Multi-Head Self-Attention and BGRU for Online Arabic Grapheme Text Segmentation

Yahia Hamdi, Besma Rabhi, Thameur Dhieb and Adel M. Alimi

February 14, 2023

Multi-head Self-attention and BGRU for Online Arabic Grapheme Text Segmentation

Abstract. The segmentation of online handwritten Arabic text into graphemes/characters is a challenging task for the recognition system due to the nature of this script. For this, it is better to employ the dependency in the context of segments written before and after it. In this paper, we introduce Multi-head self-attention (MHSA) and Bidirectional Gated Recurrent Units (BGRU) models for online handwritten Arabic text segmentation that simulate our previous grapheme segmentation model (GSM). The proposed framework consists of word embedding and the combination of complementary Multi-head self-attention and BGRU which help to detect the control points (CPs) for handwritten text segmentation. The CPs delimit each grapheme composed of three main geometric points: the starting point (SP), ligature valley point (LVP), and ending point (EP). To show the effectiveness of our MHSA-BGRU model for online handwritten segmentation and its comparison with GSM, the mean absolute error (MAE), and word error rate (WER) evaluation metrics are used. Experimental results on benchmark ADAB and online-KHATT datasets show the efficiency of our model which achieves 3.17% and 5.28% for MAE, 12.25% and 25.13% for WER respectively.

Keywords: Online handwriting trajectory, Grapheme segmentation, Transformer, Multi-head self-attention, BGRU.

1 Introduction

In recent years, online handwriting recognition topics mending more important with the progression (emergence) of digital devices such as smartphones, Tablet PCs, digital pens, electronic whiteboards, etc. It remains a challenging task due to the variability introduced by several people for their various writing styles. Things get tricky in an unconstrained realm like in Arabic script because individuals often write more than one stroke without a pen lift such as Arabic words 'سلمي', 'محمد', 'سمير', etc. Further, there may be many types of junction patterns between pseudo words or strokes. The handwriting segmentation process plays an important role in breaking these junctions and collecting individual elemental strokes/graphemes for another processing which subsequently involves its recognition. This technique is important because the success of the later stages of the recognition process powerfully depends on it [1, 2]. Their gain is being able to work on a large or free lexicon. Investigated approaches for online Arabic handwritten text segmentation have realized satisfactory results. However, the segmentation of the trajectories into characters/graphemes is a still difficult task for existing systems [3], especially for Arabic script due to its cursive nature.

Generally, the handwriting segmentation technique represents the different operations accomplished to produce the main handwriting entities that will have studied by the recognition systems. It classified into two categories: the first systems deal with the whole text and concentrated on line detection [4] based on temporal order and spatial zones, while, the second focused on the decomposition of the input data into elementary characters or even into sub-units like graphemes or strokes such as presented in [5, 6]. Among theme, many segmentation techniques are developed for on-line Arabic handwriting trajectory like described in [7] which decompose the input trajectory into elementary segments called convex/concave. [8] segment the pseudo words in graphemes based on the detection of significant points. Also, the segmentation of the word into elementary components located between pen-down and pen-up was investigated by [9]. For other scripts, [10] presented a lexicon-free segmentation strategy for online handwritten Tamil words to deal with the under and over-segmentation problem. Also, a segmentation of online handwritten Bangla word into strokes based on their positional information was proposed by [11]. At stroke-level, a busy zone is employed to find the segmentation points.

As the latest in emerging technology, deep learning has developed swiftly, and some studies have been introduced to solve problems of traditional recognition architecture, reducing processing time and dictionary size of recognition systems [12, 13, 14]. It allows a handwritten text recognition system to work efficiently and reliably on digital smart devices. In this context, a good performance of on-line handwritten text segmentation technique [15, 16] is achieved using SVM model which has been enormously used in several classification tasks. Also, RNN neural network has achieved tremendous progress in online handwriting sequence modeling [17]. The segmentation can be improved by integrating both forward and backward contexts. Indeed, an enhanced version of BLSTM [18] shows its effectiveness in many sequence classification tasks which allows to access long-range context.

Recently, transformer has attracted the attention of more researchers in many axes such as handwriting recognition [19], handwriting recovery [20], etc. It aimed to treat input data like RNN networks. However, transformers process the entire input all at once and provide context for any position in the input sequence, unlike RNNs. In this paper, we explore the performance of Transformer model on handwritten segmentation task. The main contribution of this paper is to introduces a novel method for online handwritten Arabic text segmentation simulate and replace our previous work [2]. It based on the combination of powerful multi-head self-attention and BGRU models for graphemes text segmentation. Given an input sequence $S = \{(x_1, y_1), (x_2, y_2),.., (x_n, y_n), \}$ at time-stamp t, our proposed model proceeds by a pre-treatment step that converts the sequence S to word embeddings which studied by using stacked Multi-head self-attention and BGRU models followed by a fully connected layer aims to find the validate grapheme segmentation points.

