Temporal Residual Network Based Multi-Head Attention Model for Arabic Handwriting Recognition

Ramzi Zouari
National Engineering School of Sfax
University of Sfax, Tunisia
ramzi.zouari@gmail.com

Dalila Othmen
Faculty of Sciences of Gafsa
University of Gafsa, Tunisia
othmendalila4@gmail.com

Houcine Boubaker
National Engineering School of Sfax
University of Sfax, Tunisia
boubaker-houcine@ieee.org

Monji Kherallah
Faculty of Sciences of Sfax, University of Sfax, Tunisia
monji.kherallah@fss.usf.tn

Abstract: *In this study, we developed a new system for online Arabic handwriting recognition based on temporal residual networks with multi-head attention model. The main idea behind the application of attention mechanism was to focus on the most relevant parts of the data by a weighted combination of all input sequences. Moreover, we applied beta elliptic approach to represent both kinematic and geometric aspects of the handwriting motion. This approach consists of representing the neuromuscular impulses involving during the writing act. In the dynamic profile, the curvilinear velocity can be fitted by an algebraic sum of overlapped beta functions, while the original trajectory can be rebuilt by elliptic arcs delimited between successive extremum velocity instants. The experiments were conducted on LMCA database containing the trajectory coordinates of 23141 Arabic handwriting letters, and showed very promising results that achieved the recognition rate of 97,12%.*

Keywords: *Beta stroke, velocity, residual, skip connection, attention, vanishing gradient.*

Received March 31, 2023; accepted May 10, 2023
https://doi.org/10.34028/iajit/20/3A/6

1. Introduction

Nowadays, handwriting recognition has become one of the most important applications in the document analysis field. It enables the digitization of information contained in different types of documents such as text, invoice, bank check and forms. Compared to Optical Character Recognition systems (OCR), handwriting recognition is faced with several challenges allowing it to remain an active research area. In fact, handwriting has a great deal of intra-class variability, as each person has their own writing style [7]. Moreover, the writing act is sensitive to the psychological and emotional state of the writer, which can degrade the quality of writing and consequently altering the recognition model performances. On the other hand, handwriting style is generally cursive when the letters of a word are joined together in a continuous flow [14]. This property makes the handwriting recognition task more difficult, given the need to implement an efficient algorithm to segment a word into letters.

For a few decades, handwriting recognition is done in offline mode, where the scanned document is considered as an image, and therefore only the pixel values were considered [12]. With the increased use of data entry devices such as smartphones and tablets, handwriting recognition can be also done in online mode [22]. In this case, the trajectory coordinates in time are available and subsequently dynamic information can be extracted like velocity, temporal order of generation and the stylus pressure. In this acquisition mode, the recognition model processes sequences of input vectors rather than images [13].

In this work, we are focused on online Arabic handwriting recognition task, which is among the most difficult scripts to process. In addition to its cursive style, almost of Arabic letters change their shape according to their position in the word. Furthermore, we note the presence of diacritics or delayed strokes (dots, chadda, nabra, fatha) which can be written either below or above the main body of the letter [5]. The proposed model proceeds by segmenting the handwriting trajectory into strokes, and we extracted their dynamic and geometric parameters according to the beta elliptic approach. Thereafter, a temporal residual network based on these parameters was trained with the addition of attention mechanism.

The rest of this paper is organized as follow: the next section presents an overview of the existing online handwriting modeling and recognition systems. In the section 3, beta elliptic modeling approach is described. The proposed recognition system will be detailed in the fourth section. The section 5 presents and discusses the experiments results, and we finish by a conclusion and some prospects.

2. State of the Art

Unlike offline handwriting recognition, the online acquisition mode allows to have dynamic information about the handwriting generation process. For this purpose, several works were interested in the study of biological phenomena related to the trajectory formation. In fact, handwriting movement is considered as a neuromuscular act, ordered by the central nervous system that activate the proper nerve and muscle pairs responsible of generating the desired trajectory shape [29]. In this context, several attempts have been made to represent the neuromuscular impulses involved during the writing act., Plamondon *et al*. [20] suggested that each neuromuscular impulse converges to a gaussian function shape. This approach falls in the case of asymmetric movement where the acceleration and deceleration periods are not equals. As alternative, they proposed to use the concatenation of two half gaussian functions to represent the asymmetric effect. In other work, delta lognormal function was proposed to describe the generation of rapid human movement. The writing act is considered as the synergy between a pair of agonist and antagonist neuromuscular networks [21]. Each group of muscles tends towards a lognormal curve. This method allows representing both symmetric and asymmetric movement, but the lognormal curve is unbounded where the zero velocity is asymptotically reached. Then, the neuromuscular impulse effect continues to exist after the end of the movement. Alimi *et al*. [4] demonstrates that each neuromuscular impulse obeys to a beta function shape. In fact, beta function can represent both symmetric and asymmetric movements. Moreover, it is bounded and subsequently it has no effect once the movement is finished. Bezine *et al*. [8] presented an extended version of beta model, called beta elliptic model. The geometric proprieties of the involved muscles have been studied. In addition to the kinematic characteristics, each neuromuscular impulse delimited between two successive extrema speed times is represented in the spatial domain by an elliptic arc. This approach has proven its effectiveness for several handwriting applications like Arabic and Latin scripts recognition, script identification [26], and signature verification [13].

