A Deep Learning Approach for Efficient Palm Reading

Conference Paper · December 2020

DOI: 10.1109/TAAI51410.2020.00039

Suvajit Acharjee
National Chiao Tung University

Sirapop Nuannimnoi
National Chiao Tung University

A Deep Learning Approach for Efficient Palm Reading

Suvajit Acharjee
EECS International Graduate Program
National Chiao Tung University
Hsinchu, Taiwan
saisuvajit9.eic05g@nctu.edu.tw

Sirapop Nuannimnoi
EECS International Graduate Program
National Chiao Tung University
Hsinchu, Taiwan
sirapop.eic07g@nctu.edu.tw

Ching-Yao Huang
Electronics Engineering
National Chiao Tung University
Hsinchu, Taiwan
cyhuang@mail.nctu.edu.tw

***Abstract*— Palmistry or palm reading is the art of foretelling and characterizing persons through the study of palm lines and patterns. However, this field is still not much technically developed and a palmist has to analyze palms personally and manually. In this paper, we have proposed a deep learning approach to automatically detect patterns inside palm images. Our proposed automated palm reader can effectively detect and classify a user's palm according to our predefined labels.**

Keywords—Palmistry; Palm reading; Deep Learning; Object detection; Convolutional Neural Network; Multi-label classification

I. Introduction

Palmistry is an art to forecast people's future regarding their behavior, characteristic, career, wealth etc. This field is also called as Chiromancy also popularly known as Palmist. It is said from the ancient scriptures that it has been originated from India and then it took a rise in interest and spreading through other Asian and Western countries mostly China, Tibet etc. There are many famous palmists in this modern era. This practice tells us many things from the lines on the palm like the head/mind line (the logical and intelligence), heart line (the emotions and love), life line (native will be healthy enough or not), Sun line (success, career, luck), some extra ordinary signs and symbols on one's palm. The lines, shape of the hands, the size of the fingers and the mounts can describe one's individual characteristics and show the graph of one's unconscious mind. Although it is a controversial topic, it is a science in which it predicts the past, present and the future of the native from the unconscious mind and the lines and mounts of a person.

There are many automated palms reading, but most of them are not efficient enough. The existing solutions cannot provide the accurate analysis like a palmist can do. In this paper, we have proposed a deep learning-based palm reading application which can analyze the palm of an individual and tell about his / her native accurately.

The rest of this paper are as follows. Section II reviews some recent automated palm reading techniques based on image processing algorithms. Section III describes our proposed method. Section IV explains the experimental setups and performance metrics used in this study. Section V shows results and discussion. Finally, Section VI concludes this paper with some possible future works.

II. Literature Review

Over the past decade, there have been a few attempts to use computer vision and image processing techniques to perform automated palm reading. Vishwaratana et al. [1] used Canny edge detector and Hough Transform to detect palm outlines inside palm image samples. On the basis of their palm and finger lengths, the outlines of palms extracted were analyzed with a ratio system to characterize persons into 4 groups, including Jupiter-ruled, Sun-ruled, Saturn-ruled, and Mercury-ruled persons. Leung and Law [2] proposed an adaptive thresholding for segmentation of the palm image to separate foreground (palm) from the background in order to extract fingers, and the three principal palm lines. A regression model was applied for producing connected and continuous palm lines. Then, based on 2D:4D ratio principle according to traditional Chinese "Feng Shui", they determined the persons' personalities and health. Tin [3] also developed a palm reading system with principal lines detection using Canny Edge detection algorithm and Hough transform.

As mentioned earlier, these existing solutions rely on obsolete image processing techniques, which are not effective for automated palm reading. Some patterns cannot be recognized and differentiated through edge detection and Hough transform techniques. The accurate analysis of the present and the future with a high accuracy is also not found in those current applications. In this work, we explore the potential of deep learning algorithms to recognize such patterns in palm reading and provide more accurate prediction results.

III. Our Proposed Method

We propose a unique deep learning framework for automated palm reading using deep learning algorithms. Our proposed solution aims to provide more detailed and accurate results than existing solutions and applications.

