Quality of Pre-trained Deep-Learning Models for Palmprint Recognition

Valentin Rosca and Anca Ignat

September 7, 2020

Quality of Pre-trained Deep-Learning Models for Palmprint Recognition

Valentin Roșca
Faculty of Computer Science
University "Alexandru Ioan Cuza" of Iaşi
Iași, România
Email: valentin.b.rosca@info.uaic.ro

Anca Ignat
Faculty of Computer Science
University "Alexandru Ioan Cuza" of Iaşi
Iași, România
Email: ancai@info.uaic.ro

***Abstract*—This paper presents a comprehensive study of deep learning methods and datasets used for solving the palmprint recognition problem. The quality of image embeddings provided by deep neural networks, pre-trained on the ImageNet dataset, are evaluated on palmprint recognition in the visible spectrum task. In our tests, we used twelve publicly available datasets obtained with different types of acquisition procedures: constrained, partially constrained and unconstrained. Sixteen convolutional neural networks (two from the VGG family, six from ResNet, three from Inception, two from MobileNet and three from DenseNet) were evaluated. We analyzed the results from the point of view of specialization potential, dataset difficulty and general parameter tuning. For evaluation, EER (Equal Error Rate) was employed. We ranked the datasets and appraised the feature vectors computed by the pre-trained networks using this metric. The best results, on average, were provided by the deep neural networks from the MobileNet family. The distances used for comparing the feature vectors were Euclidean, Cityblock, cosine and correlation. The best results were obtained with the cosine family distances.**

***Keywords*-convolutional neural networks; deep learning; transfer learning; palmprint recognition; palmprint datasets;**

I. Introduction

Since the advent of technology, biometric identification systems have become an integral part of our daily lives. Unlike keys and tokens, they cannot get lost, and, unlike passwords, they cannot be forgotten. The development of such a system is a matter of daily convenience and security. There are multiple types of biometric features used for the recognition of humans. The most popular ones are, in order, fingerprints, facial, iris, palmprints and voice biometric data [1]. In this article, the focus is on the increasingly popular palmprint biometrics.

Palmprint recognition is a highly accurate, efficient and promising biometric technique for identity authentication that has recently gained more popularity and user acceptance due to privacy issues brought up by other biometric identification systems. The palmprint covers a greater area of the human body than the fingerprint, making it much easier to capture and is considered more robust [17]. The identification process allows the use of any low-resolution camera present in most mobile devices, laptops, and even some desktops instead of specialized sensors. Palmprints recognition is less intrusive than more popular methods of recognition like voice or face recognition as it is harder to identify from public sources without consent, e.g., CCTVs or public videos, because of obstructions, resolution and viewing angles; they are more difficult to counterfeit than faces as they are not as publicly displayed.

Palmprint identification can be easily integrated into existing systems like Augmented and Virtual Reality headsets, driving systems, home biometrics, gesture tracking systems, and any other system that requires the use of the hands. Therefore, their potential use is excellent as long as the system in place is reliable.

In this paper, we evaluate several neural network architectures pre-trained on the ImageNet dataset to obtain both baselines for the recognition rate, as well as insights into the current state of the research of palmprint recognition in the visible spectrum (spectral images at different wavelengths or 3D images are not considered). We also analyze the degrees of difficulty of the available datasets used in the literature and assess the importance of different training parameters. The article is organized as follows. Section 2 reviews datasets available in the literature and presents the state of the art methods. Section 3 describes the evaluation pipeline. Section 4 provides the evaluation results and discusses the importance of each dataset and each parameter. Finally, in Section 5, we derive a conclusion and in Section 6, we present the direction for our future work.

II. Datasets

A biometric recognition system can either validate a person's biometric or identify a person based on its biometric. Such a system typically creates a biometric template, either implicitly or explicitly, for each person, using a database. Then, using the biometric template, the system can either verify the authenticity or identify a person based on its biometric. The database is created from a dataset containing multiple biometric features, each associated with a person. In public datasets, the person's identities are anonymized.

A. Dataset Types

Palmprint recognition datasets are divided into three categories, as described in [30], based on the restrictions imposed by the acquisition process.

The first category is constrained acquisition. In this category, restrictions are imposed on hand placement, background and devices. The fingers must be straightened and separated, the background light should be as dim as possible and the photo has to be taken from only one device. Furthermore, the acquisition is either touch-based, using the hand to touch a scanner, or touch-less, with digital devices. Therefore, these datasets represent a rigorously controlled environment, where feature extraction is the primary goal.