Apart of studying the phenomena related to the handwriting generation process, several works have focused on the recognition model. The existing systems can be classified into holistic and analytic approaches depending on whether the recognition is made without or with a preliminary step of segmentation. Generally, systems based analytic approaches are more efficient because they deal with sub-parts of the trajectory, called strokes. In [2], the stroke is delimited between two consecutive points of the trajectory, and it is classified according to its slant direction. Thereafter, Discrete Cosine Transform (DCT) is applied to extract frequency domain features. Wilson-Nunn et al. [24] proposed to segment the handwriting trajectory into strokes of equal length. For each segment, a linear interpolation is applied to calculate its dyadic signature. This approach was evaluated on Online KHATT character dataset with a fixed number of five segments. Long-Short Term Memory (LSTM) model was applied for classification, and they obtained the accuracy of 92.5%. Hamdi *et al*. [16] presented a recognition system using grapheme segmentation. This technique is based on detecting baseline and median zone to reveal bottom of ligature valleys and angular points. Each grapheme is characterized by multiple features combining beta-elliptic modeling, Fourier descriptors, and geometric localization parameters. This approach was assessed on ADAB database using both LSTM and Bidirectional LSTM (BLSTM) and achieved the recognition rates of 96.65% and 98.73% respectively. Abdelaziz *et al*. [1] defined the grapheme as a continuous trajectory segment delimited between pen-down and pen-up moments. Depending on the grapheme shape complexity, mono-grapheme, and tri-grapheme Hidden Markov Models (HMM) were used in the training stage. In [25], the proposed system proceeds by segmenting the trajectory into strokes according to the extremum points of velocity. Thereafter, both dynamic and geometric parameters have been extracted with respect to beta-elliptic modeling. The proposed approach was tested on LMCA database using Time Delay Neural Network (TDNN). This system was improved by applying TDNN with Support Vector Machines (SVM) classifier in the recognition engine [28]. The recognition rate achieved the accuracy of 99.52%.

Recently, deep learning-based approaches have shown highly successful in online handwriting recognition task. In this context, Recurrent Neural Networks (RNN) and their variants LSTM, BLSTM and Multi-Directional LSTM are the most used. Hamdi *et al*. [17] presented hybrid BLSTM-SVM recognition system. They combined offline features extracted from CNN with dynamic parameters and reached the accuracy rate of 99.11% and 93.98% for LMCA and Online-KHATT databases respectively. Alwajih *et al*. [6] introduced a transformer-based model according to the encoder-decoder architecture. Connectionist Temporal Classification (CTC) network is implemented in the last stage to recognize word without needing of explicit segmentation. This model achieved 5% of Character Error Rate (CER) and 18% of Word Error Rate (WER) on CHAW dataset, and 22% CER and 56% WER on Online-KHATT database. Othmen *et al*. [19] applied CNN based temporal convolution to process the input data sequentially. Their model was evaluated on LMCA dataset and achieved the recognition rate of 96.87%.

According to the literature review, we applied the beta elliptic approach for online handwriting modeling. In the recognition model, we developed a new version of residual neural networks (ResNet) capable of treating sequential input data. Moreover, we applied attention mechanism to the ResNet model to concentrate the processing in the most relevant parts of data at each timestamp and improving the model effectiveness.

3. Online Handwriting Modelling

The beta elliptic approach has the advantage of modeling simultaneously the kinematic and geometric profiles of the handwriting movement. In fact, the generation of handwriting movement is the result of overlapped neuromuscular subsystems, which are mobilized to deputize progressively the dynamic of the motion drive. In the geometric profile, each neurophysiologic beta impulse leads to design an elliptic allure shape.

