Occluded Multi-Lingual Offline Handwriting Inpainting Based on Multi-Head Attention and Stacked-LSTM Decoder

Besma Rabhi, Yahia Hamdi, Abdelkarim Elbaati, Habib Dhahri and Adel M. Alimi

February 14, 2023

Occluded Multi-lingual Offline Handwriting Inpainting Based on Multi-head Attention and Stacked-LSTM Decoder

Abstract. The encoder-decoder with attention model has become a common framework to online handwriting recovery. Convolutional Neural Network (CNN) and Long Short Term Memory (LSTM) with attention mechanism are respectively used as encoder and decoder. Inspired by the current success of transformers in many tasks, we introduce in this paper a novel recovery-inpainting framework, named Temporal Order with Multi-head Attention Network and stacked-LSTM decoder (TO-MultiNet), to denoise the corrupted offline handwriting and to obtain its online counterpart signal characterized by dynamic features. First, TO-MultiNet framework is trained to generate the temporal order and the pen velocity from offline handwriting. Then, the obtained model is further used to inpaint the occluded handwriting images. This work is validated by the Beta-GRU recognition system that is applied on Arabic, Latin and Indian On/Off dual handwriting datasets. Experimental results prove the effectiveness of Multi-head attention with stacked-LSTM decoder which increases the quality of the obtained uncorrupted image and improves the recognition rate based on the novel Beta-GRU model.

Keywords: Occluded offline handwriting, Transformer, Multi-head attention, LSTM, Beta Elliptic Model, GRU.

1 Introduction

For several decades, Handwriting Recognition (HR) has been successfully studied. According to the type of handwriting, there exist online HR and offline one. The online recognition systems involve the automatic transformation of online handwriting on a special digital assistant, which can be employed for computer aided education, writing boards, etc. The offline recognition systems implicate the automatic transformation of offline handwriting into letter code, which can be implemented for signature check, bank check reading, book transcription, etc.

Among different offline handwriting analysis, handwriting denoising can be an important field. In recent decades, it has achieved great interest. Offline handwriting denoising is a sub-issue of image inpainting domain, which is utilized to delete objects or fill lacking areas based on known image features. After inpainting the occluded image, a damaged image can be recovered to its original counterpart image according to a human visual test. In fact, inpainting Occluded handwriting have different advantages, including: (a) improved recognition performance, where, the overall recognition rate of the writing can be increased by inpainting the areas of the handwriting that are masked. Because the inpainted regions can give further information that could enhance the recognition. (b) improved readability: Inpainting can improve text clarity and readability, which can be helpful for projects like preserving or digitizing old documents. (c) improved user experience, when reading obscured text in digital documents that are displayed on screens, inpainting can make the experience for the user more enjoyable and less unpleasant.

Image inpainting process is widely utilized in old photos, the protection of valuable historical document, corrupted handwritten recovery, and handwriting recognition. In recent years, different methods have been proposed based on wavelet representation [2], Algorithmic Probability (AP) [5], structural features [4], variational analysis [3] and etc.

For example, Total variation approaches consider the smoothness of original images, so, which can eliminate spurious noises and small holes.

Some subsets of images can contain certain properties, such as having a low-rank property [13] or being planar [14]. Although they are able to fill holes by local pixels, they are incapable to introduce the semantic information in lost parts, specifically, when the desired information is missed in corrupted images.

There exist other methods, for example, Hareesh et al. [22] used a simple gradient function technique and fractional derivatives to estimate the priority order of filling step. The authors in [13] stated image inpainting problem as a low-rank matrix approximation issue and achieved a good approximation to the matrix rank.

Xi et al. [21] collected the features indicated by the two-dimensional entropy gray distribution in the priority formula to decrease the influence caused by the decline of the confidence term in the image inpainting process.

Contrary to the previous hand-crafted approaches, existing models, such as Convolutional Neural Network (CNN) [11, 16, 17] and Deep Recurrent Neural Network (DRNN) [9], achieve good results when treating the uncorrupted handwriting Chinese characters based on its structural characteristics.

The Stacked AutoEncoder (SAE) [10] is one of the methods used for automatic feature learning of unlabeled samples. This method is interesting in the field of image inpainting, because SAE encodes the features in the hidden units and then reconstructs the pattern by its decoder model. But, this method does not solve the handwriting image inpainting issue directly. However, it can be extended to implement it [1]. Sparse coding was used to reconstruct occluded images according to a learned dictionary [15]. GAN-based techniques have emerged as a good model for image inpainting. Initial works [23] used CNNs for inpainting and denoising small areas. Firstly, context encoders [16] trained Deep Learning models for denoising larger regions. More recently, the authors in [24] ameliorated this model by implementing global and local discriminators like adversarial losses and utilized dilated convolutions in the denoising model to replace the entirely channel-connected layers adopted in the context encoders.

Recently, the authors in [8] proposed to inpaint occluded chinese handwriting character utilizing Convolutional Generative Adversarial Network (CGAN) and recognize the obtained handwriting with GoogLeNet.