The second category is partially constrained acquisition. This category extends the former by removing one of the restrictions, i.e., either the background is unrestricted, e.g., the pictures may be taken outdoor, either the positioning of the hand is unrestricted, e.g., the fingers may be close to each other and the hand may be rotated, either more than one device must be used. Therefore, these types of datasets represent a partially restricted environment, where the goal is not only the feature extraction but also asserting the robustness of the algorithm under different environmental changes, a more realistic setting.

Finally, the third category is unconstrained acquisition. This category represents a real-world scenario with an unconstrained background, hand positioning, and multiple devices for acquisition, an actual unconstrained environment.

B. Available Datasets

The oldest dataset, released in 2003, publicly available for research is PolyU (version 2) from the Hong Kong Polytechnic University [31], a constrained dataset. The background of the images is dark, with little to no lighting for each image. This dataset is one of the two touch-based datasets available. No left/right images are provided. Samples of the dataset are available in Figure 1.

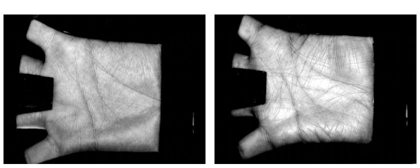
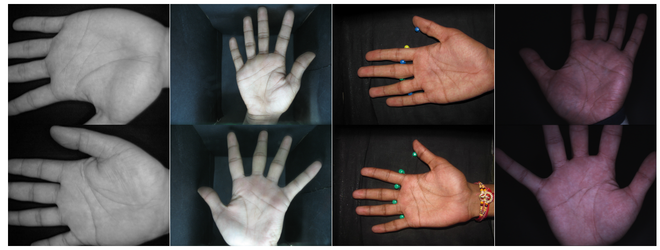


Figure 1: PolyU dataset samples from two individuals

Newer constrained databases are based on touch-less acquisition. The background lighting is more noticeable than the touch-based ones, therefore the image is much noisier, closer to a real life scenario. Some similar datasets of this category are Casia [23] (2005), IITD (version 1) [24] (2006), COEP [5] (2010), GDPS [7] (2011) are Tongji [32] (2017). In Figure 2 samples of each dataset are shown.



(a) Casia (b) IITD (c) COEP (d) Tongji

Figure 2: Samples from similar constrained datasets

Casia samples are similar to PolyU samples, but they also preserve hand geometry. IITD has much more lighting in the background, i.e., more noise is present. COEP is similar to Casia, with added landmarks for easier identification of the hand positioning for smoother ROI segmentation. Tongji provided images with poor lighting conditions and illumination, similar to real-life scenarios.

In 2017, one of the first large-scale partially constrained dataset was released. This dataset, 11k-Hands, provides both front and back pictures of hands, as well as gender information. The dataset provided both left/right hand images. The hand positioning is unconstrained. In this thesis, we used only the images of palmprints, frontal images. Samples of the dataset are available in Figure 3.

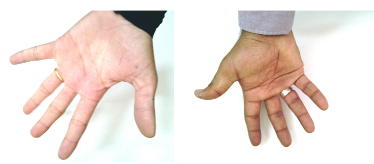


Figure 3: 11k-Hands dataset samples from two individuals

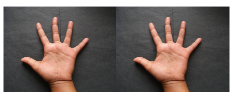
One of the recently released datasets (2019) is the PolyU-IITD dataset. This dataset is made up of two parts, treated as individual datasets. The first part is the constrained dataset PolyU-IITD-Original. Palmprints and hand geometry are cleaner than the other constrained datasets. The second part is the PolyU-IITD-TwoSessionChallenge dataset, a small dataset created in two sessions. Images from the second session are either taken many years apart or present paint spots over them. Samples of the dataset are available in Figure 4.

Birjand University Mobile Palmprint Database (BMPD) [15] and Sapienza University Mobile Palmprint Database (SMPD) [16], both released in 2019, are two partially constrained mobile databases for palmprint recognition. Both datasets have an unconstrained background. SMPD is richer than BMPD as it provides four different orientations of the hand instead of two, but BMPD provides

