Case Study: Fine Writing Style Classification Using Siamese Neural Network

Alaa Abdalhaleem
Department of Computer Science
Ben-Gurion University of the Negev
Beer-sheva, Israel
alaaabd@cs.bgu.ac.il

Berat Kurar Barakat
Department of Computer Science
Ben-Gurion University of the Negev
Beer-sheva, Israel
Berat@cs.bgu.ac.il

Jihad El-Sana
Department of Computer Science
Ben-Gurion University of the Negev
Beer-sheva, Israel
el-sana@cs.bgu.ac.il

***Abstract*—This paper presents an automatic system for dividing a manuscript into similar parts, according to their similarity in writing style. This system is based on Siamese neural network, which consists of two identical sub-networks joined at their outputs. In the training the two sub-networks extract features from two patches, while the joining neuron measures the distance between the two feature vectors. Patches from the same page are considered as identical and patches from different books are considered as different. Based on that, the Siamese network computes the distances between patches of the same book.**

***Keywords*-Siamese-network; Deep-learning; Writing-style; Writer-identification, Supervised-learning;**

I. Introduction

Dividing a manuscript into similar parts based on writing style has been attracting the interest of researchers over the last decade. This problem finds its application in various document processing tasks, such as signature identification and writer identification historical manuscripts. It is a complex and time consuming task and often done manually, which illustrate the demand for developing an automatic technique to divide a manuscript into similar parts according to writing style.

Current writer identification systems can detect different writing style, but they are not sensitive enough to detect delicate changes in the writing style. The application of such algorithm are numerous, which include, writing fraud, detecting writer state of mind, and tracking the evolution of an individual writing style over time. For historical document detecting delicate changes could shade a light on the writing practice. Owning books was a sign of elite status and for that reason manuscripts were in high demand for many periods. Detecting delicate changes in writing style can expose part of the work condition of the scribes. For examples, if they work many hours in tough labor conditions, their writing style may deteriorate over the hours of the day.

In this work we present a novel approach to measure small changes of the writing styles of the same person, which is based on Siamese convolutional neural networks. A Siamese network can rank similarity between two inputs.

We trained Siamese network on a supervised data to learn the tiny changes in person writing style. Patches taken from the same page of a manuscript are labeled as identical in writing style and patches from different manuscripts are labeled as different writing style. Then we validate on samples that have not been seen during training. Once validation and training results get close, we use the resulting model to measure the similarity between two patches from the same book.

Building a training set for Siamese network is not an easy task. The training set should be large enough to represent the input space well. For preventing overfitting, the training set size should be much bigger then the total number of weights in the network. The proposed method does not need text line segmentation but only patches of text regions that are weakly labeled as from the same book or from different books. In addition, we perform experiments on the same data with different resolutions to monitor affect of resolution on performance.

The rest of this paper is organized as follows: In Section II, we summarize related methods for writer identification using deep neural network. In Section III, we present data preparation. In Section IV,we explain our algorithm for dividing a sample book, and we present the structure of the Siamese neural network we have used in our method. The experimental study is shown in Section V. Finally, in section VII we present our conclusion and direction for future work.

II. Related work

In forensic and historical community, writer identification has attracted significant attention in the recent years. Survey of the early works in writer identification can be found in [18]. Another survey [4] presents some works on identifying writers of Arabic manuscripts. Recent researches on automatic writer identification can be found in [13]. Some machine learning algorithms were used to classify handwriting styles [20]. They proposed training a single Hidden Markov Model (HMM) per writer using data written only by this writer. In other words, for each writer there is a different HMM. In their next work, they used Gaussian Mixture Models (GMMs) to model the handwriting style [19]. They concluded that a GMM method is better than HMM in terms of simplicity, quickness and able to obtain higher identification rates. In recent research study [8], GMM supervectors were better than other GMM-based encoding algorithms. The authors used GMM supervectors for encoding and SIFT descriptors as underlying features. They proved that using Exemplar-SVMs can improve the results. In recent years, Many researchers used convolutional neural networks for handwritten documents, [16] [23] [12]. A comprehensive survey of using neural network for classification problems can be found In [24]. In [10], the authors used Convolutional Neural Networks (CNN) to generate a feature vector for each writer, which is then compared with the pre calculated feature vectors stored in the database. For the generation of the vector, the CNN is trained on a database with known writers and after training the classification layer is cut off and the output of the last fully connected layer is used as feature vector. For the identification, they used the nearest neighbor classifier. Siamese neural network include a pair of identical networks that united at the final phase by a function. This function compute a metric for deciding the difference between two images.The first time has been used was to solve signature verification as an image matching problem [6].In recent years it has been used in multiple fields, such as object tracking [5], face verification [7] and image recognition [15].

