Building A Keystroke Dynamic Recognition System Using An Improved Accelerated Method

Ula Tarik Salim [1], Sahar Lazim Qaddoori [2], Noor Mowafeq Allayla[3]

[1] Computer Engineering Department, Engineering Collage, Mosul University, Mosul /Iraq
ORCID No: https://orcid.org/0000-0002-3128-7134
[2] Electronic Engineering Department, Electronics Engineering Collage, Ninevah University, Mosul/ Iraq
ORCID No: https://orcid.org/0000-0001-6631-5329
[3] Computer Engineering Department, Engineering Collage, Mosul University, Mosul /Iraq
ORCID No: https://orcid.org/0000-0002-2137-1199

Keywords

Deep Learning, Keystroke Dynamic, Residual CNN, Parallel Computing, Authentication System

Abstract

A trustworthy security application in the actual world uses the keystroke capability of typing recognition software. Despite being straightforward, it requires a quick and precise method of character analysis. In this manuscript, a keystroke dynamic recognition method to identify and block illegal users is proposed by using deep learning of convolutional neural networks (CNNs) which it can efficiently distinguish legitimate users from illegitimate users. Where, two proposed networks are built based on 1D-CNN to increase and accelerate the recognition abilities. The first network improves the system performance by modifying the kind of activation function utilized, whereas the second network improves the system performance by employing the residual scheme. The findings display that the suggested CNN model with the swish function can deny all illegitimate users with an average equal error rate (EER) of 0.0066. Furthermore, by Graphic Processing Unit (GPU), the model performance is accelerated by approximately 2 times. Based on the outcomes, the suggested CNN model with swish function significantly outperforms other models in the literature.

Research Article
Submission Date : 14.04.2023
Accepted Date : 19.09.2023

1. INTRODUCTION

The advancement of digital information systems and the usage of biometrics today have significantly improved how to handle the variety of user data and analyze it quickly and efficiently[1] for many enterprises. Keystroke dynamic is one of the categories of behavioral measurements for users, and it relies on identifying the user based on how they behave when using a keyboard. To do this, it specifies a set of temporal characteristics that correspond to the user's behavior when using the keyboard, such as

[1] Resp author; e-mail: ula.tariq@uomosul.edu.iq

dwell time, flight time, or the time taken for keys to transition between each other, as well as their typing speed. Because this technique relies heavily on software, it may be used as a different approach to assessing user data in a variety of contexts, such as the identification of mental weariness (Maheshwary, Ganguly & Pudi, 2017; Kim & Kang, 2020). Although it would appear that there are very few distinctions between people's typing styles, this is not the case. When typing using the keyboard, users may have varied typing patterns. These patterns, which may include using the left or right shift keys or varying the duration between two key presses, make typing more unique. One can identify between authorized and unauthorized users by taking use of such distinctive behaviors. These biological traits, however, are not likely to be discernible with the human eye or with standard programmed decision-making methods (Lin, Liu & Lee, 2018). As a result, in this paper, deep learning's potent capabilities are utilized. Through a variety of learning processes, the suggested system may finally discriminate between authorized and unauthorized users. Artificial intelligence in the form of deep learning identifies a preliminary answer to a problem using induction and reasoning, then updates the solution based on learning experiences so that it eventually approaches the proper value (Lemley, Bazrafkan & Corcoran, 2017). The deep learning model has been examined in this paper: Convolutional Neural Networks (CNNs). Because the CNN cascades several features to allow features to associate with one another while doing feature analysis, which can considerably improve the accuracy of the prediction or analysis. The authentication system based on the proposed CNN model can be used in multiple applications like Electronic Health Record Systems (EHR) (Wesołowski, Porwik & Doroz, 2016), Online Learning Environments (OLE) (Muniasamy, 2019), etc.

A keystroke dynamic recognition method is proposed by using convolutional neural networks (CNNs) which can efficiently distinguish legitimate users from illegitimate users. The contributions of this paper are as follows:

- A CNN architecture is built by replacing the Rectified Linear Unit (ReLU) with Swish activation to increase the propagation of a few negative weights across the CNN architecture.
- One layer residual CNN design is developed to aid in assisting the Graphic Processing Unit (GPU) in expediting training. process.

The structure of this paper is as follows: In section 2, several previous works concerning deep learning are presented. The next section displays the methodology used to build the proposed system. In section 4, the results of the proposed system are analyzed, discussed, and compared with related works. Finally, the conclusions and future works are mentioned in section 5.

2. RELATED WORKS

This section offers an overview of the studies applying keystroke dynamics as a biometric technique for the verification and identification of users, emphasizing the work by Killourhy and Maxion, conducted on the CMU dataset (Killourhy & Maxion, 2009).