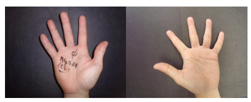
Year	Dataset	Acquisition	Extra	Device	Hands	Images	Format	Resolution
2003	PolyU (V2)	C	touch-based	scanner	386	7752	bmp	4:3
2005	Casia (V1)	C	touch-less	digital camera	624	5502	jpg	4:3
2006	IITD (V1)	C	touch-less	digital camera	470	3290	jpg	4:3
2010	COEP	C	touch-less	digital camera	168	1344	jpg	4:3
2011	GDPS	C	touch-less	two webcams	100	1000	bmp	4:3
2017	Tongji	C	touch-less	digital camera	1200	12000	tiff	4:3
2017	11k-Hands	PC	hand position	digital camera	380	11076	jpg	4:3
2018	PolyU-IITD-Original	C	touch-less	two digital cameras	700	12326	jpg	4:3
2018	PolyU-IITD-TwoSessionChallenge	UC	difficult	two digital cameras	35	1126	-	-
2019	BMPD	PC	background	smartphone camera	41	1640	jpg	4:3
2019	SMPD	PC	background	smartphone camera	100	4400	jpg	4:3
2019	MPD (V2)	UC	mobile	two smartphone cameras	200	16000	jpg	4:3

C - constrained PC - partialy constrained UC - unconstrained

Table I: Available datasets



(a) PolyU-IITD-Original

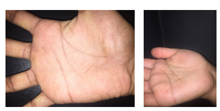


(b) PolyU-IITD-TwoSessionChallenge

Figure 4: PolyU-IITD dataset samples from two different individuals

both left/right-hand images. Samples of the dataset are available in Figure 5.

(a) BMPD



(b) SMPD

Figure 5: Samples from mobile partially constrained datasets

Lastly, the MPD dataset is an unconstrained dataset, the only dataset of this type that we analyzed. The images are taken using two mobile devices. Left/right-hand images are available. Samples of the dataset are available in Figure 6.

Figure 6: MPD dataset samples from two individual

A list of different statistics for all the mentioned datasets are listed in Table I.

III. Literature review

The topic of palmprint recognition has been treated many times by different authors. The compressive study of [33] describes results from many palmprint recognition methods in great detail. In their work, the authors separate those methods into five categories, namely, encoding, structure-based, statistical-based, subspace, and deep learning. We refer to the first four as conventional methods.

A. Conventional methods

Encoding type methods encode the palmprint image to a code that is robust to translations and rotations and use the code for recognition. Structure-based methods are more standard methods that use information such as the orientation and position of ridges and minutiae and other feature points such as palm lines. Statistical-based methods use the mean and variance and other moments in the image as hallmarks for recognition. Subspace methods use the projection of high-dimensional data, the image in this case, to low dimensional data. Some of the top methods are presented in Table II.

Method	Type	Dataset	EER	RR
CR_CompCode [19]	Encoding	Custom	-	98.78%
LLDP [22]	Structure	PolyU-V2	-	100%
High order moments [3]	Statistics	IITD-V1	7.89%	92.32%
		Casia	1.81%	98.66%
		PolyU-V2	1.0%	99.29%
LPDP [8]	Subspace	PolyU-V2	-	99.7%

Table II: Conventional methods

There are two main issues concerning conventional methods. The first one is their dependence on a good ROI (Region of Interest) segmentation, otherwise known as detecting the palm's location in the image. The second one is that conventional methods tend to be highly tuned for a single particular dataset.

B. Deep learning

The more recent category of methods uses neural networks and deep learning techniques. These techniques are gaining momentum in the domain of palmprint recognition as well as other domains for their flexibility and representational ability. In [14], the ROI and feature extraction is done using transfer learning by AlexNet networks. The final classification is done using KNN, RandomForest and SVMs.

Recent papers use neural network architecture to train an end-to-end recognition system. In [6], the authors use a smaller MobileNet ([12]) architecture trained from scratch on the PolyU database and in [34], the authors use siamese neural networks ([18]) with the VGG-16 ([27]) architecture to train the system. Lastly, a top solution that uses neural networks is provided in [21]. The authors use a slightly modified triplet network ([11]) model with a loss function that takes translations of the images into account during training by attempting to compute the minimum distance between multiple translated feature maps obtained from one image and the original from the other. Most importantly, the training is done on the left hands of IITD and the network is then tested on multiple datasets. The paper mentioned earlier is the only paper we could find where cross-database testing is used, which provides much more valuable insights into the system's ability.