Recent research had been applied to determine how a person's writing style changes with time considering a children's writing dataset is [9]. The authors considered two words were written by the same person as a discriminative label for word-level feature training. Then, based on word-level features, they define writing similarity between passages. This similarity not only shows the distinction between writing styles of different people, but also the development of style of the same person. Performance with several hidden layers in the neural network are evaluated. Since there is rich databases in historical documents archives, are available and and robust processing methods. This observation motivates us to suggest an automatic system to divide a book for similar parts, to determine how the writing style changed over time.

III. Data Preparation

We evaluated the proposed method on three manuscripts from WAHD dataset [2] and one manuscript from VML-HD dataset [14]. Both datasets are written in Arabic. WAHD dataset contains 333 manuscripts from Islamic Heritage Project (IHP) and 20 manuscripts from the National Library in Jerusalem (NLJ). We select three manuscripts of 139, 239 and 392 pages from IHP set. VML-HD dataset contains 5 manuscripts. We select one manuscript of 140 pages from VML-HD dataset. Figure 2 shows example pages from the four manuscripts. Decorations and page margin notes are not included in our experiments.

A manuscript is a document written by hand. May it be written by an author or a writer. Author is the person who originates the work being written whereas the writer is a person who writes the script. The author and the writer can be the same person or may the author let his students write different parts of the manuscript with similar fonts. Each of the 4 manuscripts in our dataset belongs to one author and consistent in its font type and size. However number of writers for each manuscript is not known.

A. Pair generation

The Siamese network learns the filters that extract descriptors from pairs of image patches. Therefore the train, validation and test sets are consist of image patch pairs. A pair of patches from the same page is labeled as positive and a pair of patches from different books is labeled as negative. This generation principle ensures at maximum similarity for positive pairs and maximum dissimilarity for negative pairs. We included positive pairs as much as needed and a negative pair per positive pair. Consequently there are equal number of positive pairs and negative pairs. As a result network learns them equally and does not overfit.

We detected main text region using color differences of decorations and sizes of page margins. Then we cropped patches from the extracted main text regions. All patches from the same page were put in the same folder. We created patches of 150×150, 120×120, 75×75 and 40×40 pixels each with approximately 3 lines of text (Figure 3). Then we decreased the page image resolution and created patches of 150×150 pixels each with approximately 4 lines of text (Figure 4).

IV. Writing Style System

Since we need a model based on pairs of patches that are labeled as same or different, to measure the distance between two patches, we chose a Siamese neural network. The architecture of this network contains a pair of identical sub-networks that share the same structure, weights and parameters, which are distributed across these sub-networks. First we prepared the dataset, which is an important phase in neural network training. We partition the dataset into three disjoint sets: training set, validation set and test set. Training set is used to compute the gradient and update the weights and biases. During training we monitored the validation error. When the validation error stop decreasing, the training is stopped and the best model is chosen. We used the best model to measure the distance between pairs of patches from the same book.

A. Siamese Neural Network

Our system is based on Siamese neural network that is presented in [15] with some changes to avoid overfitting. We tried the original network on our data. After a number of epochs the training loss continued decreasing while the validation loss started to increase. Consequently the validation accuracy dropped dramatically while the training accuracy increased overly. This was in opposite to their results where they got impressive results. Hence we made some changes on the network. We added some convolutional and pooling layers. We changed the kernel size of the convolutional layers to 3×3 instead of 5×5, Because small filters has been found to improve the results in deep convolutional networks[21]. We also changed the number of filters per layer to begin with 64, and multiplying their count by two, up to 256 features for the last convolutional layer. These changes mainly tackle: (1) The difference in the input images dimension (2) The

