Gazi Üniversitesi
Fen Bilimleri Dergisi
PART C: TASARIM VE TEKNOLOJİ

Gazi University
Journal of Science
PART C: DESIGN AND TECHNOLOGY

GU J Sci, Part C, 11(3): XX-XX (2023)

Recognition of Online Turkish Handwriting using Transfer Learning

Esma Fatıma BİLGİN TAŞDEMİR[1*]

[1]*Istanbul Medeniyet University, Faculty of Arts and Humanities, Department of Information and Document Management, Istanbul, Turkey*

Article Info

Research article
Received: 06/07/2022
Revision: 21/12/2022
Accepted: 21/12/2022

Keywords

Online Handwriting Recognition
Deep Learning
Transfer Learning,
Turkish Handwriting Recognition CNN-BLSTM

Makale Bilgisi

Araştırma makalesi
Başvuru: 06/07/2022
Düzeltme: 21/12/2022
Kabul: 21/12/2022

Anahtar Kelimeler

Çevrimiçi El Yazısı Tanıma
Derin Öğrenme
Transfer Öğrenme
Türkçe El Yazısı Tanıma
CNN-BLSTM

Graphical/Tabular Abstract (Grafik Özet)

Using Transfer learning for training a Deep Neural Network helps overcoming data scarcity problem in recognition of Turkish online handwritings. This work propose a CNN-BLSTM network which is pretrained with a large English dataset and fine-tuned with a small Turkish dataset. / Türkçe çevrimiçi el yazılarının tanınmasında kullanılacak derin bir Yapay Sinir Ağını eğitmek için Transfer öğrenimini kullanmak, veri kıtlığı sorununa bir çözüm sunabilir. Bu çalışma, önce büyük bir İngilizce veri seti ile eğitilmiş bir CNN-BLSTM ağına küçük bir Türkçe veri seti ile ince ayar yapılmasını önermektedir.

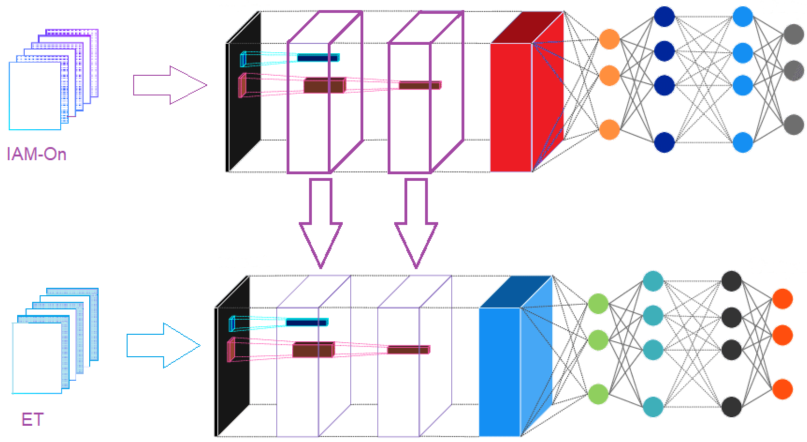

***Figure A**: Transfer Learning on a CNN-BLSTM Network /**Şekil A:**.Bir CNN-BLSTM Ağında Transfer Öğrenme*

Highlights (Önemli noktalar)

- *The Transfer Learning with fine tuning technique provides a solution to the data scarcity problem of Turkish handwriting recognition. / İnce ayarlı Transfer Öğrenme tekniği Türkçe el yazısı tanıma problemi için veri azlığı sorununa bir çözüm sunmaktadır.*
- *IAM-On dataset English handwriting dataset can be used for pretraining of a CNN-BLSTM network. / İngilizce el yazısı veri seti IAM-On , bir CNN-BLSTM ağının ön eğitimi için kullanılmıştır.*
- *Fine tuning the system with Turkish samples from ET dataset increases recognition accuracy on 2,041 Turkish samples from 49% to 85% / Ön eğitimli sistemi Türkçe ET veri seti ile ince ayar yaparak karakter tanıma başarısı %49'dan %85'e çıkarılmıştır.*

Aim (Amaç): *To develop an online handwriting recognition system for Turkish / Türkçe çevrimiçi el yazısı tanıma sistemi geliştirmek.*

Originality (Özgünlük): *The first work in the literatıre to use Transfer Learning for online Turkish handwriting recognition / Literatürdeki ilk Transfer Öğrenme kullanılan Türkçe çevrimiçi el yazısı tanıma sistemi .*