The remainder of this paper is organized as follows. In section 2, we present our proposed approach. Experimental results are described and discussed in section 3. Finally, section 4 concludes the present paper.

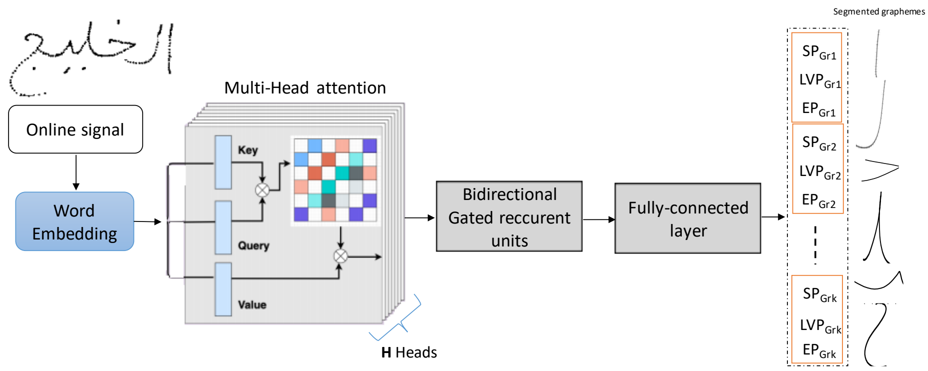


Fig. 1. Framework architecture.

2 Our Approach

The architecture of our proposed neuronal approach for online handwriting text segmentation contains four components: word embedding, Multi-Head Self-Attention (MHSA), Bidirectional Gated Recurrent Units (BGRU), and fully connected layer as illustrated in Fig. 1. The first step is to convert the input text into word vectors, which is necessary to train a deep learning model and improve the quality of the segmentation process. The constructed word vectors are used as input to multi-head self-attention module which chooses words providing the best grapheme segmentation position identified by ligature valleys points. The output of MHSA is passed to BGRU which helps to learn the long-range relationships between these highlighted words and produce a single feature vector that encodes the entire sequence. Finally, a fully-connected layer with regression activation function is employed to detect the control points (CP) for graphemes text segmentation. Below, we describe each part of the proposed architecture in detail.

2.1 Graphemes segmentation model (GSM)

Handwriting segmentation technique is among the most challenging task for online handwriting recognition cursive script [21]. The graphemes describe the graphic shapes collection composing the handwriting trajectory. In our case, grapheme can describe a whole character like ‘ح’ or a set of graphic units as ‘ـد’, ‘ى’, and ‘ں’ building the character ‘س’. As shown in Fig. 2, the segmentation of handwritten Arabic text into graphemes relies on the detection of specific points such as the ligature valleys and angular points. The former represents the segment point closest to the baseline with a tangent aligned to its direction, while the latest denotes the extremum point of a vertical shaft trajectory turning around.

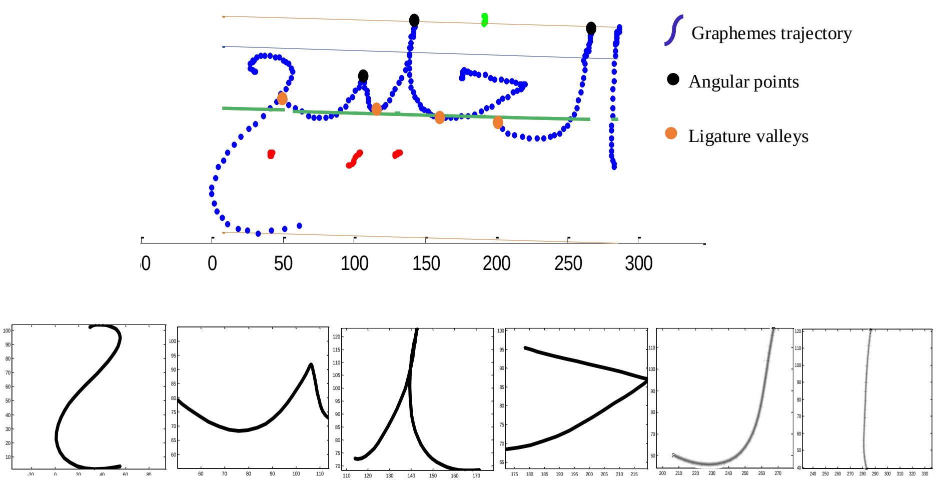


Fig. 2. Grapheme segmentation of Arabic word 'الخليج' using GSM model.