Method	Dataset	EER	RR
AlexNet Convolutions + K-NN[14]	PolyU	0.0156%	-
	IITD	0.0328%	-
AlexNet Convolutions + RF [14]	PolyU	0.0625%	-
	IITD	0.0889	-
AlexNet Convolutions + SVMs[14]	PolyU	0.0125%	-
	IITD	0.0276%	-
MobileNet [6]	PolyU	-	99.96%
VGG-16 Siamese [27]	PolyU	0.28%	-
	XJTU	4.559%	-
Triplet CNN [11]	IITD right hands	0.6%	99.2
	PolyU-IITD	0.267%	98.6
	Casia	0.51%	-

IV. Evaluation of Datasets and Pre-trained Networks

Transfer learning is a machine learning technique in which parts of a model trained for one task (the base task) are used to train on another task (the target task). The scope of the technique is transfer learning. If a model is trained on a general task, then the information it encodes should be useful for a more particular task. Also, a model trained on a similar task should encode useful general information. Therefore, the base task should be either a general task or a task related to the target task. In the context of deep-learning, transfer learning is used by initializing part of a neural network with weights from another neural network trained on another dataset, typically a more nonspecific dataset. Both the network for the target task and the base task should have similar architectures, at least for the initialized part. The biggest and most popular dataset is ImageNet [25], with hundreds and thousands of images depicting various objects. It is the most nonspecific image dataset available. A lot of the popular neural network architectures are trained on ImageNet and the weights are made public. Therefore, neural network frameworks offer easy integration with pre-trained models.

In this section, the evaluation of different image embeddings created using ImageNet pre-trained models is presented. The image embeddings or the weights of the pre-trained models can be further used for training on the palmprint recognition task.

There are two general approaches to transfer learning. One approach uses a part of a pre-trained model, typically the last convolutional layer of one of the fully connected ones, to obtain image features from the target dataset, then train a classifier on the features obtained. The other approach to add fully connected layers on top of a pre-trained model, typically on top of either the last convolutional layer or the fully connected layer before the final one. Then, part of the network or the whole network is trained on the target dataset.

For both approaches, several decisions can either make or break the target model. One such decision is the base neural network architecture that is used. This architecture depends mainly on the target dataset and not on the base one, as each architecture has its drawbacks. Another design consideration is the similarity between the target dataset and the base. If there is no similarity, transfer learning will often provide poor results.

Besides transfer learning design consideration, when building a neural network for new datasets, parameters such as image resolution, distance metric and prepossessing are essential, as well as which target dataset to choose. In this section, we conduct experiments to compare the performance on the task of palmprint identification of features extracted from the last convolutional layer of sixteen neural network architectures pre-trained on the ImageNet dataset on the twelve datasets previously explored, without any training. We also analyze the discrimination ability of the image embeddings under different dataset parameters and network parameters. The main objectives of our experiments are:

- evaluating the specialization potential and knowledge transfer ability of pre-trained neural networks on palmprint datasets
- ranking the difficulty of datasets for our future work
- setting baselines for deep learning models for our future work
- asserting the importance of general parameters

The experiments are designed as follows. Multiple image resolutions are selected and each one of the twelve datasets is resized accordingly. Then two different prepossessing schemes are applied to the resized images. Then, sixteen neural networks are loaded and the feature maps of the last convolutional layer are fetched. Then, a popular pooling method is applied to the feature maps to obtain the final image embeddings. From the image embeddings, by choosing a distance metric, a pairwise distance matrix is created. A boolean pairwise label matrix is also computed from the initial labels. A truth value (1) in the pairwise label matrix on column i and j is given if the label at position i is equal to the label at position j. The final score is the equal error rate computed using the lower triangular portion of each matrix. Samples from the same person should be as close as possible and samples from different persons should be as far as possible in the feature space. This scoring method allows us to evaluate the potential use of the image embeddings for training neural networks.

The evaluation process is designed as a staged pipeline with the following stages:

1) Dataset resize
   Resize, center and pad images in the dataset to the following resolutions:
   - 224x224
     - Native resolution of the majority of the pre-trained models
   - 299x299
     - Native resolution of the rest of the models
   - 320x240, 360x270, 400x300
     - Resolution that preserves the aspect ration of majority of the datasets while also keeping the computational cost of the networks reasonable. See Section II for details.
2) Preprocessing
   Normalize the images
   - standard
     - Rescale of the pixels from [0, 255] to [-1, 1]
   - original
     - Apply the preprocessing used on ImageNet of the network
3) Network selection
   Use the convolutional layers, with the weights from ImageNet, to extract features for each resolution and preprocessing. Five families of architectures were used:
   - VGG family
     - VGG16 [27], VGG19 [27]
   - ResNet family
     - ResNet50 [9], ResNet101 [9], ResNet152 [9]
     - ResNet50V2 [10], ResNet101V2 [10], ResNet152V2 [10]
   - Inception family
     - Xception [4], InceptionV3 [29]
     - InceptionResNetV2 [28]
   - MobileNet family
     - MobileNet [12], MobileNetV2 [26]
   - DenseNet family
     - DenseNet121 [13], DenseNet169 [13], DenseNet201 [13]
4) Pooling
   Extract the final image embedding by pooling the feature maps. The following pooling were used:
   - GlobalMaxPooling [20]
     - Pool the max value for each channel
   - GlobalAveragePooling [20]
     - Pool the average value for each channel
   - GlobalMaxPooling + GlobalAveragePooling
     - Apply both poolings and concatenate the results
5) Distance
   Compute the distance matrix of the image embeddings for each dataset. The distance functions used have the non-negativity, identity of indiscernibles and similarity properties.
   - ℓ^p family - euclidean, cityblock
   - cosine family - cosine, correlation
6) Evaluation
   Compute the equal error rate (EER) using the lower (or upper) triangular portion (excluding the diagonal) of the distance matrix. Samples from the same person should be as close as possible and samples from different persons should be as far as possible in the feature space. This scoring method allows us to evaluate the potential use of the feature maps for training neural networks.

In total, 18432 configurations were tested. Each step was run individually, on input from the previous step. The results offer meaningful information about the quality of ImageNet embeddings for palmprint recognition. Those embeddings are especially useful in the context of Siamese or Triplet networks because the equal error rate correlates with the quality of the embedding space.

V. Results

The TensorFlow [2] library was used for loading and running the pre-trained models. The distance and equal error rate evaluation were computed using our custom CUDA implementations. The evaluation pipeline took six days to complete on an FX-8120 system with a single GTX 1080 Ti GPU.

One of the evaluations' objectives is to evaluate the quality of the image embeddings obtained from the pre-trained networks. The mean and standard deviations of the equal error rates obtained are presented in Figure 7, using data from our experiments. For each configuration, the preprocessing that had provided the best results were selected. The MobileNet family has the best results across

Figure 7: Point plot of the equal error rates obtained for each network on the available datasets. The MobileNet family has the lowest overall error rate, followed by the VGG family. The ResNet and Inception families of models provide much higher error rates on constrained datasets. On partially constrained and unconstrained datasets, the differences start to diminish.

the board, followed by the VGG family. One reason for this trend is the simplicity of the convolutional kernel, imposed by the *separable convolution* layer, used in the MobileNet family of architectures. Similar kernels are used in conventional methods for feature extraction. Furthermore, from Figure 7, the difficulty of each dataset can be inferred. Unconstrained datasets are more complicated, in our context, than partially constrained ones, while partially constrained datasets are more complicated than constrained datasets. This result is motivated by the restrictions in the acquisition process. The MobileNet family's results on the hardest partially constrained and unconstrained analyzed datasets, plotted by the image resolution, distance and best pooling, are presented in Figure 8. Empirically, the best overall distance metric, on all the datasets, is the cosine distance, closely followed by the correlation distance. However, image resolution depends on the particular dataset. For our single unconstrained dataset, MPD, a larger input image did not yield any decrease in the equal error rate. On the other hand, for constrained and partially constrained datasets, an increase in image resolution yields a small decrease in the error rate at the expense of computational time. The pooling method is entirely network and dataset dependent. As for preprocessing, this is mostly architecture-dependent, with the ResNet preferring the original preprocessing, while the other architectures yielded better results using a standard preprocessing. The best configuration for each dataset is presented in Table III.

