International Journal of Recent Technology and Engineering (IJRTE)
ISSN: 2277-3878 (Online), Volume-10 Issue-6, March 2022

Third Eye – A Writing Aid for Visually Impaired People

M.S Divya Rani, T K Padma Gayathri, Sreelakshmi Tadigotla, Syed Abrar Ahmed

Abstract*: In olden eras, Braille technology was developed to eradicate the darkness of visually impaired people (Divyanjan) which made them to gain knowledge for proper interaction with the world. Students who read braille can also write braille. Using a variety of low- or high-tech devices, the trainer of students with visually challenged can use braille translation software, which converts the braille into conventional text and prints it. But recognizing the dots on the braille slate and writing the letters is a difficult task faced by many blind kids. Hence the major scope of this proposed project work is to develop a Smart writing system for blind that overcomes many complications faced by a blind child between the age 3-8 years and also to helps them to read and write a standard alphanumeric characters. Henceforth helping them to read and write like a person without disability. Therefore, the project aims to design a writing and reading system for visual impaired person mainly alphanumeric characters. The proposed idea is implemented on Raspberry Pi 4 model interfaced with Speaker and a display unit. Deep Learning algorithms namely CNN is used to recognize the scribbled handwritten text/digits. Overall Performance analysis is made by using MNIST dataset.

***Keywords*: CNN (Convolutional Neural Networks), SVM(Support Vector Machine) KNN (K- nearest neighbor)**

I. INTRODUCTION

According to the statistics provided by WHO(World Health organization), 230 million people worldwide are predicted to be visually impaired, among which 93% live in developing countries and 48 million blind individuals world-wide [1]. People born with sight are very fortunate enough to absorb any information by ourselves. However, there are people who are born without sight or have lost their sight during their course of life. It is beyond the constraints of possibility for these people to extract information from the media, especially books [2]. As the technology has advanced, it has become easy for such types of people to grasp information from the web by using e- readers or any type of computers, which have the capability to read. In recent development of technology, we see more powerful smart phones built in mobile chips, processors in people's hand, which has more computing power than normal desktop PC. This means a many smart applications can be built on mobile devices that could make people's life stress-free. Especially, nowadays mainstream mobile operating systems including iOS and Android had introduce to accesses the functions build in. With these functions, blind people can use a smart phone almost as appropriately as fair sighted people. However, there has to be a best app which helps them to "see" the world, so the main motivation behind this project is to help visually challenged people to write the text on the screen and make the person understand without the ability of sight [3]. Braille is a tactile code that is used by the visually challenged people to read and write. Each character or "cell" comprises of 6 dots, and many of the braille text are in general use. The "double feeds" and "Dot Overlapping" in codes is difficult during automatic translation processing. Recognition of the dot matrix is the main component of the optical Braille recognition system. A literature review was conducted to identify the different disadvantages of using Braille[4]. The braille system of reading documents and books requires more practice and time recent approach includes the Computers that are designed to interact with visually impaired people by reading books or documents. We also have many devices and mobile apps that scans the documents to allow the blind to sense the scanned documents on the screen either in Braille or Anusha Bhargava et al, which is also a Reading Assistant for the blind people to read by using the shape of the letters itself with help of vibrating pegs. But the developed system fails to convert the text to speech and also the computer source can read the document which has the higher font size, and it doesn't have the capability to read the smaller sized text [5]. The Google API and Raspberry Pi based aid for the blind deaf and dumb people was developed, the proposed device enables visually challenged people to read by taking an image. Further, Image to text conversion and speech synthesis is done, converting it into an audio format that reads out the extracted text translating documents, books and other available materials in daily life. For the audibly challenged, the input is in form of speech taken in by the microphone and recorded audio is then converted into text, which is displayed in the form of a pop-up window for the user in the screen of the device [6].

In the recent advancement of AI and ML,digit recognition (handwritten) has become a major topic of interest among many researchers.

Manuscript received on January 31, 2022.
Revised Manuscript received on February 10, 2022.
Manuscript published on March 30, 2021.

