Extreme Learning Machines for Signature Verification

Leonardo Espinosa Leal[1⋆], Anton Akusok[1,2], Amaury Lendasse[3], and Kaj-Mikael Björk[1,2]

[1] Arcada University of Applied Sciences, Jan-Magnus Janssons plats 1, 00560 Helsinki, Finland
{leonardo.espinosaleal,anton.akusok,kaj-mikael.bjork}@arcada.fi
[2] Hanken School of Economics, Arkadiagatan 22, 00100, Helsinki, Finland.
[3] University of Houston, Houston, TX 77004, USA.
alendass@central.uh.edu

Abstract. In this paper, we present a novel approach to the verification of users through their own handwritten static signatures using the extreme learning machine (ELM) methodology. Our work uses the features extracted from the last fully connected layer of a deep learning pre-trained model to train our classifier. The final model classifies independent users by ranking them in a top list. In the proposed implementation, the training set can be extended easily to new users without the need for training the model every time from scratch. We have tested the state of the art deep neural networks for signature recognition on the largest available dataset and we have obtained an accuracy on average in the top 10 of more than 90%.

Keywords: Signature verification · Deep Learning · Extreme Learning Machines.

1 Introduction

Despite the significant advances in personal identification using different digital methods, biometric features are still a common way to create an unique strategy for the identification of users [21]. Among all the different types, handwritten signatures are still a fashionable way to identify individuals for legal or representative purposes. Depending of how these are acquired, signatures can be divided in two categories: static (offline) and dynamic (online). In the first one, the person uses a normal pen to create the signature, therefore the geometric features of the signature are the only information recorded. In the second form, the person signs on a electronic table or similar device and then, the systems records the geometry and the speed of the signature, in some cases, also the pressure, speed and angles related to a specific user. It is interesting to note that handwritten signatures are mostly important in the western cultures. In other

⋆ Corresponding author.

cultures, digital signatures are already use in daily basis or more classical seals are employed as a mean of personal identification.

The use of handwritten signatures as a mean of identification have given rise to the use of forgeries or falsifications for fraudulent purposes. In general, there are two types of forgeries, trained (or skilled) forgeries and untrained (or random) forgeries. In the first type, the forger has previous knowledge of the user's signature, for instance, by having access to an original visual sample. The level of accuracy in the forgery (i.e. how close it looks like to the original one) can vary depending of the skills of the forger and the amount of time invested in rehearsing the forgery. In the second the forger does not have any reference of the original signature, so she creates her own version. The capacity of the forger to create an accurate forgery here can depend of the knowledge that she can have of the original users, for instance, if she knows the name, there is a much higher probability that the forgery matches the genuine signature. This last one is much easy to detect for any automatic personalized classifier, even by one trained with a low amount of samples. The challenge is evident in the case of detection of trained or skilled forgeries and this topic is nowadays an active area of research, despite the lack of publicly open available datasets for training and testing algorithms.

There are different strategies for addressing the aforementioned problem from the perspective of the machine learning: train an universal or personalized classifier. The study from the point of view of an universal classifier is the least tackled issue mainly because of the amount of signatures obtained in most of the cases are of the order of tens per user. The most common path consists in creating a personalized classifier, it means, a classifier per user from a specific database. In some cases, the classifier is build with a highly imbalance class dataset, where the positive class is a given user and the signatures of the rest of the users are taken as samples of the negative class. This method is the most extended because it allows the classifier to learn from a much larger set of available negative samples, however it possesses two technical issues: first, it is necessary to train the classifier with each user as class, which means a large multiclass space and second, if a new user is added to the dataset, as is expected for any practical purpose, the whole process must be done every time from the scratch.

In this work we present an alternative study to the verification of users via their static signatures using an extreme learning machine (ELM) classifier. We show that using the ELM method, we can address the two technical issues of having a continuous large set of users and when new users are being added to the classifier.