3.1. Velocity Profile Modelling

During the handwriting act, N neuromuscular subsystems are mobilized. Their effects on the velocity domain can be modeled by overlapped beta impulses. The idea behind the overlapping is to guarantee a non-null velocity between two successive strokes, which reflects a change of direction in the trajectory shape [9]. Then, the curvilinear velocity $V_\sigma(t)$ can be approximated by the algebraic sum of the successive beta functions in accordance with Equations (1) and (2). The parameter K_i represents the amplitude of i^{th} beta impulse and (p_i, q_i) are intermediate parameters which have an influence on the symmetry and the width of beta function shape. For the i^{th} beta function, the parameters t_{0i}, t_{ci} and t_{1i} represent respectively the starting time, the instant of the culmination amplitude and the ending time.

$$V_\sigma(t) = \sqrt{\left(\frac{dx}{dt}\right)^2 + \left(\frac{dy}{dt}\right)^2} \tag{1}$$

$$V_\sigma(t) \approx \sum_{i=1}^{n} K_i \times \beta_i\,(t, q_i, p_i, t_{0i}, t_{1i}) \tag{2}$$

With

$$\beta_i\,(t, q_i, p_i, t_{0i}, t_{1i}) = \begin{cases} \left(\dfrac{t - t_{0i}}{t_{ci} - t_{0i}}\right)^{p_i} \left(\dfrac{t_{1i} - t}{t_{1i} - t_{ci}}\right)^{q_i}, & \text{t} \in [t_{0i}, t_{1i}] \\ 0 & \text{elsewhere} \end{cases}$$

$$t_{ci} = \frac{(p_i \times t_{1i})(q_i \times t_{0i})}{p_i + q_i}$$

3.2. Geometric Profile Modelling

In the space domain, each stroke delimited between two successive extrema speed times can be assimilated by an elliptic arc. In this context, three methods of reconstruction were proposed: quarter of ellipse, two-tangents endpoints, and arc oblique projection [11]. For each script, we have chosen the method that return the least error of reconstruction. Figure 1 shows the application of beta elliptic modeling on the Arabic letter (ـم). This letter is composed of 11 beta strokes relative to the extremum velocity points. We obtained a good fitting of both curvilinear velocity and the original trajectory when applying beta elliptic approach.

According to this modeling approach, each stroke delimited between two successive extrema of velocity is characterized by a vector of 8 parameters: 5 dynamic parameters relative to the beta impulse, and 3 geometric features representing the major and minor axes length, and the incline angle of the elliptic arc [27] (Table 1).

Table 1. Beta elliptic parameters.

	Parameters	Explanation
Dynamic profile	K	Beta impulse amplitude
	$\Delta t=(t_1-t_0)$	Beta impulse duration
	Rap= p/(p+q)	Rapport of beta impulse asymmetry or culminating time
	P	Beta shape parameters
	K_i/K_{i+1}	Rapport of successive Beta impulse amplitude
Geometric profile	a	Ellipse major axis half length
	b	Ellipse small axis half length
	θ	Ellipse major axis inclination angle

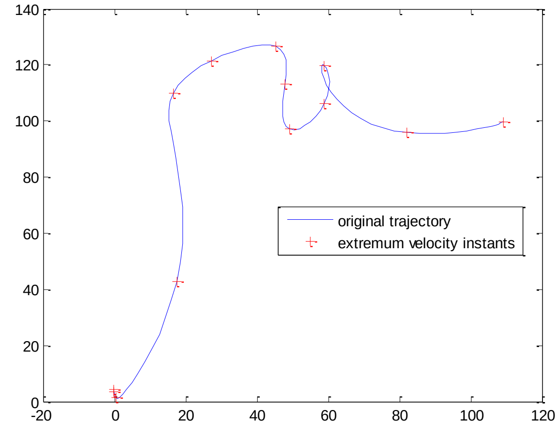
4. Handwriting Recognition Model

The recognition model is based on Residual Neural Networks which is a class of CNNs mainly applied for computer vision tasks [18]. Unlike plain networks, ResNet can support more hidden layers without falling into the vanishing gradient problem thanks to the introduction of residual blocks. These blocks are based on shortcut connections that skip one or more layers to keep a non-null value of gradient during the training stage. In this work, we tried to adapt ResNet for sequential input data, and especially beta-elliptic parameters. So, we omitted the conventional convolution layer and substituted it by a temporal convolution function.