They combine between the generator and discriminator to produce realistic characters from its counterpart occluded images. The advantage of using the contextual loss is to surpass the use of concrete positions of occluded regions.
Despite the advances in inpainting corrupted handwriting characters based on unsupervised GAN, there are some problems remain unsolved.
For example, existing methods utilize masks to localize the occluded positions and the CNNs-based methods use also the masks to guide the training process. Here, the main drawback is the difficulty of annotation and the time-consuming.
CNNs-based methods utilize the local features of occluded areas to inpaint handwriting images. However, if the occluded regions increase the CNNs might get small receptive fields, producing blurry images.
The question here is how to locate the occluded regions automatically, how we use more information from realistic areas of handwriting images and inpaint the images with large corrupted parts to finally produce more realistic results?
Otherwise, GANs-based methods consider the distribution of handwriting characters to inpaint them. This process can cause other problems: the GAN can produce some fake handwriting characters, where, it just generates images as handwriting forms which are not a truth category of character.
In this context, the authors in [9], used an end-to-end model composed by the GAN with attention mechanism, cyclic loss, a VGG-19 as perceptual loss and adversarial classification loss. They generated more realistic handwritten Chinese characters and evaluated the obtained results by applying the recognition system GoogleNet.
Attention model can capture global and local dependencies [32,26,8,27], and estimate the response at a specific position in a sequence paying attention to all positions within the sequence.
However, it is difficult to obtain dynamic features from the inpainted characters and the recognition will be reduced, thus, it will be a hard task and sometimes leads to lowest performance.
To remedy these problems to an extent, we exploit Online Handwriting Recovery (OHR) to denoise the occluded offline handwriting. Thus, we have recovered not only significant writing, but also its online counterpart with dynamic features.
In fact, OHR [32, 10, 30, 12] is an interesting enabling process to empower various high-value applications, such as handwriting recognition, signature verification, writer identification and also can solve the problem of handwriting denoising.
Compared with traditional handwriting denoising issue, OHR is a more challenging handwriting inpainting task.
The authors in [32] proposed to use attention mechanism , where the CNN-GRU is used as encoder and a BGRU with attention is used as decoder to recover the temporal order with velocity directly. The results on multi-lingual datasets outperformed traditional systems significantly. In addition, attention mechanism can force the framework to focus on semantic handwriting features.
Previous works propose that simple attention cannot raise local dependencies [25, 26]. To overcome this issue and inspired by vision transformers [26], we use a Multi-head attention that gives greater strength to encode nuances for each character.
The main contributions of this work are presented below:

- We began by investigating a novel recovery framework based on Temporal Order with Multi-head Attention Network and stacked-LSTM decoder (TO-MultiNet) of transformer to inpaint the corrupted handwriting.
- This TO-MultiNet framework is able to reconstruct the temporal order and the pen velocity of multilingual offline handwriting character.
- A set of random corruptions are added onto the realistic handwriting to produce a series of occluded samples used as the input of the trained model TO-MultiNet.
- We inpaint occluded handwriting with large corrupted regions effectively and produce more truth images. For the first time in the field of occluded handwriting inpainting, we obtain the effective image and its counterpart online characterized by dynamic features such as the temporal order and the pen velocity, which can improve the recognition rate. The evaluation is based on a new offline/online recognition system based on the Beta-GRU model.
- We combine two challenging issues which are: a) The inpainting of corrupted handwriting based on a transformer's TO-MultiNet trained for OHR. b) The Beta-GRU model for online handwriting recognition. Consequently, we obtain a novel recognition framework for occluded multilingual offline handwriting.

The second section presents the general framework architecture of the study. Section 3 analyses and discusses the obtained results. Finally, the last section presents the conclusion.

2 Inpainting realistic character via TO-MultiNet

Inspired from the success of transformer, we choose to design a specific model to improve Attention-based encoder-decoder model for Online Handwriting Recovery (OHR). Thus, we propose to use multi-head attention and stacked-LSTM decoder.
we investigate handwriting denoising on occluded handwriting characters and digits using transfer learning. Since the goal of OHR is to convert the offline handwriting to its corresponding online signal, it can be formulated as an image-to-sequence issue and implemented by the encoder-decoder with attention model [12, 32].

In this paper, we introduce a novel recovery-denoising framework named Temporal Order with Multi-head Attention Network (TO-MultiNet) to recover the online signal from its corresponding offline handwriting, then, to inpaint the corrupted handwriting. So, the input is offline handwriting character and the output will be the online signal. Here, we use the corrupted character as input to the obtained TO-MultiNet model.
The flowchart of our proposed framework is shown in Fig. 1.

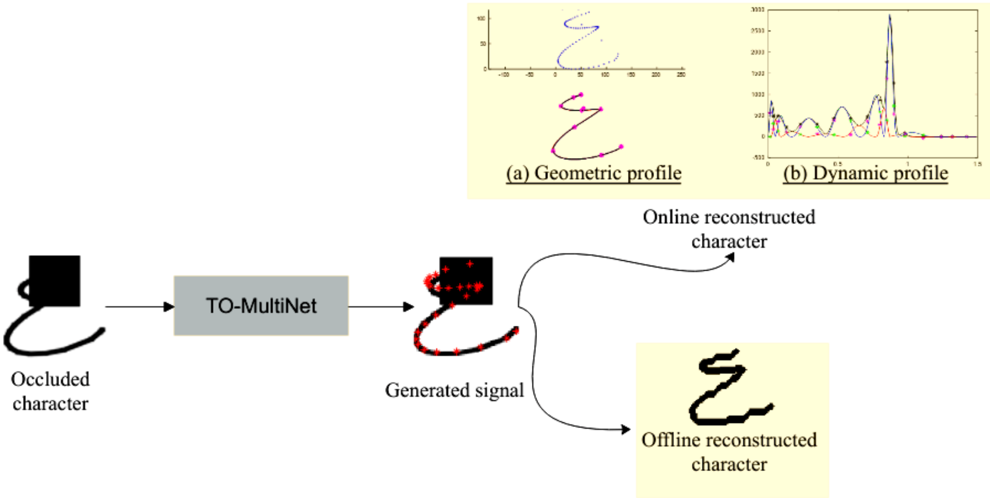

Fig. 1. The flowchart of the proposed framework

2.1 TO-MultiNet Model for OHR

Given an input handwriting character image (64*64), let $F \in R^{D\times H\times W}$ indicates the extracted feature maps from CNN, where H, W and D denote height, width and channel number, respectively. These features can be observed as a sequence of feature vector $f = \{f_1, f_2,..., f_N\}$, where $f_i \in R^D$ and $N = H \times W$. Let $p = \{p_1, p_2, \cdot\cdot\cdot, p_i\}$ represents the target output, where p_i indicates the last point corresponding to <x_i,y_i> coordinate. As depicted in Fig. 2, this baseline framework employs two LSTM layers and a mutli-head attention which work as follows:

$$h_t^1 = LSTM^1\left(\left[y'_{t-1;}\, h_{t-1}^2\right];\ h_{t-1}^1\right) \tag{1}$$

$$c_t = Att(h_t^1, f) \tag{2}$$

$$h_t^2 = LSTM^2(c_t;\ h_{t-1}^2) \tag{3}$$

For single head attention, the first LSTM layer needs to use the previous $y'_{t-1;}$, the output of the second LSTM h_{t-1}^2 and its previous hidden state h_{t-1}^1. Where y'_{t-1}

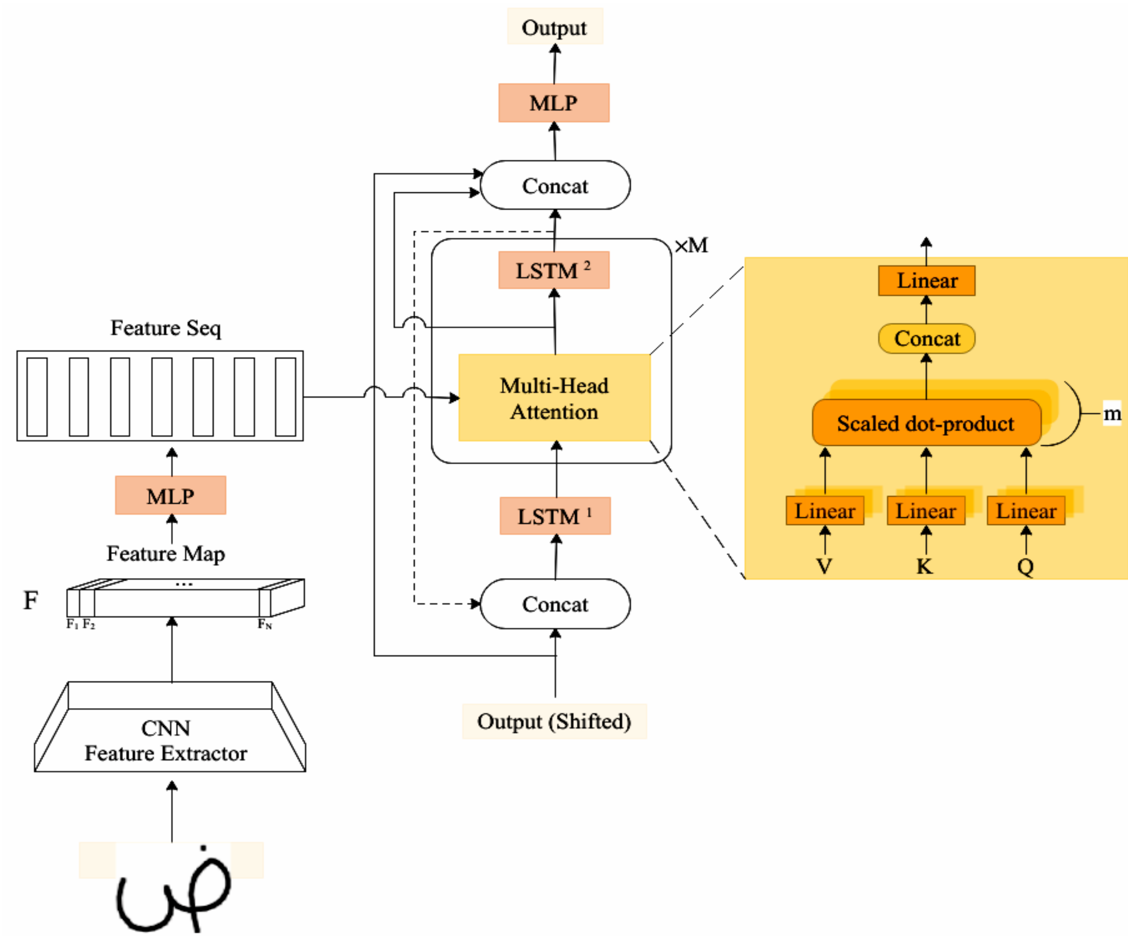

Fig. 2. The proposed architecture

denotes a trainable data of y_{t-1} . The first LSTM layer generates $h_t^1 \in R^{dLSTM}$ which represents the query (Q) of the attention mechanism. The sequence of features f works as key (K) and value (V). Then the attention mechanism outputs the context vector noted by $c_t \in R^{out}$. The second LSTM layer utilizes the context vector as input and generates $h_t^2 \in R^{dLSTM}$. The context vector is calculated as the following equation :

$$e_{t,i}=Align(h_{t-1}^1 , f_i)=h_{t-1}^1 \odot f'_i \quad (4)$$

$$\alpha_{t,i}=\frac{\exp(e_{t,i})}{\sum_{l=1}^{N}\exp(e_{t,l})} \quad (5)$$

$$c_t=\sum_{i=1}^{N}\alpha_{t,i}f_i \quad (6)$$

formula (4) calculates the align value between the image feature vector f_i and the hidden state of the first LSTM layer h_t^1 which represents the query (Q). Equation (5) is the attention weights that designate the importance of the key value f_i at time step i to produce the value (V) at time step t. The softmax function normalizes the vector e_i as attention mask. Formula (6) presents the final context c_t.
Online Handwriting Recovery need to generate complex temporal output sequence including the pen order and the other temporal information taking into account a set of pixels that represent the offline handwriting. MHA is able to predict this type of complex signal specially when different attention heads combine and capture different context feature. With multi-head attention, different attention heads capture different feature successfully.
Compared with simple attention, MHA utilizes M attention heads that are independent and run in parallel. As shown in the right of Fig. 2.
The main objective of Scaled dot-product attention (SDPA) is to select the most significant information from the input sequence for such target sequence.
The output of the latter algorithm is presented in Equation 7.

