

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 3, March 2020

Design for Converting Non-Digitized Health Data to Digital Format

Rakesh Moorthy¹, Darshan Mistry², Gaurav Sawardekar³, Sarita Rathod⁴.

Student, Department of Information Technology, K. J. Somaiya Institute of Engineering & IT, Sion, India.¹

Student, Department of Information Technology, K. J. Somaiya Institute of Engineering & IT, Sion, India.²

Student, Department of Information Technology, K. J. Somaiya Institute of Engineering & IT, Sion, India.³

Professor, Department of Information Technology, K. J. Somaiya Institute of Engineering & IT, Sion, India.⁴

ABSTRACT: The main aim of our project is to convert medical records in the traditional non-digitized format to a more useful digitized format using Handwritten Character Recognition (HCR). Health records are important documents which generally has information about patient's identity, medical history and various examination results. With this handwritten character recognition method, it shows the ability of a computer to receive and recognize handwritten data in medical records. This paper mainly focuses on the recognition of handwritten English characters. Given the importance of Health records it is necessary to always keep them updated so that optimal care can be provided whenever necessary. However the huge amount of records makes it impossible to manually update them. Fortunately, HCR for health records can do this task very much faster and accurately.

KEYWORDS: HCR (Handwritten Character Recognition), Deep learning, Neural Network, Keras, Tensorflow, CNN(Convolutional Neural Network) etc..

I. INTRODUCTION

HCR for medical documents is useful because this software provides invaluable benefits in terms of cost savings and even increases productivity. When used in hospitals or health centers such software can be used to perform a variety of functions. A use of HCR can be scanning patient identity cards and capturing data from them. After this the identity card is stored as a scanned image that is fully text searchable. This means that when processed with HCR for medical documents the resultant scanned images along with the text searchable version can be stored easily in databases, and located with a simple text search[2]. When HCR for medical documents software is used many other medical documents such as patient records, test results and active and historical medical records can all be scanned and converted into text searchable format[8]. This results in faster conversion of documents into text searchable format, as well as makes them easier to find in databases, which is not possible in the case of scanned document[7]. Before the use of HCR for medical documents manual data entry was used throughout the world when capturing data from medical records. This was a time consuming and costly process as data entry teams spent thousands of hours on capturing data and ensuring its accuracy[1]. This manual process also meant that large data entry teams had to be maintained and paid. Most manual data entry work has been done away with since HCR for medical documents software has come into use. In fact, this software can provide the same accuracy rates as manual data entry but in a fraction of the time. HCR for medical documents operates at high speeds and often can process batches of medical documents in different formats. The installation of this software can sometimes even process documents faster than an entire data entry team, and since it can operate 24/7 data documents can enter a database faster.

In healthcare, the efficacy of a treatment plan mainly depends on the information a patient can divulge on his own medical history, but our documentation is largely dependent on our memories and at best on previous medical

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 3, March 2020

reports, which largely compromises the accuracy of the history. When HCR for medical documents software is used, many medical documents such as patient records, test results and active and historical medical records can be scanned and converted into searchable text format [1]. This ultimately results in faster conversion of documents into searchable text format, and also makes it easier to find in databases, which is not possible in the case of today's system. HCR for medical documents operates at high speeds and often can process batches of medical documents in different formats. The installation of this software can sometimes even process documents faster than an entire data entry team, and since it can operate 24/7 data documents can enter a database faster. Such is the power of Deep Learning, which we are going to implement in our project. To develop the system, Python libraries for deep learning purposes are utilized. The classifiers which are useful for the classification are K-NN, CNN etc^[3].

II. RELATED WORK

In [1] Datt and Amin analyzed Gujarati handwritten characters using the Artificial neural network. Errors were corrected by RBPNN (Radial Basis Probabilistic Neural Network) method. Also, the HCR is the technology which provides the fast and automated data capture which helps to save time as well as cost too. Due to high noise tolerance, the artificial neural network is most commonly used. Image acquisition, processing, segmentation, feature extraction, classification and recognition, post-processing these are the basic steps which are involved in the HCR.

In [2] Chandarana and Kapadia compared handwritten optical characters of different fonts with standard ones by using the techniques Feature extraction, Segmentation, Template matching and correlation. Recognition accuracy observed was 91.16% when the number of images was tested under the experiment which was better as compared to previous ones. For monotypes and different types of fonts the recognition of a system is strongest.