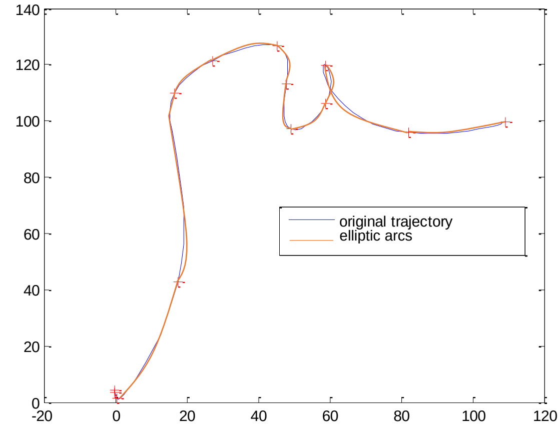
4.1. Temporal Convolution Concept

Temporal convolution requires reshaping the input data along two axes representing the timestamp and the features respectively. In our case, the horizontal axis corresponds to the number of beta strokes (timestamp), while the vertical axis contains the beta elliptic parameters relative to each stroke. Moreover, we defined a sliding window that shift along the vertical axis. This window defines the number of beta strokes to process simultaneously. At each step, we took the dot product of features within the sliding window with a kernel tensor of learned weights (Figure 2). This process is done with multiple kernels in order to project the input features into another vector space of higher dimension. The output tensor shape according to the vertical axis is relative to the number of applied kernels, while the length along timestamp axis can be calculated with respect to the length of input tensor *inp*, the size of the sliding window *w*, and the shift *s* between two consecutive sliding windows (Equation 3).

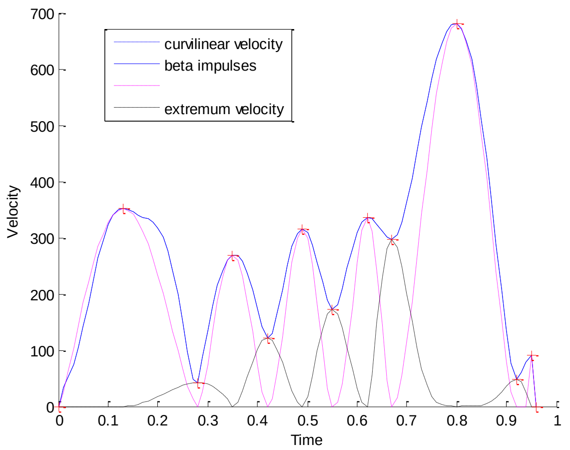
$$output_{length} = \frac{input - w}{s} + 1 \quad (3)$$



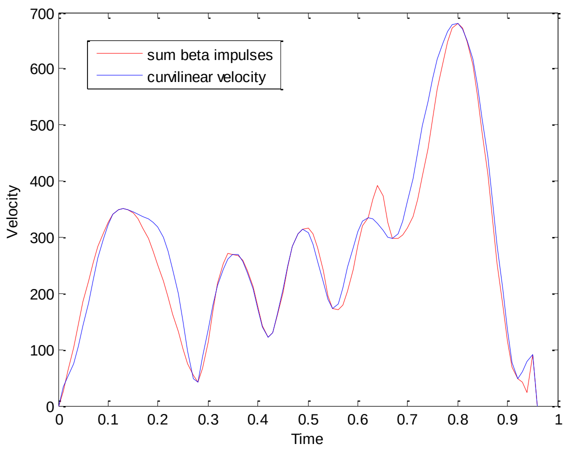
a) Original trajectory.



b) Trajectory fitting by elliptic arcs.



c) Velocity fitting by overlapped beta impulses.



d) Sum of overlapped beta impulses.

Figure 1. Beta elliptic modeling of the Arabic letter “Mim”.

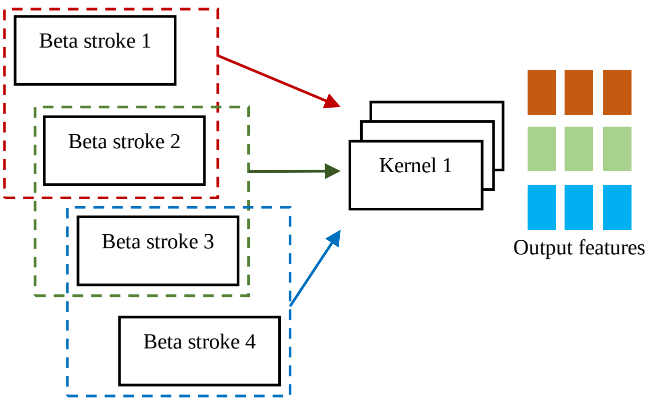


Figure 2. Temporal convolution layer concept.

4.2. Temporal Residual Blocks

The proposed system is based on two types of temporal residual blocks, called identity and convolution blocks respectively. The identity block is composed of two times temporal convolution, batch normalization and non-linear activation layers. In the third iteration, an algebraic sum between the initial input and the output of the temporal and batch normalization layers is done. Thanks to the shortcut connection, we can add more layers without decreasing the model performance, contrariwise it could increase slightly. The convolution block concept is identical to the identity block, with the except that there are temporal convolution and batch normalization layers in the shortcut path (Figure 3). Applying residual blocks requires that the outputs of the main path and the shortcut connection have the same shape. The identity and convolution blocks are used in alternance in our proposed model.