2 Related works

Previous research on the identification of handwritten signatures have been focused in using different methods for the extraction of visual features. Some of these includes techniques programmed to extract information from the attributes of the signature itself such us stroke, pressure or velocity, others tackle the fea-

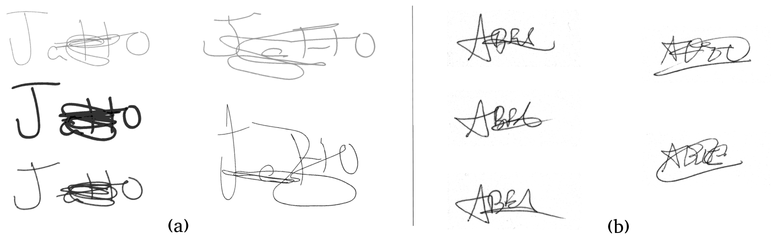

Fig. 1. Samples of signatures used in this work. In **(a)** from the GPDS Synthetic OnLine OffLine Signature (GPDSS10000) [14] database and **(b)** from the MCYT-75 database [25]. In both cases, the signatures on the left correspond to *genuine* signatures and on the right, to *skilled* forgeries.

ture extraction by means of mathematical transformations such us wavelets, cosine transformations or more recently deep neural networks. This last method has been used successfully to create representations that classifiers can use for training and have give promising results with high accuracy. For a more comprehensive review see [11] and [18] and references therein.

2.1 Feature learning for signature verification

In this paper we build our findings upon the work of *Hafemmann et al.*[17, 16, 15]. They have proposed a new family of neural network architectures for the feature representation of static signatures known as SigNet (see Fig. 2). The architectures are trained to use both, the resized and original size at different resolutions (300 dpi and 600 dpi) of the images of the signatures. Here, we have implemented three of these networks and for comparison, as baseline, we have implemented an openly available pre-trained network known as *Inception21k*, which has been used previously with success in other computer vision research areas [22, 13].

We have tested our ideas on the GPDS synthetic OnLine OffLine Signature database [14], called GPDSS10000 from now on, which is the largest dataset openly available for research purposes. The dataset contains the images of signatures for ten thousand users including twenty four genuine and thirty skilled forgeries. We also tested our method on a much smaller, but classical dataset, the MCYT-75 dataset [25], this contains the signatures for seventy five users with fifteen genuine signatures and fifteen skilled forgeries. In this work, we tested our method only using the genuine signatures, therefore regarding the size of the datasets, we have two hundred forty thousand signatures from the GPDSS10000 and one thousand one hundred twenty five from the MCYT-75.

3 Extreme Learning Machine

Extreme Learning Machine (ELM) is an extremely powerful and fast method for analysis within the realm of machine learning. It is a universal approximator and

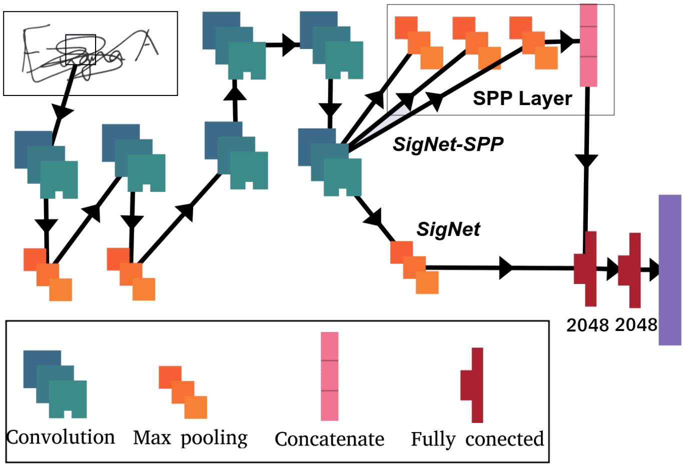

Fig. 2. Scheme of the SigNet [17] and SigNet-SPP [15] Neural Networks. Both networks share the same initial structure, they differ in that the SigNet-SPP includes an additional step with a Spatial Pyramid Pooling algorithm. The scheme was depicted using the ENNUI toolset [24].

can be viewed as a shallow layered extension of a linear model with the capacity of learn the non-linear dependencies inherent to certain kind of data [23, 19, 9]. The method has been extensively studied and successfully extended and applied in other fields of research such as visualization of data [5], mislabeled data [6, 7], computer vision [4, 13], multiclass classification [12], mobile computing [2], time series [26] among others.

