Albanian Handwritten Text Recognition using Synthetic Datasets and Pre-Trained Models

HAKIK PACI*, DORIAN MINAROLLI, EVIS TRANDAFILI, STELA PATURRI
Computer Engineering Department,
Polytechnic University of Tirana,
Bul. "Dëshmorët e Kombit", "Nënë Tereza", sheshi nr. 4, Tirana
ALBANIA

*Corresponding Author

Abstract: - Handwritten Text Recognition (HTR) has continuously attracted the focus of researchers to enable the integration of technology into our daily lives. Handwritten text recognition (HTR), a technology of considerable importance, takes a leading role in the analysis and digitization of various documents. This technology is important in facilitating the efficient use of handwritten documents, especially within academic, historical, and cultural contexts. The use of artificial intelligence in handwriting recognition offers a very good opportunity to achieve satisfactory results in this field, but to achieve good results a large dataset is needed. Creating a large dataset to train different AI models is a challenge for languages with limited resources such as the Albanian language. This paper aims to present a novel approach to the development of an HTR system for the Albanian language using an attention-based encoder-decoder architecture. The dataset used in the experiments is a synthetic dataset generated using deep learning techniques based on the English language dataset as they are both variants of the Latin alphabet. We enhanced the dataset with two letters specific to Albanian, (*"ë" and "ç"*). The usage of pre-trained English models for handwriting recognition improved our model's performance. The results of the experiments are very promising and prove that our approach is efficient in recognizing handwriting in the Albanian language. This shows that the attention-based encoder-decoder architecture can be adopted for different languages with limited resources.

Key-Words: - HTR (Handwritten Text Recognition), Albanian language, Synthetic dataset, HTR Models, Machine learning, Deep learning.

Received: June 29, 2023. Revised: February 16, 2024. Accepted: April 9, 2024. Published: May 15, 2024.

1 Introduction

According to the "Ethnologue Guide" in our world, there are more than 7,000 languages including dialects, and some of them are used by a few people and only 23 languages are used by half of the population of the world. This means that some of those languages in the long term are going to be forgotten.

The Albanian language is one of the oldest languages in the world and the only surviving representative of the Albanoid branch, which belongs to the Paleo-Balkan group, [1]. From a grammatical perspective, it has major differences compared to the other European languages. The current Albanian alphabet has 36 letters and is based on the Latin alphabet with the addition of letters "ë", "ç", and nine digraph letters: "dh", "gj", "ll", "nj", "rr", "th", "sh", "xh", and "zh".

Handwritten documents are a valuable resource of information especially when trying to make handwritten content like manuscripts, personal correspondences, legal documents, and scientific studies stored in archives, accessible and usable by NLP systems.

Handwritten text recognition (HTR) is a technique and ability of a computer to read data from paper documents written by hand, [2].

The process of extraction of digital data from paper documents can be achieved using Optical Character Recognition (OCR) and Intelligent character recognition (ICR) techniques, [3], [4]. The OCR technique usually is used when the text is printed and is a well-established technology. The ICR is used to convert/extract data from images of handwritten texts, and it is more complex than OCR technology because it can also detect and recognize different handwriting styles. Both technologies focus on the recognition of individual characters, [4] and they do not check if the generated characters are real words in a linguistic and semantic context.

The technology for interpreting handwritten text (not only the set of characters or words) into machine-readable data has become increasingly important in the current Artificial Intelligence era. The recognition of handwritten text has always attracted researchers to work on real-life applications in healthcare, banking, insurance, online libraries, etc. However, for the Albanian language, there is a significant gap in this research field. As far as we know, there is no publicly available system, nor a dataset for supporting the recognition of handwritten text in Albanian.

Nevertheless, we must notice that several efforts have been made to address the handwriting recognition problem from an NLP (Natural Language Processing) perspective. The first attempts to deal with HTR were based on Hidden Markov Models, [5], dating back 10 to 15 years ago. Recent advancements in neural networks have shifted the focus towards the architectures based on neural networks which excel with sequential data.