* Correspondence Author

Dr. M.S Divya Rani*, Assistant Professor, Department of Electronics & Communication Engineering, Presidency University, Bengaluru, Karnataka India. E-mail: divyarani@presidencyuniversity.in

Ms. T K Padma Gayathri, Assistant Professor, Department of Electronics and Communication Engineering, Sir M Visvesvaraya Institute of Technology, Bengaluru, Karnataka, India. Email: padmagayathri_te@sirmvit.edu.

Ms. Sreelakshmi Tadigotla, Assistant Professor, Department of Electronics and Communication Engineering, Sir M Visvesvaraya Institute of Technology, Bengaluru, Karnataka, India. Email: sreelakshmi_te@sirmvit.edu.

Syed Abrar Ahmed, Assistant Professor, Department of Electronics and Communication Engineering, Sir M Visvesvaraya Institute of Technology, Bengaluru, Karnataka, India. Email: syedabrarahmed@presidencyuniversity.in

In various research was found that the handwritten text/digit recognition models have been developed using Deep Learning algorithm like multilayer CNN which gives the maximum accuracy in image classification problems. Hence Deep learning algorithms are most widely used to develop the model compared any machine learning algorithms like SVM, KNN & RFC. It is quite challenging to get a better performance of the model using less parameters for the development of large-scale neural network. Improving the accuracy with less error to build the model using CNN is quite challenging [7] Using MNIST datasets and CIFAR, many applications-based models are emerging day by day to minimize error rates for image recognition by the Researchers [8].

A survey was made by visiting Rakum Blind school. According to the trainers, the tool used by the blind children to edit on the Braille Slate is quite hard because they require a more pressure to print. Under such circumstances many blind schools have introduced only vocabulary activities for the students from Lower grade to 4th grade. Though there are various existing solutions to the problem of assisting the blind to read and write the Braille text, our solution brings out a smart and autonomous system that assist the blind to write and read the conventional text (alphabets and digits). Hence we propose an electronic aid to help the visually impaired school kids to write and practice the conventional text.

In this work we have proposed a camera mount assistive device to help blind person in reading the text present on the products, text labels and printed notes and also helps them to write the conventional alphabets and digits. The proposed work involves Text Extraction from the captured image and convert the written text to Speech , a process which makes visually impaired to read and hear, one of the most important requirement in developing an application for recognizing the real world. Therefore a Deep Learning algorithm is developed and High end processors with camera are embedded to help the blind children to practice writing and reading at their pre-learning stages like a normal school children do.

II. MODELING OF ALPHABET RECOGNITION SYSTEM

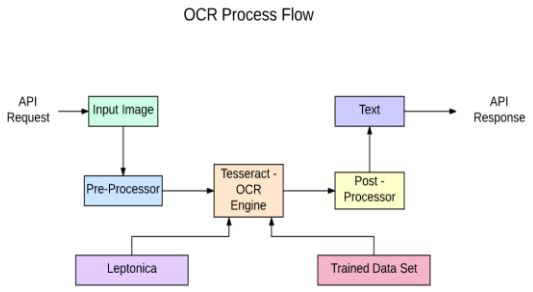

Fig.1 OCR Process Flow

To develop a prototype that helps in recognizing the text or alphabets, an open source optical character recognition (OCR) engine namely Python-tesseract is used which is available under the Apache 2.0 for python. It can be used directly to build the application or using an API to extract the printed or written text. OCR transform a 2D image of text, that could contain handwritten or system printed text into machine-readable text. OCR software engine consists of numerous sub-processes to perform the text recognition as accurately as possible [9]. Tesseract 4.00 comprises a new neural network subsystem designed as a text recognizer. The neural network sub system in Tesseract pre-dates TensorFlow with Variable Graph Specification Language (VGSL) that is also available for TensorFlow [10]. With camera enabled OCR system, the handwritten text images can be recognized even when visually impaired person write the text in a jotted way. Hence the OCR engine supports the complete system. The OCR process flow is shown in Fig.1