Despite the simplicity and extreme accuracy of a single ELM layer populated with thousand of non-linear neurons, the application in the field of computer vision cannot be done directly, an intermediary strategy where a set of features with a fix length obtained from the images must be applied. The most extended technique consists in the use of deep neural networks (pre-trained or not). In this work, we tested four networks in combination with a new and optimized ELM implementation known as Scikit-ELM [3], a toolbox developed in the python language with fully integration to the well known Scikit-learn[4] python library. This toolbox was born as an improvement of a previous openly available GPU compatible ELM implementation [1].

For the ELM algorithms, the order of data samples is irrelevant for the computation of the covariance matrix. The library is designed to compute by batch update, and the randomness of the data is irrelevant for the calculations. This differs the closed form solution of ELM from an iterative solution, and greatly simplifies the code and in addition, it allows the update of the trained model with more data, this is the main advantage of the ELM in comparison to other machine learning methods. Once all the data has been processed and the final

[4] https://scikit-learn.org

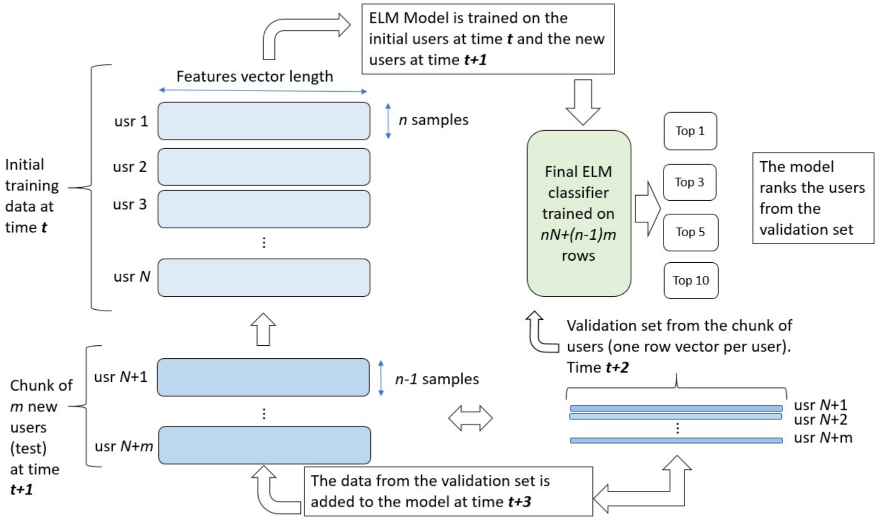

Fig. 3. Scheme of the proposed method. At time ***t*** the ELM classifier is trained with a starting set of N samples of users from the dataset. Then, a chunk of m new users is added at time ***t+1***, removing one sample per user and adding it to the validation set. The ELM classifier is retrained with the new set of $N+m$ users. In the next step, the classifier is used to rank the samples from the validation set at time ***t+2***. Finally, at ***t+3*** the samples used in the validation set are added back to the classifier, then is retrained and ready for the next chunk of users. In each step the classifier is saved in a pickle file.

covariance matrices are available, they are dumped back to the main memory and the ELM solution is computed via the Cholesky decomposition method[8].

4 Methodology

Here, we propose a novel strategy for the classification of static signatures. We train a ELM classifier using half of the samples from the signatures. Then, we save the model and later, this is used as the starting point for training new samples added in chunks of users of a given size. For each chunk of new users, one sample is retained an included in the set used to validate the accuracy of the classifier (see Figure 3 for a detailed scheme of the methodology). The accuracy of the trained classifier is measured by counting the position in which the sample is predicted. Within this method, We study the validated signatures at four different positions: top 1, top 3, top 5 and top 10. Once the whole chunk of users in analyzed, the data used as validation is included back into the trained model and then, a new chunk of users is studied. In the case of the GDPSS10000 dataset, the starting training set consisted of five thousand users (half of the dataset, it means one hundred twenty thousand images), and the chunks of new users consisted of one hundred users (two thousand three