Deep Learning (DL) is a useful subset of machine learning methods based on artificial neural networks. DL architectures such as deep neural networks (DNN) and convolutional neural networks [7 - 9] (CNN) have been applied heavily in the field of computer vision and machine vision with outstanding performance. CNNs are similar to ordinary Neural Networks in the fact that they consist of many layers of neurons with some learnable weights and biases. A CNN consists of two important building blocks, which are convolutional layers and fully-connected layers. They can classify images, detect objects in images and perform object segmentation. They also have abilities to process images of various sizes. These are the reasons why we intend to apply these useful DL architectures in our work.

There are three phases in our approach of deep learning-based automated palm reading:

A. Semantic Segmentation of Palm images

First, palms needed to be taken out from the input images. We scaled all images to 1024 pixels × 2048 pixels, and used labelme [4] to prepare segmentation masks for all images. In order to segment the foreground palm image from the background, we trained our palm image dataset using a

978-1-6654-0380-1/20/$31.00 ©2020 IEEE
DOI 10.1109/TAAI51410.2020.00039

2020 International Conference on Technologies and Applications of Artificial Intelligence (TAAI) | 978-1-6654-0380-1/20/$31.00 ©2020 IEEE | DOI: 10.1109/TAAI51410.2020.00039

semantic segmentation technique. Semantic segmentation [5], also known as scene parsing, is a group of machine learning tasks whose goals are to give each pixel of an image a proper object category label. Fast-SCNN [6] or fast segmentation convolutional neural network is a real-time semantic segmentation model on high resolution image data (1024 × 2048 pixels) suitable for efficient computation on devices with low memory. The algorithm was claimed to achieve at least 68 % of mean Intersection over Union (IoU) in real-time segmentation of Cityscapes dataset. In addition, Fast-SCNN does not require large pre-training. The overall architecture of Fast-SCNN is as shown in Fig. 1.

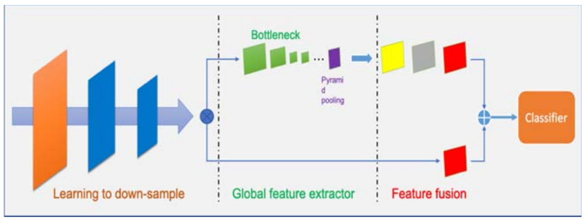

Fig. 1. Fast-SCNN architecture

Fast-SCNN, as shown above, is constructed using 4 major building blocks including Learning to Down-sample, Global Feature Extractor, Feature Fusion, and finally the classifier. All building blocks are built using depth-wise separable convolution. In the first building block, low level features such as edges and corners from the image are extracted by deep CNNs. After feature fusion, two depth-wise Separable convolutional layers followed by one Point-wise convolutional layer are introduced. At the end of each layer, both normalization layers and ReLu activations are applied.

In this work, we labeled a palm in each input image as class 1, and the rest part of the image or background as class 0. The labeling tool, labelme, converted these label files into PNG files for our semantic segmentation model. An example of these resulting files is shown in Fig. 2.

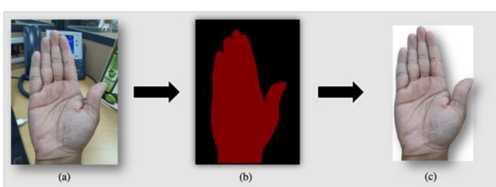

Fig. 2. Palm image segmentation: (a) an original sample, (b) a labeled segmentation mask, and (c) the expected ending result before applying grids in the next step

B. Palm regions with Multiple-Grid Approach

Second, new palm images with white background were divided by 5 x 5 grids, as shown in Fig. 3 below. The order of these grids is from the top to the bottom, and from the left to the right. The index starts from 0 to 24. These grids were put together to make each region, which would be fed into each convolutional neural network for further classification of personalities, lifestyles and future directions in life. Each of these regions is treated as an image, which is re-scaled depending on how many grids it includes.

These regions along with their descriptions and list of class labels are summarized in Table 1.