III. HANDWRITTEN DIGITS CLASSIFICATION MODEL

A 6-layered convolutional neural network (CNN) is developed to distinguish the handwritten digits in our model. The CNN comprises of a single input layer followed by four consecutive hidden layers and a single layer at the output for the design. The layer at the input contains 28 x 28-pixel size, which means that it holds 784 neurons. The pixels at this layer are classified as grayscale images, for black pixel it is set to one and for white pixel it is set to zero. The feature extraction from the input data namely the digit is acquired by CN layer1, which is the first layer of the hidden section. Which performs the convolution operation with the previous with the help of filters. The activation function used in our model is ReLU which is used at the end of each convolution layer and in the fully connected layer to enhance the model performance. The layer followed by CN layer1 I is Pooling layer1. Here, we have used max pooling to subsample the dimension of each extracted feature from the image. The other two hidden layers are CN layer2 and Pooling layer2, which performs the same operation as CN layer1 and Pooling layer1. But differs in feature representations and kernel size. Next layer to Pooling layer is Flatten layer, which transforms a two-dimensional matrix of featured image into a one-dimensional feature vector and allows the output to be handled by the Fully Connected layers known as Dense layer.

To elevate the model to perform better, SOFTMAX activation function [11] is used at the output, which transforms the featured vector into values between 0 and 1. The MNIST dataset for handwritten digits is used for the model development to recognize the digits between 0 to 9. In MNIST dataset, 70,000 stored images of handwritten digits are there, for training data set 60,000 patterns of digits and for test data 10,000 patterns digits are used to build the model.

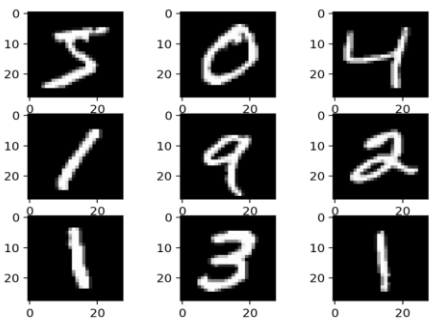

Fig.2 MNIST Dataset

The images are represented in a grayscale with a size of 28×28 pixels. Few sample data of MNIST is shown in Fig.2. We build our model by using Keras API which uses TensorFlow on the backend and hence helps us create a CN networks with high level knowledge.

IV. RESULTS AND DISCUSSION

The components used in this proposed model include Raspberry pi 4 with 6 inch touch screen, a portable Pi camera, power supply unit. The portable Pi camera act has a image recognizing unit, with the top level Machine Learning algorithm namely CNN (Convolution Neural Networks) the Raspberry Pi 4 identify the written images on new grid based slate fixed with a sheet of writing paper. The working of this project is explained with the help of a block diagram shown in the Fig.3.

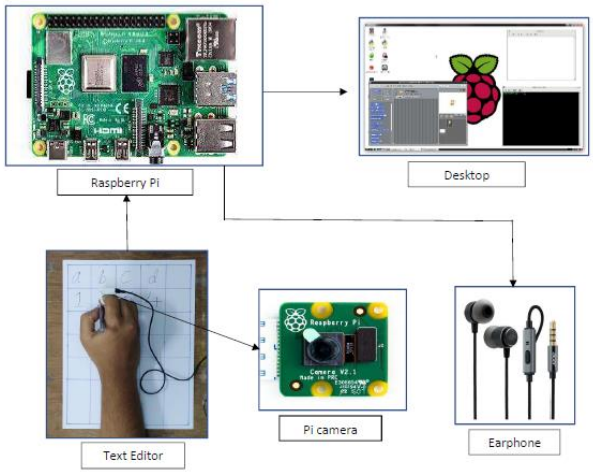

Fig.3 Hardware Setup of The Model

The hardware setup comprises a Pi Camera that recognises the written/reading text on the drawing board by the blind kids. The written image is processed further by Raspberry Pi4 to identify the written text. An efficient machine-learning algorithm called namely CNN is developed to recognize the correct text (Numbers and alphabets).Further the text to speech conversion unit is used to convert the recognized text into speech. The touch screen enables to print the edited text for further interpretation. With the help of a Wi-Fi module, the data can be transmitted to the receiver section i.e. a database for further evaluation if necessary.