Results (Bulgular): *The proposed system achieves 85% character recognition accuracy on ET dataset / Geliştirilen sistem ET veri seti üzerinde %85 karakter tanıma başarısı göstermiştir.*

Conclusion (Sonuç): *The data scarcity of Turkish handwriting recognition problem can be overcome by using The Transfer Learning with fine tuning technique / Türkçe el yazısı tanıma problemi için veri azlığı sorunu İnce ayarlı Transfer Öğrenme tekniği ile çözülebilmektedir.*

*Corresponding author, e-mail: esmabilgin.tasdemir@medeniyet.edu.tr
DOI: 10.29109/gujsc.1141508

GU J Sci, Part C, 11(3): XX-XX (2023)

Gazi Üniversitesi
Fen Bilimleri Dergisi
PART C: TASARIM VE TEKNOLOJİ

Gazi University
Journal of Science
PART C: DESIGN AND TECHNOLOGY

http://dergipark.gov.tr/gujsc

Recognition of Online Turkish Handwriting using Transfer Learning

Esma Fatıma BİLGİN TAŞDEMİR[1]

[1]*Istanbul Medeniyet University, Faculty of Arts and Humanities, Department of Information and Document Management, Istanbul, Turkey*

Article Info

Research article
Received: 06/07/2022
Revision: 21/12/2022
Accepted: 21/12/2022

Keywords

Online Handwriting Recognition
Deep Learning
Transfer Learning,
Turkish Handwriting Recognition CNN-BLSTM

Abstract

We present a recognition system for online Turkish handwriting using transfer learning. Training deep networks requires large amounts of data. Since such a sufficiently large collection of Turkish handwriting samples is not available, so we adopt the transfer learning approach and train and optimize a CNN-BLSTM recognition system first using the standard IAM-On dataset of English handwriting. Then, we fine tune it with Turkish handwriting samples from a smaller dataset. Fine tuning increases the character recognition rate of the final system which is evaluated on 2,041 samples of isolated Turkish words from the initial value of 49% to 85%. The results show that transfer learning can be a solution to the data scarcity problem of online Turkish handwriting.

Transfer Öğrenmesi Kullanarak Çevrimiçi Türkçe El Yazısı Tanıma

Makale Bilgisi

Araştırma makalesi
Başvuru: 06/07/2022
Düzeltme: 21/12/2022
Kabul: 21/12/2022

Anahtar Kelimeler

Çevrimiçi El Yazısı Tanıma
Derin Öğrenme
Transfer Öğrenme,
Türkçe El Yazısı Tanıma
CNN-BLSTM

Öz

Bu çalışmada, Türkçe çevrimiçi el yazısı için geliştirilen bir tanıma sistemi anlatılmaktadır. Çok katmanlı Yapay Sinir Ağlarını eğitmek için gereken çok miktardaki verinin temini amacıyla Transfer Öğrenme yöntemi kullanılarak bir CNN-BLSTM ağı eğitilmiştir. Bu amaçla ilk olarak standart veri setlerinden olan IAM-On İngilizce el yazısı veri seti ile eğitilen modele, ardından daha küçük boyutlu bir Türkçe el yazısı veri seti ile ince ayar yapılmıştır. 2,041 Türkçe kelime üzerinde yapılan testlerde, ince ayar sayesinde karakter tanıma oranının %49 dan %85 e yükseldiği görülmüştür. Çalışmada elde edilen sonuçlara göre Transfer Öğrenme yöntemi Türkçe el yazısı tanıma problemi için uygun bir çözüm sayılabilir.

1. INTRODUCTION (GİRİŞ)

Handwriting recognition is an active research area where symbols, characters, words or lines of words written by human writers are recognized by computer systems. Online handwriting is a digital form of handwriting generated by a pen tip moving on a special digitizer surface. Traces of the movements are represented as a time series of coordinates. Timing of the movements, pressure of the pen and pen status as up or down can be captured along with the coordinates. In contrast, offline handwriting is a modality where handwritten text is represented as image data. Increasing use of mobile devices, interactive whiteboards and other handwriting capturing devices make online handwriting a daily modality to a wider audience. Although the performance of recognition systems is improving steadily, there are still many challenges to be faced before declaring the online handwriting recognition problem to be solved. In the earlier years, much of this research was focused on the recognition of derivations of Latin alphabet, and especially of English, but other scripts started to gain attention in recent years [1], [2], [3], [4]. However, research on Turkish script is still very limited. This is mostly due to lack of sufficiently large datasets for both online and offline Turkish handwriting recognition. A possible solution to the problem of insufficient training data can be Transfer Learning where data from a similar domain is used for pretrainign a system and then fine-tuning the

*Corresponding author, e-mail: esmabilgin.tasdemir@medeniyet.edu.tr
DOI: 10.29109/gujsc.1141508

system with available scare data of the target domain. In this work, we develop a recognition system for online Turkish handwritten words. We use the Transfer Learning approach to overcome the data scarcity problem of Turkish handwriting samples and to achieve better recognition accuracy. The system obtains recognition rates comparable to state-of-art using some Deep Learning (DL) techniques like Convolutional Neural Networks (CNN) and Bidirectional Long-Short Memory Networks (BLSTM).