Table 1. Averaged Results for the MCYT-75 dataset. The normalized score for the top 1, top 3 top 5 and top 10 users verified using the Inception21k, SigNet, SigNet 300 dpi and SigNet 600 dpi and different values of ELM neurons (n), with hyperbolic tangents.

n	Inception21k				SigNet				SigNet 300dpi				SigNet 600dpi			
	top1	top3	top5	top10	top1	top3	top5	top10	top1	top3	top5	top10	top1	top3	top5	top10
4	0.0	0.12	0.16	0.44	0.04	0.28	0.36	0.48	0.04	0.28	0.28	0.56	0.04	0.12	0.24	0.48
8	0.16	0.32	0.44	0.72	0.16	0.32	0.32	0.68	0.16	0.44	0.56	0.72	0.2	0.32	0.44	0.76
16	0.28	0.44	0.48	0.76	0.44	0.6	0.72	0.84	0.44	0.76	0.92	1.0	0.36	0.6	0.72	0.88
32	0.48	0.64	0.76	0.84	0.64	0.92	0.92	1.0	0.84	0.96	0.96	1.0	0.52	0.8	0.96	1.0
64	0.84	0.96	1.0	1.0	0.92	1.0	1.0	1.0	0.8	0.96	1.0	1.0	0.96	0.96	0.96	1.0
128	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table 2. Averaged results for the GDPSS10000 dataset. The normalized score for the top 1, top 3 top 5 and top 10 users verified using the Inception21k, SigNet, SigNet 300 dpi and SigNet 600 dpi and different values of ELM neurons (n), with hyperbolic tangents.

n	Inception21k				SigNet				SigNet 300dpi				SigNet 600dpi			
	top1	top3	top5	top10	top1	top3	top5	top10	top1	top3	top5	top10	top1	top3	top5	top10
16	0.00	0.01	0.02	0.03	0.00	0.01	0.02	0.04	0.00	0.01	0.02	0.03	0.00	0.01	0.02	0.03
32	0.02	0.04	0.06	0.09	0.02	0.05	0.07	0.11	0.01	0.03	0.04	0.07	0.01	0.03	0.05	0.07
64	0.04	0.10	0.14	0.21	0.09	0.16	0.20	0.28	0.04	0.08	0.11	0.16	0.05	0.11	0.14	0.20
128	0.14	0.27	0.34	0.45	0.22	0.35	0.42	0.51	0.13	0.22	0.27	0.35	0.16	0.27	0.34	0.42
256	0.31	0.5	0.58	0.69	0.4	0.55	0.62	0.71	0.29	0.43	0.5	0.59	0.34	0.5	0.57	0.66
512	0.51	0.7	0.77	0.84	0.53	0.68	0.75	0.82	0.45	0.61	0.68	0.74	0.52	0.68	0.74	0.81
1024	0.64	0.8	0.85	0.91	0.64	0.78	0.83	0.88	0.58	0.73	0.78	0.84	0.64	0.78	0.83	0.88
2048	0.7	0.84	0.89	0.93	0.7	0.83	0.87	0.91	0.66	0.79	0.84	0.89	0.72	0.84	0.88	0.92
4096	0.75	0.88	0.92	0.95	0.75	0.86	0.9	0.94	0.7	0.83	0.87	0.91	0.74	0.87	0.9	0.93

hundred images), which means that the validation dataset had of the same size. The total number of chunks was fifty. For the MCYT-75 dataset, the starting training data consisted of fifty users (seven hundred fifty images), with chunks of five users (seventy images) for a total of five chunks.

The data used for training the classifier were the vector representation obtained by means of the last fully connected layer of four pretrained neural networks: Inception21k, SigNet, SigNet 300dpi and SigNet 600dpi. For the Inception21k network, the number of features is of 1024 and for the SigNet family, of 2048 in length. The feature extraction was done in a first step for all the signatures, therefore the whole information per user was saved in a *numpy* file format, previously to the ELM model training part. The ELM classifier was trained using different number of non-linear (hyperbolic tangent) neurons and the calculations were run on double precision.