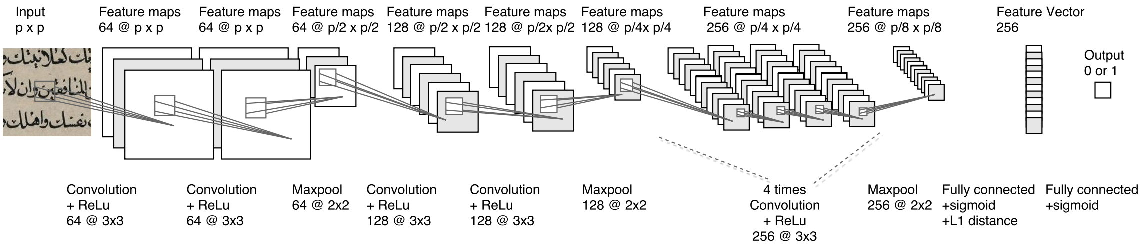

Figure 1. Siamese convolutional neural network architecture. The Siamese twin connects after the last fully connected layer, where L1 distance is calculated from the feature vectors and the result is provided by the Sigmoid function.

Figure 2. Example pages from manuscripts that are used in the experiments. First image is from VML-HD and the rest are from WAHD.

Table I
SUMMARY OF THE RESULTS.

Training Size	Number Of Line	Patch Size	Training Accuracy	Test Accuracy	Validation Accuracy
8,266	3	150 × 150	99.3830	98.5813	97.6017
8,266	3	120 × 120	99.2430	98.1481	97.5627
8,266	3	75 × 75	99.1895	94.7468	94.2951
8,266	3	40 × 40	97.9555	94.4041	94.3689
8,266	4	150 × 150	99.7766	98.6913	98.1132
4,270	3	150 × 150	99.9297	96.4026	96.9388
2,568	3	150 × 150	98.9486	95.1247	94.9128

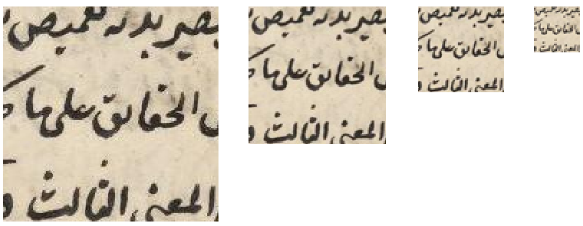

Figure 3. Example patches of 150 × 150, 120 × 120, 75 × 75 and 40 × 40 pixels with approximately 3 lines of text.

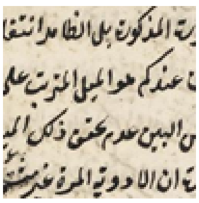

Figure 4. Example patch of 150 × 150 pixels with approximately 4 lines of text.

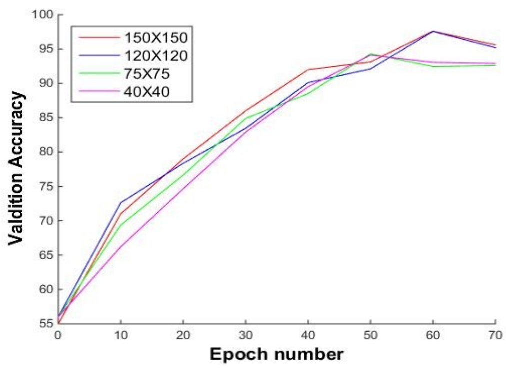

Figure 5. validation accuracy, as a function of the epoch number, in different resolution patches

overfitting problem because the input is different. We saw that this changes were necessary the convergence of the network. The complete network architecture is shown In Figure 1. We calculated the distance between two samples using absolute linear distance and joined them using sigmoid activation function in the same way as the original network.

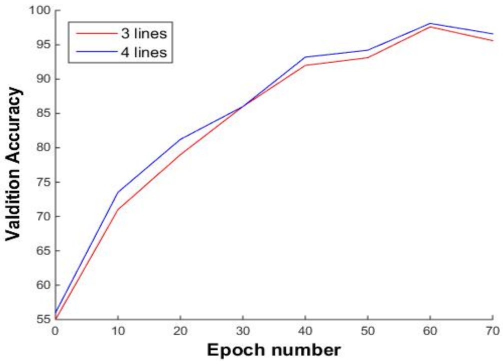

Figure 6. validation accuracy, as a function of the epoch number, in different number of line in patch