In this research (Çeker & Upadhyaya, 2017), the convolutional neural network and the Gaussian data augmentation approach are used to look at efficacy of the deep learning when applied on three different datasets. Authors revised up the precision of previous

approaches with 10%. They also reduced the equal error rate (EER) by 7.3%, which is 10% better than the one proposed earlier. Because of that, the analysis suggests that the crucial part of a network which learned a keyboard input is a fully-connected layer conjugated by a convolution operation as they are the one that mainly affect the network's accuracy and speed of learning.

The research (Andrean, Jayabalan & Thiruchelvam, 2020) applied to the model using the CMU benchmark dataset a deep learning model based on Multilayer Perceptron (MLP) for keystroke typing dynamics for the purpose of user verificiation. With regard to the other reference classifiers, the MLP`s optimal EER was only 4.45%, while the outlier count`s result was 9.96%, Mahalanobis Nearest Neighbor`s was 9.96%, and for the scaled Manhattan`s it was even 9.6%. We could also say that the results of the Outlier Count classifier (Altwaijry, 2020) took it one step further by introducing CNN-Detect, a convolutional neural network that processes characteristics of typical human typing patterns in an attempt to recognize culprits. Having evaluated the data with appropriate feature engineering, CMU keystroke dynamics dataset which can be accessed publicly, the suggested model was tested with an average EER of 0.009 and ZM-FAR of 0.027 being found.

In this study (Sahu & Banavar, 2021), the authors enrich the model that counts the users and assigns their place in the queue to those who had similar typing rhythms. Therefore, an idea about eliminating outliers was a quantile transformation that was supposed to turn unprocessed keystroke characteristics to a uniform distribution. Last, the rotated the feature space is recreated using projects on the lower-dimensional space through principal component analysis and other methods. Using the k-means clustering to examine the number of people utilizing the system in a smaller feature space was the algorithm selected for this experiment. Our findings were served mathematically through the vector search approach and the labeling of tightness clusters. The method was verified to be more than 93% accurate and were validated through the use of two standards datasets namely MobiKey and the CMU keystroke benchmark dataset.

The research (Mao, Wang & Ji, 2022) described the development of a key stroke dynamics authentication mechanism for identifying users based on deep learning. Bi-directional LSTM-based model (BI-LSTM), CNN, and the attention mechanism become an interlaced structure. This model embraces both character's typing speed and writing context. First of all, the CNN conducts the operations with the characteristic vectors that comprise data. The normalized sequence goes into the bi-LSTM network to be trained. Through the use of the Buffalo Open Data Collection to judge, the model is assessed. The findings demonstrate that the False Accept Rate (FAR), Equal Error Rate (EER), and False Reject Rate (FRR) are correspondingly 3.03%, 4.23%, and 3.09%.

3. METHODOLOGY

This section includes the experimental settings and details of the two suggested networks. The two proposed networks are built based on 1D-CNN to increase and accelerate the recognition abilities. The first network improves the system performance by modifying the kind of activation function utilized, as given in Figure 1, whereas the second network improves the system performance by employing the residual scheme, as given in Figure 2.

The CMU Keystroke Dynamics Dataset (Killourhy, 2009) was used for the tests to assess the performance of the suggested network topologies. The 50 people are utilized as illegitimate users while one person is chosen to be a legitimate user. The dataset is divided into training samples (80%) and testing samples (20%). The two networks receive a series of 50 user samples (N) and 31 feature vectors (Vi), and they output either legal or illegal users as two outcomes.

The first network's proposed structure (as given in Figure 1) consists of four convolutional blocks: a pair of 1D-convolution layers, an activation function, and a layer for batch normalization. A 1D-convolution layer extracts the features from input data using a collection of filters that perform a convolution operation with a stride of 1 and the same padding. The first and second convolutional blocks apply 16 and 32 filters, respectively, to the data. The network is training two times, one employing the Swish activation function and another using the ReLU activation function. To improve classification accuracy and speed up the training process, batch normalization adds a transformation to the network. The findings are then sent to a one-dimensional average pooling layout with a stride of two to reduce memory usage and improve network translation, distortion, and scaling robustness. The network employed the dropout and early stopping regularization techniques to lessen system overfitting. The probability values for the needed two classes were then produced using two fully connected (FC) layers and a softmax classifier. On the other hand, the second suggested network is made up of two fully connected layers, a SoftMax classifier, one residual convolutional block, and 1D-average pooling strided with two. The convolutional block's structural components include a layer of 1D-convolution strung together in a stack, 64 filters of the same padding, a layer of ELU activation function, and a layer of batch normalization with the residual connection. In this study, residual connections enable features information to pass from the convolutional layer as the input layer to the batch normalization layer as the output layer and apply addition operation, as given in Figure 2. For knowledge, the type of the penultimate FC layer in the three methods is the ReLU type.