2. RELATED WORK (İLİŞKİLİ ÇALIŞMALAR)

Emergence of studies on online handwriting recognition dates backto the 1990s [1]. Starting from isolated characters and symbols, recognition systems working at word, line and even paragraph levels have been developed in the course of time [5], [6], [7]. Although Latin-alphabet based scripts have been getting much of the attention, the number of studies on recognition of many other writing systems increased recently [8], [2], [3], [42]. 2 Different machine learning techniques like Hidden Markov Models (HMM), Support Vector Machines (SVM) and Artificial Neural Networks (ANN) and their combinations are employed in online handwritten text recognition systems in the literature [9], [10], [11], [12], [13], [14], [15], [16]. HMM based recognizers have been particularly very popular due to their capability of modelling time series effectively [17], [18], [19], [20]. Recognition performances improved dramatically by introduction of the Deep Learning methods especially in problems where a large amount of training data is available [21], [22], [23]. Yet, HMM based systems are still viable in cases of limited data and computational resources [24] and particular scripts like Arabic [25], [26]. Recurrent Neural Networks and their variants are successfully used for tasks regarding sequential data like online handwriting or speech where data is represented as time series [27], [28], [41]. Addition of memory units to RNNs solved the so-called "vanishing gradient problem" and improved their capability of modeling temporal dependencies in data. In particular, long shortterm memory neural networks (LSTMs) and their variants have been very successful in both online and offline handwritten and machine printed text recognition problems in recent years [7], [21], [22]. Research about recognition of handwritten Turkish text is very limited. There are studies on offline Turkish character recognition with some constraints applied on the style or the case of writing [29], [30], [31]. In [32], a HMM system which was previously developed for English, is used for offline handwritten Turkish text recognition. Some characters with dots, cedilla and breve which are specific to the Turkish alphabet are mapped to their English counterparts. A Turkish prefix parser to detect non-Turkish word prefixes during decoding is employed instead of using a lexicon. A 56% top-10 word recognition rate is reported using a 17,000-word lexicon. Recognition accuracy decreases to around 40% when the Turkish prefix parser is used. [33] uses HMMs in a character-based word recognition system for offline lowercase mixed-style handwritten Turkish words. The reported recognition rate is 84% using a lexicon of size 2,500. [31] proposes a machine printed character recognizer developed using ANNs. The character recognition rate is reported as 95.2% for a proprietary dataset. Another character recognizer for offline Turkish Handwriting is proposed in [29]. Using a classifier based on Size-Dependent Negative Log-Likelihood, a recognition rate of 93.4% is achieved on a test set of 6,322 samples. A CNN based recognizer is proposed for Turkish handwritten character recognition in [34]. Its recognition accuracy is reported as 96.07%. In a recent work [34], a CNN-based system is trained to recognize Turkish handwritten characters and achieved 96.0% recognition accuracy on a test set of 5165 samples. [35] presents a comprehensive evaluation of various HMM architectures and parameters for online handwriting recognition tasks. A word recognition rate of 94% is achieved using a 1,000-word lexicon with character HMMs. Another HMM system is proposed in [24] where the data scarcity problem is overcome by using a larger English dataset along with a Turkish dataset containing words taken from elementary school textbooks. 91.7% word recognition accuracy is reported for a middle-sized, 1,950-word lexicon task and 800 test samples. When the lexicon size is increased to 12,500, recognition accuracy is measured as 67.9%, using a bi-gram language model based on word stems and suffixes. In a recent study, a CNN-BLSTM network which is pre-trained with a synthetic dataset and fine-tuned with the Turkish dataset used in [24] achieved 88% character recognition accuracy in an open dictionary recognition task on that Turkish dataset [40].

3. METHODOLOGY (METODOLOJİ)

3.1. Datasets (VERİSETLERİ)

There are several online handwriting datasets which are publicly available.