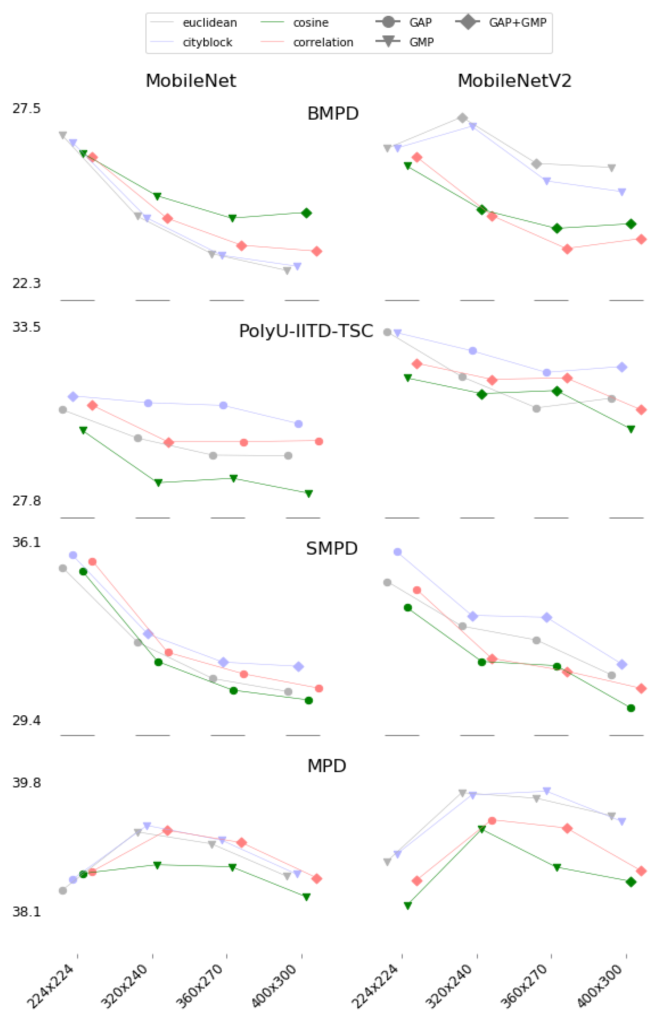


Figure 8: The lowest equal error rates of the MobileNet family for each image resolution and distance on the hardest datasets

dataset	PolyU-IITD-O	IITD	COEP	Casia	GPDS	Tongji	PolyU	11k-Hands	BMPD	PolyU-IITD-TSC	SMPD	MPD
Network	VGG16	DenseNet201	MobileNet	MobileNet	MobileNet	MobileNet	DenseNet201	MobileNetV2	VGG16	VGG16	MobileNetV2	ResNet101
Image Resolution	400x300	400x300	400x300	400x300	360x270	400x300	224x224	400x300	360x270	320x240	400x300	224x224
Preprocessing	original	original	standard	standard	original	standard	standard	original	original	standard	original	original
Pooling	GAP+GMP	GAP	GAP	GAP+GMP	GAP+GMP	GAP+GMP	GAP	GMP	GMP	GMP	GAP	GAP
Embedding size	1024	1920	1024	2048	2048	2048	1920	1280	512	512	1280	2048
Distance	cosine	cityblock	cosine	cityblock	cityblock	cityblock	cosine	cosine	cosine	euclidean	cosine	cityblock
Equal Error Rate	**2.94**	**5.15**	**5.62**	**5.65**	**6.91**	**9.44**	**11.07**	**14.18**	**22.22**	**25.57**	**29.78**	**37.14**

Table III: Best configuration for each dataset

VI. Conclusion and Future work

Deep-learning frameworks already reached a point of maturity in the frame of conventional deep-learning models, with many architectures and pre-trained networks available. However, the transferability of knowledge embedding in deep neural networks trained on large image classification datasets is not yet assured. In this work, we analyzed in-depth the behavior of sixteen convolutional neural networks on twelve palmprint datasets, publicly available, using weights from pre-trained models on the large ImageNet dataset. Pre-trained models provide image embeddings of a higher quality for datasets subject to various constraints in the acquisition process than less constrained palmprint recognition datasets. Therefore, pre-trained deep-learning models can extract palmprint-related features without any additional training but are highly susceptive to noise added by the ambient background, translations and rotations. The least noise susceptive architecture was the MobileNet architecture, which, similarly to conventional methods, uses less complicated convolutional kernels via separable convolutions. Based on the observations presented in our paper, we plan to do a large scale training of Deep Bayesian Neural Networks for Palmprint Recognition. All the empirical data obtained in our experiments will be used to decide on significant parameters such as image resolution, pooling, network architecture and datasets.

References

[1] URL: https://identity.utexas.edu/assets/uploads/publications/Current-Biometric-Adoption-and-Trends.pdf.

[2] Martín Abadi, Ashish Agarwal, Paul Barham, Eugene Brevdo, Zhifeng Chen, Craig Citro, Greg S. Corrado, Andy Davis, Jeffrey Dean, Matthieu Devin, Sanjay Ghemawat, Ian Goodfellow, Andrew Harp, Geoffrey Irving, Michael Isard, Yangqing Jia, Rafal Jozefowicz, Lukasz Kaiser, Manjunath Kudlur, Josh Levenberg, Dandelion Mané, Rajat Monga, Sherry Moore, Derek Murray, Chris Olah, Mike Schuster, Jonathon Shlens, Benoit Steiner, Ilya Sutskever, Kunal Talwar, Paul Tucker, Vincent Vanhoucke, Vijay Vasudevan, Fernanda Viégas, Oriol Vinyals, Pete Warden, Martin Wattenberg, Martin Wicke, Yuan Yu, and Xiaoqiang Zheng. TensorFlow: Large-scale machine learning on heterogeneous systems, 2015. Software available from tensorflow.org. URL: https://www.tensorflow.org/.