The suggested networks are learned using the following parameters: 32-batch size, Adam optimizer, learning rate = 0.001, momentums = 0.9 and 0.99, categorical cross entropy as loss function, 100 epochs, and EarlyStopping to halt training when the parameter updates do not improve performance on the validation set.

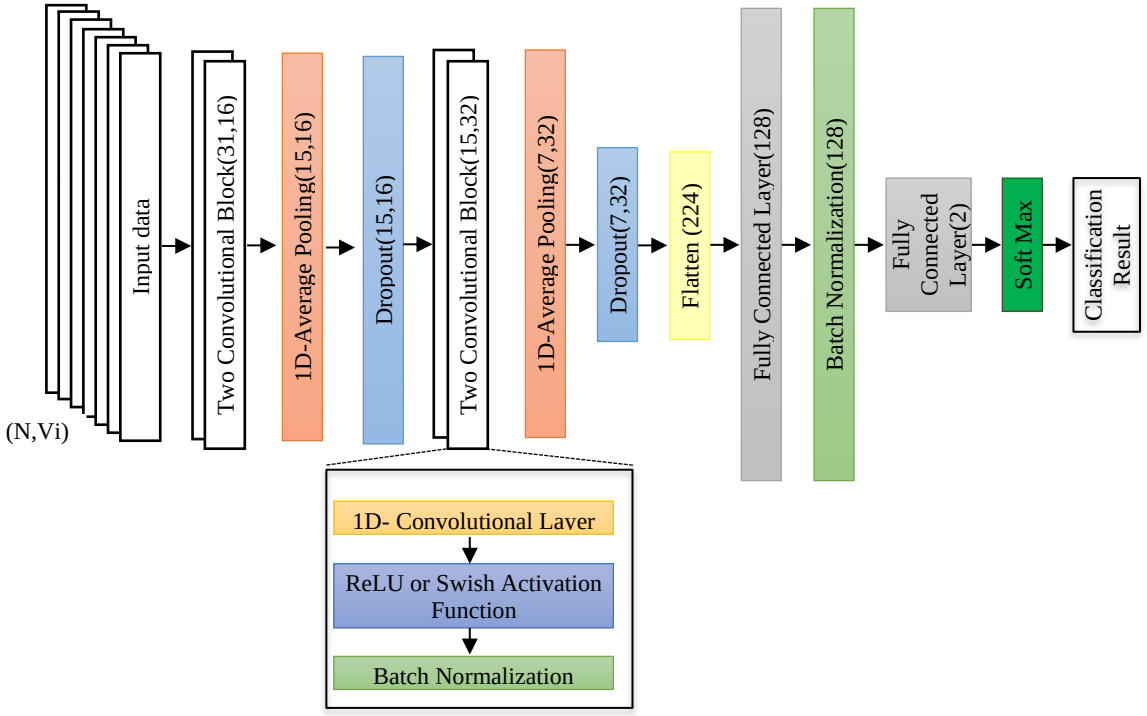

Figure 1. The proposed architecture of CNN

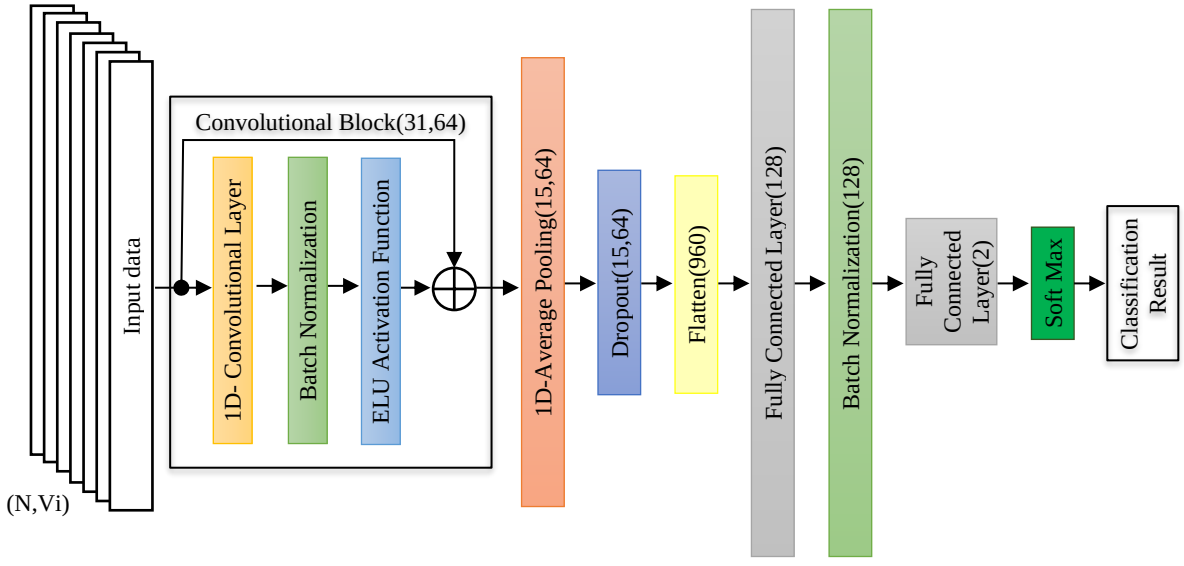

Figure 2. The proposed architecture of one layer Residual CNN

4. RESULTS AND DISCUSSION

The recognition findings with their discussion are presented in this section. The proposed networks are implemented on two computers with the characteristics listed in Table 1 and constructed in Python software using the frameworks of TensorFlow and Keras.

Table 2 provides a summary of many average findings after modifying numerous different hyperparameters for the suggested CNN and the residual CNN models. The proposed CNN architecture with the Swish activation function produces the best EER value. This is because the Swish function has a smoothness feature that allows a relatively small number of negative weights to get through, whereas ReLU limits all negative weights to zero. Nevertheless, despite being, the architecture-based Residual connection produces an adequate performance.

Table 1. Machines specifications

	Equipment	Machine 1	Machine 2
CPU	Computing name	Intel i7-7500U	Intel i7-10700
	#core	2	2
	#thread	4	4
	Clock frequency	2.7GHz	2.9GHz
GPU	Architecture	Maxwell	Turing
	Computing name	940MX	RTX 2060 super
	CUDA Driver Version	10.2	11.2
	CUDA Capability	5.0	7.5
	# Multiprocessors	3	34
	# CUDA Cores	384	2176
	Clock frequency	1.242GHz	1.68GHz
	Memory bandwidth	14.4 GB/s	448 GB/s

Table 2. Average results of recognition

Evaluation measure	Architecture based ReLU	Architecture based Swish	Architecture based Residual connection