In [3] Patel and Desai used cross types approach predicated on tree classifier and k-Nearest Neighbour (k - NN) for identification of handwritten Gujarati characters. Combination of structural features along with statistical features is used for classification and identification of characters. The features are relatively simple to derive. By studying the appearance of various handwritten characters the structural features are selected. Combination of structural features and statistical features are used for classification and identification of characters. The structural features were selected by studying the appearance of various handwritten characters of Gujarati script. A success rate of 63% was achieved using this approach.

In [4] Sojitra and Dhakad preferred the algorithm Neural Network, Self-organizing map and the classifier used is a nearest neighbour classifier. The analysis was carried out with ten input samples and five different fonts. 50 samples are collected for a single character of Gujarati script. Accuracy found was 97.78% and recognition rate for slight changes in the 3 training set was found 98.83%. The results state that pre-processing of data before giving input to K-NN has given the highest recognition rate. Merging Neural with existing methods for recognition has given optimum results and best recognition rate. The comparison is performed on a pixel by pixel basis.

In [5] Prasad and Kulkarni have stated that the recognition rate image is highly affected by the similarity of various characters, and these characters which can be verified with some recognition rates of Class 2, 3 and 4. In these classes, there are more similar characters which in turn degrade and recognition rate. It classified the Gujarati Character set into 6 classes, which is based on their geometrical shapes. An average overall Recognition Efficiency of 71.66 %. Each letter was evaluated in a learning set and efficiency of recognition was evaluated.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 3, March 2020

III. PROPOSED SYSTEM

The workflow can be divided into 3 steps.

Step 1: the CNN layers are fed the input image to extract features. The output obtained is a feature map.

Step 2: through the implementation of Long Short-Term Memory (LSTM), the RNN is able to propagate information over longer distances and provide more robust features to training.

Step 3: with RNN output matrix, the Connectionist Temporal Classification (CTC) [7] calculates loss value and also decodes into the final text.

However, For all architectures presented, step 3 (CTC) is the same, then the Vanilla Beam Search [8] method is used, since it doesn't require a dictionary for its application, unlike other known methods such as Token Passing [9] and Word Beam Search [10]. Therefore, the architectures presented in the following sections only act in steps 1 and 2.

In addition, all datasets use the same **charset** for encoding text. So, 95 printable characters from ASCII table (Figure 1) by default are used in the list.

```
!"#$%&'()*+,-./0123456789:;<=>?@ABCDEFGHIJKLMNPOQRSTUVWXYZ[\]^_`abcdefghijklmnopqrstuvwxyz{|}~
```

Fig. 1: charset adopted to encoding text

Through this paper purposefully, we wanted to bring a more uniform approach between the models and databases. Since the use of specific dictionaries and charsets can bring better results in some scenarios, but not so much in others. For example, a charset with only 60 characters is "easier to hit" (1/60) than another with 125 (1/125). In fact, the approach here becomes more difficult for text recognition by the models.

Datasets

For the project, we used the free segmentation approach of text lines of Bentham and IAM datasets. The methodology of partitioning (training, validation and test) of the data was kept original for each dataset

Bringing a brief detailing of each dataset, Bentham [11] is a collection of manuscripts written by English philosopher Jeremy Bentham (1748–1832). This historical dataset (Figure 2) is the more complex among the two datasets adopted. It also has a considerable amount of punctuation marks in the texts.

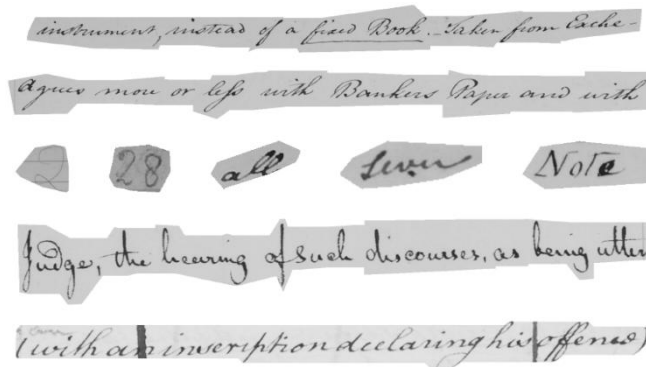


Fig. 2: Sample images from Bentham and data numbers per partition

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 3, March 2020

The Institut für Informatik und Angewandte Mathematik (IAM) database [12] contains forms with English manuscripts, which can be considered as a simple base, since it has a good quality for text recognition (Figure 3). However, it brings the challenge of having several writers, that is, the cursive style is unrestricted.