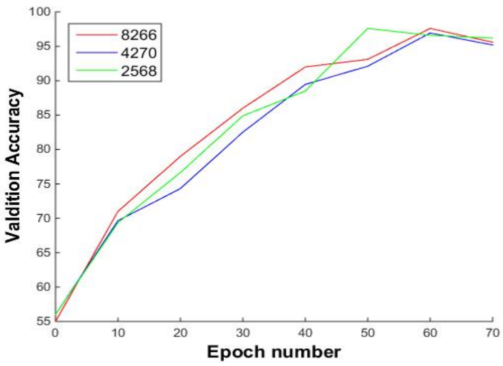

Figure 7. validation accuracy, as a function of the epoch number, in different training set size

V. Experiments

In this section we present the experiments we have conducted with the Siamese neural network. We will start with training and testing the network on labeled samples, and after that showing the results on specific manuscripts.

A. Training and test phase

We initialized the network weights in the same way mentioned in [15], both convolutional and fully connected layers are initialized from a normal distribution with zero-mean. We trained the network using the training sets of various sizes on a single NVIDIA 1080GTX. We have used Keras [11] front-end and Tensorflow [1] back-end. The input is a pair of gray-scale images. The neural network performance was monitored on a validation set of size $2,752$. Once the training finished we tested it on the test set of size $2,646$. Three types of experiments were performed: *(a)* Modification in the training set size: 8266, 4270, 2568. *(b)* Modification in sample resolution: 150×150, 120×120, 75×75, 40×40. *(c)* Modification in the number of text lines in each sample: $3,4$. We trained the neural network over 70 epochs for every training sub-set input size, and chose the model of the epoch providing the best result for each set. The Figures 5, 6 and 7 present the results in the three type of experiments. Results summary can be found in table I.

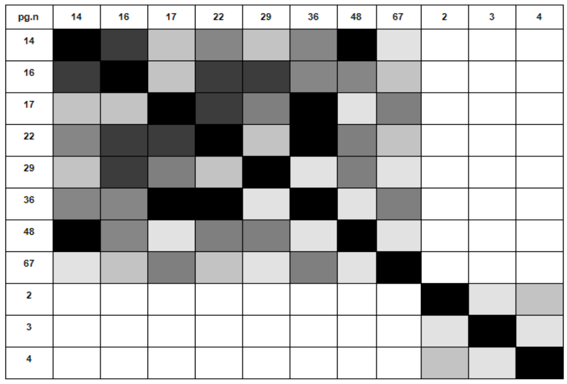

Figure 8. The distances between random pages, (the last three pages are from different manuscripts), more bright more different

B. Dividing manuscript phase

We chose randomly 8 paragraphs from the same manuscript and 3 paragraph from different manuscripts and measure the distance between them based on the results from the neural network, as shown in Figure 8. It can be seen that the similarity is not constant in the same book, and it changes randomly, as in real life, the writer have different styles, but the styles is chosen by the situation.

VI. Discussion

Our purpose in this paper is to examine the in-writer variations in a manuscript. This approach in this scenario is the first work in state of art, so we do not a way to compare our approach with other works.

As it work in researches that use neural networks, the researcher do not know which features to extract, but with comparison with ground truth he know that the network converge in the right direction, We go in this direction and prove that our results is right.

With regard to writing style number for each righter, So the is no constant number for each righter, it depend on other things.

VII. Conclusion

In this work we have introduced a novel system for dividing a book into similar parts in writing style. The system takes as input paragraphs from the same book and classify them by similarity. These paragraphs were used to build data for the neural network. We trained a model to decide if a pair of images has paragraphs with similar writing style. In our experiments we have shown that the neural network achieved high accuracy. For future work, we plan to experiment with more historical data sets and on modern handwritten data sets.

Acknowledgment

This research was supported in part by the Lynn and William Frankel Center for Computer Sciences at Ben-Gurion University, Israel, and wed like to thank them for their support.