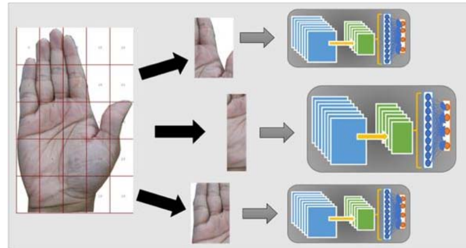

Fig. 3. Grid cells defined in each palm image (5 x 5 grids)

In the first region, class 0 depicts that a person will be wealthy, very famous, and successful because the line as shown and marked in red in Fig. 4 (left) represents the sign of success without issues or obstacles. This line appears very long in this region, which represents long-term success. Class 1 is similar to class 0. The wealth line, as marked in red, is shown but the blue diagonal line is also shown which signifies struggle at that particular age. This is a sign of defects, which represents a person's needs to overcome some obstacles. The pattern of class 2 shows a lot of issues in life because more defects are obviously found in this pattern than class 1. For class 3, a person has many more of red-marked lines for success and wealth as well as vertical blue-marked lines. This shows the division of good and bad energies in an individual. His / her life will be considered "normal", which means he / she will not be very successful, famous, and wealthy.

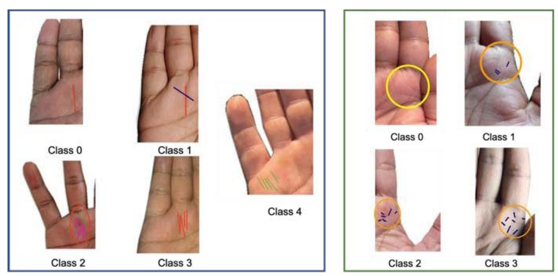

Fig. 4. Region 1 (left) and Region 2 (right)

Region ID	**Region detail**		
	Grid list	***Class***	***Description***
1	1, 6, 2, 7	0	Will be wealthy, successful, and have a good name
		1	Native will earn good name and will be successful but with great difficulty
		2	Will not a good wealth, failure in life
		3	Normal wealth and success
		4	Have good impressive, creative skills
2	11, 16, 12, 17	0	Will be wealthy, successful, and have a good married life
		1	Will be wealthy but not able to keep stability of his/her income
		2	Unsuccessful
		3	Problem in family, stressful, obstacles will be there
3	7, 8, 9	0	Good fate
		1	Average
		2	Bad luck

Table I List of grid-based palm regions

In the second region, class 0 shows the person's good marriage life and wealth due to the appearance of a prominent mount shown in Fig. 4 (right). Class 1, as seen on the mount, some blue-marked lines appear as defects. These are not good for retaining wealth as there will be more expenses despite good incomes. Class 2 shows the weak ambition and life goals as the mount appeared is not prominent enough to draw success and express leadership qualities. Class 3 shows more defects which signify relationship stress, unhappy married life and tension of career.

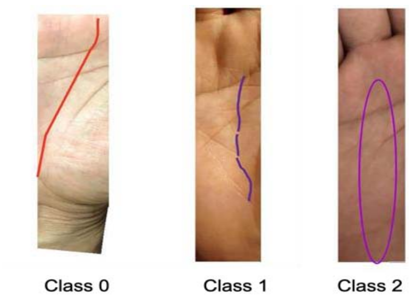

Fig. 5. Region 3

Finally, in the last region as shown in Fig. 5, a pattern for class 0 shows that the person has a clear fate line, as marked in red. There is no defect found in this pattern, therefore a person will be very lucky. If this line is missing (in case of class 2), the person will not have good lucks in life and will struggle a lot. For class 1, although the fate line, marked in blue, is found on the palm region, he / she will still face some issues due to the break inside the line.

C. Multi-Class classification

Third, each region as defined above by our proposed grid system was fed into a convolutional neural network for training. The architecture detail of our deep CNN is described as follows.