Due to the infinite human writing styles, the implementation of a Handwritten Text Recognition system is laborious, and requires a very large amount of training data, often resulting in poor performance. In this work, we are looking to modestly address this gap by proposing an HTR system specifically designed for the Albanian language and assessing and comparing its accuracy and efficiency with state-of-the-art data. We aim to contribute to the development of an HTR system for the processing of Albanian texts by introducing a novel approach, employing an encoder-decoder architecture. For the creation of our Albanian HTR model, we relied on the paper [2], who proposed an attention-based sequence-to-sequence model for English with an encoder composed of ResNet for feature extraction and bidirectional LSTM for sequence modeling.

There are no contributions to HTR in Albanian therefore it was not possible to employ an existing dataset of images of handwritten text prepared to train our HTR model. To address this gap, we generated a synthetic dataset through a deep learning model able to replicate human handwriting in different styles. Furthermore, the training process was sped up by employing a transfer learning technique with several modifications to adapt the model to the peculiarities of the Albanian language. The utilization of pre-trained English models proved to be an effective solution in scenarios where there is a limited availability of training data, providing an enhanced performance for HTR models designed for low-resource languages.

The paper is organized as follows: Section 2 covers the proposed approaches; Section 3 investigates the synthetic generation of the Albanian dataset used to train the HTR model; Section 4 analyses the performance and evaluates the effectiveness of the synthetic dataset in the Attention HTR model; and finally, we conclude in Section 5.

2 Proposed Approaches

In this section, we will explain the architecture of the Attention HTR model. There are three main components of the model: ResNet, bidirectional-LSTM, and Transformer.

- **Residual Network (ResNet)** is a breakthrough in the field of deep learning, especially in Convolutional Neural Networks. The seminal paper entitled "Deep Residual Learning for Image Recognition", [6], [7] addressed the challenges faced by deep neural network training—the vanishing gradients, which become an issue when the network depth increases. The residual blocks in ResNet allow deeper networks to train without the problem of the vanishing gradient. This can be achieved by using the output of the previous layer through a skip connection as an input to a new layer. This technique enables the network to learn better residual functions, concerning the input layer, and, therefore, facilitates better training of deep networks. In the HTR problem, we use ResNet in the feature extraction phase. The feature extraction phase is a very important and critical step in text recognition and transcription of HTR because handwritten text is quite complex and different from a person's style of writing. Therefore, a network's ability to learn the needed features and encode those in the input data helps in the performance of the HTR.
- **LSTM** is an RNN type but with memory cells that can store or retrieve information over long sequences. The memory cells enable the model to capture and remember, two characteristic features of sequential data. As such, LSTMs are a great option for this type of data processing, be it NLP or HTR, [8]. BiLSTM: BiLSTM is an extension of the LSTM that introduces bidirectional processing, [9]. The BiLSTMs can capture both past and future contextual information in a sequence because they can process the input sequence in two directions, forward and reward. Thus, with BiLSTM in HTR, a model may not only recognize

characters/words but grasp the relationships between them within a sentence or paragraph. This is important because many people's writings are context-dependent.

- The Transformer architecture was initially developed for machine translation tasks, its ability to address multiple challenges has transformed it into the model, [10], [11] used by most NLP applications. Transformer uses self-attention to define the relevance of different sections of the input sequence to the prediction. Instead of processing data in series, other sequence models do, including RNNs and LSTMs, [12], [13], [14]. As a result, for training and inference, allowing the Transformer to reach faster and higher efficiency, the model's speed is drastically increased. The Transformer architecture is utilized in HTR with both word and character-level options. It processes sequence data from images of handwritten text, effectively modeling context and dependencies between characters or words, [15], [16]. This contextual understanding is especially important in cases involving complex handwriting styles or context-dependent character variations.

The Attention HTR model, [2], that we employed in for our Albanian HTR, uses the attention encoder-decoder architecture for distinguishing human handwriting. To decompose and summarize the main elements of words, and feature extraction, the model relies on ResNet, whereas for sequence modeling it is based on bidirectional LSTM [9], [17], [18] and for making accurate predictions on content-based attention mechanism. To address the challenge presented by the limited amount of data available in the Albanian language dataset, we applied transfer learning techniques within the model, [2]. Consequently, the system pre-trained using the deep-learning model Attention HTR will enable us to train the system using an Albanian language dataset.