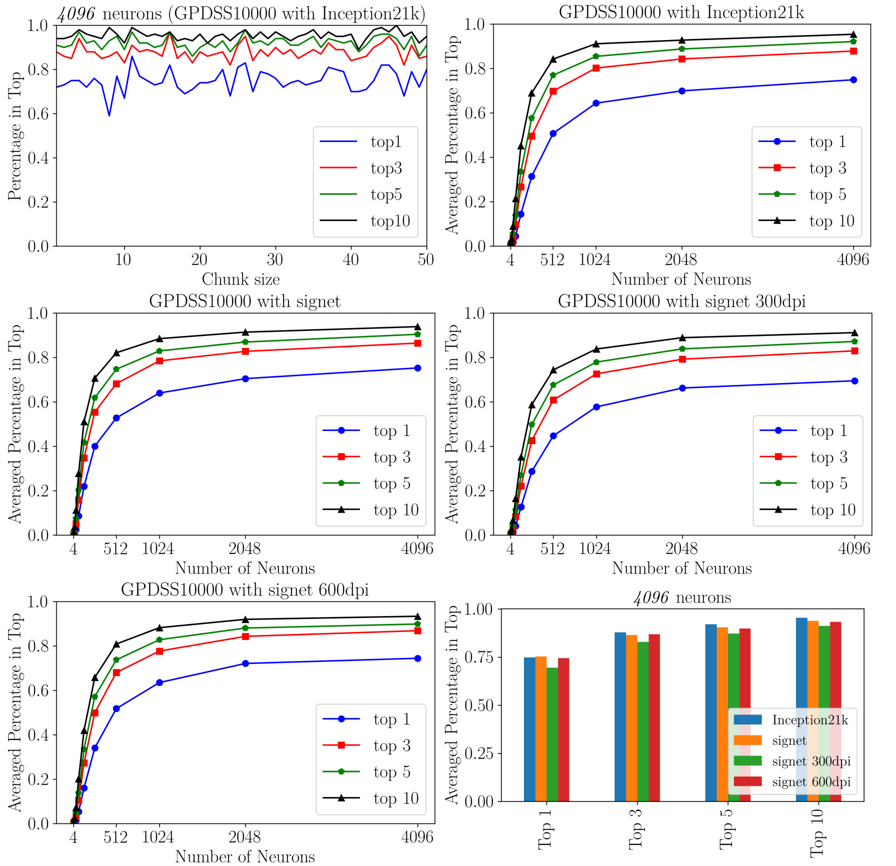

Fig. 4. Results for the GPDSS10000 dataset. Signatures in the top 1 (in blue), top 3 (in red), top 5 (in green) and top 10 (in black). *Top-left:* Percentage of signatures in top as a function of the chunk size with 4096 non-linear neurons using Inception21k as feature extractor.

5 Results and Discussion

The results of our experiments are consigned in the tables 2 and 1 and a set of representative results are presented in the Figure 4. There we consigned the results for the top 1, top 3, top 5 and top 10 for the two datasets with the four neural networks as feature extractors and different number of non-linear ELM neurons. In the case of the MCYT-75 dataset the number of non-linear neurons ranged from 4 to 128. Here is clear that the overfitting of the ELM classifier was reached for a number of non-linear neurons larger than 128. This result is expected because the amount of users present in the dataset. In the case of the GDPSS10000 dataset, the number of non-linear neurons studied ranged from 16 to 4096. The experiments show that for all four neural networks the results are similar in average, despite the fact that the SigNet family was specifically trained for signature recognition and the Inception21k was trained for the general purpose of image classification, and was trained on the full imagenet dataset [10] and based on Inception-BN network [20] with but with more capacity. In general, the average in accuracy in the four networks is of 73% for the top 1, 86% for the top 3, 90% for the top 5 and 93% for the top 10. This result shows that the signature of a given user inside of the dataset can be verified with an accuracy of more than 90% within the top 10. On the *top-left* of the Figure 4 we can see the variation in the different tops for 4096 non-linear neurons per chunk of users added into the trained model. For the other number of non-linear neurons the results were similar. The accuracy oscillates around a certain range with the addition of every chunk of users into the model. In the *top-right*, *center* and *bottom-left*, the results of the averaged normalized percentage of users within a specific top is depicted for the four neural networks used as feature extractors as a function of the number of non-linear neurons. From the results is clear that the accuracy increases with the number of non-linear neurons in the model and that there is a slightly room for improvement of the models by including more non-linear neurons. It is interesting to note here that the separation in accuracy among the tops 3 (in red), 5 (in green) and 10 (in black) is small in comparison to the distance between these and the top 1 (in blue). This result highlights one alternative strategy for the verification of users using static signatures, instead of suggesting if the signature corresponds or not to a given user, a trained classifier can suggest if the signature belongs with certain probability to a pre-established top of signatures. This method will help to eliminate the false positives caused by the well known fact that signatures of users are not so similar one from the other due to many random factors (time, mood, surface, pen design, quality of the paper among others). In the *bottom-right* the condensed results for the four networks with 4096 non-linear networks is depicted. Here it is clear that the performance of the networks are similar in accuracy, which is an indication that the accuracy can be improved regardless of the employed feature extractor.