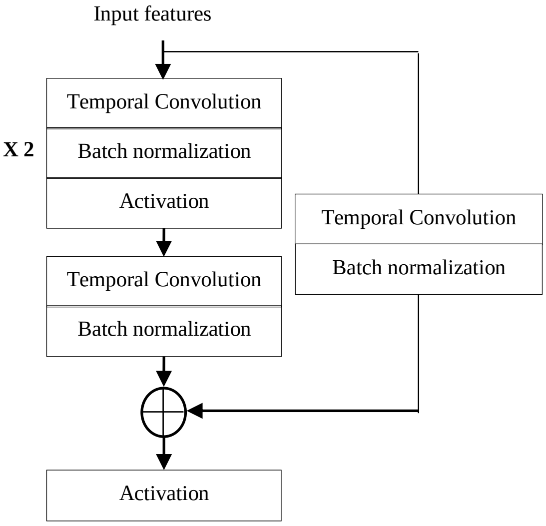


Figure 3. Temporal residual block architecture.

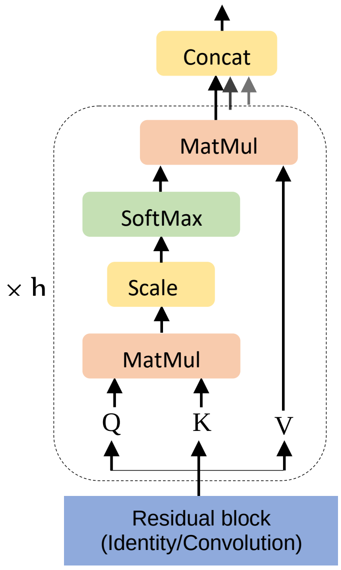


Figure 4. Proposed multi-head attention model.

4.3. Multi Head Attention Model

Multi-Head Attention (MHA) is a mechanism used to provide an additional focus on a specific component in the data. It enables the network to concentrate on a few aspects at a time and ignoring the rest [23]. MHA consists of several attention layers running in parallel, instead of performing one single attention function. In particular, the input consists of queries and keys of dimension d_k (Q and K respectively), and values of dimension d_v (V). The output of the attention model is done by computing the scaled dot product of the queries with all keys and applying a SoftMax function to obtain the weights on the values V (Equation 4). The attention mechanism is linearly projected h times with different learned weights (W^Q, W^K, W^V). These different representation subspaces are concatenated into one single attention head to form the final output result (Equation 5).

$$Attention(Q,K,V) = softmax\left(\frac{Q \times K^T}{\sqrt{d_k}}\right) \times V \tag{4}$$

$$\begin{cases} MultiHead(Q,K,V) = Concat(head_1,\ldots,head_h) \\ head_i = Attention(QW^Q, KW^K, VW^V) \end{cases} \tag{5}$$

In our study case, we applied MHA to the output of both identity and convolution temporal residual blocks (Figure 4). We used a particular version of attention model called self-attention, in which query, key and value inputs are the same. The calculation process follows these steps: First, we made the dot product (MatMul) of query and keys tensors and scale the obtained scores. Next, we apply a SoftMax function on these scores to obtain attention probabilities. Finally, we take a linear combination of these distributions with the value input tensors and concatenate them into one channel.

5. Experiments and Results

5.1. Database Description

The experiments were conducted on the public Arabic dataset LMCA [10]. It was developed by the Research Group on Intelligent Machines at the National Engineering School of Sfax (Tunisia). LMCA is a multi-writer database, produced by 55 persons (37 male and 18 female) whose age is between 8 and 66 years old. It consists of 212 files containing the coordinates (x, y) of 23141 characters relative to the 56 Arabic letter shapes. These files are stored according to the UNIPEN format file. We divided the database into two subsets: two-third of samples have been used for the model training and the rest for the inference stage. When applying beta-elliptic model, every character has not got the same strokes number, which is variable and takes a value from 2 to 13 (Table 2). Thus, we applied zero padding technique to bring all the input tensors to the shape (13×8). On the other hand, we noticed that beta-elliptic parameters differ greatly between their ranges, and they are defined in different unit of measures. Then, we standardized the input features before we integrate it into the recognition model. So, all features contribute equally when fitting the model and avoid it from being biased [9]. The beta-elliptic parameters have been scaled in such a way that they have a mean of 0 and a standard deviation as 1.

Table 2. Beta strokes numbers.

Arabic letter	**Number of beta strokes**						
	2	**3**	**4**	…..	**11**	**12**	**13**
ا	X						
ب			X				
د			X				
ـــح					X		
ــم					X	X	
س						X	X
ك						X	X

5.2. Experimental Setup

The recognition model is set up as follow: in the network head, we set the values of kernel size and window shift to 2 and 1 respectively. In both identity and convolution blocks, the number of filters is set to 64 for the main path and the shortcut connections (Table 3). In the multi-head attention model, we set the number of attention layers to 8. Thus, we obtained a total of 472250 trainable parameters relative to the kernel weights and the multi-head matrices weights. In the classification part, we applied global average pooling layer to convert the stack of feature maps into one dimensional vector. After that, we used one hidden layer of 256 neurons and 56 neurons in the output layers which corresponds to the total number of classes.