$$f_{SDPA}(Q,K,V) = Softmax\left(\frac{QK^T}{\sqrt{dk}}\right)V \tag{7}$$

where dk is the dimension of keys and queries. The dot products of the query Q with all keys K^T is computed, and divide each one by $\sqrt{dk}$ and use the softmax operation to get the best weights on the values.

To use MHA in the context of OHR, we basically replace simple attention in Eq. (3) with MHA as the equation bellow:

$$c_t = Multi - head\ (h_t^1, f) \tag{8}$$

For each attention head there are different linear transformations for the known Query (Q), Key (K) and Value (V). The trainable parameters W for each linear projection of Q, K and V are different.
For each M time, the scaled dot-products of attention output are concatenated. The obtained value after a linear transformation is used as MHA output. Thus, the final attention output is a concatenation between the output context vectors, as shown in the following formula:

$$\mathrm{Head}_m = Att(QW_m^Q, KW_m^K, VW_m^V) \tag{9}$$

$$\text{Multi-head}\ (Q,\mathrm{K},\mathrm{V}) = Concat(Head_m, \ldots, Head_M)\mathrm{W}^1 \tag{10}$$

where M is the number of different stacked LSTM-decoder blocks.
In stacked LSTM-decoder, the output of the first LSTM-decoder block is utilized as a query for the second LSTM-decoder block. Stacked LSTM-decoder increases the framework efficiency and can realize good result than a single LSTM-layer decoder.

For OHR, we choose this model insight and propose a novel decoder using stacked LSTM-decoder with MHA. the main function is as follow:

$$h_t^1 = LSTM^1\left(\left[y'_{t-1;}\, h_{t-1}^M\right];\ h_{t-1}^1\right) \tag{11}$$

$$\begin{cases} c_t^m = Multi-head\ (h_t^{m-1}, f) \\ h_t^m = LSTM^m(c_t^m;\ h_{t-1}^m) \end{cases} \quad \text{for } 1 \leq \text{m} \leq \text{M}, \tag{12}$$

The stacked decoder generates a probability distribution of two outputs $<x,y>$ coordinates.

To calculate the point coordinates probability, the context vector c_t, the output of the second LSTM h^2 and the previous predicted point y_{t-1} are concatenated together and used as input to a simple MLP network. The generated 50 point coordinates are updated using the loss L1 as follow:

$$L_1 = \frac{1}{i}\sum_{t=1}^{i} |P(t) - P'(t)| \tag{13}$$

where i is the number of points, $P' = [<x'_1,y'_1>,...,<x'_i,y'_i>]$ presents the target sequence and P denotes the predicted output sequence.

In the test step, given the corrupted handwriting image as input to the trained model, the proposed framework produces a sequence of points corresponding to human-like writing.

2.2 Transfer learning for inpainting corrupted handwriting

The main idea of existing image inpainting methods is to get the probability density function of different images and then change the problem of treating lost pixels to the problem to find the most likelihood function. In general, it is difficult to procure the best likelihood function from image pixels, that encouraged researchers to utilize GAN to produce new samples from an unknown data distribution, then find sample pixel values to reconstruct corrupted characters.

In the sequence-to-sequence with attention mechanism for OHR [32], given an image I, the first step is to extract the corresponding features by an encoder. Then, produce the point coordinates in a step-by-step mode by a decoder model and attention mechanism collects appropriate features.

we introduce a novel system that use the transfer learning of our proposed recovery framework named TO-MultiNet described in the previous section.

Transfer learning is to re-use the training data to generate the online handwriting signal from its counterpart occluded one.

TO-MultiNet framework has been trained on online handwriting recovery from offline handwriting.

After obtaining a trained deep learning model, we transfer the online recovery learning to the inpainting of occluded handwriting.

Here, just the input is changed from original offline handwriting to an occluded image. Then, passed it (as test step) to the learned model to finally generate its counterpart online one. For more clarification, Fig. 2 depicts an example of the occluded Arabic letter "ain" which used as input to the OHR model and the online signal is generated effectively corresponding to the realistic character. As shown in the right of figure 2 our novel framework can generate both online and offline handwriting. Thus, from occluded handwriting we can obtain the dynamic feature from the reconstructed character.

2.3 Recognizing Reconstructed chars via Beta-GRU

The evaluation of the inpainting process is done based on the recognition step. Originally, the recognition systems used the preprocessing, the feature extraction and finally the classification step to recognize the character handwriting.

Existing works are based on hidden markov model (HMM) [19], support vector machines (SVM) [18], and etc.

In fact, traditional machine learning algorithms have reached a bottleneck in handwriting recognition, in addition, the training is more complicated.

With the success achieved by Deep Learning (DL) models, the solution to handwriting recognition has been changed from hand-crafted approaches to CNN [20], LSTM [27] models.

