IJIRCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798

# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

## IN COMPUTER & COMMUNICATION ENGINEERING

Volume 8, Issue 8, August 2020

INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 7.488

9940 572 462 | 6381 907 438 | ijircce@gmail.com | www.ijircce.com

# Optical Character Recognition Using Deep Learning on Basis of Convolution Neural Networks

Harshith Kumar Paruchuri[1], Mohammad Shoeab Papasaheb[2], M.Mahendran[3]

UG Student, Dept. of CSE, Panimalar Engineering College, Chennai, India[1,2]

Assistant Professor, Dept. of CSE, Panimalar Engineering College, Chennai, India[3]

**ABSTRACT:** Optical character recognition (OCR) has become one of the most important techniques in computer vision, given that it can easily obtain digital information from various images. However, existing OCR techniques pose a big challenge in the recognition of the of Government documents, tablets, product sheets. In order to solve the problem. This paper proposes a deep learning-aided OCR technique for improving recognition accuracy.First, we generate a database of the uppercase characters to train four neural networks: a convolution neural network a visual geometry group, a capsule network, and a residual network. Second, the four networks are tested on the generated dataset in terms of accuracy, network weight, and test time.Finally, in order to reduce test time and save computational resources. a convolution neural network (CNN), a visual geometry group, a capsule network, and a residual network. Second, the four networks are tested on the generated dataset in terms of accuracy, network weight, and test time. Finally, in order to reduce test time and save computational resources, we also develop a lightweight CNN method to prune the network weight by 96.5% while reducing accuracy by no more than 1.26%.

**KEYWORDS:** Optical character recognition (OCR), convolution neural network (CNN), visual geometry group, residual network, capsule network, pruning network.

## I.INTRODUCTION

Optical character recognition (OCR) is a technology that uses computer software to automatically recognize optical characters in the applications of internet of things (IoT). It is essentially a form of image classification. It is one of the most important techniques in computer vision and has attracted a large amount of attention across different applications put forward a set of algorithms for license plate segmentation and recognition, and obtained high character recognition accuracy . Inoue *et al*. proposed a method of combining classifiers using nonlinear discriminant analysis to improve the accuracy of hand-written character recognition Kokawa proposed a Japanese text classification method based on the language features, which can greatly improve the accuracy of the text classification Using Poisson foil and an edge-enhanced maximum stable extreme value area for text recognition, the text area can be accurately separated from the image, which improves the reliability of recognition Text recognition can be applied not only to images, but also to a wide range of applications in video The main method of text recognition from video involves dividing the video into individual frames. When it comes to the recognition of Arabic text, the authors in pro- posed an effective end-to-end trainable hybrid architecture. Their model is able to recognize Arabic text in high accuracy. Existing OCR techniques perform very well in the recog- nition of English words as well as Arabic numerals. However, the accuracy of these techniques is high for recognizing English characters due to different language families characters are very similar. As a result, conventional methods cannot achieve high recognition accuracy.

OCR is an important part of pattern recognition while deep learning has good performance in pattern recognition. Deep learning is therefore considered to be an effective method for handling big data and improving identification perfor- mance. Recently, it has been successfully applied in different applications . In, a deep learning-aided non-orthogonal multiple access (NOMA) scheme was proposed to improve achievable rate and access performance. The system uses the long short-term memory (LSTM) network for the NOMA system and trains the LSTM network with data under different channel conditions so that the proposed scheme can automatically detect channel characteristics while ensuring its robustness. The authors proposed a deep learning-aided super-resolution channel estimation and direction-of-arrival estimation technique by using a deep neural network (DNN) for both offline and online learning . In addition, pilot allocation is also an important part in multiple-input multiple- output techniques. A new deep learning-based pilot design scheme was proposed in , which uses a multi-layer per- ceptron to infer the optimal pilot allocation scheme. The Internet of Things (IoT) has high requirements for energy and resource efficiency, and it is impossible to directly implement edge computing on the IoT . In order to improve the spectrum efficiency of the IoT, NOMA technology was introduced to solve the problem of energy-saving resource allocation, and a recurrent neural network (RNN) was introduced to optimize resource allocation

OCR technique with the aid of a deep convolutional denois- ing autoencoder. In , the proposed deep learning model achieved better accuracy on poor quality text images and an overall reduction of 21.5% in error rate compared to the existing OCR technologies.