The ELM method used here shows an excellent accuracy and more importantly, it can be re-used and extended to an arbitrary number of users without the need of retrain the whole classifier. This feature is a property of the ELM implementation and it will allows to build an universal classifier for the users because, within our proposed scheme, it is possible to combine different datasets by using a general purpose deep neural network as feature extractor.

6 Conclusion and Future Research

In this work we presented a set of results regarding the verification of users using static signatures. We proposed a new metric in which the classifier ranks the users in a top list of users instead of the classical equal error rates (EER). In general, the research on verification of signatures focuses in improving the results following this and other similar metrics, however a more versatile measure of the accuracy when users are verified could improve the trust in the algorithms presented in the scientific literature. This method allows a more versatile means of identification of users which can have a range of non-similar handwritten static signatures. We also showed that the verification of static signatures using Extreme Learning Machines is an efficient and accurate strategy that allows the incorporation of a large number of new users without the need of retrain the the classifier every time. In addition, we found that the state of the art neural networks and a general pre-trained network perform similarly as feature extractors. Therefore, further investigations could shed light on more efficient and light models constructed without degrading the accuracy of the classifier with the possibility of deployment in mobile devices.

7 Acknowledgments

The authors wish to acknowledge CSC – IT Center for Science, Finland, for computational resources.

References

1. Akusok, A., Björk, K.M., Miche, Y., Lendasse, A.: High-performance extreme learning machines: A complete toolbox for big data applications. Access, IEEE **3** (2015). https://doi.org/10.1109/ACCESS.2015.2450498
2. Akusok, A., Espinosa Leal, L., Björk, K.M.: High-performance elm for memory constrained edge computing devices with metal performance shaders. In Proceedings of the ELM2019 (2019)
3. Akusok, A., Espinosa Leal, L., Björk, K.M., Lendasse, A.: Scikit-elm: an extreme learning machine toolbox for dynamic and scalable learning. In Proceedings of the ELM2019 (2019)
4. Akusok, A., Grigorievskiy, A., Lendasse, A., Miche, Y., Villmann, T., Schleif, F.: Image-based classification of websites. Machine Learning Reports **2**, 25–34 (2013)
5. Akusok, A., Miche, Y., Björk, K.M., Nian, R., Lauren, P., Lendasse, A.: Elmvis+: improved nonlinear visualization technique using cosine distance and extreme learning machines. In: Proceedings of ELM-2015 Volume 2, pp. 357–369. Springer (2016)
6. Akusok, A., Veganzones, D., Miche, Y., Björk, K.M., du Jardin, P., Severin, E., Lendasse, A.: Md-elm: originally mislabeled samples detection using op-elm model. Neurocomputing **159**, 242–250 (2015)
7. Akusok, A., Veganzones, D., Miche, Y., Severin, E., Lendasse, A.: Finding originally mislabels with md-elm. In: ESANN (2014)
8. Burian, A., Takala, J., Ylinen, M.: A fixed-point implementation of matrix inversion using cholesky decomposition. In: Circuits and Systems, 2003 IEEE 46th Midwest Symposium on. vol. 3, pp. 1431–1434. IEEE (2003)
9. Deng, C., Huang, G., Xu, J., Tang, J.: Extreme learning machines: new trends and applications. Science China Information Sciences **58**(2), 1–16 (Feb 2015)
10. Deng, J., Dong, W., Socher, R., Li, L.J., Li, K., Fei-Fei, L.: Imagenet: A large-scale hierarchical image database. In: 2009 IEEE conference on computer vision and pattern recognition. pp. 248–255. Ieee (2009)