Table 3. Recognition model configuration.

Model part	Parameters	Value
Network head	Temporal convolution	Kernels = 32 Window size = 2 Stride = 1
	Activation	Relu
	Temporal convolution	Pool size = 2
Identity block	Temporal convolution	Kernels = 64 Window size = 2 Stride = 1
	Activation	Relu
Convolution block	Temporal convolution	Kernels = 64 Window size = 2 Stride = 1
	Activation	Relu
Multi-head Attention model	Number of heads	8
Classification	Hidden layer	512 neurons
	Output layer	56 neurons

5.3. Experimental Results

The proposed model was trained using Adam optimizer with the initial learning rate 0.001. This value is reduced by a factor of 0.5 once learning stagnate. Moreover, categorical cross entropy method was applied to compute the loss between the actual and the predicted outputs. We divided the training dataset into batches of size 64 with shuffling option, to obtain different mini-batch samples in each epoch and ensuring the model generalization. In the experiments, we have implemented four models. The first is called Plain network which does not contain residual blocks, the second model is a Residual network containing only two identity blocks, the third model contains both identity and convolution blocks, and the last model contains temporal residual blocks with the addition of multi-head attention mechanism. The best performance was obtained using the last model architecture (Table 4). When applying multi-head attention model, the validation accuracy improves from 96.87% to 97.12%. Moreover, precision, recall and f1-score metrics were calculated to measure the ability of the model to label the negative and positive samples (Table 5).

Table 4. Experimental results.

	Validation accuracy	**Validation loss**
Plain network (Without residual blocks)	0.9291	0.216
Residual network (Only identity block)	0.9418	0.198
Residual network (identity & convolution blocks)	0.9687	0.130
Residual network with multi-head attention model	**0.9712**	**0.105**

Table 5. Classification report.

		Precision	**Recall**	**F1-score**
Residual network without attention	Macro avg	0.96	0.95	0.95
	Weighted avg	0.97	0.97	0.97
Residual network without attention	Macro avg	0.971	0.962	0.962
	Weighted avg	**0.9802**	**0.802**	**0.803**

5.4. Results Comparison and Discussion

The proposed model achieved the state-of-the-art performances on LMCA database (Table 6). First, this is due to efficiency of beta elliptic approach in modeling the dynamic and geometric aspects of the handwriting trajectory. Moreover, temporal convolution was applied to process the input data sequentially. The concept is different from recurrent layers and allows training the model with less trainable parameters. However, we presented some novelties compared to the existing works. In fact, we turned the conventional residual network to handle with sequential input data. Also, we have experimented the utility of using multi attention mechanism with temporal residual network for improving the model performances.

Table 6. Performance comparison on LMCA dataset.

Authors	Approach	Model	Accuracy
[10]	Beta-elliptic	MLP	0.9414
[25]	Beta-elliptic	TDNN	0.9516
[17]	Beta-elliptic	DBLSTM	0.9492
[19]	Beta-elliptic	Temporal ResNet	0.9687
Our work	Beta-elliptic	Temporal ResNet with attention	**0.9712**

7. Conclusions

We proposed in this work a new system for Arabic letters recognition based on beta elliptic modelling and Temporal Residual Networks. We have demonstrated the effectiveness of using temporal convolution to process sequential input data. To prevent network from vanishing gradient descent, we applied skip connection in both identity and convolution blocks. Moreover, we have demonstrated the effectiveness of applying multi-head attention mechanism in the context of online handwriting recognition. As perspective, we intend to apply transformers to recognize the Arabic handwriting script.

References

[1] Abdelaziz I., Abdou S., and Al-Barhamtoshy H., "A Large Vocabulary System for Arabic Online Handwriting Recognition," *Pattern Analysis and Applications*, vol. 19, no. 4, pp. 1129-1141, 2016. https://doi.org/10.48550/arXiv.1410.4688

[2] Abuzaraida M., Elmehrek M., and Elsomadi E., "Online Handwriting Arabic Recognition System Using K-Nearest Neighbors Classifier and DCT Features," *International Journal of Electrical and Computer Engineering*, vol. 11, no. 4, pp. 3584-3592, 2021. http://doi.org/10.11591/ijece.v11i4.pp3584-3592

[3] Ahsan M., Mahmud M., Saha P., Gupta K., and Siddique Z., "Effect of Data Scaling Methods on Machine Learning Algorithms and Model Performance," *Technologies*, vol. 9, no. 3, pp. 52-59, 2021. https://doi.org/10.3390/technologies9030052

[4] Alimi A., "An Evolutionary Neuro-Fuzzy Approach to Recognize On-Line Arabic Handwriting," *in Proceedings of the 4th International Conference on Document Analysis and Recognition*, Ulm, pp. 382-386, 1997. 10.1109/ICDAR.1997.619875

[5] Altwaijry N. and Al-Turaiki I., "Arabic Handwriting Recognition System Using Convolutional Neural Network," *Neural Computing and Applications*, vol. 33, no. 7, pp. 2249-2261, 2021.