The IAM 3 On-Line Handwriting Database (IAM-On DB) is a large online handwriting dataset containing forms of handwritten English text acquired using a special system which traces movements of pen on a whiteboard [36]. It is used to train and test handwritten text recognizers and to perform writer identification and verification experiments in many studies [23], [7], [22], [37]. There are 13,049 lines written by 221 writers in the dataset. It contains 86,272 word instances from a 11,059-word dictionary. Pen coordinates, timing and pen-up/down status for each stroke of each sample are stored in XML files. Ground-truth for each sample is given in a text file separately. Elementary Turkish (ET) dataset is a collection of around 10,000 isolated words written by 113 writers, including children [24]. Words are selected from a 2,089-word lexicon derived from 1st and 2nd Grade Turkish textbooks. It is split into three sets where writers are not overlapping. The train set contains 7,360 samples from a 1956-word lexicon by 79 writers whereas the test set contains 2,500 samples from a 2089-word lexicon written by 34 writers. The train set lexicon covers the test set lexicon. Each sample is represented with its coordinates and pen status as up or down in the ET dataset.

3.2. Data (VERİ)

Each sample of the IAM-On is made of a line composed of strokes. Each stroke is represented with a collection of points. Points of a stroke are defined with their x and y coordinates, creation time with respect to the creation time of the first point of that stroke and pen status which always takes the up value except for the last point of that stroke, which takes the down value. We apply some basic preprocessing to eliminate variations in data. First we use linear regression to detect the baseline of a sample and de-skew line by rotating according to baseline angle. Points are normalized by subtracting mean coordinates values and division by standard deviation of y coordinates. We resample strokes to obtain equi-distant points and finally up-sample strokes which are shorter than a threshold. After the preprocessing step we use the following elements to represent points in a stroke:

- x- and y-coordinates,
- differences from the x- and y-coordinates of the next point
- pen status.

3.3. Recognition System (TANIMA SİSTEMİ)

We propose a CNN-BLSTM network for recognition of online Turkish handwriting. Hybrid systems where a CNN is used for feature extraction and a RNN or its variant is used as a classifier are widely used for sequence learning tasks. We take the same approach and use a CNN-BLSTM network for recognition of the online Turkish handwriting. Since size of the ET dataset is not large enough to train a deep network, we use transfer learning approach to make use of data from a larger dataset i.e. IAM-On. Transfer learning (TL) is a technique for exploiting information obtained during a learning process to improve another learner from a different but related domain. TL is preferred especially in cases where training data is limited. TL has been in use for various problems like text sentiment classification, image classification, and text recognition [38], [39]. Once a DL network is trained with data from a related domain, TL can be applied by freezing some of the layers of which knowledge is accumulated as weights. Freezing prevents update of the weights. Other layers which will continue learning canbe reset or keep learned weights before the network is trained with new data from the target domain. Another method is to fine tune the system with the new data. Here, weights of all layers are updated according to the information 4 learned from the new data. When the domains are similar, fine tuning becomes a more preferable method.

4. EXPERIMENTS (DENEYLER)

4.1. Experimental setup (DENEY ORTAMI)

In [21], IAM-On dataset is split into four subsets as a training set, a test set, a validation set to be used during training (validation-1), and a validation set to be used for language modeling (validation-2). However, the names of samples in each subset are not published, only information about sizes of subsets are available. So, we split the IAM-On dataset randomly according to given subset sizes. As for the ET dataset, we use approximately 65% of the data in transfer learning and fine tuning. 10% of the samples is used for validation during training and 25% of the samples is used to evaluate the final system. Table 1 shows our dataset split sizes for the two datasets.

Table 1. Subset sizes after split of the IAM-On and ET datasets (IAM-On ve ET veri setlerinin bölümlenmesi sonrası büyüklükleri)

Dataset	Train	Test	Val.-1	Val.-2
IAM-On (lines)	5,364	3,859	1,438	1,518
ET (words)	5,124	2,041	1,000	-

Performance evaluation metrics of the proposed system are Character Error Rate (CER) and Word Error Rate (WER) percentages that are based on the Edit Distance. Edit distance is calculated as the minimum number of edits with substitution, insertion and deletion of characters from the reference string to the output, normalized by the number of reference characters. WER is computed in a similar way.

We conducted a series of experiments with different network architectures and parameters to find the system with highest recognition accuracy on the IAM-On testset. Accordingly, our network and its parameters are decided as explained below.