REFERENCES

[1] Abadi, M., Agarwal, A., Barham, P., Brevdo, E., Chen, Z., Citro, C., Corrado, G.S., Davis, A., Dean, J., Devin, M., et al.: Tensorflow: Large-scale machine learning on heterogeneous distributed systems. arXiv preprint arXiv:1603.04467 (2016)

[2] Abdelhaleem, A., Droby, A., Asi, A., Kassis, M., Al Asam, R., El-sanaa, J.: Wahd: A database for writer identification of arabic historical documents. In: International Workshop on Arabic Script Analysis and Recognition, pp. 140–145 (2017)

[3] Arampatzis, T., Lygeros, J., Manesis, S.: A survey of applications of wireless sensors and wireless sensor networks. In: Intelligent Control, 2005. Proceedings of the 2005 IEEE International Symposium on, Mediterrean Conference on Control and Automation, pp. 719–724. IEEE (2005)

[4] Awaida, S.M., Mahmoud, S.A.: State of the art in off-line writer identification of handwritten text and survey of writer identification of arabic text. Educational Research and Reviews **70**(20), 445–463 (2012)

[5] Bertinetto, L., Valmadre, J., Henriques, J.F., Vedaldi, A., Torr, P.H.: Fully-convolutional siamese networks for object tracking. In: European Conference on Computer Vision, pp. 850–865. Springer (2016)

[6] Bromley, J., Guyon, I., LeCun, Y., Säckinger, E., Shah, R.: Signature verification using a" siamese" time delay neural network. In: Advances in Neural Information Processing Systems, pp. 737–744 (1994)

[7] Chopra, S., Hadsell, R., LeCun, Y.: Learning a similarity metric discriminatively, with application to face verification. In: Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, vol. 1, pp. 539–546. IEEE (2005)

[8] Christlein, V., Bernecker, D., Hönig, F., Maier, A., Angelopoulou, E.: Writer identification using gmm supervectors and exemplar-svms. Pattern Recognition **63**, 258–267 (2017)

[9] Chu, J., Srihari, S.: Writer identification using a deep neural network. In: Proceedings of the 2014 Indian Conference on Computer Vision Graphics and Image Processing, p. 31. ACM (2014)

[10] Fiel, S., Sablatnig, R.: Writer identification and retrieval using a convolutional neural network. In: International Conference on Computer Analysis of Images and Patterns, pp. 26–37. Springer (2015)

[11] Francois Chollet: Keras (2015)

[12] He, K., Zhang, X., Ren, S., Sun, J.: Spatial pyramid pooling in deep convolutional networks for visual recognition. In: European Conference on Computer Vision, pp. 346–361. Springer (2014)

[13] Idicula, S.M.: A survey on writer identification schemes. Writer **25**(12), 15 (2011)

[14] Kassis, M., Abdalhaleem, A., Droby, A., Alaasam, R., El-Sana, J.: Vml-hd: The historical arabic documents dataset for recognition systems. In: 1st International Workshop on Arabic Script Analysis and Recognition. IEEE (2017)

[15] Koch, G., Zemel, R., Salakhutdinov, R.: Siamese neural networks for one-shot image recognition. In: ICML Deep Learning Workshop, vol. 2 (2015)

[16] Krizhevsky, A., Sutskever, I., Hinton, G.E.: Imagenet classification with deep convolutional neural networks. In: Advances in neural information processing systems, pp. 1097–1105 (2012)

[17] Kubrusly, C.S., Malebranche, H.: Sensors and controllers location in distributed systems?a survey. Automatica **21**(2), 117–128 (1985)

[18] Plamondon, R., Lorette, G.: Automatic signature verification and writer identification the state of the art. Pattern Recognition **22**(2), 107 – 131 (1989)

[19] Schlapbach, A., Bunke, H.: Off-linewriter identification using gaussian mixture models. In: Proceedings of the International Conference on Pattern Recognition, pp. 992–995. Hong Kong, China (2006)

[20] Schlapbach, A., Bunke, H.: A writer identification and verification system using hmm based recognizers. Pattern Analysis and Applications **10**(1), 33–43 (2007)

[21] Simonyan, K., Zisserman, A.: Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)

[22] Tönshoff, H.K., Inasaki, I.: Sensors applications, volume 1, sensors in manufacturing. Sensors Applications, Volume 1, Sensors in Manufacturing, by Hans Kurt Tönshoff (Editor), Ichiro Inasaki (Editor), pp. 426. ISBN 3-527-29558-5. Wiley-VCH, April 2001. p. 426 (2001)

[23] Zeiler, M.D., Fergus, R.: Visualizing and understanding convolutional networks. In: European conference on computer vision, pp. 818–833. Springer (2014)

[24] Zhang, G.P.: Neural networks for classification: a survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews) **30**(4), 451–462 (2000)