In this work, we extract beta elliptical parameters [7] from the reconstructed online signal, since it has demonstrated a great success, particularly in recognition tasks [27, 7]. Each stroke of word is limited by a rectangle regrouping a sequence of points. The major characteristics used for the classification are as follow:

- The position of reference points: terminal points and crossing points, which are detected based on the variation of pen velocity.
- The form of trajectory: the deviation of angles is calculated based on the elliptical arc located in each reference point.
- The measure of Curve deviation: in such cases, there are analoguous strokes that we can differentiate over them utilising the measure of Curve deviation based on the angle.

To achieve a good recognition rate, GRU with three layers is employed and each layer has 512 units. As shown in the right of figure 2.

3 Experiments

3.1 Datasets

We validate our framework using Arabic, Latin and Indian scripts. We specifically employ the dual on/off IRONOFF dataset [28] for Latin data. We use the isolated lower and capital letters, and numeric data that have been divided into 26 and 10 categories, respectively.

We use the dual on/off Dhad [30] dataset for Arabic script. It has 389 handwritten classes created by 40 Tunisian scriptors. Where 28 for single characters. For Indian script, we utilize Telugu[1] dataset which contains 116 characters with 500 samples per character.

Due to the small size of the various datasets, we increased the number of samples applying a data augmentation method [31]. This approach is used on the online script. We transform the augmented scripts into their offline handwriting counterparts. The image skeleton is created by concatenating the pen points. The training, testing, and validation samples are listed in Table 1.

Table 1. Details of datasets

Scripts	Training	Testing	Validation
Dhad Chars	70,000	20,000	10,000
IRONOFF Chars	56,000	17,800	10,200
IRONOFF Digits	70,000	20,000	10,000
Telugu	58,000	17,400	11,600

On each data script, we added a black/white rectangular block randomly with the area scale of 5%, 15% and 25% of the original handwriting, respectively. Figure 3a-d depicts different examples of the original data and its occluded counterpart with different noise block. Fig. 3.a shows the original data, Fig. 3.b-d show various occluded characters with round noise, rectangular noise and removed pixels, respectively.

The recovery framework was implemented utilizing the GPU-accelerated GeForce GT 650M and the Tensorflow library. The Adam Stochastic trainer algorithm is used. The learning rate is 0.0001 and the batch size is 32. We used training data at random inside each of the 400 iterations. The implementation took 16 hours to complete the recovery process. Utilizing Matlab, the Beta-GRU recognition system is constructed. The network is tested by the generated signal and trained using the online data. The training process took one hour. GRU is trained in the recognition experiments by reducing the Softmax loss. The settings are updated using the Adam algorithm. The learning rate is set at 0.0002 per second. After 100 epochs, the training process is terminated. We use ten features of Beta.

3.2 Visual results of inpainting via TO-MultiNet

Based on 50 test samples from each dataset, the visual inpainting results on a variety of test images are shown in Fig. 4 and 5. Those figures show that, regardless of how many pixels are missing from the character, inpainting works well when there is a small or moderately sized occluded block. Even while the results can sometimes be rather disorganized and the corrupted sections are larger, we can still differentiate which letters can be, and the key details can sometimes be inpainted to some extent.

[1]http://lipitk.sourceforge.net/datasets/teluguchardata.htm

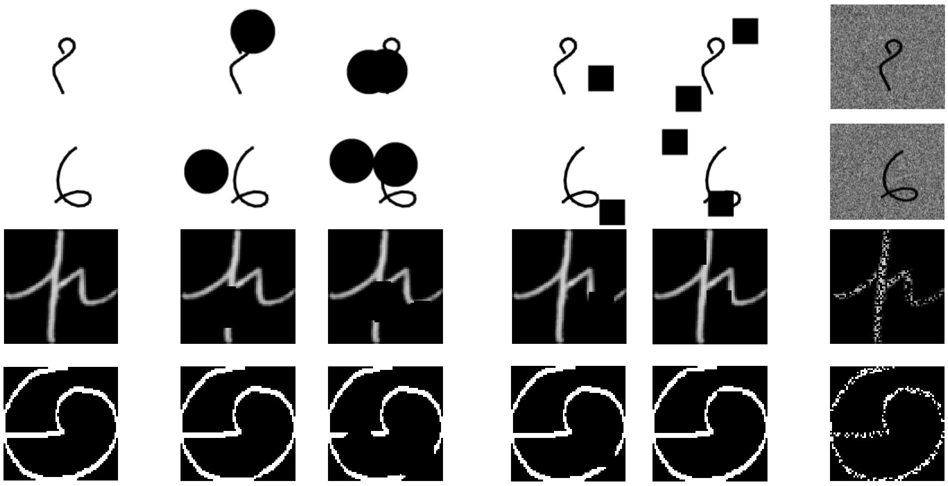

Fig. 3. Truth offline handwriting and its occluded with different noise block

With intent to compare our results with the model [32], we use the same occluding region. Thus, for an occluding region of a particular type, the inpainting results for the various occluded zones of a specific character change. For example, in Fig. 4, when recovering the Arabic letter “sad” with the model [32], the output is the Arabic letter “lam”. In addition, in Fig. 5, when reconstructing the Latin letter "b" with the model [32], the output is another character "f". This problem is due to the use of a simple attention head. However, Thanks to the multi-head attention, our model is able to reconstruct the same letter but can exist in the training dataset.