Accuracy	99.3%	99.3%	99.1%
Recall	98.6%	98.7%	98.3%
Precision	100%	99.9%	99.9%
F1	99.3%	99.4%	99.1%
EER	0.0070	0.0066	0.0091

Table 3 displays the training time (per epoche) and testing time when applied on two distinct machines. Because GPUs parallel computing has a greater number of cores, the implementation is faster than using CPUs and is suited for the suggested designs. Furthermore, because of the distinct network architectural design process and its confinement on one layer, the residual network was faster than other networks.

Table 4 shows a comparison of the number of parameters and FLOPs for the proposed models, where the residual model requires a higher design cost than the two proposed architectures (Architecture-based ReLU and Architecture-based Swish function) because it contains more filters as well as the use of residual connections.

Table 3. Time of different equipment of GPU and CPU

Machine	Equipment	ReLU-model		Swish-model		Residual-model	
		Training (ms)	Testing (ms)	Training (ms)	Testing (ms)	Training (ms)	Testing (ms)
Machine 1	Intel i7-7500U	4060	281	4184	301	3286	230
	940MX	3330	190	4105	200	3130	160
Machine 2	Intel i7-10700	1953	170	2093	284	1162	124
	RTX 2060 super	1876	125	1969	126	960	78

Table 4. The architectural and computational complexities of the proposed models

Model	Architecture based ReLU	Architecture based Swish	Architecture based Residual connection
Parameters	35,474	35,474	124,290
FLOPs	255 K	257 K	270 K

The average EER result of the best proposed architectural design was compared to several recent works and outperformed them, as given in Table 5, because of building the suggested network differently and using the swish function.

Table 5. Comparative average EER of the proposed system with the related works methods over CMU dataset.