Our deep CNN receives the input of fixed size of (512, 512, 3), which represents the shape of an image, which in this case it is a cropped region from a full palm image, and its corresponding RGB values. For any images whose sizes are bigger than 512 x 512 pixels, they are all scaled down using a high-quality down-sampling filter. For any images whose sizes are smaller than 512 x 512 pixels, we pad them with white pixels. The first convolutional block consists of 25 2D filters of size 5 x 5 pixels, followed by a ReLu activation and a max pooling layer of size 2 x 2. The second convolutional block consists of 50 2D filters of the same size. In order to stabilize the learning process and accelerate the training of this architecture, a batch normalization is applied before the output of this block enters the next convolutional block. The third convolutional block is similar to the second one, but it has 70 filters of only 3 x 3 pixels. After all the features from an input image are extracted by the convolutional layers, the resulting output is flattened into a 1D vectors which are then fed into two layers of fully-connected layers of 100 hidden units. Finally, the output layer varies based on the number of classes specified for each region. Softmax activation function is applied to calculate the class probabilities. The optimization algorithm used in the models is Adam.

IV. EXPERIMENTATION

We collected our own private palm datasets by providing actual palm reading services. With permission from our sample groups, we collected 553 palm images in total for this experiment. 85 % of this dataset was used for training semantic segmentation model and multi-class convolutional neural networks. The rest was used for testing the performance of the trained models. We set the training epochs of both models at 100 iterations. In this study, we focus on right palms.

The order of our proposed approach to study the performance, mean Intersection-over-Union (IoU, Jaccard Index) is used as performance metrics for semantic segmentation of palm images. The calculations of these metrics are given as

$$mIoU = (\sum_k \frac{|X \cap Y|}{|X \cup Y|})/K \tag{1}$$

where X is the predicted pixel set values and Y is the ground truth pixel set values.

Accuracy scores and F1 scores are used as performance metrics for multi-label multi-class classification models. The calculations of these metrics are given as

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \tag{2}$$

$$Precision = \frac{TP}{TP+FP} \tag{3}$$

$$Recall = \frac{TP}{TP+FN} \tag{4}$$

$$F1 = 2 \cdot \frac{Precision \times Recall}{Precision + Recall} \tag{5}$$

where TP, TN, FP, and FN denote true positive, true negative, false positive, and false positive, respectively.

To evaluate the performance of our deep CNN algorithm for palm reading, we compare its classification performance with well-known complex CNN architectures including AlexNet [8], and ResNet [10]. These two sophisticated architectures were successfully used for image classification tasks such as MNIST and CIFAR. We keep every hyperparameter of both models by default.

V. RESULTS AND DISCUSSION

Table II and Table III report the performance of the semantic segmentation and classification models used in this study, respectively.

Dataset	***Mean IoU***
Train set	72.8 %
Test set	68.0 %

Table II Performance of segmentation algorithm on our dataset

As shown in Table II above, Fast S-CNN still delivers good enough mean IoU in the task of semantic segmentation on our dataset. An IoU of more than 0.5 is normally considered a "good" prediction. However, it will be nicer if our predictions of mask areas are averagely closer to 1. Therefore, we could avoid any mistakes when we crop the palm out of each image. This will eventually lead to our solution becomes applicable.

Table III illustrates the classification performance of our CNN. The accuracy scores and F1 scores for the three regions have been obtained from the models. We obtained 98.29% and 93.97% accuracy score from region 1 from the train set and test set, whereas from region 2 accuracy score of 99.57% and

97.59% obtained from train set and test set, respectively. The region 3 has less accuracy score in both the train set and the test set which is 94.88% and 83.14%, respectively.

The less accuracy scores and F1 scores in the region 3 may be resulted from unclear patterns of the region found in the dataset. It is also possible that more training images or better AI architecture could improve the predictive performance of this region.

Region ID	Classification Performance			
	Accuracy scores		*F1-score*	
	Train	*Test*	*Train*	*Test*
1	98.29 %	93.97 %	0.9156	0.9253
2	99.57 %	97.59 %	0.9759	0.9786
3	94.88 %	83.14 %	0.8313	0.8252

Table III Classification performance of our CNN on each region

Table IV and V shows the classification performances of AlexNet and ResNet on each region. ResNet shows very low accuracy scores on validation sets, compared to our more simple model. It is very likely that ResNet model is overfit to our training set as well. AlexNet does not seem to converge after a long time training.