[3] G. S. Badrinath, Naresh K. Kachhi, and Phalguni Gupta. Verification system robust to occlusion using low-order zernike moments of palmprint sub-images. *Telecommunication Systems*, 47(3):275–290, 2011. doi:10.1007/s11235-010-9318-y.

[4] Francois Chollet. Xception: Deep learning with depthwise separable convolutions. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul 2017. URL: http://dx.doi.org/10.1109/CVPR.2017.195, doi:10.1109/cvpr.2017.195.

[5] Pune-411005 (An Autonomous Institute of Government of Maharashtra) College of Engineering. Coep palm print database. Dataset accessible at: https://www.coep.org.in/resources/coeppalmprintdatabase. URL: http://www.coep.org.in/resources/coeppalmprintdatabase.

[6] Xueqiu Dong, Liye Mei, , and Junhua Zhang. Palmprint recognition based on deep convolutional neural networks. 2018.

[7] Miguel Ferrer, J. Vargas-Bonilla, and Aythami Morales. Bispectral contactless hand based biometric identification device. 05 2011. Dataset accessible at: http://www.gpds.ulpgc.es/downloadnew/download.htm.

[8] Jie Gui, Wei Jia, Ling Zhu, Shu-Ling Wang, and De-Shuang Huang. Locality preserving discriminant projections for face and palmprint recognition. *Neurocomputing*, 73(13):2696 – 2707, 2010. Pattern Recognition in Bioinformatics Advances in Neural Control. URL: http://www.sciencedirect.com/science/article/pii/S0925231210002213, doi:https://doi.org/10.1016/j.neucom.2010.04.017.

[9] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2016. URL: http://dx.doi.org/10.1109/cvpr.2016.90, doi:10.1109/cvpr.2016.90.

[10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Identity mappings in deep residual networks. *Lecture Notes in Computer Science*, page 630–645, 2016. URL: http://dx.doi.org/10.1007/978-3-319-46493-0_38, doi:10.1007/978-3-319-46493-0_38.

[11] Elad Hoffer and Nir Ailon. Deep metric learning using triplet network. *Lecture Notes in Computer Science*, page 84–92, 2015. URL: http://dx.doi.org/10.1007/978-3-319-24261-3_7, doi:10.1007/978-3-319-24261-3_7.

[12] Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications, 2017. arXiv:1704.04861.

[13] Gao Huang, Zhuang Liu, Laurens Van Der Maaten, and Kilian Q. Weinberger. Densely connected convolutional networks. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jul 2017. URL: http://dx.doi.org/10.1109/CVPR.2017.243, doi:10.1109/cvpr.2017.243.

[14] Mahdieh Izadpanahkakhk, Seyyed Razavi, Mehran Taghipour-Gorjikolaie, Seyyed Zahiri, and Aurelio Uncini. Deep region of interest and feature extraction models for palmprint verification using convolutional neural networks transfer learning. *Applied Sciences*, 8(7):1210, Jul 2018. URL: http://dx.doi.org/10.3390/app8071210, doi:10.3390/app8071210.

[15] Mahdieh Izadpanahkakhk, Aurelio Uncini, Seyyed Zahiri, Mehran Taghipour, and Seyyed Razavi. Novel mobile palmprint databases for biometric authentication. *International Journal of Grid and Utility Computing*, 1:1, 01 2019. doi:10.1504/IJGUC.2019.10019524.

[16] Mahdieh Izadpanahkakhk, Aurelio Uncini, Seyyed Zahiri, Mehran Taghipour, and Seyyed Razavi. Novel mobile palmprint databases for biometric authentication. *International Journal of Grid and Utility Computing*, 1:1, 01 2019. doi:10.1504/IJGUC.2019.10019524.

[17] Anil Jain, Arun Ross, and Salil Prabhakar. An introduction to biometric recognition. *Circuits and Systems for Video Technology, IEEE Transactions on*, 14:4 – 20, 02 2004. doi:10.1109/TCSVT.2003.818349.