A pre-trained HTR model is adjusted and adapted to the specifics of the Albanian language and can be used to compare the performance with state-of-the-art models in the field of HTR.

The archiving of good results means a big step in the domain of HTR for the Albanian language.

3 Dataset

To train the HTR model it is necessary to provide a dataset that contains different images of handwritten texts. Finding a dataset with images of handwritten texts in the Albanian language wasn't possible because the number of studies in HTR is limited. To train a HTR model for this purpose we decided to use a synthetic dataset, [19], [20]. The dataset was generated by a deep learning technique developed for the English language. This technique can be modified, [21], for the Albanian language because almost all the letters in the Albanian language are the same as the letters of the English alphabet except for two letters, 'ë' and 'ç', [22]. To make the dataset more diverse, we generated it using six different handwriting styles.

The dataset's primary goal is to support the implementation of the Albanian language, accommodating unique characters 'ë' and 'ç', which differentiate Albanian from English. Precisely, positioning these characters in word images required image processing techniques, [20]. To attain this objective, we placed diacritical marks over 'e' as shown in Figure 1, and a line under 'c' as shown in Figure 2. By accurately positioning the mentioned two characters in the images processed by our synthetic writing model, we created a dataset of more than 12,000 Albanian words, prepared for training, testing, and validation purposes, [23]. The distribution graph of those word counts per number of letters is shown in Figure 3.

Fig. 1: Sample Word Generated by Synthetic Model with 'ë'

Fig. 2: Sample Word Generated by Synthetic Model with 'ç'

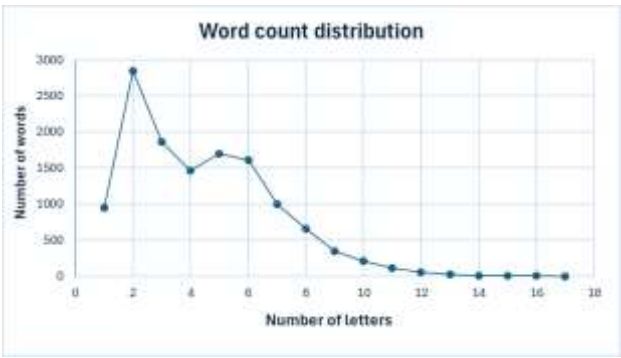

Fig. 3: Word count per number of letters

As shown in the graph in Figure 3, words with lengths from two to six are more widespread in the generated dataset. Words with two characters are usually connectors in the Albanian language and contain the letter 'ë' in most cases. For this reason,

these words are also a good way of adapting the letter 'ë' in the training of the system.

During the generation of images of this dataset in the Albanian language we developed an algorithm to correct the diacritical points on the letter 'ë' and the inclusion of a line distinctive below the letter 'ç'. The algorithm calculates the coordinates of the location of each 'ë' and 'ç', [24]. Then it modifies the letters 'e' and 'c' because those are the letters that are different from the English alphabet which the algorithm uses to generate the image of handwriting.

Figure 4 illustrates the calculation of the coordinates where it is necessary to modify the image to convert the letter 'e' to look like 'ë'.

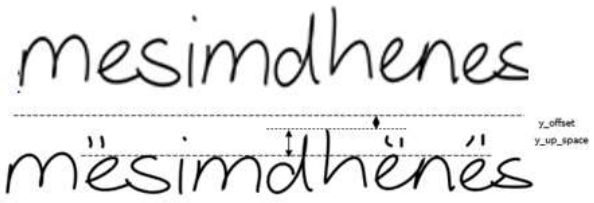

Fig. 4: Location and offset calculation for the 'ë' letter

4 Experimental Evaluation

The focus of this section is the evaluation of the effectiveness of the synthetic dataset within the modified Attention HTR, [15], [25] model designed for the symbol system of the Albanian language. We performed training and validation using three different datasets employing a case-sensitive model. Using those different datasets, we evaluate the accuracy rate, and compare the performances of each experiment.