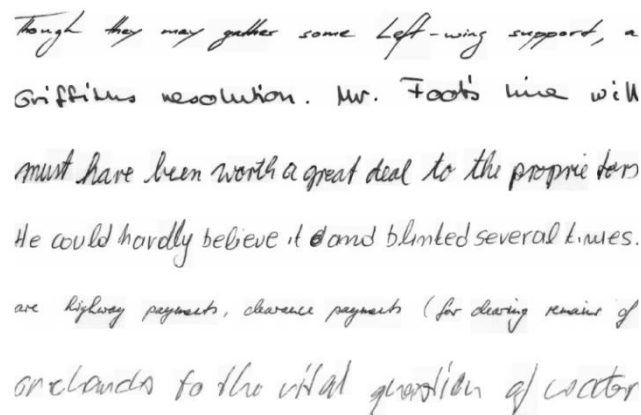
The image displays a sample of cursive handwriting from the IAM database. The text is written in a fluid, slanted script. The words are: "Though they may gather some Left-wing support, a Griffiths resolution. Mr. Foot's line will must have been worth a great deal to the proprietors He could hardly believe it and blinced several times. are highway payments, clearance payments (for drawing remains of orlands to the vital provision of water". The text is arranged in several lines, with some words highlighted in a light grey box.

Fig. 3: Sample images from IAM and data numbers per partition

In short, both explored databases bring a certain challenge. Even though there are specific techniques to solve each of these challenges, but the purpose of the study is to bring methods and approaches that are common to both of the datasets presented..

Finally, to make an easy input of the data into the model, the project was prepared to transform each dataset. The transformation process is simple. It put already preprocessed images, ground truth texts and encoded texts into only one HDF5 file.

Architectures

We chose two model approaches from state-of-art as inspiration for this study. The architectures we chose, correspond to approaches that use BLSTM within the recurrent layer (RNN) and that were eventually compared with MDLSTM [13] as an alternate to computational cost and high complexity.

In this way, the first architecture (CNN-BLSTM), introduced by Puigcerver [14], features a high level of recognition rate and a large number of parameters (around 9.6 million). The Figure 4 shows intimately the distribution of parameters and hyperparameters through 5 convolutional layers and 5 BLSTM.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 3, March 2020