[18] Gregory R. Koch. Siamese neural networks for one-shot image recognition. 2015.

[19] A. W. . Kong and D. Zhang. Competitive coding scheme for palmprint verification. In *Proceedings of the 17th International Conference on Pattern Recognition, 2004. ICPR 2004.*, volume 1, pages 520–523 Vol.1, 2004.

[20] Min Lin, Qiang Chen, and Shuicheng Yan. Network in network, 2013. arXiv:1312.4400.

[21] Yang Liu and Ajay Kumar. A deep learning based framework to detect and recognize humans using contactless palmprints in the wild, 2018. arXiv:1812.11319.

[22] Yue-Tong Luo, Lan-Ying Zhao, Bob Zhang, Wei Jia, Feng Xue, Jing-Ting Lu, Yi-Hai Zhu, and Bing-Qing Xu. Local line directional pattern for palmprint recognition. *Pattern Recognition*, 50:26 – 44, 2016. URL: http://www.sciencedirect.com/science/article/pii/S0031320315003131, doi:https://doi.org/10.1016/j.patcog.2015.08.025.

[23] The Chinese Academy of Sciences. Casia palmprint database. Dataset accessible at: http://biometrics.idealtest.org/dbDetailForUser.do?id=5. URL: http://biometrics.idealtest.org.

[24] Indian Institute of Technology Delhi. Iit delhi touchless palmprint database (version 1.0). Dataset accessible at: http://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Palm.htm. URL: https://www4.comp.polyu.edu.hk/~csajaykr/IITD/Database_Palm.htm.

[25] Olga Russakovsky, Jia Deng, Hao Su, Jonathan Krause, Sanjeev Satheesh, Sean Ma, Zhiheng Huang, Andrej Karpathy, Aditya Khosla, Michael Bernstein, Alexander C. Berg, and Li Fei-Fei. Imagenet large scale visual recognition challenge, 2014. arXiv:1409.0575.

[26] Mark Sandler, Andrew Howard, Menglong Zhu, Andrey Zhmoginov, and Liang-Chieh Chen. Mobilenetv2: Inverted residuals and linear bottlenecks. *2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Jun 2018. URL: http://dx.doi.org/10.1109/CVPR.2018.00474, doi:10.1109/cvpr.2018.00474.

[27] Karen Simonyan and Andrew Zisserman. Very deep convolutional networks for large-scale image recognition, 2014. arXiv:1409.1556.

[28] Christian Szegedy, Sergey Ioffe, Vincent Vanhoucke, and Alex Alemi. Inception-v4, inception-resnet and the impact of residual connections on learning, 2016. arXiv:1602.07261.

[29] Christian Szegedy, Vincent Vanhoucke, Sergey Ioffe, Jon Shlens, and Zbigniew Wojna. Rethinking the inception architecture for computer vision. *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Jun 2016. URL: http://dx.doi.org/10.1109/CVPR.2016.308, doi:10.1109/cvpr.2016.308.

[30] Adrian-Stefan Ungureanu, Saqib Salahuddin, and Peter Corcoran. Toward unconstrained palmprint recognition on consumer devices: A literature review. *IEEE Access*, 8:86130–86148, 2020. URL: http://dx.doi.org/10.1109/ACCESS.2020.2992219, doi:10.1109/access.2020.2992219.

[31] The Hong Kong Polytechnic University. Polyu palmprint palmprint database. URL: http://www.comp.polyu.edu.hk/~biometrics/.

[32] Lin Zhang, Lida Li, Anqi Yang, Ying Shen, and Meng Yang. Towards contactless palmprint recognition: A novel device, a new benchmark, and a collaborative representation based identification approach. *Pattern Recognition*, 69:199 – 212, 2017. URL: http://www.sciencedirect.com/science/article/pii/S0031320317301681, doi:https://doi.org/10.1016/j.patcog.2017.04.016.

[33] Dexing Zhong, Xuefeng Du, and Kuncai Zhong. Decade progress of palmprint recognition: A brief survey. *Neurocomputing*, 08 2018. doi:10.1016/j.neucom.2018.03.081.

[34] Dexing Zhong, Yuan Yang, and Xuefeng Du. Palmprint recognition using siamese network. In Jie Zhou, Yunhong Wang, Zhenan Sun, Zhenhong Jia, Jianjiang Feng, Shiguang Shan, Kurban Ubul, and Zhenhua Guo, editors, *Biometric Recognition*, pages 48–55, Cham, 2018. Springer International Publishing.