The first dataset employed is the dataset generated using the deep-learning model, [26], adapted explicitly for the generation of word images in the Albanian language. In this dataset, we used approximately 2,000 words, [19], and using six different handwriting styles we generated a dataset with more than 10,000 records.

The second dataset is generated with approximately 16,000 unique words. The images of words are written in six different handwriting styles and this dataset offers the possibility to test the performance for lexical variety.

The third dataset is generated from the second dataset and the existing English IAM dataset, [15]. The selection of words from the dataset is random and contains 25% of its data. In this way, we generate a dataset with two linguistic areas, the Albanian and English IAM.

The evaluation of results provides accuracy analysis based on the formula as follows:

$$performance = \frac{correct\ results}{Total\ of\ words} * 100 \quad (1)$$

4.1 Attention HTR Model Trained with Different Size Datasets

The synthetic two datasets generated by deep learning with six different handwriting styles, [27], [28] are used in this experiment. This experiment aims to investigate the effects of dataset size and diversity of the words in the dataset on the system's performance. Both datasets encompass six distinct handwriting styles, each utilized across different paragraphs within the datasets. The 10,000-word dataset includes Albanian literary excerpts, while the second dataset comprises alternative words sourced, [29], from additional Albanian literature works. We trained the HTR model with each of these datasets and the results are shown in Table 1. The outcomes from training the attention HTR model with these datasets underscore the same pattern: *the larger the training dataset, the more effectively the system performs during training.*

Table 1. Performance comparison between datasets of different sizes

Dataset size	Train accuracy
10.000	83.2%
16.000	92.4%

4.2 Training the Attention HTR Model with Multilingual Datasets

In the second experiment, we used the entire dataset, consisting of 16,000 Albanian word instances, alongside the IAM dataset, [15], which includes English language instances. Given the dataset's volume, we randomly selected a subset of around 25,000-word instances for further analysis, [30]. The testing was performed on 30% of the data in both datasets. To evaluate the performance of the Albanian model with non-synthetic images, we tested it with both Albanian and English language test datasets.

The results shown in Table 2, present higher accuracy in English language text validation due to the size of the dataset and customized network weights. Nevertheless, the Albanian language model performed well, considering dataset limitations and training.

Regarding language testing, the English test dataset, comprising about 7,500 examples, produced satisfactory results. In contrast, the Albanian test dataset with 1,300 examples demonstrated a higher score than English.

Additionally, cross-testing with the opposite language datasets revealed challenges. English model couldn't train with the dataset created in Albanian by 16,000 words due to the presence of the distinct characters ‘ë’ and ‘ç’. Conversely, evaluating the Albanian model's performance using English data was feasible, given the subset nature of English characters within the Albanian language. However, these assessments yielded suboptimal outcomes. The model, primarily trained on synthetic images, struggled to handle human handwriting images from the English test cases and leaned towards recalibrating for synthetic handwriting. This highlights a key challenge addressed in the following experiment.

Table 2. Comparison of Albanian and English Language Datasets

Dataset	Train accuracy	Test accuracy with Albanian test cases	Test accuracy with English test cases
Albanian Language	92.4%	94.7%	2.7%
English Language	97.5%	-	83.1%

4.3 Training the Attention HTR Model with the Hybrid and Albanian Language Datasets

In this experiment, we addressed the issue identified in the previous study where the system trained with the Albanian language dataset, had difficulty accommodating the human handwriting images from the English model. To resolve this, [21], [24], [31], we explored creating a combined dataset that includes both Albanian and English. This approach allows the model to learn both languages, not only from synthetic images but also from real handwritten data. The results from training the model with a hybrid dataset, alongside the results from the model trained solely with the Albanian dataset, are presented in Table 3.

The results unambiguously reveal the enhanced performance of the hybrid model in predicting handwritten English words, [32], with a notable 66% improvement when compared to the performance of the Albanian language dataset in tests involving human handwriting images. Furthermore, the hybrid model achieves commendable accuracy in synthetic Albanian language tests.

After the analyses of data in the tests the performance decrease was minimal even when we used a hybrid dataset of 22,000 training data and about 3,100 instances of validation data.