In practice, we retain the particular points verifying the empiric conditions described by Eq.1 for ligature valleys points and Eq.2 for angular points:

$$\begin{cases} R_{\Delta y} = \dfrac{\Delta y}{h_{LM}} < R_{\Delta y\max} = \dfrac{1}{2} \\ \Delta\alpha < \Delta\alpha_{\max} = \dfrac{\pi}{6} \end{cases} \tag{1}$$

$$\begin{cases} \Delta\theta > \Delta\theta_{\min} = \dfrac{3\pi}{4} \\ \theta_{med} = \left|\theta_{med} - \dfrac{\pi}{2}\right| < \theta_{\max} = \dfrac{\pi}{5} \end{cases} \tag{2}$$

Where $R_{\Delta y}$ and Δy are respectively the ratio of the point M position in the handwritten trajectory and the distance with respect to the baseline. h_{LM} denotes the width of the baseline. $\Delta\alpha$ is the tangent deviation angle of the point M respect to the baseline. $\Delta\theta$ is the neighborhood's deviation angle of the point M, and θ_{med} is the deviation angle with respect to their vertical median direction.

To simplify the complexity degree of finding grapheme segmentation points, we have used only LVP points which are sufficient to train our model.

2.2 Multi-Head Self-Attention

To identify the trajectory control points for grapheme segmentation in such sentence S, it is meaningful to use specific words having a length $\leq L$, that provide the best validate control points (CPs). L is determined empirically after quantitative analysis of used databases. Indeed, we utilize a typical tokenizer on a text T_x to generate E dimensional embeddings for each word in the sentence. These embeddings from the input to our model are $T_x = \{e_1, e_2,..., e_N\}$, and $T_x \in R^{N \times E}$. We use multi-head self-attention to boost our suggested model and extract these words from the input text.

In order to solve the task ahead, attention is a technique for identifying patterns in the input. Self-attention [22] is a sequence-specific attention mechanism employed in deep learning that aids in learning the task-specific relationships between various sequence elements to generate a better series representation. Three linear projections of the given input sequence are created in the self-attention module: Query (Q), Key (K), and Value (V) where Q, K, and $V \in R^{N \times E}$. According to Eq.3, the learnt softmax attention (QK^t) and the output of this module, $A \in R^{N \times E}$ denotes scaled dot products that are used to calculate the attention map.

$$A = softmax\left(\frac{QK^t}{\sqrt{E}}\right) V \tag{3}$$

In multi-head self-attention, multiple copies of the self-attention module are used in parallel. Each head captures different relationships between the words in the input text and identifies those candidate words that aid in classification. In our model, we use a series of multi-head self-attention layers (L) with multiple heads (H) in each layer.

2.3 Gated Recurrent Units

Self-attention mechanism finds similar words that have the same length vector in the input text. While BGRUs is used to learn long term dependency between these words. It formed by units which designed to automatically remember and forget information over long time based on Reset (r_t) and Update (f_t) gates which solve the vanishing gradient problem of the standard recurrent neural networks.

In our model, we use BGRU of single layer to treat the input sequence S, since these units use contextual information from both back and forth directions. At each time step, BGRU produces a hidden states $H = \{h_1, h_2, ..., h_N\}$, $H \in \mathbb{R}^{N \times D}$ for each element of a given input sequence $S \in R^{N \times D}$ expressed as:

$$r_t = \sigma\left(W_r S_t + U_r h_{t-1} + b_r\right) \tag{4}$$

$$f_t = \sigma\left(W_f S_t + U_f h_{t-1} + b_f\right) \tag{5}$$

$$\tilde{h} = tanh\ \left(W_c S_t + \mathrm{U}_h\left(r_t \odot h_{t-1}\right) + b_h\right) \tag{6}$$

$$h_t = f_t \odot h_t + \left(1 - f_t\right) \odot \tilde{h}_{t\text{-}1} \tag{7}$$

where σ(.) denote the element-wise sigmoid function. W, U, are the vector weights and b is the biases. r_t, f_t, h_t, $\tilde{h_t} \in \mathbb{R}^d$, where d is the size of the produced hidden state vector.