Traditional networks of deep learning that are used to recognize characters have huge network weight and thus their computational burden is very high. To solve these problems and accelerate the development of edge computing, we con- sider pruning the networks. Some typical applications of deep reinforcement learning (DRL) in network slicing are described in [23], and the possible challenges of DRL in net- work slicing resource management are discussed. Kato et al. proposed a heterogeneous computing platform based on deep learning and intelligent routing development [24]. Compared to existing deep learning methods, the proposed methods can ensure more stable network performance when the network topology changes.

**NETWORKS OF DEEP LEARNING**

Deep learning is an important division of machine learning and has been successfully applied in many fields such as speech recognition and natural language processing (NLP). In this section, we will train four networks to recognize kinds of English uppercase characters and a special sym- bol. Both the test set and the verification set have images. We made data enhancements to the datasets, including rotat- ing, cropping, and denoising the original image before entering the network.

**CONVOLUTION NEURAL NETWORK**

The CNN is the most famous network of deep learning and is widely applied in computer vision and NLP. It is a class of feedforward neural network with convolutional computation and deep structure

For general large-scale image classification problems, a CNN can be used to construct hierarchical classifiers, and can also be used in fine classification recognition to extract discriminant features of images for other classifiers to learn. It enjoys excellent performance in image classification and has some advantages over traditional technologies like good adaptive performance and high resolution. It integrates the feature extraction function into the multi-layer percep- tron via structural reorganization and reducing the weight, and omits the complicated image feature extraction process before recognition.

A CNN consists of three main parts. The first part is the input layer. The second part consists of a combination of convolutional layers and pooling layers. The third part consists of fully-connected layers [28]. The convolutional layer is the core component of a CNN. It consists of multiple filters. The parameters of each filter are optimized using a backpropagation algorithm. The purpose of convolution is to extract the features of the input image

**RESIDUAL NETWORK**

It is generally believed that the deeper the network layer is, the higher the features are extracted, and the better the final effect. However, the main problems encountered by deep learning for network depth are gradient disappearance and gradient explosion. The traditional corresponding solution is the initialization and regularization of data. Although this solves the problem of gradient, when the depth is deepened, it will bring another problem which is the degradation of network performance. The depth is deepened, and the error rate is increased

**LIGHTWEIGHT CONVOLUTION NEURAL NETWORK**

The storage and computation of neural networks on embed- ded devices has become a huge challenge due to storage space and power constraints. In general, the deeper the number of layers of the neural network and the more parameters, the more accurate the conclusions are. However, accurate results mean more computing resources are consumed. For mobile devices, speed and accuracy are of equal importance. In order to solve the problem, pruning technology came into being.

Pruning is the removal of parameters that do not contribute much to the output.

The neurons of the model are first sorted according to the order of their contribution to the final result

In deep neural networks, most of the neurons are activated to zero, and the neurons with 0 activation are redundant. Eliminating them can greatly reduce the model size and the energy computation. We use the Average Percentage of Zeros (ApoZ) algorithm to measure the number of values activated by 0 in each filter as a criterion for evaluating whether a filter is important

## II.LITERATURE SURVEY

We extract different layer features from the deep fully convolution network and the last layer features from the shallow fully convolution network, and concatenate them together. After that we add two de-convolution layers to make the output image have the same resolution with the input image.
In particular, our approach only uses the first layer features of the CNN pre-trained offline on Image Net
Based on the Recurrent Neural Network. Besides, we also have collected two datasets and labeled them. Extensive experiments have been implemented

## III. EXISTING SYSTEM

- Optical Character Recognition (OCR) is a type of document image analysis where a scanned digital image that contains machine printed script is input into an OCR software engine and translating it into an editable machine readable digital text format (like ASCII text). Existing system provide less accuracy. Less number of input taken.

OCR - PreProcessing

- **Binarization** – Usually presented with a grayscale image, binarization is then simply a matter of choosing a threshold value.
- **Morphological Operators** – Remove isolated specks and holes in characters, can use the *majority* operator
- **Segmentation** – Chexxck connectivity of shapes, label, and isolate. e.g. the letter *i*, a semicolon, or a colon (; or :). Segmentation is by far the most important aspect of the pre-processing stage. It allows the recognizer to extract features from each individual character. In the more complicated case of handwritten text, the segmentation problem becomes much more difficult as letters tend to be connected to each other.