Table 3. Comparative Analysis: Hybrid and Albanian Language Datasets

Dataset	Train accuracy	Test accuracy with Albanian test cases	Test accuracy with English test cases
Albanian Language	92.4%	94.7%	2.7%
Hybrid Dataset	82.1%	93.5%	68.4%

5 Conclusions

In this paper we explored a methodology to generate a synthetic dataset, mimicking human handwriting for texts in the Albanian language using deep learning techniques based on the English language. The difference of letters between Albanian and English alphabets is just two letters, exactly the letters 'ë' and 'ç' and those letters can be generated [21], [33], using the letters 'e' and 'c' by calculation of letter position and modifying them. The dataset trained a Handwritten Text Recognition (HTR) system. Furthermore, the trained model has been enhanced using a pre-trained English model, [21], showing promising results.

One of our main aims was to test the model's performance in various scenarios. Most importantly, we generated a hybrid dataset, including text in Albanian and English, aiming to develop a system that can recognize handwritten and synthetic texts.

The results achieved from the experiments confirmed our idea to establish an HTR system for the Albanian language based on a model for the English language was the right one.

Finally, we notice that our research presents initial efforts in developing hybrid HTR systems for different languages even if they are limited in resources, motivating us for further work in the field.

The developed models require optimization to be more accurate and effective, especially with different handwriting styles. Future work will be focused on the usage of larger nonsynthetic datasets, with different sources. The Albanians speak two dialects that have the same sentence structure, the Gheg and Tosk dialects. The first one is spoken mostly in the north of Albania, Kosovo, and the Albanian community in North Macedonia, and the Tosk dialect which is spoken in the south of Albania. The research on how to detect and analyze

based on dialectic elements will be part of our future work in this field.

Further research in this field will improve existing HTR systems, simplifying the implementation of those systems, and increasing their performance and accuracy rate.

References:

[1] Stefano Coretta, Josiane Riverin-Coutlée, Enkeleida Kapia, and Stephen Nichols. "Northern Tosk Albanian." *Journal of the International Phonetic Association,* vol.53, Issue no. 3, pp 1122–44, 2023, DOI: 10.1017/S0025100322000044.

[2] Dmitrijs Kass and Ekta Vats, "AttentionHTR, Handwritten Text Recognition Based on Attention Encoder-Decoder Networks", *Document Analysis Systems: 15th IAPR International Workshop, DAS 2022,* La Rochelle, France, pp 507–522, DOI: 10.1007/978-3-031-06555-2_34.

[3] Ray Smith Daria Antonova Dar-Shyang Lee, "Adapting the Tesseract Open-Source OCR Engine for Multilingual OCR", *The International Workshop on Multilingual OCR (2009)*, Barcelona, Spain, 2009, Article No.: 1, Pages 1–8, DOI: 10.1145/1577802.1577804.

[4] Minghao Li, Tengchao Lv, Jingye Chen, Lei Cui, Yijuan Lu, Dinei Florencio, Cha Zhang, Zhoujun Li, Furu Wei, "TrOCR: Transformer-based Optical Character Recognition with Pre-trained Models", *The Thirty-Seventh AAAI Conference on Artificial Intelligence*, Washington DC, USA, 2023, pp. 13094-13112, DOI: 10.48550/arXiv.2109.10282

[5] Bianne-Bernard, Anne-Laure and Menasri, Fares and Al-Hajj Mohamad, Rami and Mokbel, Chafic and Kermorvant, Christopher and Likforman-Sulem, Laurence, "Dynamic and contextual information in hmm modeling for handwritten word recognition", *IEEE transactions on pattern analysis and machine intelligence*, vol. 33, no. 10, 2066– 2080, 2011. DOI: 10.1109/TPAMI.2011.22

[6] Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun. "Deep Residual Learning for Image Recognition.", *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* Las Vegas, NV, USA, pp. 770-778, 2016, DOI: 10.1109/CVPR.2016.90

[7] Kartik Dutta, Praveen Krishnan, Minesh Mathew and. Jawahar C. V, "Improving CNN-RNN Hybrid Networks for Handwriting Recognition," *16th International Conference on Frontiers in Handwriting Recognition (ICFHR)*, Niagara Falls, NY, USA, 2018, pp. 80-85, DOI: 10.1109/ICFHR-2018.2018.00023.