We consider the final hidden state, h_N, which encrypts all the information of the sequence, at the output of this module.

2.4 CPs prediction values

The final output is computed using a single fully connected layer with a regression activation function. Indeed, this layer makes as input the feature vector h_N generated by GRU module to produce the predicted value y_t calculated as:

$$y_t = \sigma\left(Wh_n + b\right) \tag{8}$$

where $W \in \mathbb{R}^d$ denote the weights of the fully connected layer and b is the bias. The model is trained using '*relu*' activation function. The outputs are the sequence of *SP*, *EP*, and CPs defined by LVP describing each grapheme of the online handwriting trajectory.

3 Experiments

In this section, we employ a thorough study on handwritten graphemes segmentation using public benchmark ADAB and online-KHATT datasets spanning wide graphemes. In addition, we outline evaluation results on response time and recognition of Arabic script between MHSA-BGRU and GSM models.

3.1 Datasets

We use benchmark ADAB (Arabic DAtaBase) database [23] encompass more than 21,000 words collected by 166 writers from 937 Tunisian town/village names. It is considered the most widely used dataset to evaluate online Arabic handwriting recognition.

Table 1. ADAB dataset description.

Sets	Number of words	Number of pseudo words	Writers
Set 1	5037	40296	56
Set 2	5090	25450	37
Set 3	5031	15093	39
Set 4	4417	22085	25
Set 5	1000	4000	6
Set 6	1000	8000	3
Total	21575	114924	166

This dataset is divided into six different parts from the ICDAR 2011 competition on online Arabic handwriting recognition [24] as shown in Table 1. For training our model, we use sets 1, 2, and 3 comprise more than 150.000 graphemes, and set 4 for validation. We tested our model using more than 100.000 graphemes of sets 5 and 6.

The second dataset is Online-KHATT, a new open-vocabulary Arabic database developed by [25]. It is made up of 10,040 lines of online Arabic text from 40 books, collected by approximately 623 writers using Android and Windows devices. This database consists of many problems such as dots number, position, thickness, and writing styles which represent a challenging tasks.

3.2 Experimental Setup

To investigate the impact of our purpose of online Arabic text segmentation, we performed two groups of experiments. The first test aims to validate the effectiveness of the proposed MHSA-BGRU for graphemes text segmentation and come upon the influence of using MHSA for grapheme control point detection. The second test evaluates the performance and robustness of our segmentation approach on word recognition rate.
We implement our models using a deep-learning framework developed in Python TensorFlow.

To extract word embeddings, we use a standard tokenizer [26] to convert input words to tokens and then to word embeddings. The constructed word embeddings of the input text are passed through multi-head self-attention layers *L* composed of multiple heads *H*. The output from the self-attention layer is transmitted to a single BGRU hidden layer with a dimension of 256 units. Afterward, the output feature vector of the BGRU is studied by the fully connected layer to yield a 1-dimensional output containing the control points of each segment (grapheme).

During training, we use AdaDelta optimizer algorithm [27] with decay rate $\rho = 0.9$, weight decay of 1e-4, and batch size of 64. The training set is augmented using different geometric augmentation methods presented in [28] in order to improve the performance of MHSA-BGRU segmentation model.

3.3 Segmentation

To clarify the validity of MHSA-BGRU for grapheme segmentation, we measure its output's similarity degree with those using GSM algorithm. The segmentation of the input text into grapheme consists to find valid segmentation points. To do this, we use three main criteria such as CP_t, $X(t)$, and $Y(t)$ to make this evaluation. Indeed, CP_t denotes the control *LVP* point detection time of segmented grapheme. $X(t)$, and $Y(t)$ are the coordinates points of each grapheme of the handwriting trajectory. Table 2 shows an example of grapheme segmentation of Arabic word ‘الحامة’ using GSM and MHSA-BGRU models.

Table 2. Comparison of grapheme segmentation process between GSM and MHSA-BGRU of online Arabic word ‘الحامة’.