**MODULES**:
dataset collection
Word beam search decoding
Model Training

## IV.MODULE DESCRIPTION

### DATASET COLLECTION

Handwritten Text Recognition (HTR) system implemented with TensorFlow (TF) and trained on the IAM off-line HTR dataset. This Neural Network (NN) model recognizes the text contained in the images of segmented words as shown in the illustration below. As these word-images are smaller than images of complete text-lines, the NN can be kept small and training on the CPU is feasible. 3/4 of the words from the validation-set are correctly recognized and the character error rate is around 10%. I will give some hints how to extend the model in case you need larger input-images or want better recognition accuracy.

- **WORD BEAM SEARCH DECODING**

Besides the two decoders shipped with TF, it is possible to use word beam search decoding [4]. Using this decoder, words are constrained to those contained in a dictionary, but arbitrary non-word character strings (numbers, punctuation marks) can still be recognized.

- **MODEL TRAINING**

The parameters are loaded from the IAm Dataset. After each epoch of training, validation is done on a validation set (the dataset is split into 95% of the samples used for training and 5% for validation as defined in the class DataLoader). The validation given a trained Neural network. The model [1] is a stripped-down version of the HTR system I implemented for my thesis [2][3]. What remains is what I think is the bare minimum to recognize text with an acceptable accuracy.

## V.SYSTEM ARCHITECTURE

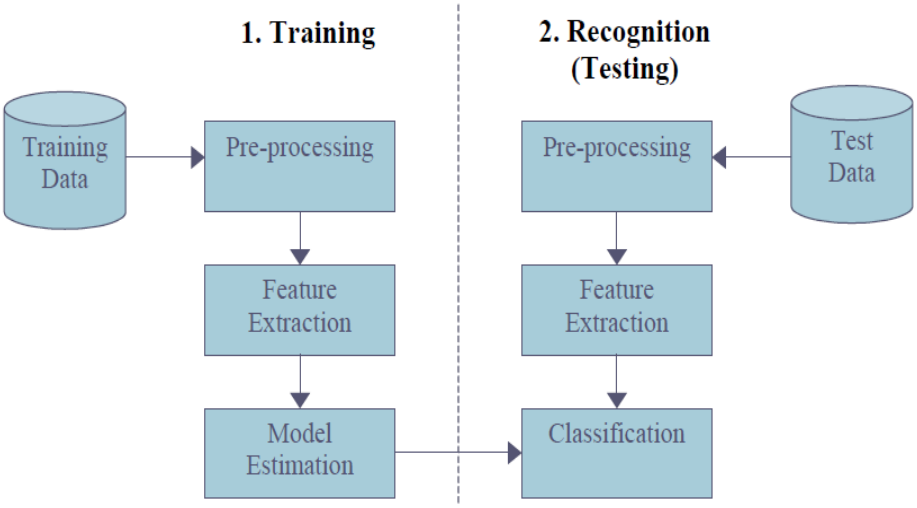



*Figure 2: The pattern classification process*

**Enhanced model**

This model is based on Deep learning, we can provide a lot of data set as an Input to the software tool which will be recognized by the machine and similar pattern will be taken out from them. We can use Octave as a building tool for this product but Octave is recommended in initial state as it's free and easy to use. The Implementation of such a tool depends on two factors – Feature extraction and classification algorithm. So you can use various classifiers available online and also read about basic feature extraction algorithm. The basic version of the product can be implemented in Octave with limited training data set and simple component analysis

