HANDWRITTEN CHARACTER RECOGNITION USING CNN

S. Anandh Kishan, UG Scholar, Sri Venkateswara College of Engineering, anandhrockstar@gmail.com

J. Clinton David, UG Scholar, Sri Venkateswara College of Engineering, clintondavid412@gmail.com

P.Sharon Femi, Assistant Professor, Sri Venkateswara College of Engineering, sharon@svce.ac.in

ABSTRACT-- Handwritten character recognition has been one of the active and challengingresearch areas in the field of image processing and pattern recognition. It hasnumerous advantages such as reading aid for bank cheques, recognizing characterfrom form applications etc. An attempt is made to recognize handwrittencharacters for English alphabets using CNN. EMNIST dataset which consists of English alphabets and numbers are made useof to train the neural network. EMNIST balanced dataset consist of 131,600images of characters and 47 classes .The feature extraction technique is obtainedby normalizing the pixel values. Pixel values will range from 0 to 255 whichrepresents the intensity of each pixel in the image and they are normalized torepresent value between 0 and 1. Convolutional neural network is used as aclassifier which trains the EMNIST dataset. The work is extended by adding somemore dataset to EMNIST dataset of characters from Tamil language and trainingthe model. The prediction for the given input image is obtained from the trainedclassifier.

Keywords: CNN, Handwritten character recognition, Feature extraction

I. INTRODUCTION

Handwritten character recognition is a field of research in artificial intelligence, computer vision, and pattern recognition. A computer performing handwriting recognition is said to acquire and detect characters in paper documents, pictures, touch-screen devices and other sources and convert them into machine-encoded form. Its application is found in optical character recognition and more advanced intelligent character recognition systems. Most of these systems implement machine learning mechanisms such as neural networks.

After the extraction of individual characters occurs, a recognition engine is used to identify the corresponding computer character. Several different recognition techniques are currently available.Neural network recognizers learn from an initial image training set. The trained network then makes the character identifications. Each neural network uniquely learns the properties that differentiate training images. It then looks for similar properties in the target image to be identified. Neural networks are quick to set up; however, they can be inaccurate if they learn properties that are not important in the target data.

Handwriting recognition has been one of the most fascinating and challenging research areas in field of image processing and pattern recognition in the recent years. It contributes immensely to the advancement of automation process and improves the interface between man and machine in numerous applications. Several research works have been focusing on new techniques and methods that would reduce the processing time while providing higher recognition accuracy.

The rest of this paper is organized as follows: Chapter 2 presents the related work. Chapter 3 presents the proposed work. Chapter 4 presents the module description. Chapter 5 presents the experimental results for handwritten character recognition. Chapter 6 presents the performance evaluation.

II. RELATED WORK

Toru [6] addresses the problem of reinforcing the ability of the k-NN classification of handwritten characters via distortion-tolerant template matching techniques with a limited quantity of data. Three kinds of matching techniques, namely, Conventional Simple Correlation, the Tangent Distance and the Global Affine Transformation (GAT) correlation are compared. The k-NN classification method consumes a lot of time. Therefore, to reduce the computational cost of matching in k-NN classification, the GAT correlation method was accelerated by reformulating its computational model and adopting efficient lookup tables. Recognition experiments performed on the IPTP CDROM1B handwritten numerical database show that the matching techniques achieved recognition rates of 97% to 98% .The computation time ratios of the tangent distance and the accelerated GAT correlation to the simple correlation were 26.3 and 36.5 to 1.0 respectively for each technique.

Nasien [4] proposed a recognition model for English handwritten character recognition which includes upper and lower cases and also letters. Freeman chain code (FCC) was used as the representation technique of an image character. Chain code representation gives the boundary of a character image in which the codes represent the direction of the location of the next pixel. An FCC method that uses 8-neighbourhood that starts from direction labelled as 1 to 8 is used. Randomized algorithm was used to generate the FCC which builds the features vector.The criteria of features to input the classification is the chain code that converted to 64 features. Support vector machine (SVM) was chosen for the classification step. NIST Databases are used as the data in the experiment. By applying the proposed model, a relatively high accuracy for the problem of English handwritten recognition was reached.