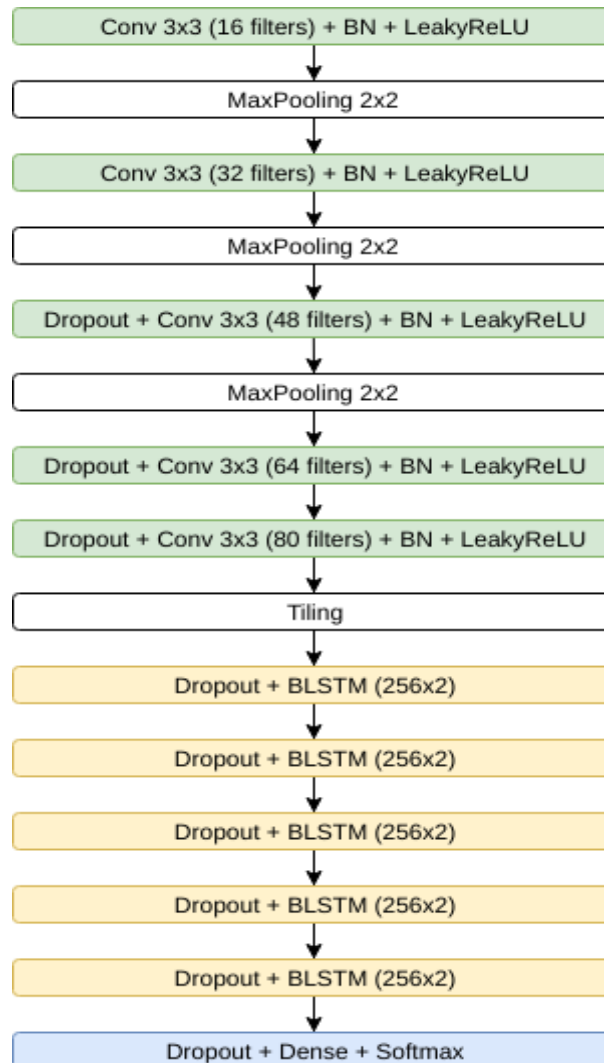


Fig. 4: Workflow of the Puigcerver architecture

The second architecture (Gated-CNN-BLSTM), introduced by Bluche and Messina [15], has the highlight of the Gated-CNN approach. In summary, this system aims to extract more relevant features of the image. Rather than the Puigcerver approach, this model has only a few parameters (around 730 thousand), making it more compact and faster. The distribution of parameters and hyperparameters through 8 convolutional layers (3 gated included) and 2 BLSTM is shown in figure 5.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 3, March 2020

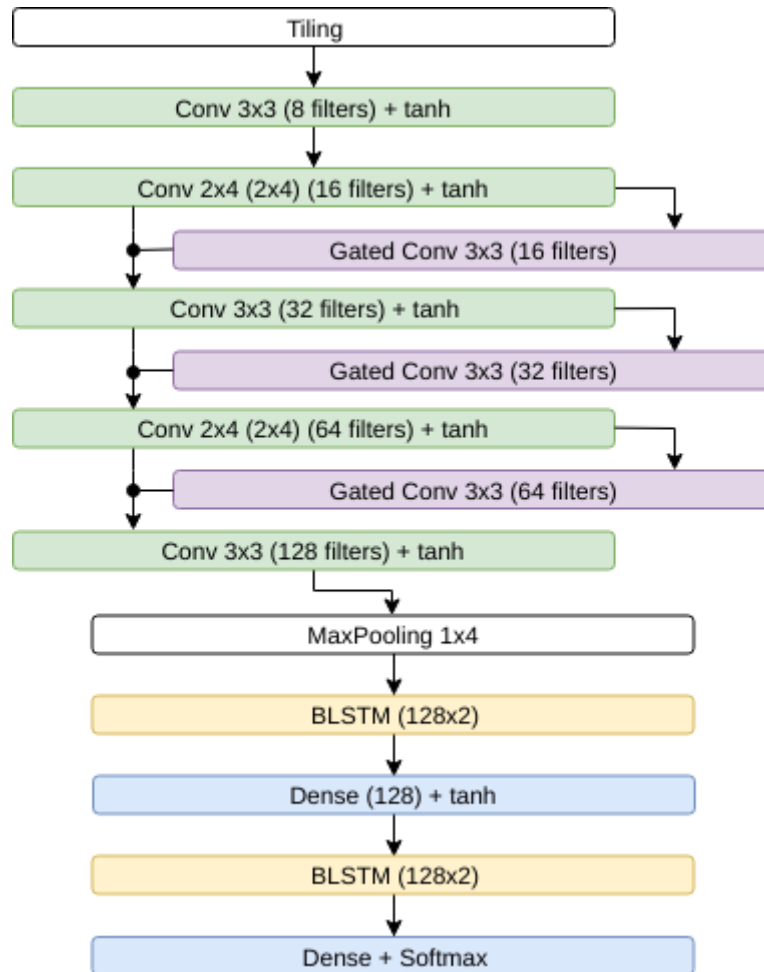


Fig. 5: Workflow of the Bluche et al. architecture

Lastly, the proposed architecture (Gated-CNN-BLSTM) that was inspired by [14] and [15], aiming at: (i) to achieve results compatible with Puigcerver model; and (ii) to keep the small number of parameters (not exceeding one million), such as Bluche et al. model. It was a great challenge to combine the two advantages of each architecture into a single one, however, in the developmental, we applied another Gated approach, once presented by [16] in the context of Language Model. Other methods that also brought differences in results (improvements) were the Batch Renormalization [17] and the Parametric Rectified Linear Unit (PReLU) activator [18].