[6] Alwajih F., Badr E., and Abdou S., "Transformer-based Models for Arabic Online Handwriting Recognition," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 5, pp. 898-905, 2022. 10.14569/IJACSA.2022.01305102

[7] Amudha J., Thakur M., Shrivastava A., Gupta S., Gupta D., and Sharma K., "Wild OCR: Deep Learning Architecture for Text Recognition in Images," *in Proceedings of International Conference on Computing and Communication Networks*, Kanpur, pp. 499-506, 2022. DOI:10.1109/ICCCNT45670.2019.8944593

[8] Bezine H., Alimi A., and Sherkat N., "Generation and Analysis of Handwriting Script with the Beta-Elliptic Model," *in Proceedings of 9th International Workshop on Frontiers in Handwriting Recognition*, Kokubunji, pp. 515-520, 2004. 10.1109/IWFHR.2004.45

[9] Boubaker H., Chaabouni A., Tagougui N., Kherallah M., and Alimi A., "Handwriting and Hand Drawing Velocity Modeling by Superposing Beta Impulses and Continuous Training Component," *International Journal of Computer Science Issues*, vol. 10, no. 5, pp. 57-63, 2013.

[10] Boubaker H., Elbaati A., Tagougui N., El Abed H., Kherallah M., and Alimi A., *Guide to OCR for Arabic Scripts*, Springer, 2012.

[11] Boubaker H., Rezzoug N., Kherallah M., Gorce P., and Alimi A., "Spatiotemporal Representation of 3D Hand Trajectory Based on Beta-Elliptic Models," *Computer Methods in Biomechanics and Biomedical Engineering*, vol. 18, no. 15, pp. 1632-1647, 2015. doi: 10.1080/10255842.2014.940331

[12] Chikmagalur A., Devaraj S., Hebbakavadi P., Rajasekhar K., and Reddy U., "Off-Line Signature Confirmation based on Cluster Representations of Geometrical and Statistical Features through Vector Distance, Neural Network and Support Vector Machine Classifiers," *The International Arab Journal on Information Technology*, vol. 19, no. 4, pp. 667-673, 2022. https://doi.org/10.1016/j.patrec.2005.04.013

[13] Dhieb T., Boubaker H., Njah S., Ben Ayed M., and Alimi A., "A Novel Biometric System for Signature Verification Based on Score Level Fusion Approach," *Multimedia Tools and Applications*, vol. 81, no. 6, pp. 7817-7845, 2022. https://doi.org/10.1007/s11042-022-12140-7

[14] Diaz M., Moetesum M., Siddiqi I., and Vessio G., "Sequence-based Dynamic Handwriting Analysis for Parkinson's Disease Detection with One-Dimensional Convolutions and BiGRUs," *Expert Systems with Applications*, vol. 168, pp. 1-34, 2021. https://doi.org/10.1016/j.eswa.2020.114405

[15] Eltay M., Zidouri A., and Ahmad I., "Exploring Deep Learning Approaches to Recognize Handwritten Arabic Texts," *IEEE Access*, vol. 8, pp. 89882-89898, 2020. DOI: 10.1109/ACCESS.2020.2994248

[16] Hamdi Y., Boubaker H., and Alimi A., "Online Arabic Handwriting Recognition Using Graphemes Segmentation and Deep Learning Recurrent Neural Networks," *in Proceedings of the Enabling Machine Learning Applications in Data Science*, Springer, pp. 281-297, 2021. DOI:10.1007/978-981-33-6129-4_20

[17] Hamdi Y., Boubaker H., Dhieb T., Elbaati A., and Alimi A., "Hybrid DBLSTM-SVM based Beta-Elliptic-CNN Models for online Arabic Characters Recognition," *in Proceedings of the International Conference on Document Analysis and Recognition*, Sydney, pp. 545-550, 2019. DOI: 10.1109/ICDAR.2019.00093

[18] He K., Zhang X., Ren S., and Sun J., "Deep Residual Learning for Image Recognition," *in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 770-778, 2016. 10.1109/CVPR.2016.90