Region ID	AlexNet's Performance			
	Accuracy scores		*F1-score*	
	Train	*Test*	*Train*	*Test*
1	53.47 %	34.93 %	0.5132	0.3411
2	55.82 %	36.14 %	0.5253	0.3646
3	56.06 %	33.37 %	0.5337	0.3605

Table IV Classification performance of AlexNet on each region

Region ID	ResNet's Performance			
	Accuracy scores		*F1-score*	
	Train	*Test*	*Train*	*Test*
1	91.47 %	54.22 %	0.9059	0.5479
2	97.65 %	74.70 %	0.9569	0.7484
3	88.91 %	79.52 %	0.8137	0.7477

Table V Classification performance of ResNet on each region

VI. CONCLUSION AND FUTURE WORKS

This research work is the first one to develop a deep learning-based palm reading technique that could segment the palm from the background, and predict future life directions of individuals. The proposed solution seems to have potential to be used with acceptable accuracy on mobile applications in the future. However, there are still some important problems to address. Our algorithm still needs more improvement in terms of predictive performance. Semantic segmentation performance of Fast-SCNN is acceptable but it will be better to achieve mean IoU closer to 1. For image classification task like this, model complexity and amount of data samples used for training also play significant roles in improving predictive performance. Our results suggest that we need more amount of training and test data samples to make our proposed solution more reliable.

Possible directions for future works include the increase of data samples in order to avoid overfitting, improvement and development of palm segmentation algorithms and life direction and human personality classification algorithms, more comprehensive comparison study with other detection algorithms and classification algorithms, and the deployment on mobile platforms. Other directions such as end-to-end classification methods for automated palm reading should also be considered. Last but not least, this approach could also be applied on the left palm images in our future work as well. A combination of prediction results from both palms might also enhance the foretelling of a person.

REFERENCES

[1] N. Vishwaratna, Y. Divakar, and T. Manisk, "A Novel Approach for Hand Analysis Using Image Processing Techniques,"International Journal of Computer Science and Information Security. May, 2010.

[2] K. Leung and N. F. Law, "An efficient automatic palm reading algorithm and its mobile applications development," 2016 IEEE International Conference on Signal Processing, Communications and Computing (ICSPCC), Hong Kong, 2016, pp. 1-6, doi: 10.1109/ICSPCC.2016.7753706.

[3] K. H. H. Tin, "An effective method of a person's character or future using the palmprint images," *International Journal of Research and Scientific Innovation*. IV. 09-12. Dec, 2016.

[4] K. Wada, "labelme: Image Polygonal Annotation with Python," 2016.

[5] N. Atif, M. Bhuyan and S. Ahamed, "A Review on Semantic Segmentation from a Modern Perspective," *2019 International Conference on Electrical, Electronics and Computer Engineering (UPCON),* ALIGRAPH, India, 2019, pp. 1-6, doi: 10.1109/UPCON47278.2019.8980189.

[6] R. PK Poudel, S. Liwiki, and R. Cipolla, "Fast-SCNN: Fast Semantic Segmentation Network," in 30th British Machine Vision Conference, BMVC 2019, Cardiff, Welsh, UK, September 10-12, 2019.

[7] Ezat, W., Dessouky, M. and Ismail, N., 2020. Multi-class Image Classification Using Deep Learning Algorithm. *Journal of Physics: Conference Series*, 1447, p.012021.

[8] Krizhevsky, A., Sutskever, I. and Hinton, G., "ImageNet classification with deep convolutional neural networks," 2017 *Communications of the ACM*, 60(6), pp. 84-90.

[9] F. Sultana, A. Sufian and P. Dutta, "Advancements in Image Classification using Convolutional Neural Network," 2018 Fourth International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN), Kolkata, India, 2018, pp. 122-129, doi:10.1109/ICRCICN.2018.8718718.

[10] K. He, X. Zhang, S. Ren and J. Sun, "Deep Residual Learning for Image Recognition," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.