Fig.4 Modified Hardware setup

For a better demonstrable purpose, the setup is modified as shown in Fig.4.

A. Discussion of the Obtained Simulated Results of Text Recognition Models

Fig.5 shows the character recognition system showing the arrangements made to capture the written text on the blank sheet and then process it through the OCR engine. When the code is executed, the camera module focused on the text editor containing the input (which can be a text segment or a combination of text and numbers, handwritten or printed). The code uses the tesseract engine to perform the text recognition and uses the python espeak module [12] to generate the required sound output. Fig.6 shows the captured image, which is handwritten and converted into text that is displayed.

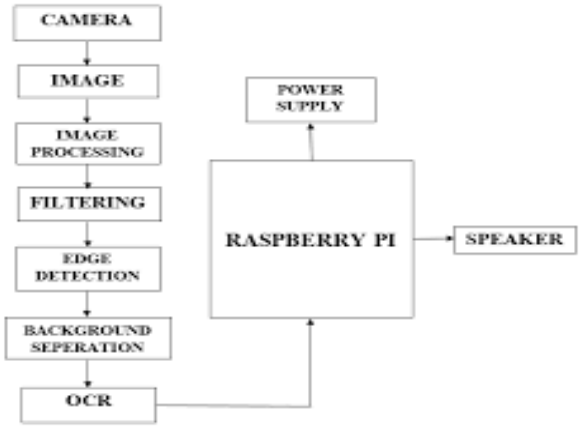

Fig.5 Character Recognition System

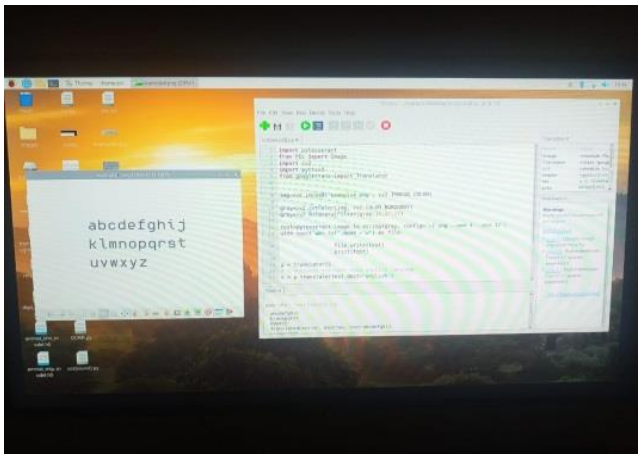

Fig.6 Display of Text through OCR

B. Discussion of the Obtained Simulated Results of Digit Recognition Models

In this section, we have developed a Deep learning algorithm namely CNN algorithm which produces a better performance in recognizing the handwritten digits. MNIST dataset has been used to train model. The algorithm is executed on Keras and Tensorflow platform using python. 10 epochs were used to observe the Training and validation accuracy by taking the batch size of 512.

The batch defines the number of samples the CNN model looks at before updating the internal model parameters.

Table 1 shows the simulation results of accuracies of training and validation data that is found by varying the number of Epochs for the recognizing the handwritten digits.

Out of four hidden layers, the first hidden layer is CN layer1, comprises of 32 filters with the optimal kernel size of 5×5 pixels, mainly used to extract the feature of the image and ReLU is the activation function used to elevate the overall model performance. Maximum pooling is used with 2×2-pixel size, hence the spatial size of the output of CN layer1 is reduced. The second hidden layer is CN layer2 comprises of 64 filters with a kernel size of 5×5 pixels and ReLU activation function followed by the pooling layer 2. We pad the image with a padding size of 1 so that the input and output dimensions are same. Softmax is the activation function used at the output layer to produce the output from digit 0 through digit 9. In the overall performance, the validation accuracy is found around 99.25%. Initially at the first epoch the training accuracy (min) and validation accuracy is found to be 98.06% and 97.92%. At the 10th epoch, the training accuracy (max) is around 99.88% and validation accuracy as 99.25%. Therefore, the total validation loss is found approximately 0.049449 as shown in Fig.7.