Thus, the proposed Gated-CNN-BLSTM architecture preserves the low number of parameters (around 820 thousand) and high recognition rate. The Figure 10 shows in detail the distribution of parameters and hyperparameter through 11 convolutional layers (5 gated included) and 2 BLSTM.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 3, March 2020

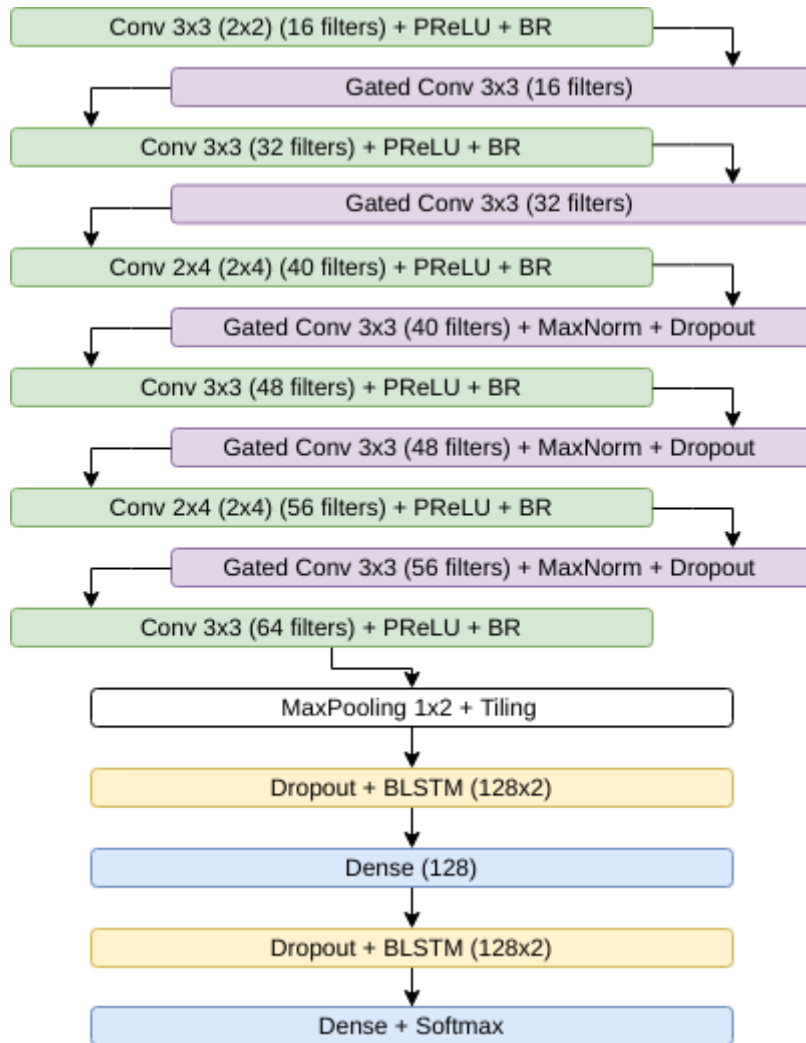


Fig. 10: Workflow of our model's architecture

For training, we have used the RMSProp optimizer [19] with the standard learning rate of each model to incrementally update the parameters using CTC loss gradients.

IV. FUTURE SCOPE

94% of the words from the IAM dataset and Bentham datasets are correctly recognized by NN. In future the accuracy can be further improved by:

- Data augmentation: We can increase the size of our dataset by applying random transformations to the input images. Currently, we are only performing random distortions.
- Convert the cursive writing style to normal writing style..
- Replace LSTM with 2D-LSTM.

International Journal of Innovative Research in Science, Engineering and Technology

(A High Impact Factor, Monthly, Peer Reviewed Journal)

Visit: www.ijirset.com

Vol. 9, Issue 3, March 2020

- Replace optimizer: Although the Adam optimizer increases the accuracy, it also increases the number of training epochs.
- Decoder: Token passing or Word beam Search can be used to constrain the words to the ones in the dictionary.

V. CONCLUSION

We have proposed a design for Handwriting Character Recognition (HCR) Technique which will help in converting handwritten text into a digital text. The Proposed Design will help in detecting the text which is present in the given input and help that input to convert to the digital format which can be stored in a desired format. With the help of Deep Learning, we are able to give the efficient design for the HCR technique.