[8] Sepp Hochreiter and Jürgen Schmidhuber, "Long Short-Term Memory." *Neural Computation*, vol. 9, pp. 1735-1780, 1997, DOI: 10.1162/neco.1997.9.8.1735.

[9] Mike Schuster and Kuldip Paliwal, "Bidirectional Recurrent Neural Networks." *Signal Processing, IEEE Transactions*, vol. 45, pp. 2673 – 2681, 1997, DOI: 10.1109/78.650093.

[10] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser and Illia Polosukhin, "Attention Is All You Need.", *31st Conference on Neural Information Processing Systems (NIPS 2017),* Long Beach, CA, USA, 2017, DOI: 10.48550/arXiv.1706.03762.

[11] Alex Graves, "Generating Sequences with Recurrent Neural Networks", *ArXiv*, vol. abs/1308.0850, 2014.

[12] Karen Simonyan, Andrew Zisserman "Very Deep Convolutional Networks for Large-Scale Image Recognition", *3rd International Conference on Learning Representations, {ICLR} 2015,* San Diego, CA, USA, 2015, abs/1409.1556.

[13] Tao Wang, David J. Wu, Adam Coates, Andrew Y. Ng. "End-to-End Text Recognition with Convolutional Neural Networks", *21st International Conference on Pattern Recognition (ICPR2012),* Tsukuba, Japan, 2012, pp. 3304-3308.

[14] Rakesh Kumar Mandal, N. R. Manna, "Handwritten English Character Recognition Using Column-wise Segmentation of Image Matrix (CSIM)", *WSEAS Transactions on Computers*, vol. 11, pp.148-158, 2012.

[15] Urs Victor Marti and H. Bunke, "The iam-database: an English sentence database for offline handwriting recognition". *International Journal on Document Analysis and Recognition* vol. 5, no. 1, pp. 39–46, 2002, DOI:10.1007/s100320200071.

[16] Aiquan Yuan, Gang Bai, Lijing Jiao, and Yajie Liu, "Offline handwritten English character recognition based on convolutional neural network", *10th IAPR International Workshop on Document Analysis Systems, DAS 2012*, Washington, DC United States 2012. DOI: 10.1109/DAS.2012.61.

[17] Ioannis Giachos, Eleni Batzaki, Evangelos C. Papakitsos, Michail Papoutsidakis, Nikolaos Laskaris, "Developing a Natural Language Understanding System for Dealing with the Sequencing Problem in Simulating Brain Damage", *WSEAS Transactions on Biology and Biomedicine*, vol. 21, pp. 138-147, 2024, https://doi.org/10.37394/23208.2024.21.14.

[18] Feng Li, Chenxi Cui, Yashi Hu, Lingling Wang, "Sentiment Analysis of User Comment Text based on LSTM," *WSEAS Transactions on Signal Processing*, 2023, vol. 19, pp. 19-31, https://doi.org/10.37394/232014.2023.19.3.

[19] Max Jaderberg, Karen Simonyan, Andrea Vedaldi, and Andrew Zisserman. "Synthetic data and artificial neural networks for natural scene text recognition.", *The Workshop on Deep Learning, NIPS*, Montréal 2014, DOI: 10.48550/arXiv.1406.2227.

[20] Ankush Gupta, Andrea Vedaldi, Andrew Zisserman, "Synthetic data for text localization in natural images", *IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV, USA 2016, pp. 2315–2324, DOI:10.1109/CVPR.2016.254.

[21] Hoo-Chang Shin, Holger R. Roth, Mingchen Gao, Le Lu, Ziyue Xu, Isabella Nogues, Jianhua Yao, Daniel Mollura, and Ronald M. Summers, "Deep Convolutional Neural Networks for Computer-Aided Detection: CNN Architectures, Dataset Characteristics and Transfer Learning", *IEEE Transactions on Medical Imaging*, vol. 35, pp. 1285-1298, 2016, DOI:10.1109/TMI.2016.2528162.