TCP order		1	2	3	4	5	6	7
CP_t time	GSM	0.01	0.18	0.24	0.39	0 .58	0.79	0.92
	MHSA-BGRU	0.03	0.19	0.26	0.43	0 .61	0.82	0. 93
X(t)	GSM	54.22	65.17	66.42	68.43	77.33	69.18	40.98
	MHSA-BGRU	50.51	63.60	62.07	69.20	81.15	67.01	44.18
Y(t)	GSM	78.48	60.29	53.66	30.22	6.17	3.55	4
	MHSA-BGRU	73.71	56.90	52.00	29.16	6.35	3.28	4

Indeed, seven TCPs identified which therefore generates seven graphemes. We can see a heavy similarity with a +/-0.02 deviation of the achieved values in terms of graphemes CP_t points using both models. Also, we observe a small alteration of [0, 3.5] between the graphemes coordinates $X(t)$ and $Y(t)$ points. Through reported results, we pretend the effectiveness of our proposed deep-learned segmentation model that well simulates the convolutional approach. The performance of the proposed neuronal approach is evaluated utilizing the regression Mean Absolute Error (MAE) metric for each mentioned criterion as expressed in the following equation:

$$MAE = \frac{\sum_{i=1}^{n} |p_i - y_i|}{n} \tag{9}$$

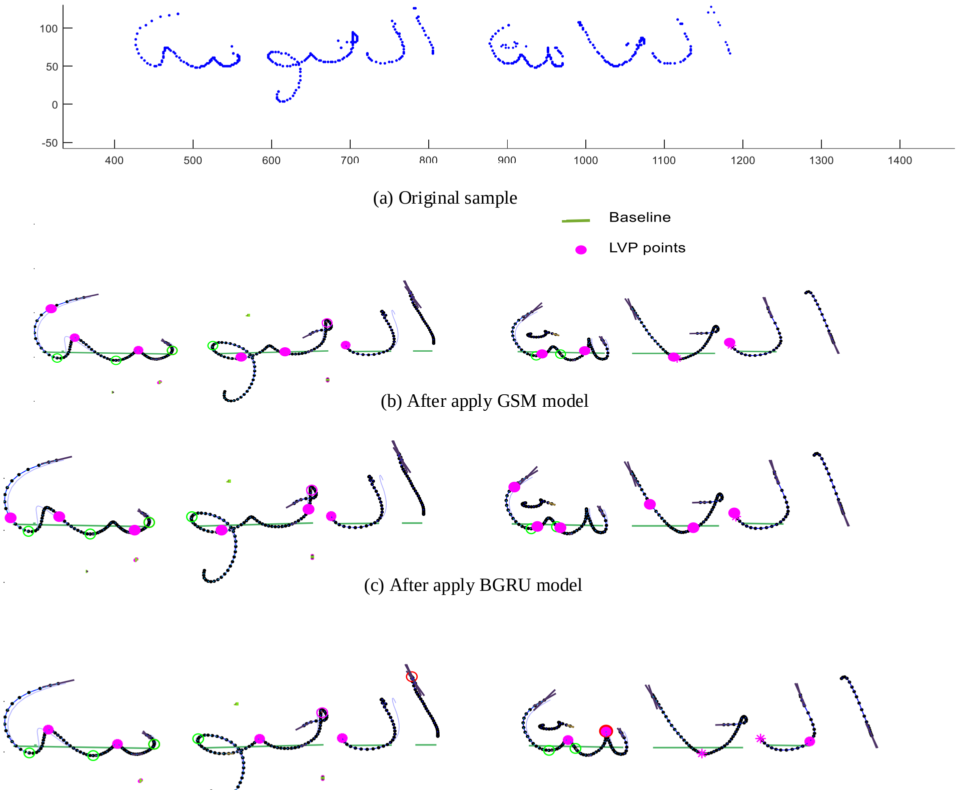
The MAE is computed between n observations, where p_i is the predicted values of MHSA-BGRU model, and y_i denotes the real values obtained employing GSM model.

Table 3. MAE error measure of grapheme control points between GSM and MHSA-BGRU model

Evaluation metrics		MHSA-BGRU			BGRU		
		CPs time	X(t)	Y(t)	CPs time	X(t)	Y(t)
MAE	ADAB	**3.17%**	3.59%	2.33%	6.59%	4.90%	5.91%
	Online-KHATT	**5.28%**	5.90%	4.15%	9.60%	8.66%	9.54%

To determine the impact of multi-head self-attention model and its success to find the validated CP_t points, we have also described the MAE values of only BGRU which is trained directly using the raw input text, and its combination with MHSA.

Table 3 depicts the obtained results of each mentioned criterion for graphemes text segmentation compared to GSM algorithm. The evaluation results are reported using ADAB and online-KHATT datasets. The obtained results present a high similarity between the two models of grapheme segmentation. It demonstrates the effectiveness of the proposed approach for CPs localization that delimits each handwritten grapheme.