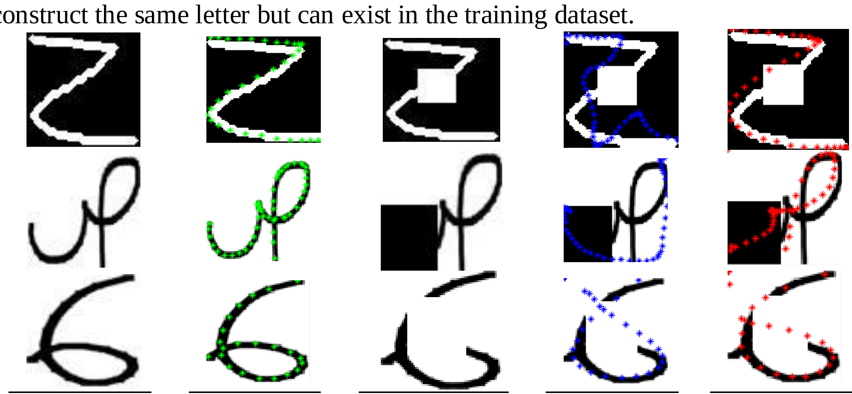

Fig. 4. The inpainting results of rectangular noise on Arabic characters and digit

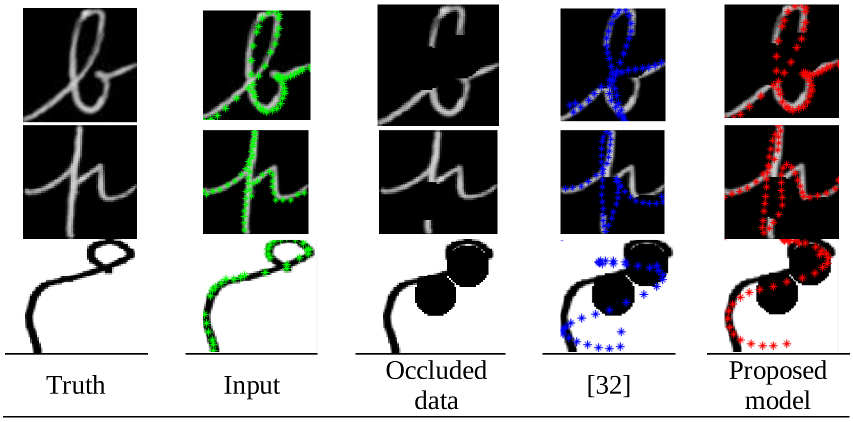

Fig. 5. The inpainting results of round noise for Latin and Arabic scripts

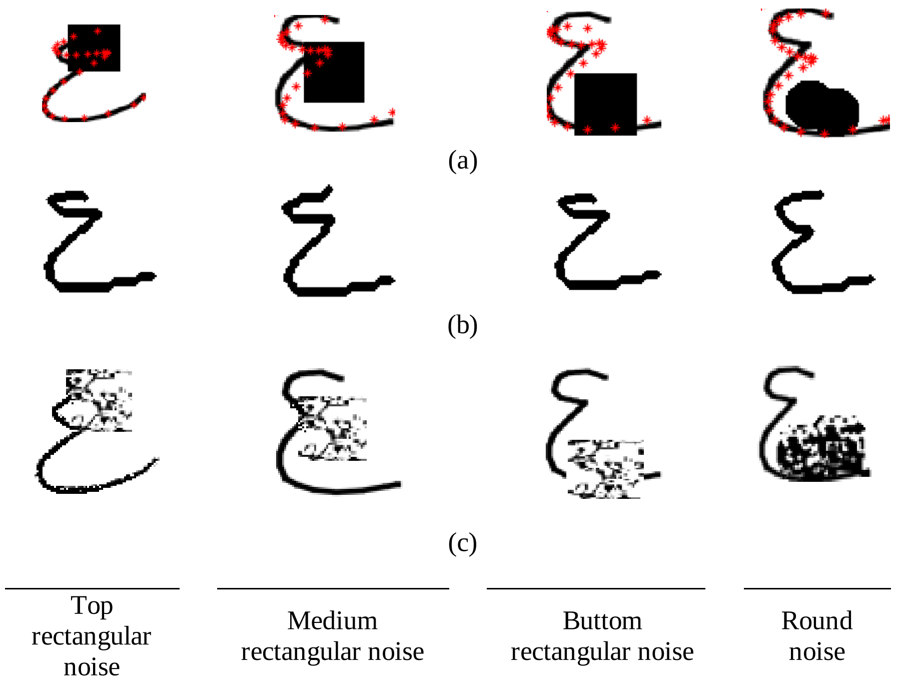

Fig. 6. The inpainting results of different type of noise for the Arabic character "ain". (a) the online denoising results with our proposed model, (b) our reconstructed offline generated from the online inpainted handwriting, and (c) the denoising results with [33]

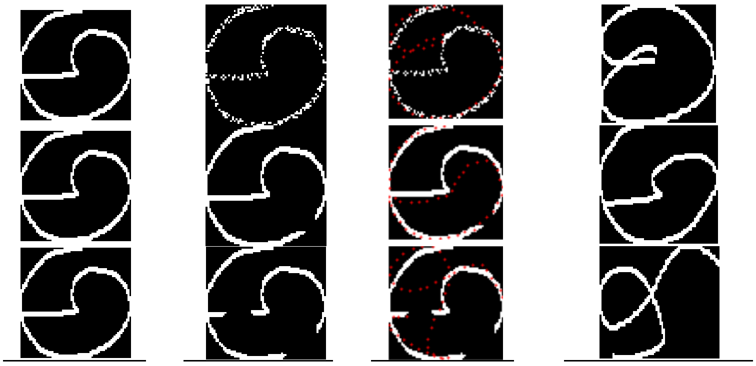

Fig. 7. The inpainting results of pixel removal for Indian script.

Authors in [33] utilize low-rank minimization to reconstruct corrupted photos. To evaluate its performance, we use the code from [33]. However, in comparison to our suggested model, the model of [33] is not strong enough. For example, the black zone of noise maintains the same but becomes a little blurry when we recover the Arabic letter "ain , as shown in Fig. 6.