[19] Othmen D., Zouari R., Boubaker H., and Kherallah M., "Temporal Convolution based Skip Connections for Online Arabic Handwriting Recognition," *in Proceedings of the International Arab Conference on Information Technology*, Abu Dhabi, pp. 1-5, 2022. DOI: 10.1109/ACIT57182.2022.9994095

[20] Plamondon R., Alimi A., Yergeau P., Leclerc F., "Modelling Velocity Profiles of Rapid Movements: A Comparative Study," *Biological cybernetics*, vol. 69, no. 2, pp. 119-128, 1993. DOI: 10.1007/BF00226195

[21] Plamondon R., Guerfali W., The Generation of Handwriting With Delta-Lognormal Synergies, *Biological Cybernetics*, vol. 78, no. 2, pp. 119-132, 1998. DOI:10.1007/s004220050419

[22] Sankara Babu B., Nalajala S., Sarada K., Muniraju Naidu, V., Yamsani, N., and Saikumar K., "Machine Learning Based Online Handwritten Telugu Letters Recognition for Different Domains," *in Proceedings of the A Fusion of Artificial Intelligence and Internet of Things for Emerging Cyber Systems*, pp. 227-241, Springer, 2021. DOI:10.1007/978-3-030-76653-5_12

[23] Vaswani A., Shazeer N., Parmar N., Uszkoreit J., Jones L., Gomez A., Kaiser L., and Polosukhin I., "Attention is All You Need," *in Proceedings of the 31st Conference on Neural Information Processing Systems*, Long Beach, pp. 1-11, 2017.

[24] Wilson-Nunn D., Lyons T., Papavasiliou A., and Ni H., "A Path Signature Approach to Online Arabic Handwriting Recognition," *in Proceedings of the 2nd International Workshop on Arabic and Derived Script Analysis and Recognition*, London, London, pp. 135-139, 2018. DOI: 10.1109/ASAR.2018.8480300

[25] Zouari R., Boubaker H., and Kherallah M., "A Time Delay Neural Network for Online Arabic Handwriting Recognition," *in Proceedings of the International Conference on Intelligent Systems Design and Applications*, Portugal, pp. 1005-1014, 2016. DOI:10.1007/978-3-319-53480-0_99

[26] Zouari R., Boubaker H., and Kherallah M., "RNN-LSTM based Beta-Elliptic Model for Online Handwriting Script Identification," *The International Arab Journal on Information Technology*, vol. 15, no. 3A, pp. 532-539, 2018.

[27] Zouari R., Boubaker H., and Kherallah M., "Two Staged Fuzzy SVM Algorithm and Beta-Elliptic Model for Online Arabic Handwriting Recognition," *in Proceedings of the International Conference on Artificial Neural Networks*, Alghero, pp. 450-458, 2017. DOI:10.1007/978-3-319-68612-7_51

[28] Zouari R., Boubaker H., and Kherallah M., "Hybrid TDNN-SVM Algorithm for Online Arabic Handwriting Recognition," *in Proceedings of the 16th International Conference on Hybrid Intelligent Systems*, Marrakech, pp. 113-123, 2016. DOI:10.1007/978-3-319-52941-7_12

[29] Zouari R., Boubaker H., and Kherallah M.., "Towards Online Handwriting Recognition System Based on Reinforcement Learning Theory," *in Proceedings of the International Conference on Neural Information Processing*, Bangkok, pp. 562-570, 2020.

Ramzi Zouari He is a Phd in Engineering of informatic systems. He obtained the doctorate diploma from the national school of engineering of Sfax, Tunisia in 2020. His research interests include handwriting trajectory modeling and recognition. Moreover, he focuses his research on the application of deep neural networks models on smart applications.

Dalila Othmen She is a Phd student at the faculty of sciences of Gafsa, Tunisia. She obtained a master's degree in computer science for the University of Gafsa in 2022. Her main research interest concerns the application of deep learning models for online handwriting recognition.

Houcine Boubaker Graduated in Electrical Engineering in 1995, obtained a master's degree in systems Analyses and Digital Signal Processing in 1997. He is a researcher in Electrical and Computer Engineering at the University of Sfax and affiliate to the Research Groups in Intelligent Machines laboratory (ReGIM). His research interest includes trajectories modeling and pattern recognition. He focuses his research on drawing, handwriting and arm-hand movements modeling and Analyses.

Monji Kherallah Received the Ing. Diploma degree and the PhD in electrical engineering, in 1989 and 2008, respectively from University of Sfax (ENIS), Tunisia. He is a member in Research Group of Intelligent Machines: REGIM. His research interest includes the following projects: "i-Bag", "i-House" and "i-Car". The techniques used are based on methods intelligent, such as neural network, logic fuzzy, genetic algorithm, etc. He is a reviewer of several international journals.