(a) Original sample

(b) After apply GSM model

(c) After apply BGRU model

(d) After apply MHSA-BGRU model

Fig. 3. Grapheme segmentation of Arabic text 'الحامة الجنوبية' using BGRU, MHSA-BGRU, and their comparison with GSM model.

Fig. 3.c) and Fig. 3.d) show the segmentation of the Arabic sentence 'الحامة الجنوبية' using BGRU and MHSA-BGRU models respectively. Here, it is exhibited that the MHSA-BGRU model significantly simulates the traditional GSM approach rather than using only BGRU. It can be explained by the powerful MHSA model that allows the BGRU neural network to control the mixing of information between the words of an input sequence and to find the appropriate geometric points for grapheme segmentation.

3.4 Recognition rate

In online handwriting recognition area, word recognition accuracy is highly dependent on the appropriate segmentation of words to obtain valid component segments. In order to confirm the effectiveness of the proposed segmentation model and its effect on word recognition process, we have also reported the comparison of both segmentation models in terms of recognition rate. We carried out this experiment on grapheme level by classifying it into four groups depending on their position in the word (Beginning grapheme (BG), middle grapheme (MG), isolated grapheme (IG), and end grapheme (EG)), and at the level of whole word using WER (Word Error Rate).

Table 4. Grapheme classification using GSM and MHSA-BGRU models.

Grapheme group	Recognition rate (%)	
	GSM	MHSA-BGRU
BG	99.30	97.42
MG	98.12	95.97
IG	99.76	98.02
EG	97.87	95.13

Table 5. Word error rate (WER) using both segmentation models.

Grapheme class	WER (%)	
	GSM	MHSA-BGRU
ADAB	10.00	12.25
Online-KHATT	22.88	25.13

Table 4 summarizes the evaluation of graphemes recognition rate using BLSTM network. The average accuracy of grapheme subgroups was about 98,76% and 96,63% using GSM and MHSA-BGRU respectively. The obtained results show the robustness and precision of the proposed neuronal segmentation model for grapheme classification.

As illustrated in Table 5, the word recognition rates obtained by our system using BLSTM-CTC architecture after applying segmentation step on ADAB and online-KHATT datasets are very close. It shows the efficiency of the proposed segmentation approach and its influence on word recognition rate.

3.5 Waiting time

The average waiting time of both GSM and MHSA-BGRU models for graphemes text segmentation is measured by modifying the number of graphemes (NG). As shown in Fig. 4, we observe a decrease in the average waiting time by about 40% using MHSA-BGRU compared with conventional segmentation algorithm. It can be explaining the customization of our developed model in terms of response and CPU times compared to the previous classic algorithm.

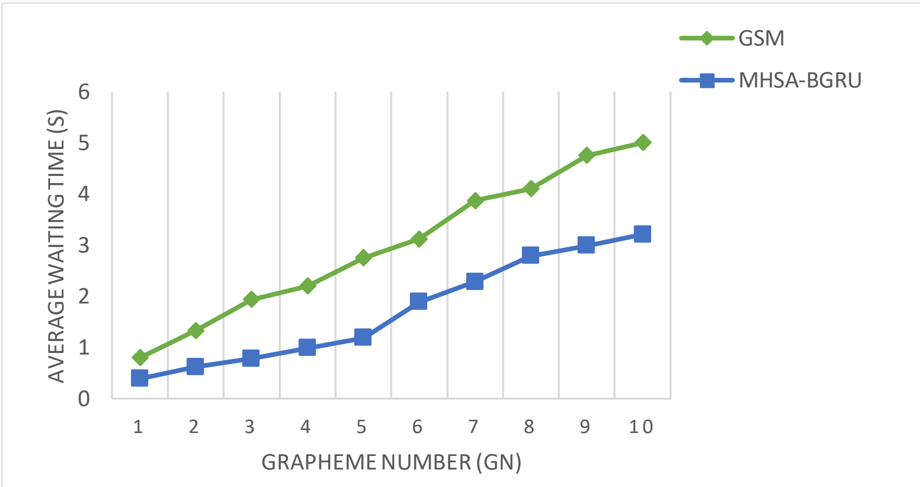


Fig. 4. Waiting Time of GSM and MHSA-BGRU models for grapheme segmentation.

This leads to faster-handwritten text segmentation and therefore recognition times up to 4 times rapids related to our previous system. Consequently, our neuronal approach is easy and swift, and can be applied in different commercial applications such as handwriting recognition, signature verification, handwritten disorder detection, etc.