In fact, the limitation of our suggested framework is that it only produces another learnt character when our model is unable to inpaint accurate data, which might be a benefit over other work [33-8] that creates new forms that are not comparable to human writing. Figure 7 depicts a failed attempt to rebuild Indian letter, but those failed results are a realistic Indian characters.

3.3 Recognition results via Beta-GRU

We vise to prove whether the proposed framework TO-MultiNet can support the effectiveness of inpainting occluded offline handwriting. Precisely, we compare five models: (1) GA [19] system is the heuristic recovery method based on the skeleton extraction and Genetic Algorithm; (2) S2S-BLSTM [12] system is a sequence-to-sequence with BLSTM model without attention. (3) Att-BGRU [32] system is an End-To-End framework based on traditional Deep Learning models GRU with attention mechanism. (4) S2S-CBGRU system is a sequence-to-sequence based on CNN and BGRU as encoder and decoder without attention mechanism, which is inspired from [32]. (5). TO-MultiNet framework is detailed in section 2.1.

All the experimental results are presented in Table 2. From this Table, the recognition rate achieved by our framework, regardless of the database, is higher than other methods.

The lower recognition rates realized by the heuristic method GA [19] are due to their system's sensitivity to noise recovering the incorrect order.

The model S2S-BLSTM [12] achieves inferior results (89.3%) when recovering the digit, due to the absence of a simple attention layer between the encoder and the decoder which recover, in some cases, the incorrect letter.

The Deep Learning system [32] applies the End-To-End concept the most closely to our framework. The degree of their precision (94.2%) is comparable to our rate on Dhad (94.8%). However, using simple attention head rather than multi-head attention of transformers may be a valid rationale to achieve an effective outcome.
These studies demonstrate the superiority of our suggested method when inpainting the corrupted handwriting over the recovery models [32] and the image inpainting method [33] in terms of point order recovery as well as time and even efficiency.

Table 2 A comparison of the recognition rate of our framework TO-MultiNet with other baseline frameworks on four datasets

Models	IRONOFF	IRONOFF digits	Dhad	Telugu
GA [19]	46.4	63.1	31.1	52.6
S2S-BLSTM [12]	89.1	89.3	90.4	89.7
Att-BGRU [32]	91.1	92.5	94.2	90.3
S2S-CBGRU	90.2	90.1	90.5	89.8
Att-BGRU(Ours)	**91.9**	**93.0**	**94.8**	**92.2**

Conclusion

This study initially suggests a TOP-MultiNet-based architecture for inpainting offline, handwritten characters that have been occluded. The inpainting process contains two main components which are: (a) The online handwriting recovery integrating the temporal order and the pen velocity. This process is utilized to train the network to generate an effective online signal from its counterpart occluded offline. (b) The study also suggests to use Beta-GRU model to recognize the inpainted characters. The employment of beta elliptic features and the dynamic characteristics can improve the recognition rate.
Previous research literature has never explored the proposed technique for recognizing occluded Multi-lingual offline handwriting. And a number of tests confirm the efficiency of our framework. Because we extract, for the first time, the dynamic features from occluded handwriting. The suggested idea cannot be applied with traditional recovery systems which are based on skeleton and contour extraction.
The following issue need to be taken into account for future work: inpaint original characters and words using the suggested method when their larger zones are corrupted. To better resolve this challenging issue, we will improve the architecture and use another deep learning models.

References

1. Shcherbakov, O., Batishcheva, V. , "Image inpainting based on stacked autoencoders," In Journal of Physics: Conference Series, Vol. 536, No. 1, p. 012020, IOP Publishing, 2014.
2. Chan T. F., Shen J., Zhou H.-M. A, "Total Variation Wavelet Inpainting Model with Multilevel Fitting Parameters," Proc. SPIE 6313, Advanced Signal Processing Algorithms, Architectures, and Implementations XVI, 2006.