Once the training and Validation of the model is completed, next will be the drawing platform to enable the blind people to write here in our project we have used Pygame [13] a multimedia library, which provides the virtual drawing area to scribble the digits. It is a Python wrapper module which contains python functions and classes that will allow the users to draw the items on the screen,play sound effects and music , and handle mouse ,keyboard and joystick input. SDL library provide cross platform access to the systems and devices. Installing a Pygame is much easier and straightforward on the Linux and Windows Platform. Iinstallation of all the addons work using that for my Python requirements. Once the number has been written through the drawing board that is enabled by Pygame SDL, the CNN model detects the input, processes and finally recognizes the number as shown in Fig.8 and hence predicts the digit which will be automatically displayed on the drawing pad as shown in Fig.9. Further the text to speech conversion module is used to convert the written digit into audible feature. Therefore, the visually impaired person can identify the written digit on the virtual drawing board.

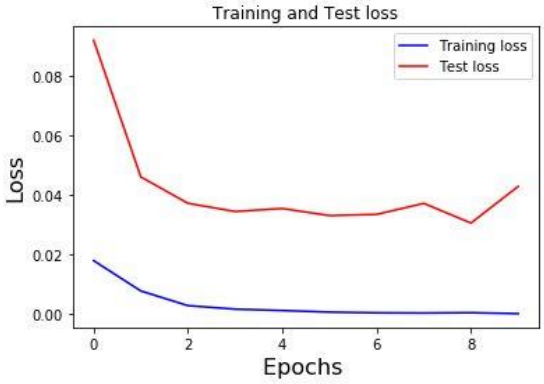

Fig.7 Loss analysis

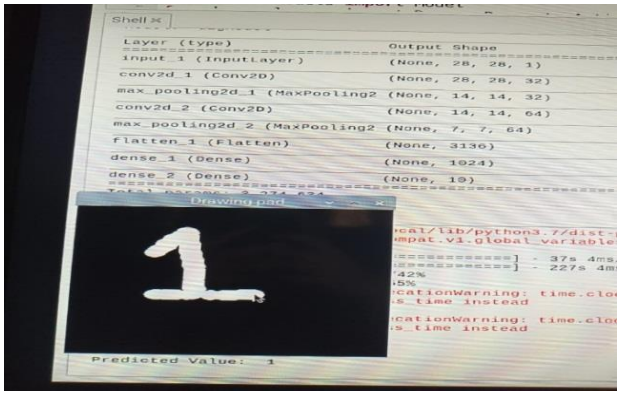

Fig.8 Drawing Pad to write Digit

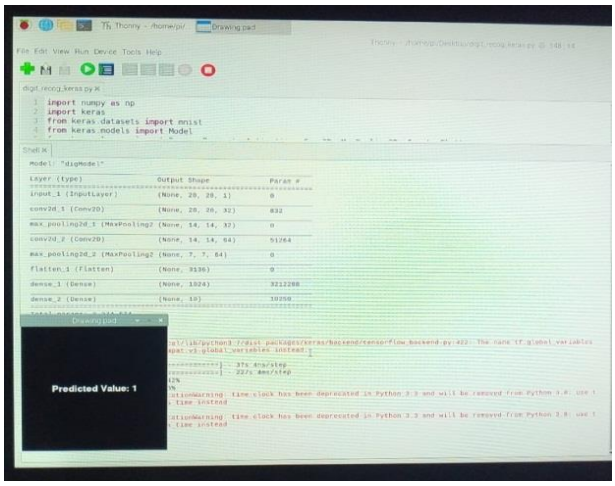

Fig.9 Predicted Value

Table I: Accuracy Evaluation

Number of hidden layers	Batch Size	Minimum Training Accuracy		Minimum Validation Accuracy		Maximum Training Accuracy		Maximum Validation Accuracy	
		Epoch	Accuracy (%)	Epoch	Accuracy (%)	Epoch	Accuracy (%)	Epoch	Accuracy (%)
4	512	1	98.06	1	97.9	8	99.82	8	99.2
4	512	2	98.95	2	98.6	9	99.76	9	99.16
4	512	4	99.49	4	99.14	10	99.88	10	99.25

