suggests that the middle finger is the best finger to use if we want a more balanced classification.

For female fingerprint classification, success rate is lowest with the thumbs and seem to be highest (combined) with the little fingers. Inversely, male fingerprint classification suffered the lowest accuracy on little fingers but scored highly on thumbs. This is somewhat in sync with the experiment results done by Gnanasivam and Muttan. [14]. This suggests that if we have a predisposition to the gender of the fingerprint owner the best fingers to use for verification are the little and thumb fingers.

V. CONCLUSION

In this project, we have successfully validated the usage of visual features in fingerprints in determining gender using Deep Learning. Convolutional Neural Networks (CNNs) have made this more convenient by eliminating the need to extract these features and automatically learn them. Various CNN architectures and methodologies have been developed in recent years with improving performance for image classification and there will probably be better ones in the future which will further improve performance of the results achieved in the experiments. Aside from visual features, the results suggest that the finger class also plays a role with gender classification.

There are still other methods known today which can probably further improve classification performance. Most image classification competitions today are dominated by ensembling several big models and performing model averaging. Also, there are preprocessing techniques as well which can be applied such as PCA whitening [11].

ACKNOWLEDGMENT

The authors would like to thank Dr. Prospero Naval for allowing us to work on this topic. This is done as a Mini Project Requirement for CS 295 Deep Learning course.

REFERENCES

[1] S. F. Abdullah, A. F. N. A. Rahman, and Z. A. Abas. Classification of gender by using fingerprint ridge density in northern part of malaysia. *ARPN Journal of Engineering and Applied Sciences*, 10(22):596 – 599, 2015.

[2] Nayak V. C., Rastogi P., Kanchan T., Yoganarasimha K., Kumar G. P., and Menezes R. G. Sex differences from fingerprint ridge density in chinese and malaysian population. *Forensic science International*, 197(1):67–69, 2010.

[3] Djork-Arne Clevert, Thomas Unterthiner, and Sepp Hochreiter. Fast and accurate deep network learning by exponential linear units (elus). *arXiv:1511.07289 [cs.LG]*, 2016.

[4] Ganesh B. Dongre and S. M. Jagade. A review and study on fingerprint based gender classification using classified techniques. *International Journal of Scientific & Engineering Research*, 6:596 – 599, 2015.

[5] Gutierrez-Redomero E., Alonso C., Romero E., and Galera V. Variability of fingerprint ridge density in a sample of spanish caucasians and its application to sex determination. *Forensic Science International*, 180(1):17–22, 2008.

[6] Gutierrez-Redomero E., Alonso M. C., and Dipierri J. Sex differences in fingerprint ridge density in the mataco-mataguayo population. *HOMO-Journal of Comparative Human Biology*, 62(6):487–499, 2011.

[7] Ian Goodfellow, David Warde-Farley, Mehdi Mirza, Aaron Courville, and Yoshua Bengio. Maxout networks. *arXiv:1302.4389v4 [stat.ML]*, 2013.

[8] S. S. Gornale, Basavanna M., and Kruthi R. Gender classification using fingerprints based on support vector machines (svm) with 10-cross validation technique. *International Journal of Scientific & Engineering Research*, 6(7):588 – 593, 2015.

[9] Wang JF, Lin CL, Chang YH, Nagurka ML, Yen CW, and Yeh C. Gender determination using fingertip features. *Internet Journal of Medical Update*, 3(2):22 – 28, 2008.

[10] Yangqing Jia, Evan Shelhamer, Jeff Donahue, Sergey Karayev, Jonathan Long, Ross Girshick, Sergio Guadarrama, and Trevor Darrell. Caffe: Convolutional architecture for fast feature embedding. *arXiv preprint arXiv:1408.5093*, 2014.

[11] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. Imagenet classification with deep convolutional neural networks. In F. Pereira, C. J. C. Burges, L. Bottou, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 25*, pages 1097–1105. Curran Associates, Inc., 2012.

[12] L. Kumar, S. Agarwal, R. Garg, A. Pratap, and V. Mishra. Gender determination using fingerprints in the region of uttarakhand. *J Indian Acad Forensic Med October-December 2013*, 35(4):596 – 599, 2013.

[13] Yann LeCun, Leon Bottou, Yoshua Bengio, and Patrick Haffner. Gradient-based learning applied to document recognition. In *Proceedings of the IEEE*, 1998.

[14] Gnanasivam P. and Muttan S. Fingerprint gender classification using wavelet transform and singular value decomposition. *arXiv:1205.6745 [cs.CV]*, 2012.

[15] Ruxin Wang, Congying Han, Yanping Wu, , and Tiande Guo. Fingerprint classification based on depth neural network. *arXiv:1409.5188v1 [cs.CV]*, 2014.