Olarik [5] used the local gradient feature descriptors, namely the scale invariant features transform key point descriptor and the histogram of oriented gradients, for handwritten character recognition. The local gradient feature descriptors are used to extract feature vectors from the handwritten images, which were then presented to a machine learning algorithm for the actual classification. As classifiers, the k-nearest neighbour and the support vector machine algorithms were used. The feature descriptors and classifiers had been evaluated on three different language scripts, namely Thai, Bangla and Latin, consisting of both handwritten characters and digits. The results showed that the local gradient feature descriptors significantly out-perform directly using pixel intensities from the images. When the proposed feature descriptors are combined with the support vector machine, very high accuracies were obtained on the Thai handwritten datasets (character and digit), the Latin handwritten datasets (character and digit), and the Bangla handwritten digit dataset.

Mubarok[1] proposed Hierarchical graph matching for handwritten character recognition. Handwritten character wastransformed into graphs based on its underlying skeletonstructure. Edges of the extracted graph were categorized intoshape types and vertices were extracted from each of the edgesusing line simplification algorithm. Matching procedure of thegraph was performed in hierarchical approach and followed sub-graphisomorphism principals. Performance evaluation of theproposed method was conducted using validated CEDAR datasetand the method reached a recognition rate of 93.40%.

Bautista[2] investigated the accuracy and precision of the proposed system by cross examining the values solved using the proposed system with the values solved manually. The feature extractor and classifier directly influenced the accuracy and precision of Optical Character Recognition, hence considered choosing the combination of Feature Extractor-Classifier combination for handwritten characters which is the Projection Histogram and Support Vector Machine (SVM) combination. The model included three stages. The Pre-Processing or Feature extraction stage then the Recognition stage using SVM. The last stage was Solving Equations and Accuracy measurement. The SVM is trained with Linear, Polynomial and RBF as its kernel, using 90 training images per each character (a total of 5580 images) and a database was created which contained the unique features that represents a specific character.

Lei [3] explains the current status of handwritten character recognition and problems for research and feature selection. On the basis of existing research results, a new feature named direction string is proposed for handwritten character recognition. It uses stroke trend and integrate the properties of both the traditional statistical features and structural features. A measure of distance between different direction strings is proposed and a classifier for handwritten character recognition is implemented using nearest neighbour matching algorithm based on the proposed direction strings and their distances. It explains the application of handwritten character recognition in handwritten calculator. Direction string is proposed to represent the key features of handwritten symbol strokes and distance between direction symbols and direction strings are also defined for nearest neighbour string matching algorithm. Experiments have shown that this method could perform quite well.

III. PROPOSED WORK

In proposed system, EMNIST dataset is extended by adding some more characters from Tamil language. First the input image is provided and is converted into a gray-scale image and normalized in such a way that it represents the same resolution (28 x 28) as that of EMNIST dataset. CNN is trained using EMNIST dataset and use it as a classifier which will yield better results when compared with other machine learning algorithms. The feature vectors are extracted from the input image and provided to the trained model of Convolution Neural Network which recognizes and provides the desired output. Figure 3.1 presents the handwritten character recognition architecture diagram.