3.6 Discussion

We proposed a novel transformer deep-learning method for online Arabic text segmentation based on the combination of MHSA and BGRU models. We evaluated its effectiveness in terms of similarity values, recognition rate, and response time. In fact, the proposed model simulates and replaces the functionalities of the previous GSM algorithm.

As a consequence:

- We calculated the MAE metric between the actual and predicted values of graphemes trajectory control points, *X(t)*, and *Y(t)* using a single BGRU and its combination with MHSA model.
- We have investigated the performance of proposed model and its influence on handwriting recognition tasks using grapheme classification and WER metrics.
- We also calculated and compared the waiting time for grapheme text segmentation of both models.

Comparison of the results obtained using ADAB and online-KHATT datasets highlight that our MHSA-BGRU model simulated well the GSM algorithm for Arabic handwriting text segmentation, allowing us to achieve lower MAE over only using BGRU model. To show the efficiency of our proposed model in practical applications, we used the output of MHSA-BGRU for Arabic script recognition keeping the same architecture. However, our proposed model has certain limitations such as the diminution of the recognition rate due to over or under-segmentation.

4 Conclusion

Motivated by the recent progress of Transformer network, we propose in this paper to improve the quality of Arabic handwriting text segmentation using an encoder network based on the combination of MHSA and BGRU models. The trained model pretends and replaces the handcrafted grapheme segmentation model in order to facilitate its manipulation in several tasks such as handwriting recognition, especially for free-lexicon context.

Experimental results on ADAB and online-KHATT datasets show the efficiency of proposed neuronal encoder model for online handwritten text segmentation. Our model achieves 3.17% and 5.28% expression MAE on ADAB and online-KHATT test sets, respectively. Moreover, we compare our model in terms of recognition accuracy and waiting time. As future work, we will continue to enhance MHSA-BGRU accomplishment by leveraging more optimization models and training methods to deal with under and over-segmentation problems. Also, it will be more interesting to apply the developed segmentation model in other works such as handwriting analysis, handwriting disorder detection, etc.

References

1. Alginahi, Y. M. A survey on Arabic character segmentation, International Journal on Document Analysis and Recognition. 16 (2) (2013) 105–126.
2. Boubaker, H., Tagougui, N., El Abed, H., Kherallah, M., and Alimi, Adel M. 2014. Graphemes segmentation for arabic online handwriting modeling. J. Inf. Process. Syst. 10, 4 (2014), 503--522.
3. Ullah, I., Azmi, S.M., Desa, I.M. Segmentation of touching Arabic characters in handwritten documents by overlapping set theory and contour tracing, International Journal of Advanced Computer Science and Applications (IJACSA). 10 (5) (2019) 155-160.
4. Elkhayati, M., Elkettani, Y., Mourchid, M. (2022). Segmentation of handwritten Arabic graphemes using a directed Convolutional Neural Network and mathematical morphology operations. Pattern Recognition, 122, 108288. DOI 10.1016/j.patcog.2021.108288.
5. Daifallah K, Zarka N, Jamous H (2009) Recognition-based segmentation algorithm for on-line Arabic handwriting. In: Proceedings of the international conference on document analysis and recognition. Barcelona, Spain, pp 877–880
6. Abdelazeem S, Eraqi H (2011) On-line Arabic handwritten personal names recognition system based on HMM. In: Proceedings of ICDAR 2011. pp 1304–1308LNCS Homepage, http://www.springer.com/lncs, last accessed 2016/11/21.
7. Izadi, Sara et al. "A New Segmentation Algorithm for Online Handwritten Word Recognition in Persian Script." (2008).
8. Eraqi, H. M. and Abdelazeem, S. "A New Efficient Graphemes Segmentation Technique for Offline Arabic Handwriting," 2012 International Conference on Frontiers in Handwriting Recognition, Bari, Italy, 2012, pp. 95-100, doi: 10.1109/ICFHR.2012.162.