## VI. RESULTS

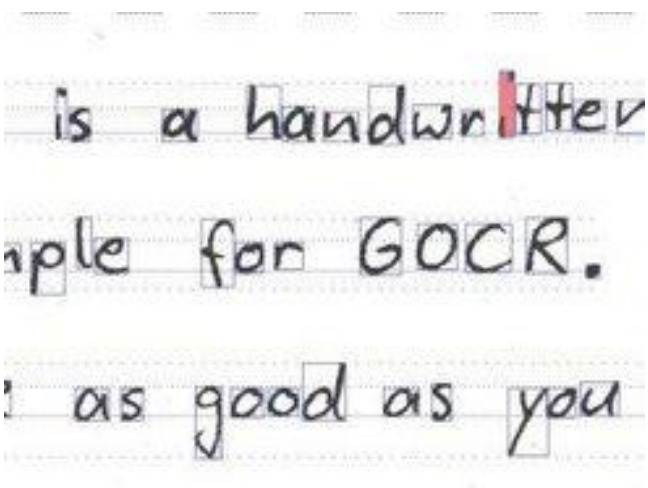


**Figure 3 OUTPUT**

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | Impact Factor: 7.488 |

|| Volume 8, Issue 8, August 2020 ||

Fig.4 INFORMATION EXTRACTED FROM THE IMAGE

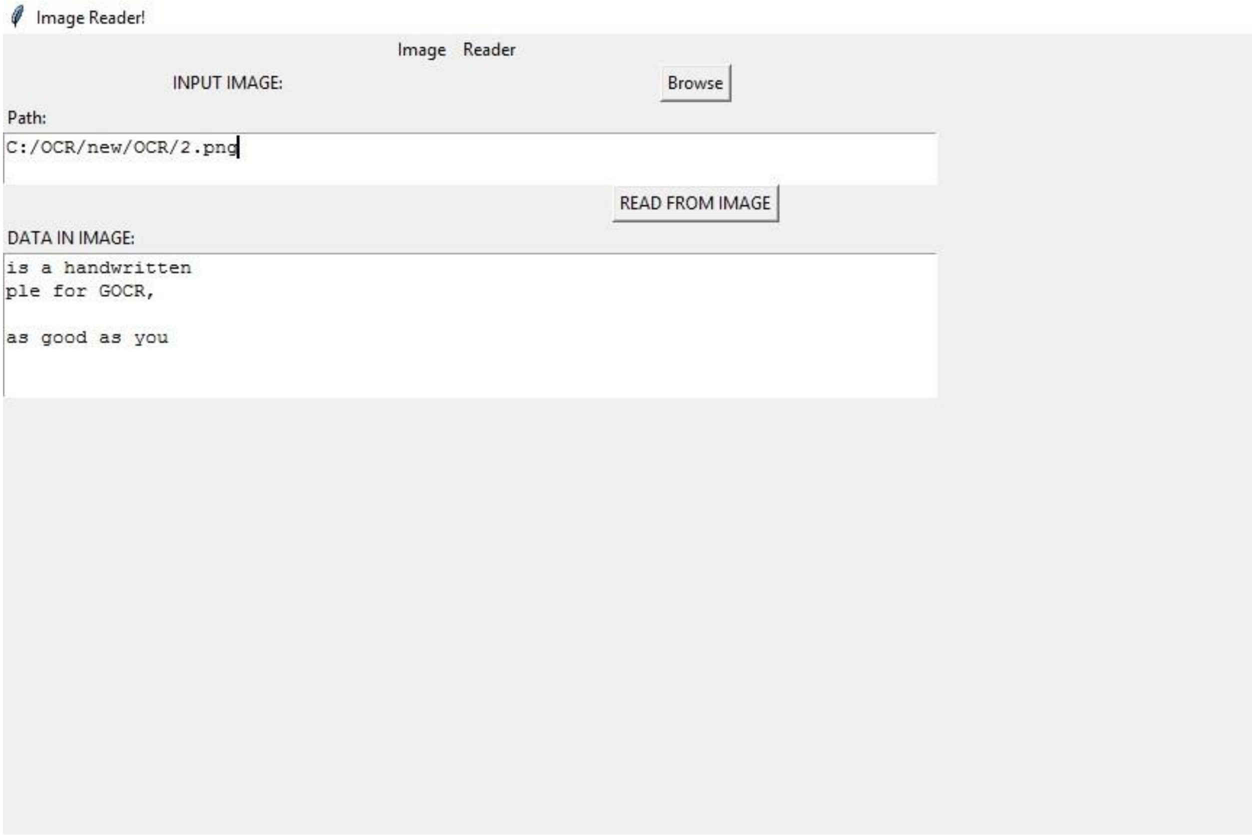
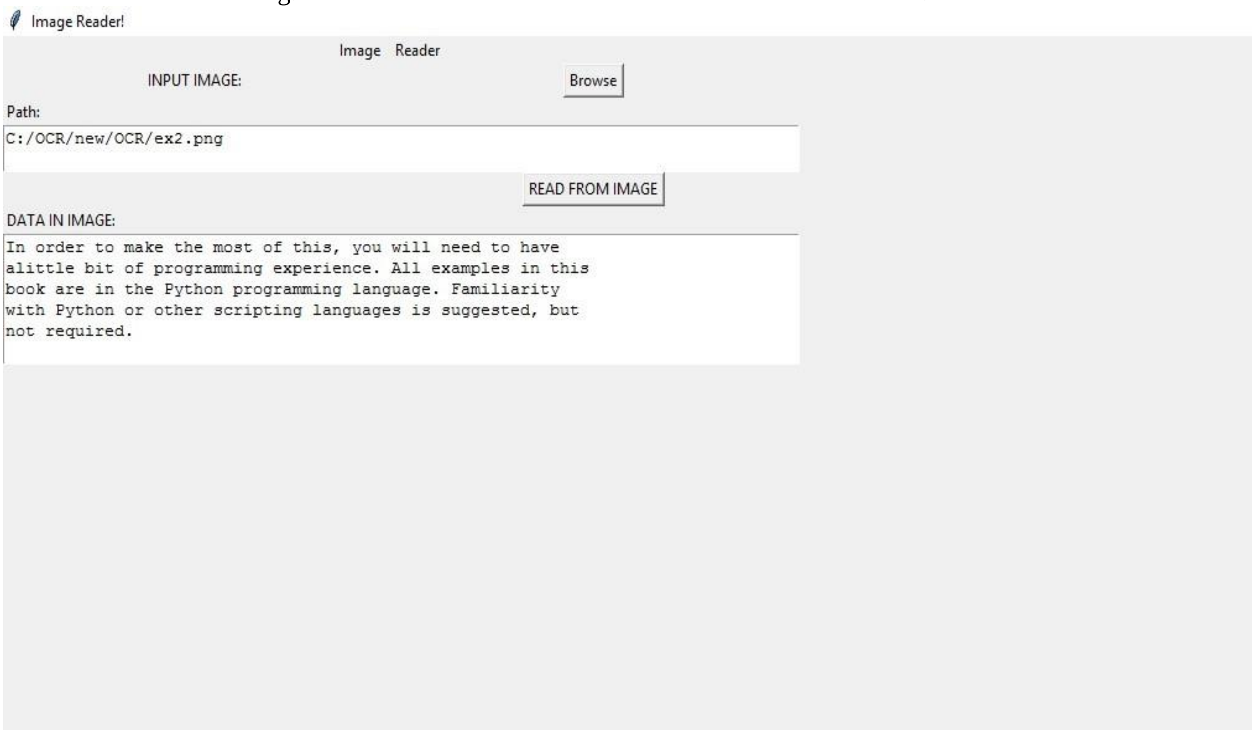
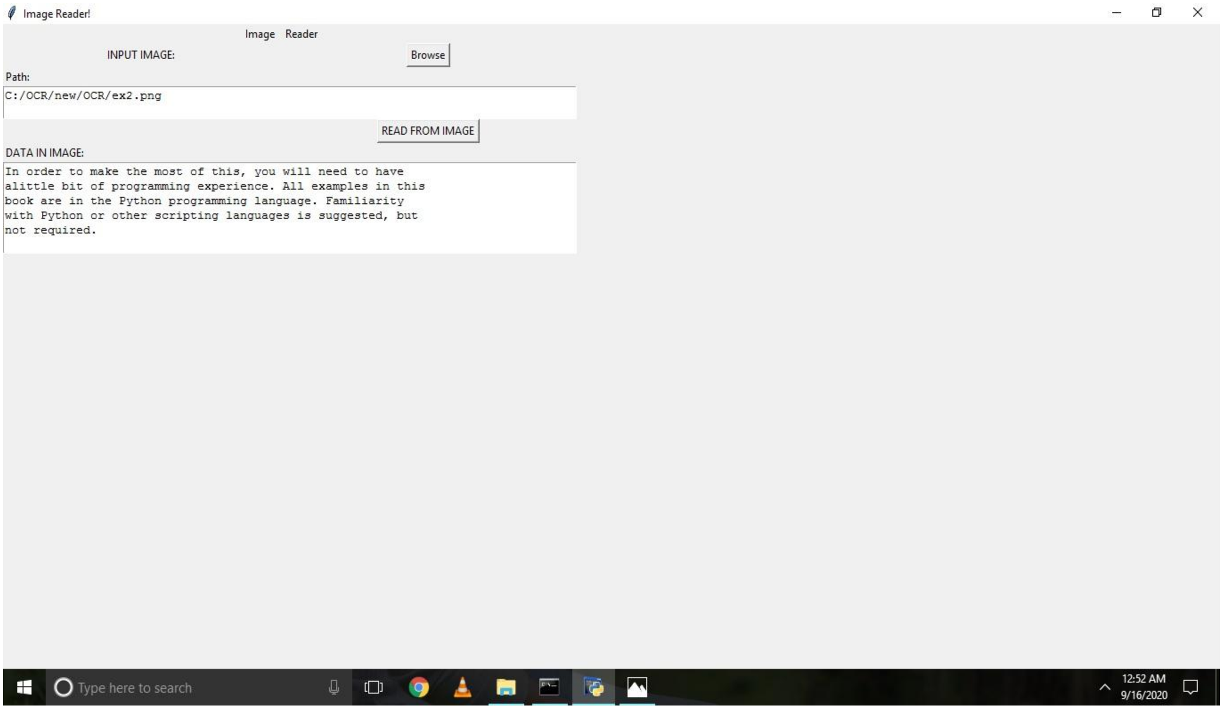