The network architecture contains two parts; i) a CNN network to extract features from the data, ii) a BLSTM network to classify the features and to recognize the text. The feature extraction network is made of two blocks containing three Convolutional layers each. Batch Normalization is applied after each Convolutional layer. After each block comes an average pooling layer to downsample data representation using a window size of 2. The classifier network has four BLSTM networks stacked together. Each BLSTM layer has one forwards working and one backward working LSTM layer outputs of which are concatenated. Finally, a softmax layer calculates the probability of each symbol in the recognition alphabet. There are 83 unique symbols used in IAM-On samples. The ET dataset contains Turkish letters, upper and lower case and two more symbols. We merge the two symbol sets to obtain a recognition alphabet of 97 symbols. CTC loss is calculated using predicted labels from the softmax layer and the ground truth labels. Activation function is ReLU for all trainable layers of the system. Kernels of layers are started using the He uniform variance scaling initializer.

Details of the network layers are presented in Table 2.

Table 2. The network architecture (Ağ mimarisi)

Layer type	Kernel size	# Kernels	# Units	Pool Size
Input	-	-	-	-
Conv1D	7	60	-	-
Batch Norm	-	-	-	-
Conv1D	5	90	-	-
Batch Norm	-	-	-	-
Conv1D	5	120	-	-
Batch Norm	-	-	-	-
Average Pooling	-	-	-	2
Conv1D	3	120	-	-
Batch Norm	-	-	-	-
Conv1D	3	160	-	-
Batch Norm	-	-	-	-
Conv1D	3	200	-	-
Batch Norm	-	-	-	-
Average Pooling	-	-	-	2
BLSTM	-	-	60	-
BLSTM	-	-	60	-
BLSTM	-	-	60	-
BLSTM	-	-	60	-
Batch Norm	-	-	-	-
Dense	-	-	93	-

The network is implemented using the TensorFlow libraries whereas the experiments are run on a NVIDIA GeForce RTX 2060 graphical processing unit (GPU) card. The network is trained to minimize the CTC loss function. We use the RMSprop optimization algorithm with an initial learning rate of 0.001 and mini batches of size 20. Training is stopped when the loss on the validation set does not improve after 20 consecutive epochs.

4.2. Experiments (Deneyler)

After the network architecture and its parameters are decided, we apply the transfer learning technique by fine tuning the trained model to obtain a recognition system for the ET dataset.

As it is explained in Section 4.1, IAM-On dataset has two validation subsets one of which is reserved for language modeling originally. We first train the network with only the training set. Then, we add the validation-2 set to the training set and obtain an extended training set containing 6,882 line samples. We report results from evaluation of these systems on both the IAM-On and ET test sets in Table 3. According to the results, the proposed system obtains 85% character recognition accuracy on IAM-On test set.Its performance on the ET test set before the fine tuning process is measured as 44% character recognition accuracy. Training with more data increases recognition rate by 3.5% for IAM-On samples by reaching 88%. Similarly, it has a positive effect on recognition rate of the ET samples.

Table 3. Recognition accuracies of systems trained with the IAM-On dataset (IAM-On ile eğitilen sistemlerin doğruluk değerleri)

Train set	Test set	Recognition errror rate	
		CER	WER
IAM-On train s et	IAM-On test set	0.15	0.49
IAM-On train s et + validation-2 set	IAM-On test set	0.12	0.41
IAM-On train s et	ET test set	0.56	1.20
IAM-On train s et + validation-2 set	ET test set	0.51	1.16

We run two more experiments where a trained network is fine-tuned with the ET training samples. After completion of the transfer learning part, the final system is evaluated on the ET test set. Results of that evaluation are presented in Table 4.

Table 4. Results from the transfer learning experiments (Transfer Öğrenme deneylerinin sonuçları)

Train set	Test set	Recognition errror rate	
		CER	WER
IAM-On train set	ET test set	0.16	0.57
IAM-On train set + validation-2 set	ET test set	0.15	0.54

After applying fine tuning to the system which is previously trained with the IAM-On extended training set, recognition rate of the ET samples is measured as 85%. The other system which is trained with only the IAM-On training set achieves almost the same accuracy.

5. DISCUSSION

According to results from Table 3, training with more data brings performance improvement on IAM-On test set, which is an expected situation. However, using the extended set has almost no effect on recognition rate of the ET test samples. Based on that, we can deduce the ET samples are significantly different from the IAM-On samples.

After the fine tuning is applied, recognition rate of the ET samples increases significantly for both networks. Actually, CER of fine-tuned networks on the ET are almost equal to CER measured in recognition of IAM-On samples by the initial networks of Table 3. Based on these, we can say that the fine tuning approach successfully increases the system performance.

Much of the errors are due to confusion of visually similar characters or character groups like k-h, r-n, ın-m, P-D. Another observation we make from the output of the recognizers is that Turkish letters that do not exist in the English alphabet are recognized considerably well. Even the least frequent characters like ü and ö are recognized correctly in many samples.