[22] In-Jung Kim, and Xiaohui Xie, "Handwritten Hangul recognition using deep convolutional neural networks", *International Journal on Document Analysis and Recognition (IJDAR)*, vol.18, pp. 1-3, 2015, DOI:10.1007/s10032-014-0229-4.

[23] Ali Asghar, Leghari Mehwish, Hakro Dil, Awan Shafique, Jalbani Dr, Pakistan Nawabshah, "A Novel Approach for Online Sindhi Handwritten Word Recognition using Neural Network". *Sindh University Research Journal SURJ (Science Series),* Vol. 48(1), pp. 213-216, 2016.

[24] Yudong Liang, Jinjun Wang, Sanping Zhou, Yihong Gong, and Namming Zheng, "Incorporating image priors with deep convolutional neural networks for image super resolution", *Neurocomputing*, vol. 194, pp. 340-347, 2016, DOI: 10.1016/j.neucom.2016.02.046.

[25] I. Khandokar, Mokhtar M. Hasan, Ferda Ernawan, Saiful Islam, and Muhammad Nomani Kabir, "Handwritten Text Recognition Using Convolutional Neural Network*", Journal of Physics: Conference Series*, 2021, volume 1918, no. 4, DOI: 10.1088/1742-6596/1918/4/042152.

[26] Chowdhury, Arindam and Lovekesh Vig. "An Efficient End-to-End Neural Model for Handwritten Text Recognition." *British Machine Vision Conference,* Newcastle, England, 2018.

[27] Ahmed El-Sawy, Mohamed Loey, Hazem EL-Bakry, "Arabic Handwritten Characters Recognition Using Convolutional Neural Network," *WSEAS Transactions on Computer Research*, vol. 5, pp. 11-19, 2017.

[28] Amin Al Ka'Bi, "A Proposed Artificial Intelligence Algorithm for Development of Higher Education", *WSEAS Transactions on Computers,* vol. 22, pp. 7-12, 2023, https://doi.org/10.37394/23205.2023.22.2.

[29] Ritesh Sarkhel, Nibaran Das, Amin K. Saha, and Mita Nasipuri, "A multi-objective approach towards cost effective isolated handwritten Bangla character and digit recognition", *Pattern Recognition*, vol. 58, pp. 172-189, 2016, DOI: 10.1016/j.patcog.2016.04.010.

[30] Manmatha, R. and Srimal, N., n.d. "Scale Space Technique for Word Segmentation in Handwritten Documents". *Lecture Notes in Computer Science*, vol 1682, pp. 22–33, Greece 1999, DOI: 10.1007/3-540-48236-9_3.

[31] Jeonghun Baek, Geewook Kim, Junyeop Lee, Sungrae Park, Dongyoon Han, Sangdoo Yun, Seong Joon Oh, Hwalsuk Lee, "What is wrong with scene text recognition model comparisons? dataset and model analysis", *IEEE International Conference on Computer Vision*, Seoul, Korea, 2019, pp. 4715–4723, DOI: 10.1109/ICCV.2019.00481.

[32] Jemimah K, "Recognition of Handwritten Characters based on Deep Learning with TensorFlow", *International Research Journal of Engineering and Technology (IRJET)*, vol. 6, Issue: 09, pp 1164-1165, 2019.

[33] Chunpeng Wu, Wei Fan, Yuan He, Jun Sun, and Satoshi Naoi, "Handwritten Character Recognition by Alternately Trained Relaxation Convolutional Neural Network", *14th International Conference on Frontiers in Handwriting Recognition, ICFHR*, Allen, TX, USA, 2014, DOI: 10.1109/ICFHR.2014.56.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

- Hakik Paci developed the algorithm for image processing, prepared the structure of the article, and the contributed to write the main part of the article.
- Evis Trandafili reviewed the related work and focused on linguistic for the implementation of the HTR in Albania language.
- Dorian Minarolli prepared the infrastructure and the dataset used in simulation.
- Stela Paturi is a student who has organized and executed the experiments while preparing her diploma thesis.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The authors have no conflicts of interest to declare.