V. COMPARISON WITH EXISTING MODELS

Several models of digit recognition is existing in the literature. The handwritten digit recognition can be improved using many neural network algorithms, A multi-layered unsupervised learning using CNN model was developed [14] by using 70,000 images from MNIST data group to eliminate the images which are blurred and also found the overall validation accuracy of 98% with the performance loss ranging from 0.1% to 8.5% [15] . Siddique et al. [16] developed an L-layered feed-forward CNN , where they have applied neural network for the handwritten digit recognition with different network layers on the MNIST database and hence found the deviations in the accuracies of the model for different combinations of hidden layers and epochs. The maximum accuracy for 4 hidden layers at 50 epochs was found 97.32% in their work Comparing with results of other models developed for handwritten digit recognition, the model we have developed achieved a better performance. In our model, we have found the maximum accuracy of training data to be of 99.88% and maximum accuracy of validation data of 99.25% both at epoch 10. The overall loss is ranging from 0.0263 to 0.049. Hence, the proposed model of Deep learning algorithm is more efficient than the other existing work for digit recognition.

VI. CONCLUSION AND FUTURE SCOPE

In this work, we have developed a deep learning model, which helps the visually impaired people to read and write the text and digits of conventional language. Raspberry Pi board with camera and a OCR is employed to perform the alphabet recognition on the localized text regions and rework into audio out for blind users.

The CNN model was developed to recognize the handwritten text written by the visually impaired people. The overall performance analysis was made by considering six-layered neural network with one input layer followed by four hidden layers and one output layer is designed. The model was estimated with maximum training accuracy to be of 99.88% and maximum validation accuracy of 99.25%.

Hence, the described proposed writing aid is a portable, low cost and efficient device, which could make the visually impaired people, understand how to write/read like a normal people. Hence, the model developed is best suitable to recognize the scribbled digits and text.

This technological solution can be implemented in blind schools for the real time performance measure. In future, the system can be improved with advanced technology by developing an Application-using machine learning algorithms on the android platform, a Smart gadget for convenience based on the feedback from the visually impaired people.

ACKNOWLEDGMENT

I Thank the entire team of Rakum Blind School, Bengaluru for supporting our project team to interact with visually impaired students and also for demonstrating the Braille scripts.

I am very grateful for their Valuable discussions made to identify the drawbacks in Braille system. My appreciation to the entire project team members for their effort in developing the idea of a writing aid for blind people.