The difference between CER and WER percentages which is observed in all of the experiments indicates that only a few characters are misrecognized in each misrecognized word.

There are several works using the IAM-On dataset for handwriting recognition in the literature. In [21], Graves et al. report 11.5% CER on the IAM-On test set using a BLSTM network with a CTC layer. Frinken et al. obtain 12.3% CER using a bi-gram language model [37], using a similar deep architecture. In [7], Carbune et al. obtain a record-low CER of 4.0% by using a BLSTM-CTC network along with a 7-gram language model and some handcrafted features and heuristics. Comparing our results (12.0% CER) with the literature, we can say that the proposed system is on par with the state-of-the-art. The positive impact of using a CNN network to extract better features can be deduced, especially when the lack of a language model is considered.

As for recognition of online Turkish handwriting, [24] trains a system and evaluates it using the ET dataset. It reports 91.7% word recognition accuracy for a middle-sized, 1,950-word lexicon task and 800-sample test set. Our work is not directly comparable to that one since it is a closed-dictionary task (lexicon size 1,956) tested on a different, smaller portion of the data.

Taking an approach similar to the method proposed in this work, a CNN-BLSTM network, which is pre-trained with a synthetic dataset and fine-tuned with the ET train samples, achieved 12% CER and 44% WER on ET test set in [40]. Our results are approximately equal to those last results while we use a less complex method.

6. CONCLUSION

This work presents an online handwriting recognition system for Turkish. The system achieves state-of-art results comparable to those published in the literature. We train CNN-BLSTM networks using the standard IAM-On dataset. Later we adopt a transfer learning approach via a fine tuning process where already trained networks are further trained with Turkish handwriting samples. Fine tuning increases recognition accuracy of the final system on 2,041 Turkish samples from 49% to 85%. We conclude that the fine tuning technique provides a solution to the data scarcity problem of Turkish handwriting recognition. Using a language model, training with even larger data and post-processing the results can improve the system performance dramatically. Using data augmentation techniques can be useful as well.

DECLARATION OF ETHICAL STANDARDS
(ETİK STANDARTLARIN BEYANI)

The author of this article declares that the materials and methods they use in their work do not require ethical committee approval and/or legal-specific permission.

Bu makalenin yazarı çalışmalarında kullandıkları materyal ve yöntemlerin etik kurul izni ve/veya yasal-özel bir izin gerektirmediğini beyan ederler.

AUTHORS' CONTRIBUTIONS (YAZARLARIN KATKILARI)

***Esma FAtıma BİLGİN TAŞDEMİR*:** He conducted the experiments, analyzed the results and performed the writing process.

Deneyleri yapmış, sonuçlarını analiz etmiş ve makalenin yazım işlemini gerçekleştirmiştir.

CONFLICT OF INTEREST (ÇIKAR ÇATIŞMASI)

There is no conflict of interest in this study.

Bu çalışmada herhangi bir çıkar çatışması yoktur.

REFERENCES (KAYNAKLAR)

[1] R. Plamondon and S. N. Srihari, "On-line and off-line handwriting recognition: A comprehensive survey", IEEE Trans. Pattern Anal. Mach.Intell., vol. 22, no. 1, pp. 63–84, 2000.
[2] B. M. Al-Helali and S. A. Mahmoud, "Arabic online handwriting recognition (AOHR): A survey", ACM Comput. Surv., vol. 50, no. c3, pp. 33:1–33:35, 2017.

[3] N. Tagougui, M. Kherallah, and A. M. Alimi, "Online Arabic handwriting recognition: a survey", IJDAR, vol. 16, no. 3, pp. 209–226, 2013.

[4] D. S. Doermann and S. Jaeger, Eds., Arabic and Chinese handwriting recognition - SACH 2006 Summit College Park, MD, USA, September 27-28, 2006 Selected Papers, ser. Lecture Notes in Computer Science, vol. 4768. Springer, 2008.

[5] A. Priya, S. Mishra, S. Raj, S. Mandal, and S. Datta, "Online and offline character recognition: A survey", in 2016 International Conference on Communication and Signal Processing (ICCSP), 2016, pp. 0967–0970.

[6] M. Liwicki and H. Bunke, "Handwriting recognition of whiteboard notes", In Proceedings of the 12th Conference of the International Graphonomics Society, 2005, pp. 118–122.

[7] V. Carbune, P. Gonnet, T. Deselaers, H. A. Rowley, A. N. Daryin, M. Calvo, L. Wang, D. Keysers, S. Feuz, and P. Gervais, "Fast multi-language lstm-based online handwriting recognition", Int. J. DocumentAnal. Recognit., vol. 23, no. 2, pp. 89–102, 2020.