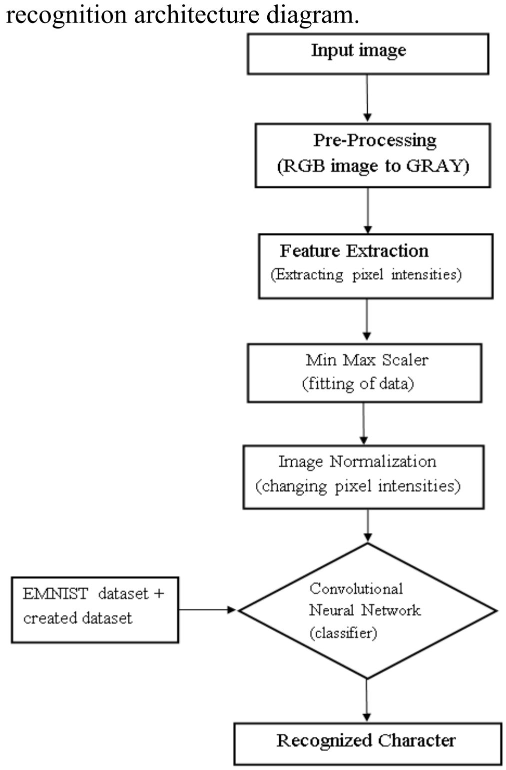


Figure 3.1 Handwritten character Recognition Architecture

IV. MODULES DESCRIPTION

The various modules of the proposed system includes Pre-processing, Feature Extraction, MinMaxScaler -fitting of data, Image normalization and Classification.

A. PRE-PROCESSING

Pre-processing of input image is carried out by converting the given image into gray-scale image. Usually a normal colored image consist of three channels- red channel, green channel, blue channel commonly known as RGB. Then the coloured image is converted it to gray-scale image which consist of single monochrome channel in order to avoid unwanted noise in the image. The given input image would be of varied size which may be lead to loss of accurate prediction when the image is compared with that of trained convolutional neural network. So the image is resized and placed upon a empty 28 x 28 pixel blank image so that the image resolution matches the resolution of EMNIST dataset.

B. FEATURE EXTRACTION

Feature extraction is the process of transforming the input data into a set of features which can very well represent the input data. Feature extraction is related to dimensionality reduction. When the input data is too large to be processed, then it can be transformed into a reduced set of features (also named a feature vector). Determining a subset of the initial features is called feature selection. The selected features are expected to contain the relevant information from the input data, so that the desired task can be performed by using this reduced representation instead of the complete initial data.After resizing the image, pixel values are obtained in the form of 1D array which represents values between 255 and 0 based on pixel intensity.

C. MIN MAX SCALER

The min-max scalar form of normalization uses the mean and standard deviation to box all the data into a range lying between a certain min and max value. It transforms features by scaling each feature to a given range. This estimator scales and translates each feature individually such that it is in the given range on the training set, i.e. between zero and one. This transformation is often used as an alternative to zero mean, unit variance scaling. It essentially shrinks the range such that the range is now between 0 and 1 (or -1 to 1 if there are negative values). The MinMaxScaler is the probably the most famous scaling algorithm, and follows the following formula for each feature:

$$(x_i - \min(x)) / (\max(x) - \min(x))$$

D. IMAGE NORMALIZATION

Normalization is a process that changes the range of pixel intensity values. Normalization is sometimes called contrast stretching or histogram stretching. In this input image the normalization is carried out by removing the background pixels and the character alone will be provided as it is in the image. This can be done by using a random value so that the background pixels will have a value certainly less than the pixel values of shades of the character. In this way the image is normalized such that the image is similar to the values in the EMNIST dataset. In this image, the pixel values are more than 0 for the region where the character 'A' is written and all other regions have pixel values 0 after image normalization.

E. CLASSIFICATION

Convolutional neural network is used as a classifier for classifying the handwritten character from the input image. A CNN consists of an input and an output layer, as well as multiple hidden layers. The hidden layers of a CNN typically consist of convolutional layers, pooling layers, fully connected layers and normalization layers. A CNN consists of three major components which are convolutional layer, pooling layer and output layer. The activation function that is commonly used with CNN is ReLU which stands for Rectified Linear Unit.

Convolution layer will compute the output of neurons that are connected to local regions in the input, each computing a dot product between their weights and a small region they are connected to in the input volume. The pooling layer is a form of non-linear down-sampling. Max pooling is the most common which partition the input image into a set of non-overlapping rectangles and, for each such sub-region, outputs the maximum. ReLU applies the non-saturating activation function . It increases the nonlinear properties of the decision function and of the overall network without affecting the receptive fields of the convolution layer. A rectified linear unit has output 0 if the input is less than 0, and raw output otherwise. Its value is obtained based on the formula which is as follows:

$$\mathbf{f(x)= max(x,0)}$$

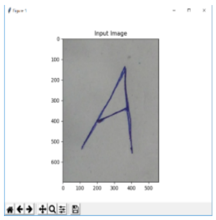
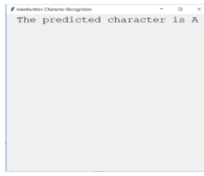
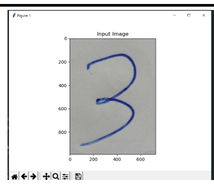
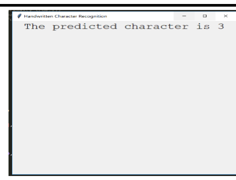
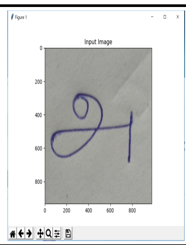
The softmax function is often used in the final layer of a neural network-based classifier. The softmax function squashes the outputs of each unit to be between 0 and 1, just like a sigmoid function. But it also divides each output such that the total sum of the outputs is equal to 1. The output of the softmax function is equivalent to a categorical probability distribution. Thus, softmax function calculates the probabilities distribution of the event over 'n' different events.

V. EXPERIMENTAL RESULTS

The proposed method is implemented in Python 3. The input may be either an alphabet, number or Tamil character. The input is pre-processed, scaled and normalized and provided to the CNN classifier. The CNN classifier is modeled with the EMNIST dataset and so could predict the input character. The experimental results are shown in the following Table 1.

TABLE 1

IMPLEMENTATION RESULTS

Language	**Input image**	**Output**
English Alphabet		
Number		
Tamil character		

VI. PERFORMANCE EVALUATION

The performance metrics that is essential for analysing the performance of the classification test are based on the true positives and false positives.

True positives are the Positive data correctly identified as positive and False positive are the Positive data incorrectly identified as negative.

TABLE II

PERFORMANCE ANALYSIS

Language	**Input data**	**True Positives**	**False Positives**
English Alphabet	26 characters	22	4
Number	10 numbers	9	1
Tamil character	12 characters	9	3

VII. CONCLUSION

Handwritten character recognition has been a challenging task in the past few years. But due to development of machine learning domain in recent years and creation of huge amount of data from our day-to-day life, image recognition for computer vision has seen enormous improvement. EMNIST dataset provides about 132,000 images of 47 characters to be trained and recognised. The convolution neural network was used to train EMNIST dataset to obtain high accuracy. The EMNIST dataset is extended to support the 12 characters from Tamil language and the recognition of these character are tested. The input image is pre-processed, standardized normalized and given to the classifier to predict the character. The model improves the true positive rate and reduces the false positive rate.

REFERENCES

[1] Al Mubarak A, HertogNugroho, " Handwritten Character Recognition using Hierarchical Graph Matching", 2016, International Conference on Advanced Computer Science and Information Systems(ICACSIS)

[2] Bautista SA, Vishnu , Navata, Aldrich H.N, Timothy S, Justine D, Edison A. Roxas, "Recognition of Handwritten Alphanumeric Characters using Projection Histogram and Support Vector Machine", 8th IEEE International Conference Humanoid, Nanotechnology, Information Technology Communication and Control, Environment and Management (HNICEM) The Institute of Electrical and Electronics Engineers Inc. (IEEE)

[3] LI Lei, ZHANG Li-liang, SU Jing-fei, "Handwritten character recognition via direction string and nearest neighbor matching", October 2012, 19(Suppl. 2): 160–165.

[4] Nasien D, Habibollah H, Siti Sophiayati Y , "Support Vector Machine For English Handwritten Character Recognition", 2010 Second International Conference on Computer Engineering and Applications.

[5] Olarik, F.Karaaba, LambertR.B, Schomaker, A.Wiering, "Recognition of handwritten characters using local gradient Feature descriptors",Engineering Applications of Artificial Intelligence45(2015)405–414.

[6] Toru Wakahara ,Yukihiko Yamashita , "k-NN classification of handwritten characters via GAT correlation", Pattern Recognition 47 (2014) 994-1001.