REFERENCES

- [1] Sagar S.Dutt, Prof. Jay D.Amin Student of Gujarat Technological University “Handwritten Gujarati text recognition using artificial neural network and Error correction using Probabilistic Neural network in recognized text” ,IEEE Fourth International Conference on Multimedia Big Data (BigMM) ,2016.
- [2] Jagruti Chandaran, Mayank Kapadia, “Optical Character Recognition”, IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2014.
- [3] Chhaya Patel, Anand, Apurva Desai, Veer Narmad, “Gujarati Handwritten Character Recognition Using Hybrid Method Based On Binary Tree-Classifer And K-Nearest Neighbour” IEEE Transactions on Pattern Analysis and Machine Intelligence ,2014.
- [4] S. Rishi Kumar, G.Madhavan, M. Naveen, S.Subash, U. Selvamalar Beulah Ponrani “Image Processing based Multilingual Translator for Travellers using Raspberry Pi”, IEEE Global Humanitarian Technology Conference (GHTC) ,2017 .
- [5] Abin M Sabu, Anto Sahaya Das ,Vimal Jyothi Engineering,” A Survey on various Optical Character Recognition Techniques”, Proc. IEEE Conference on Emerging Devices and Smart Systems (ICEDSS 2018) ,2-3 March 2018.
- [6] Kian Peymani, Mohsen Soryani” Machine Generated To Handwritten Character Recognition; A Deep Learning Approach”, 3rd International Conference on Pattern Recognition and Image Analysis (IPRIA 2017) April 19-20, 2017.
- [7] A. Graves, S. Fernández, and F. Gomez, “Connectionist temporal classification: Labelling unsegmented sequence data with recurrent neural networks”, International Conference on Machine Learning, pp. 369–376, 2006.
- [8] K. Hwang and W. Sung, “Character-level incremental speech recognition with recurrent neural networks”, 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 3 2016.
- [9] A. Graves, M. Liwicki, S. Fernández, R. Bertolami, H. Bunke, and J. Schmidhuber, “A novel connectionist system for unconstrained handwriting recognition”, IEEE transactions on pattern analysis and machine intelligence, vol. 31, pp. 855–68, 06 2009.
- [10] H. Scheidl, S. Fiel, and R. Sablatnig, “Word beam search: A connectionist temporal classification decoding algorithm”, in 16th International Conference on Frontiers in Handwriting Recognition (ICFHR). Niagara Falls, United States: IEEE Computer Society, 08 2018, pp. 253–258.
- [11] M. Diem, S. Fiel, F. Kleber, R. Sablatnig, J. M. Saavedra, D. Contreras, J. M. Barrios, and L. S. de Oliveira, “Competition on handwritten digit string recognition in challenging datasets (hdsr 2014)”, 14th International Conference on Frontiers in Handwriting Recognition (ICFHR), vol. 2014, 12 2014.
- [12] U.-V. Marti and H. Bunke, “The iam-database: An english sentence database for offline handwriting recognition”, International Journal on Document Analysis and Recognition, vol. 5, pp. 39–46, 11 2002.
- [13] B. Moysset and R. O. Messina, “Are 2D-LSTM really dead for offline text recognition?” CoRR, vol. abs/1811.10899, 2018.
- [14] J. Puigcerver, “Are multidimensional recurrent layers really necessary for handwritten text recognition?” 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), pp. 67–72, 11 2017.
- [15] T. Bluche and R. Messina, “Gated convolutional recurrent neural networks for multilingual handwriting recognition”, 14th IAPR International Conference on Document Analysis and Recognition (ICDAR), pp. 646–651, 11 2017.
- [16] Y. N. Dauphin, A. Fan, M. Auli, and D. Grangier, “Language modeling with gated convolutional networks”, 2017.
- [17] S. Ioffe, “Batch renormalization: Towards reducing minibatch dependence in batch-normalized models”, CoRR, vol. abs/1702.03275,2017.
- [18] K. He, X. Zhang, S. Ren, and J. Sun, “Delving deep into rectifiers: Surpassing human-level performance on imagenet classification”, in 2015 IEEE International Conference on Computer Vision (ICCV), 12 2015, pp. 1026–1034.
- [19] T. Tieleman and G. Hinton, “Lecture 6.5–rmsprop: Divide the gradient by a running average of its recent magnitude”, COURSERA: Neural Networks for Machine Learning, 2012.