Fig.5 EXTRACTING FROM THE PRINTED DOCUMENTS:






## VII.CONCLUSION

Thus we can conclude that we have obtained the maximum recognition rate as 94.29% by using GABN one of variant of Gabor Filter output as a Feature Extractor of dimensionality 200. The purpose of using Gabor Filters as mode of feature extractor is to promote its utility as major feature extraction technique in field of character recognition of Indian Scripts especially Gurmukhi. Very less literature is available on utilization of Gabor Filters for character Recognition. The work can be extended to increase the results by using or adding some more relevant features along with Gabor features. We can determine optimum combinations of which would yield higher recognition accuracies. We can use some features specific to the mostly confusing characters, to increase the recognition rate. We can divide the entire character set to apply specific and relevant features differently. More advanced classifiers as MQDF or MIL can be used and multiple classifiers can be combined to get better results.

## REFERENCES

[1] J. Wu, K. Ota, M. Dong, and C. Li, ''A hierarchical security framework for defending against sophisticated attacks on wireless sensor networks in smart cities,'' *IEEE Access*, vol. 4, pp. 416–424, 2016.

[2] J. Wu, S. Luo, S. Wang, and H. Wang, ''NLES: A novel lifetime exten- sion scheme for safety-critical cyber-physical systems using SDN and NFV,'' *IEEE Internet Things J.*, to be published. doi: 10.1109/JIOT.2018. 2870294.

[3] M.-Y. Day and Y.-J. Wang, ''Analysis of identifying linguistic phenomena for recognizing inference in text,'' in *Proc. IEEE 15th Int. Conf. Inf. Reuse Integr.*, Aug. 2014, pp. 607–612.

[4] N. Wang, X. Zhu, and J. Zhang, ''License plate segmentation and recogni- tion of Chinese vehicle based on BPNN,'' in *Proc. 12th Int. Conf. Comput. Intell. Secur.*, Dec. 2016, pp. 403–406.

[5] R. Inoue, H. Nakayama, and N. Kato, ''Historical hand-written string recognition by non-linear discriminant analysis using kernel feature selec- tion,'' in *Proc. 18th Int. Conf. Pattern Recognit.*, vol. 2, Aug. 2006, pp. 1094–1097.

[6] A. Kokawa, L. S. P. Busagala, W. Ohyama, T. Wakabayashi, and F. Kimura, ''An impact of OCR on automated classification of OCR Japanese texts with parts-of-speech analysis,'' in *Proc. Int. Conf. Document Anal. Recog- nit.*, Sep. 2011, pp. 543–547.

[7] J. Mol, A. Mohammed, and B. S. Mahesh, ''Text recognition using Poisson filtering and edge enhanced maximally stable extremal regions,'' in *Proc. Int. Conf. Intell. Comput., Instrum. Control Technol.*, Jul. 2017, pp. 302–306.

[8] S. Shetty, A. S. Devadiga, S. S. Chakkaravarthy, and K. A. V. Kumar, ''Ote- OCR based text recognition and extraction from video frames,'' in *Proc. IEEE 8th Int. Conf. Intell. Syst. Control*, Jan. 2014, pp. 229–232.