[8] X.-Y. Zhang, Y. Bengio, and C.-L. Liu, "Online and offline handwritten Chinese character recognition: A comprehensive study and new bench-mark", Pattern Recognition, vol. 61, pp. 348–360, 2017.

[9] S. Jager, S. Manke, J. Reichert, and A. Waibel, "Online handwriting recognition: the NPen++ recognizer", IJDAR, vol. 3, no. 3, pp. 169–180, 2001.

[10] S. Garcia-Salicetti, B. Dorizzi, P. Gallinari, and Z. Wimmer, "Maximum Mutual information training for an online neural predictive handwritten word recognition system", IJDAR, vol. 4, no. 1, pp. 56–68, 2001.

[11] E. Caillault and C. Viard-Gaudin, "Mixed discriminant training of hybrid ANN/HMM systems for online handwritten word recognition", IJPRAI,vol. 21, no. 1, pp. 117–134, 2007.
12] T. M. T. Do and T. Arti`eres, "Maximum margin training of gaussian HMMs for handwriting recognition", in 10th International Conference On Document Analysis and Recognition, ICDAR 2009, Barcelona, Spain, 26-29 July 2009, 2009, pp. 976–980.

[13] J. Schenk and G. Rigoll, "Novel hybrid NN/HMM modelling techniques for on-line handwriting recognition", in 10th International Workshop on Frontiers in Handwriting Recognition, IWFHR 2006, IAPR. , La Baule,France, Oct 2006, 2006, pp. 619—6230.

[14] N. Gauthier, T. Artieres, P. Gallinari, and B. Dorizzi, "Strategies for combining on-line and off-line information in an on-line handwriting recognition system", in 6th International Conference on Document Analysis and Recognition, ICDAR 2001, 10-13 September 2001, Seattle,WA, USA, 2001, pp. 412–416.

[15] S. Marukatat, T. Artieres, P. Gallinari, and B. Dorizzi, "Sentence recognition through hybrid neuro-markovian modeling", in 6th International Conference on Document Analysis and Recognition, ICDAR 2001, 10-13 September 2001, Seattle, WA, USA, 2001, pp. 731–737.

[16] M. Schenkel, I. Guyon, and D. Henderson, "On-line cursive script recognition using time-delay neural networks and Hidden Markov Models", Mach. Vis. Appl., vol. 8, no. 4, pp. 215–223, 1995.

[17] J. Hu, S. G. Lim, and M. K. Brown, "Writer independent on-line handwriting recognition using an HMM approach", Pattern Recognition,vol. 33, no. 1, pp. 133–147, 2000.

[18] A. Biem, "Minimum classification error training for online handwriting recognition", IEEE Trans. Pattern Anal. Mach. Intell., vol. 28, no. 7, pp.1041–1051, 2006.

[19] M. Liwicki and H. Bunke, "IAM-OnDB - an on-line English sentence database acquired from handwritten text on a whiteboard", in Eighth International Conference on Document Analysis and Recognition, ICDAR 2005, 29 August - 1 September 2005, Seoul, Korea, 2005, pp. 956–961.

[20] M. Liwicki, A. Graves, H. Bunke, and J. Schmidhuber, "A novel approach to on-line handwriting recognition based on bidirectional long short-term memory networks", in Proceedings of the 9th International Conference on Document Analysis and Recognition, ICDAR, 2007, pp.367–371.

[21] A. Graves, S. Fern´andez, M. Liwicki, H. Bunke, and J. Schmidhuber, "Unconstrained on-line handwriting recognition with recurrent neural networks", in Advances in Neural Information Processing Systems 20, Proceedings of the Twenty-First Annual Conference on Neural Information Processing Systems, NIPS, Vancouver, British Columbia, Canada,December 3-6, 2007, 2007, pp. 577–584.

[22] A. Graves, M. Liwicki, S. Fernandez, R. Bertolami, H. Bunke, and J. Schmidhuber, "A novel connectionist system for unconstrained handwriting recognition", IEEE Trans. Pattern Anal. Mach. Intell., vol. 31,no. 5, pp. 855–868, 2009.

[23] M. Liwicki, A. Graves, H. Bunke, and J. Schmidhuber, "A novel approach to on-line handwriting recognition based on bidirectional long short-term memory networks", In Proceedings of the 9th International Conference on Document Analysis and Recognition, ICDAR 2007, 2007.