9. Mandal, S., Choudhury, H., S. R. M. Prasanna and S. Sundaram, "DNN-HMM Based Large Vocabulary Online Handwritten Assamese Word Recognition System," 2018 16th International Conference on Frontiers in Handwriting Recognition (ICFHR), 2018, pp. 321-326, doi: 10.1109/ICFHR-2018.2018.00063.
10. Sundaram, S., & Ramakrishnan, A. G. (2013). Attention-feedback based robust segmentation of online handwritten isolated tamil words. ACM Transactions on Asian Language Information Processing, 12.
11. Sen S., Chowdhury S., Mitra M., Schwenker F., Sarkar R., and Roy K. 2020. A novel segmentation technique for online handwritten Bangla words. Pattern Recogn. Lett. 139 (2020), 26–33. DOI: DOI: https://doi.org/10.1016/j.patrec.2018.02.008.
12. Zhu, B., and Nakagawa, M. "Building a compact online MRF recognizer for large character set by structured dictionary representation and vector quantization technique," Pattern Recognition 47(3): 982-993 (2014).
13. Rabhi, B., Elbaati, A., Hamdi, Y., and Alimi, A. M. "Handwriting Recognition Based on Temporal Order Restored by the End-to-End System," 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, NSW, Australia, 2019, pp. 1231-1236, doi: 10.1109/ICDAR.2019.00199.
14. Hamdi, Y., Boubaker, H., Dhieb, T., Elbaati, A., and A. M. Alimi, "Hybrid DBLSTM-SVM Based Beta-Elliptic-CNN Models for Online Arabic Characters Recognition," 2019 International Conference on Document Analysis and Recognition (ICDAR), Sydney, NSW, Australia, 2019, pp. 545-550, doi: 10.1109/ICDAR.2019.00093.
15. Nguyen, C., Nakagawa, M. An improved segmentation of online English handwritten text using recurrent neural networks, in: 2015 3rd IAPR Asian Conference on Pattern Recognition.
16. Ly, N., Nguyen, C., Nguyen, K., Nakagawa, M. Deep convolutional recurrent network for segmentation-free offline handwritten japanese text recognition, in: 2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), 2017, pp. 5–9, https://doi.org/10.1109/ICDAR.2017.357.
17. Graves, A. Sequence transduction with recurrent neural networks, ICML Representation Learning Worksop, 2012.
18. Hamdi, Y., Boubaker, H., Rabhi, B., Qahtani, A.M., Alharithi, Fahd. S., Almutiry, A., Dhahri, H., Alimi, M. Adel. Deep learned BLSTM for online handwriting modeling simulating the Beta-Elliptic approach, Engineering Science and Technology, an International Journal,V35, 2022, https://doi.org/10.1016/j.jestch.2022.101215.
19. Kassem, Aly M. et al. “OCFormer: A Transformer-Based Model for Arabic Handwritten Text Recognition.” 2021 International Mobile, Intelligent, and Ubiquitous Computing Conference (MIUCC) (2021): 182-186.
20. Rabhi, B., Elbaati, A., Boubaker, H. et al. Multi-lingual character handwriting framework based on an integrated deep learning based sequence-to-sequence attention model. Memetic Comp. 13, 459–475 (2021). https://doi.org/10.1007/s12293-021-00345-6
21. Plamondon R., Srihari S. N., “On-Line and Off-Line Handwriting Recognition: A Comprehensive Survey,” IEEE Transaction on Pattern Analysis and Machine Intelligence, Vol. 22, N° 1, January 2000, pp. 63 – 84.
22. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, Ł.; Polosukhin, I. Attention is all you need. In Proceedings of the Advances in Neural Information Processing Systems, Long Beach, CA, USA, 4–9 December 2017; pp. 5998–6008.

23. Boubaker H, Elbaati A, Tagougui N, ElAbed H, Kherallah M, Alimi AM (2012) Online Arabic databases and applications. In: Guide to OCR for Arabic Scripts. Springer, pp 541–557
24. Kherallah M, Elbaati A, Abed HE, Alimi AM (2008) The on/off (LMCA) dual Arabic handwriting database. In: 11th International conference on frontiers in handwriting recognition (ICFHR). Montréal, Québec, Canada.
25. Mahmoud SA, Luqman H, Al-Helali BM, BinMakhashen G, Parvez MT (2018) Online-khatt: an open-vocabulary database for Arabic online-text processing. Open Cybern Syst J 12(1):42–59
26. Wolf, T.; Debut, L.; Sanh, V.; Chaumond, J.; Delangue, C.; Moi, A.; Cistac, P.; Rault, T.; Louf, R.; Funtowicz, M.; et al. In Proceedings of the HuggingFace’s Transformers: State-of-the-art Natural Language Processing. arXiv 2019, arXiv:1910.03771.
27. Zeiler, M.D.: ADADELTA: an adaptive learning rate method, arXiv preprint arXiv:1212.5701 (2012)
28. Hamdi Y, Boubaker H, Alimi AM (2021) Data augmentation using geometric, frequency, and beta modeling approaches for improving multi-lingual online handwriting recognition. Int J Doc Anal Recogn (IJDAR) 24(3):283–298