REFERENCES

1. World Health Organization. World Health Statistics 2019. Geneva: WHO 2019
2. Pascolini D, Marioţ SP, Pokharel GP,Global update of available data on visual impairment: “a complication of population based prevalence studies”. Ophthalmic Epidemiol 2004;11-67.
3. Pooja P. Gundewar and Hemant K. Abhyankar, “AReview on an Obstacle Detection in Navigation of VisuallyImpaired”, International Organization of ScientificResearch Journal of Engineering (IOSRJEN), Vol.3, No.1pp. 01-06, January 2013.
4. BhuvaneshArasu and SenthilKumaran, “Blind Man‟s Artificial EYE An Innovative Idea to Help the Blind”, Conference Proceeding of the International Journal ofEngineering Development and Research(IJEDR), SRM University, Kattankulathur, pp.205-207, 2014.
5. Amjed S. Al-Fahoum, Heba B. Al-Hmoud, and Ausaila a. Al-Fraihat, “A Smart Infrared Microcontroller-Based Blind Guidance System”, Hindawi Transactions on Active and Passive Electronic Components, Vol.3, No.2, pp.1-7, June 2013.
6. S.Bharathi, A.Ramesh, S.Vivek and J. Vinoth Kumar,“Effective Navigation for Visually Impaired byWearable Obstacle Avoidance System”, InternationalJournal of Power Control Signal and Computation(IJPCSC), Vol.3, No.1, pp. 51-53, January-March 2012.
7. Michael Mc Enancy "Finger Reader Is audio reading gadget for Index Finger " in July-2014. learning in a spiking convolutional neural network," International Joint Conference on Neural Networks (IJCNN), 2017, pp. 2023-2030.
8. Y. Le Cun, "The MNIST database of handwritten digits,"http://yann. lecun. com/exdb/mnist/, 1998.
9. J. Jin, K. Fu, and C. Zhang, "Traffic sign recognition with hinge loss trained convolutional neural networks," IEEE Transactions on Intelligent Transportation Systems, vol. 15, no. 5, pp. 1991- 2000, 2014.
10. M. Wu and Z. Zhang, "Handwritten digit classification using the mnist data set," Course project CSE802: Pattern Classification & Analysis, 2010.
11. Y. Liu and Q. Liu, "Convolutional neural networks with large margin softmax loss function for cognitive load recognition," in 2017 36th Chinese Control Conference (CCC), 2017, pp. 4045-4049: IEEE.
12. Character recognition algorithms A A Valke and D G LobovIOP Conf. Series: Journal of Physics: Conf. Series 1210 (2019) 012154 IOP Publishing doi:10.1088/1742-6596/1210/1/012154.
13. https://github.com/pygame/pygame.
14. B. Zhang and S. N. Srihari, "Fast k-nearest neighbour classification using cluster-based trees," IEEE Transactions on Pattern analysis and machine intelligence, vol. 26, no. 4, pp. 525-528, 2004.
15. E. Kussul and T. Baidyk, "Improved method of handwritten digit recognition tested on MNIST database," Image and Vision Computing, vol. 22, no. 12, pp. 971-981, 2004.
16. R. B. Arif, M. A. B. Siddique, M. M. R. Khan, and M. R. Oishe, "Study and Observation of the Variations of Accuracies for Handwritten Digits Recognition with Various Hidden Layers and Epochs using Convolutional Neural Network," in 2018 4th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT), 2018, pp. 112-117: IEEE.

AUTHORS PROFILE

Dr. M.S Divya Rani received B.E Degree in Electronics and Communication Engineering from SJCE, Mysore, M.Tech in Electrons from Sir MVIT, Bangalore, She is currently working as an Assistant Professor in Department of Electronics and Telecommunication Engineering, Presidency University Bengaluru. Her research interests include Wireless communication,WSN and IOT. She is a life member of ISTE.

T K Padma Gayathri received BE Degree in Electronics and Communication Engineering from Madras University, Chennai, Tamilnadu, M.Tech in Microelectronics and VLSI Design from IITM, Chennai. She is currently working as an Assistant Professor in Department of Electronics and Telecommunication Engineering, Sir MVIT, Bengaluru. Her research interests include Analog and Mixed signal circuit Design and Low power Circuit Design . She is a life member of ISTE.

Sreelakshmi Tadigotla received B.Tech degree in Electronics and Communication Engineering from S.V University, Tirupati, Andhra Pradesh, M.Tech degree in Digital Communication Engineering from M.S.Ramaiah Insittute of Technology, Bengaluru, Karnataka in 2010. She is currently working as an Assistant Professor in Electronics & Telecommunication Engineering Department, Sir M.V.I.T, Bengaluru. Her research interests include wireless sensor networks (WSN's), Computer Networks and IoT. She is a member of IEEE and life member of ISTE.

Mr. Syed Abrar Ahmed, Assistant Professor, Department of Electronics and Communication Engineering, Presidency University, Bangalore has graduated in Electronics and Communication Engineering from Ghousia College of Engineering, Ramanagaram in 2009 and took his post graduate degree on VLSI and Embedded Systems from Dr. Ambedkar Institute of Technology, Bangalore. His current research interests are in the area of VLSI and Wireless Communication.