3. Arias P., Caselles V., Sapiro G. A., "Variational Framework for Non-local Image Inpainting," Proc. EMMCVPR'09. 345–358, 2009.
4. Shibata T., Iketani A., Senda Sh., "Fast and Structure-preserving Inpainting Based on Probabilistic Structure Estimation," MVA2011 IAPR Conference on Machine Vision Applications, Nara, JAPAN 22–25, 2011.
5. Potapov A., Scherbakov O., Zhdanov I., "Practical algorithmic probability: an image inpainting example," Proc. SPIE 9067, Sixth International Conference on Machine Vision.
6. Potapov A. S., "Using multiple video-information representations in automatic image analysis systems," Journal of Optical Technology 79 689–692, 2012.
7. Y. Hamdi, H. Boubaker, T. Dhieb, A. Elbaati and A. M. Alimi, "Hybrid DBLSTM-SVM Based Beta-Elliptic-CNN Models for Online Arabic Characters Recognition," International Conference on Document Analysis and Recognition (ICDAR), Sydney, NSW, Australia, 2019, pp. 545-550, doi: 10.1109/ICDAR.2019.00093, 2019.
8. Li, J., Song, G., Zhang, M., "Occluded offline handwritten Chinese character recognition using deep convolutional generative adversarial network and improved GoogLeNet," Neural Computing and Applications, 32(9), 4805-4819, 2020.
9. Song, G., Li, J., Wang, Z., "Occluded offline handwritten Chinese character inpainting via generative adversarial network and self-attention mechanism," *Neurocomputing*, *415*, 146-156, 2020.
10. Rabhi, B., Elbaati, A., Boubaker, H., Pal, U., Alimi, A., "Multi-lingual Handwriting Recovery Framework based on Convolutional Denoising Autoencoder with Attention model," TechRxiv. Preprint. https://doi.org/10.36227/techrxiv.21347262.v1, 2022.
11. LeCun, Y., Boser, B., Denker, J., Henderson, D., Howard, R., Hubbard, W., Jackel, L., "Handwritten digit recognition with a back-propagation network," Advances in neural information processing systems, 2, 1989.
12. Rabhi, B., Elbaati, A., Hamdi, Y., Alimi, A., "Handwriting Recognition Based on Temporal Order Restored by the End-to-End System," International Conference on Document Analysis and Recognition (ICDAR), Sydney, NSW, Australia, 2019, pp. 1231-1236, doi: 10.1109/ICDAR.2019.00199, 2019.
13. Hu, Y., Zhang, D., Ye, J., Li, X., He, X., "Fast and accurate matrix completion via truncated nuclear norm regularization," IEEE transactions on pattern analysis and machine intelligence, 35(9), 2117-2130, 2012.
14. Huang, J. B., Kang, S. B., Ahuja, N., Kopf, J., "Image completion using planar structure guidance," ACM Transactions on graphics (TOG), 33(4), 1-10, 2014.
15. Mairal, J., Elad, M., & Sapiro, G. "Sparse representation for color image restoration," IEEE Transactions on image processing, 17(1), 53-69, 2007.
16. Pathak, D., Krahenbuhl, P., Donahue, J., Darrell, T., Efros, A. A., "Context encoders: Feature learning by inpainting," In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 2536-2544), 2016.
17. Xie, J., Xu, L., & Chen, E., "Image denoising and inpainting with deep neural networks," Advances in neural information processing systems, 25, 2012.
18. Cortes, C., Vapnik, V., "Support-vector networks," Machine learning, 20(3), 273-297, 1995.
19. Elbaati, A., Boubaker, H., Kherallah, M., Ennaji, A., El Abed, H., Alimi, A. M., "Arabic handwriting recognition using restored stroke chronology," In 2009 10th International Conference on Document Analysis and Recognition (pp. 411-415) IEEE, 2009.
20. Hamdi, Y., Boubaker, H., Rabhi, B., Abdulrahman, M. Q., Fahd, S., Alharithi, O., Almutiry, Dhahri, H., Alimi, A. M., "Deep learned BLSTM for online handwriting modeling

simulating the Beta-Elliptic approach," Engineering Science and Technology, an International Journal,V35, https://doi.org/10.1016/j.jestch.2022.101215, 2022.

21. X. Xi, F. Wang, Y. Liu, "Improved Criminisi Algorithm Based on a New Priority Function with the Gray Entropy," Ninth International Conference on Computational Intelligence and Security, 2013, pp. 214-218, doi: 10.1109/CIS.2013.52, 2013.
22. Sai Hareesh, A., Chandrasekaran, V., "Exemplar-based color image inpainting: a fractional gradient function approach," Pattern Analysis and Applications, 17(2), 389-399, 2014.
23. Xu, L., Ren, J. S., Liu, C., Jia, J., "Deep convolutional neural network for image deconvolution,". Advances in neural information processing systems, 27, 2014.
24. Iizuka, S., Simo-Serra, E., & Ishikawa, H., "Globally and locally consistent image completion," ACM Transactions on Graphics (ToG), 36(4), 1-14, 2017.
25. Wu, H., Xiao, B., Codella, N., Liu, M., Dai, X., Yuan, L., Zhang, L.: Cvt: Introducing convolutions to vision transformers. arXiv preprint arXiv:2103.15808, 2021.
26. Yuan, K., Guo, S., Liu, Z., Zhou, A., Yu, F., Wu, W., "Incorporating convolution designs into visual transformers," arXiv preprint arXiv:2103.11816, 2021.
27. Hamdi, Y., Boubaker, H., Rabhi, B., Ouarda, W., Alimi, A. M., "Hybrid Architecture based on RNN-SVM for Multilingual Online Handwriting Recognition using Beta-elliptic and CNN models," TechRxiv. Preprint. https://doi.org/10.36227/techrxiv.13903661.v3, 2021.
28. C. Viard-Gaudin, P. M. Lallican, S. Knerr, P. Binter, “The ireste on/off (ironoff) dual handwriting database,” In Proceedings of the Fifth International Conference on Document Analysis and Recognition. ICDAR, pp. 455-458, 1999.
29. M. Kherallah, A. Elbaati, H. Elabed, A. M. Alimi,“The on/of (LMCA) dual Arabic handwriting database,” In: 11th international conference on frontiers in handwriting recognition (ICFHR), 2008.
30. B. Rabhi, A.Elbaati, T. M. Hamdani, A. M. Alimi, “ASAR 2021 Competition on Online Signal Restoration Using Arabic Handwriting Dhad Dataset,” In: Barney Smith E.H., Pal U. (eds) Document Analysis and Recognition – ICDAR 2021 Workshops. ICDAR 2021. Lecture Notes in Computer Science, vol 12916. Springer, Cham. https://doi.org/10.1007/978-3-030-86198-8_26, 2021.
31. Y. Hamdi, H. Boubaker, A. M. Alimi,“Data Augmentation using Geometric, Frequency, and Beta Modeling approaches for Improving Multi-lingual Online Handwriting Recognition,” IJDAR 24, 283– 298,https://doi.org/10.1007/s10032-021-00376-2, 2021.
32. B. Rabhi, A. Elbaati, H. Boubaker, A. M. Alimi, “Multi-lingual character handwriting framework based on an integrated deep learning based sequence-to-sequence attention model,” Memetic Comp. 13, 459– 475,https://doi.org/10.1007/s12293-021-00345-6, 2021.
33. Lu C, Tang J, Yan S, Lin Z., "Generalized nonconvex nonsmooth low-rank minimization," In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 4130–4137, 2014.