Ref.	Method	EER	Accuracy
(Killourhy and Maxion 2009)	Manhattan (scaled)	0.096	N/A
(Çeker and Upadhyaya 2017)	CNN	0.065	94%
(Andrean, Jayabalan et al. 2020)	MLP	N/A	96.7%
(Altwaijry 2020)	CNN-Detect	0.009	N/A
(Sahu and Banavar 2021)	Nearest neighbor rule	N/A	93%
(Mao, Wang et al. 2022)	ACBM	0.0423	N/A
Our Work	Proposed method	0.0066	99.3%

5. CONCLUSION

The field of keystroke dynamics biometrics has developed in recent years. The primary motivation for most works is how easily and affordably keystroke dynamics biometrics may be included in prevailing computer security systems without user involvement. Numerous studies on data collection techniques, feature representations, classification algorithms, and experimental approaches have been conducted. The experimental technique has been emphasized in this paper. By altering the kind of activation function used, the first networgivenk enhances system performance, whereas the second network enhances system performance by utilizing the residual scheme. The experimental findings demonstrate that the suggested CNN model is capable of blocking all unauthorized users with an average EER of 0.0066.

Moreover, parallel processing (GPU) enhances system performance by a factor of two. The suggested CNN model with the swish function greatly outperforms other models in the literature based on the experimental results. More parallel architectures will be used to implement the model in the future work.

Conflict of Interest

The authors declare no conflict of interest

Contribution of Authors

All authors contributed to the conception, design, examination, resource allocation, data collection, literature review, and analysis and interpretation aspects of the research.

REFERENCES

Altwaijry, N. (2020). Keystroke Dynamics Analysis for User Authentication Using a Deep Learning Approach. International Journal of Computer Science and Network Security, 20(12), 209–216. https://doi.org/10.22937/IJCSNS.2020.20.12.23

Andrean, A., Jayabalan, M., & Thiruchelvam, V. (2020). Keystroke Dynamics Based User Authentication using Deep Multilayer Perceptron. International Journal of Machine Learning and Computing, 10(1), 134–139. https://doi.org/10.18178/ijmlc.2020.10.1.910

Ceker, H., & Upadhyaya, S. (2017). Sensitivity analysis in keystroke dynamics using convolutional neural networks. 2017 IEEE Workshop on Information Forensics and Security, WIFS 2017, 2018-Janua, 1–6. https://doi.org/10.1109/WIFS.2017.8267667

Killourhy, K. (2009). "Keystroke dynamics – benchmark dataset", Carnegie-MellonUniversity, http://www.cs.cmu.edu/~keystroke/#sec2 ,

Killourhy, K. S., & Maxion, R. A. (2009). Comparing anomaly-detection algorithms for keystroke dynamics. 2009 IEEE/IFIP International Conference on Dependable Systems & Networks, 125–134. https://doi.org/10.1109/DSN.2009.5270346

Kim, J., & Kang, P. (2020). Freely typed keystroke dynamics-based user authentication for mobile devices based on heterogeneous features. Pattern Recognition, 108. https://doi.org/10.1016/j.patcog.2020.107556

Lemley, J., Bazrafkan, S., & Corcoran, P. (2017). Deep Learning for Consumer Devices and Services: Pushing the limits for machine learning, artificial intelligence, and computer vision. IEEE Consumer Electronics Magazine, 6(2), 48–56. https://doi.org/10.1109/MCE.2016.2640698

Lin, C. H., Liu, J. C., & Lee, K. Y. (2018). On neural networks for biometric authentication based on keystroke dynamics. Sensors and Materials, 30(3), 385–396. https://doi.org/10.18494/SAM.2018.1757

Maheshwary, S., Ganguly, S., & Pudi, V. (2017). Deep secure: A fast and simple neural network based approach for user authentication and identification via keystroke

dynamics. IWAISe: First International Workshop on Artificial Intelligence in Security, August 2017, 59. https://api.semanticscholar.org/CorpusID:53459138

Mao, R., Wang, X., & Ji, H. (2022). ACBM: attention-based CNN and Bi-LSTM model for continuous identity authentication. Journal of Physics: Conference Series, 2352(1). https://doi.org/10.1088/1742-6596/2352/1/012005

Muniasamy, A. (2019). Applications of keystroke dynamics biometrics in online learning environments: A selective study. Biometric Authentication in Online Learning Environments, 97–121. https://doi.org/10.4018/978-1-5225-7724-9.ch005

Sahu, C., & Banavar, M. (2021). A nonlinear feature transformation-based multi-user classification algorithm for keystroke dynamics. Conference Record - Asilomar Conference on Signals, Systems and Computers, 2021-Octob(March 2022), 1448–1452. https://doi.org/10.1109/IEEECONF53345.2021.9723223

Wesołowski, T. E., Porwik, P., & Doroz, R. (2016). Electronic Health Record Security Based on Ensemble Classification of Keystroke Dynamics. Applied Artificial Intelligence, 30(6), 521–540. https://doi.org/10.1080/08839514.2016.1193715