[9] M. Jain, M. Mathew, and C. V. Jawahar, ''Unconstrained scene text and video text recognition for arabic script,'' in *Proc. 1st Int. Workshop Arabic Script Anal. Recognit.*, Apr. 2017, pp. 26–30.

[10] Q. Li, W. An, A. Zhou, and L. Ma, ''Recognition of offline handwritten Chinese characters using the Tesseract open source OCR engine,'' in *Proc. 8th Int. Conf. Intell. Hum.-Mach. Syst. Cybern.*, Aug. 2016, pp. 452–456.

[11] T. Wang, C.-K. Wen, H. Wang, F. Gao, T. Jiang, and S. Jin, ''Deep learning for wireless physical layer: Opportunities and challenges,'' *China Commun.*, vol. 14, no. 11, pp. 92–111, Nov. 2017.

[12] L. F. Maimó, Á. L. P. Gómez, F. J. G. Clemente, M. G. Pérez, and G. M. Pérez, ''A self-adaptive deep learning-based system for anomaly detection in 5G networks,'' *IEEE Access*, vol. 6, pp. 7700–7712, 2018.

[13] G. Gui, H. Huang, Y. Song, and H. Sari, ''Deep learning for an effec- tive nonorthogonal multiple access scheme,'' *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8440–8450, Sep. 2018.

[14] H. Huang, J. Yang, H. Huang, Y. Song, and G. Gui, ''Deep learn- ing for super-resolution channel estimation and doa estimation based massive MIMO system,'' *IEEE Trans. Veh. Technol.*, vol. 67, no. 9, pp. 8549–8560, Sep. 2018.

[15] K. Kim, J. Lee, and J. Choi, ''Deep learning based pilot allocation scheme (DL-PAS) for 5G massive MIMO system,'' *IEEE Commun. Lett.*, vol. 22, no. 4, pp. 828–831, Apr. 2018.

[16] K. Miyanabe, T. G. Rodrigues, Y. Lee, H. Nishiyama, and N. Kato, ''An Internet of Things traffic-based power saving scheme in cloud-radio access network,'' *IEEE Internet Things J.*, to be published.

[17] M. Liu, T. Song, J. Hu, H. Sari, and G. Gui, ''Anti-shadowing resource allocation for general mobile cognitive radio networks,'' *IEEE Access*, vol. 6, pp. 5618–5632, 2018.

[18] M. Liu, T. Song, and G. Gui, ''Deep cognitive perspective: Resource allocation for NOMA based heterogeneous IoT with imperfect SIC,'' *IEEE Internet Things J.*, to be published. doi: 10.1109/JIOT.2018.2876152.

[19] S. Nawaz, A. Calefati, N. Ahmed, and I. Gallo, ''Hand written characters recognition via deep metric learning,'' in *Proc. 13th IAPR Int. Workshop Document Anal. Syst.*, Apr. 2018, pp. 417–422.

[20] D. S. Joshi and Y. R. Risodkar, ''Deep learning based Gujarati handwritten character recognition,'' in *Proc. Int. Conf. Adv. Commun. Comput. Tech- nol.*, Feb. 2018, pp. 563–566.

[21] M. Rajnoha, R. Burget, and M. K. Dutta, ''Handwriting comenia script recognition with convolutional neural network,'' in *Proc. 40th Int. Conf. Telecommun. Signal Process.*, Jul. 2017, pp. 775–779.

[22] C. Wiraatmaja, K. Gunadi, and I. N. Sandjaja, ''The application of deep convolutional denoising autoencoder for optical character recognition pre- processing,'' in *Proc. Int. Conf. Soft Comput., Intell. Syst. Inf. Technol.*, Sep. 2017, vol. 25, no. 3, pp. 72–77.

[23] T. C. Wei, U. U. Sheikh, and A. A.-H. A. Rahman, ''Improved optical character recognition with deep neural network,'' in *Proc. IEEE 14th Int. Colloq. Signal Process. Appl.*, Mar. 2018, pp. 245–249.

[24] R. Li *et al.*, ''Deep reinforcement learning for resource management in network slicing,'' *IEEE Access*, vol. 6, pp. 74429–74441, 2018.

[25] Z. M. Fadlullah, B. Mao, F. Tang, and N. Kato, ''Value iteration archi- tecture based deep learning for intelligent routing exploiting heteroge- neous computing platforms,'' *IEEE Trans. Comput.*, to be published. doi: 10.1109/TC.2018.2874483.