11. Diaz, M., Ferrer, M.A., Impedovo, D., Malik, M.I., Pirlo, G., Plamondon, R.: A perspective analysis of handwritten signature technology. ACM Computing Surveys (CSUR) **51**(6), 117 (2019)
12. Eirola, E., Gritsenko, A., Akusok, A., Björk, K.M., Miche, Y., Sovilj, D., Nian, R., He, B., Lendasse, A.: Extreme learning machines for multiclass classification: refining predictions with gaussian mixture models. In: International Work-Conference on Artificial Neural Networks. pp. 153–164. Springer (2015)
13. Espinosa Leal, L., Akusok, A., Lendasse, A., Björk, K.M.: Classification of websites via full body renders. In Proceedings of the ELM2019 (2019)
14. Ferrer, M.A., Diaz, M., Carmona-Duarte, C., Morales, A.: A behavioral handwriting model for static and dynamic signature synthesis. IEEE transactions on pattern analysis and machine intelligence **39**(6), 1041–1053 (2016)
15. Hafemann, L.G., Oliveira, L.S., Sabourin, R.: Fixed-sized representation learning from offline handwritten signatures of different sizes. International Journal on Document Analysis and Recognition (IJDAR) **21**(3), 219–232 (2018)
16. Hafemann, L.G., Sabourin, R., Oliveira, L.S.: Writer-independent feature learning for offline signature verification using deep convolutional neural networks. In: 2016 International Joint Conference on Neural Networks (IJCNN). pp. 2576–2583. IEEE (2016)
17. Hafemann, L.G., Sabourin, R., Oliveira, L.S.: Learning features for offline handwritten signature verification using deep convolutional neural networks. Pattern Recognition **70**, 163–176 (2017)
18. Hafemann, L.G., Sabourin, R., Oliveira, L.S.: Offline handwritten signature verification—literature review. In: 2017 Seventh International Conference on Image Processing Theory, Tools and Applications (IPTA). pp. 1–8. IEEE (2017)
19. Huang, G.B., Zhou, H., Ding, X., Zhang, R.: Extreme learning machine for regression and multiclass classification. IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) **42**(2), 513–529 (April 2012)
20. Ioffe, S., Szegedy, C.: Batch normalization: Accelerating deep network training by reducing internal covariate shift. arXiv preprint arXiv:1502.03167 (2015)
21. Jain, A.K., Ross, A., Prabhakar, S., et al.: An introduction to biometric recognition. IEEE Transactions on circuits and systems for video technology **14**(1) (2004)
22. Leal, L.E., Björk, K.M., Lendasse, A., Akusok, A.: A web page classifier library based on random image content analysis using deep learning. In: Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference. pp. 13–16. ACM (2018)
23. Lendasse, A., Akusok, A., Simula, O., Corona, F., van Heeswijk, M., Eirola, E., Miche, Y.: Extreme learning machine: A robust modeling technique? yes! In: International Work-Conference on Artificial Neural Networks. pp. 17–35. Springer (2013)
24. Michel, J., Holbrook, Z., Grosser, S., Strobelt, H., Shah, R.: Ennui elegant neural network user interface , https://math.mit.edu/ennui/
25. Ortega-Garcia, J., Fierrez-Aguilar, J., Simon, D., Gonzalez, J., Faundez-Zanuy, M., Espinosa, V., Satue, A., Hernaez, I., Igarza, J.J., Vivaracho, C., et al.: Mcyt baseline corpus: a bimodal biometric database. IEE Proceedings-Vision, Image and Signal Processing **150**(6), 395–401 (2003)
26. Sovilj, D., Sorjamaa, A., Yu, Q., Miche, Y., Séverin, E.: Opelm and opknn in long-term prediction of time series using projected input data. Neurocomputing **73**(10-12), 1976–1986 (2010)