[24] E. F. B. Tasdemir and B. A. Yanikoglu, “Large vocabulary recognition for online Turkish handwriting with sublexical units”, Turkish J. Electr.Eng. Comput. Sci., vol. 26, no. 5, pp. 2218–2233, 2018.

[25] F. Biadsy, J. El-Sana, and J. Habash, “Online Arabic handwriting recognition using Hidden Markov Models”, in Tenth International Workshop on Frontiers in Handwriting Recognition, IWFHR 2007, IAPR. New York,USA, 2006, 2006.

[26] H. A. A. Alshafy and M. E. Mustafa, “HMM based approach for online Arabic handwriting recognition”, in 14th International Conference on Intelligent Systems Design and Applications, ISDA 2014, Okinawa, Japan, November 28-30, 2014. IEEE, 2014, pp. 211–215.

[27] T. Robinson, M. Hochberg, and S. Renals, The Use of Recurrent NeuralNetworks in Continuous Speech Recognition.Boston, MA: Springer US, 1996, pp. 233–258.

[28] M. Husken and P. Stagge, “Recurrent neural networks for time series classification”, Neurocomputing, vol. 50, pp. 223–235, 2003.

[29] A. Capar, K. Tasdemir, O. Kilic, and M. Gokmen, “A Turkish handprint character recognition system”, in Computer and Information Sciences -ISCIS 2003, 18th International Symposium, Antalya, Turkey, November 3-5, 2003, Proceedings, 2003, pp. 447–456.

[30] K. Kaplan, H. M. Ertunc , and E. Vardar, “Handwriting character recognition by using fuzzy logic”, Fırat University Turkish Journal of Science & Technology, vol. 12, pp. 71 – 77, 2017.

[31] S. U. Korkmaz, G. Kirçiç eği, Y. Akinci, and V. Atalay, “A character recognizer for Turkish language”, in 7th International Conference on Document Analysis and Recognition (ICDAR 2003), 2-Volume Set, 3-6 August 2003, Edinburgh, Scotland, UK, 2003, pp. 1238–1241.

[32] B. Yanikoglu and A. Kholmatov, “Turkish handwritten text recognition: a case of agglutinative languages”, in Document Recognition And Retrieval X, 22-23 January 2003, Santa Clara, California, USA, Proceedings, 2003, pp. 227–233.

[33] M. Şekerci, “Turkish connected and slant handwritten recognition system”, Master’s thesis, Trakya Universitesi, 2007.

[34] A. T. Kabakus and P. Erdogmus, “A novel handwritten turkish letter recognition model based on convolutional neural network”, Concurr. Comput. Pract. Exp., vol. 33, no. 21, 2021.

[35] E. Vural, H. Erdogan, K. Oflazer, and B. A. Yanikoglu, “An online handwriting recognition system for Turkish”, in Document Recognition And Retrieval XII, DRR 2005, San Jose, California, USA, January 16-20,2005, Proceedings, 2005, pp. 56–65.

[36] M. Liwicki and H. Bunke, “Iam-On DB - an on-line English sentence database acquired from handwritten text on a whiteboard”, in Eighth International Conference on Document Analysis and Recognition (ICDAR 2005), 29 August - 1 September 2005, Seoul, Korea. IEEE Computer Society, 2005, pp. 956–961.

[37] V. Frinken and S. Uchida, “Deep BLSTM neural networks for unconstrained continuous handwritten text recognition”, in 13th International Conference on Document Analysis and Recognition, ICDAR 2015, Nancy, France, August 23-26, 2015. IEEE Computer Society, 2015, pp. 911–915.

[38] K. R. Weiss, T. M. Khoshgoftaar, and D. Wang, “A survey of transfer learning”, J. Big Data, vol. 3, p. 9, 2016.

[39] S. Niu, Y. Liu, J. Wang, and H. Song, “A decade survey of transfer learning (2010-2020)”, IEEE Trans. Artif. Intell., vol. 1, no. 2, pp. 151–166, 2020.

[40] E. F. B. Tasdemir, “Online Turkish Handwriting Recognition Using Synthetic Data”, European Journal of Science and Technology, vol. 32, no. 5, pp. 649-656, 2021.

[41] B. Tombaloğlu and H. Erdem, “Turkish Speech Recognition Techniques and Applications of Recurrent Units (LSTM and GRU)”, Gazi University Journal of Science, vol. 34, no. 4, 1035-1049, 2021.

[42] C.Yao and G. Cheng, “Approximative Bayes optimality linear discriminant analysis for Chinese handwriting character recognition”, Neurocomputing, vol.207, issue C (September 346–353, 2016.