Proceedings of the Indian National Science Academy
https://doi.org/10.1007/s43538-024-00309-7

RESEARCH PAPER

Gujarati handwritten character and modifiers recognition using deep hybrid classifier

Priyank D. Doshi[1] · **Pratik A. Vanjara**[2]

Received: 25 January 2023 / Accepted: 26 April 2024
© Indian National Science Academy 2024

Abstract
In the area of handwritten character recognition many researchers have worked and still working to achieve remarkable result. For the performance improvement of Indic and non-Indic scripts recognition, the necessary condition is to acquire proper domain knowledge and its intricacies otherwise research cannot be fruitful. Here, a Deep Hybrid Learning Classifier with fusion of convolutional neural network has been proposed that learns deep features for offline Gujarati handwritten character and modifier recognition (GHCMR). The proposed model works competently for training as well as testing and exhibits a good recognition performance. The datasets comprising huge image set of offline handwritten Gujarati characters with modifiers have been employed in the present work. The testing accuracies achieved using the proposed network is 97% for characters with modifiers.

Keywords Deep convolutional neural network · Deep learning · Gujarati character recognition · Machine learning

Introduction

Now days, convolutional neural networks (CNN) are popularly used for various purposes of image processing which includes face recognition, object detection, pattern recognition, natural language processing (Harish and Rangan 2020) and image classification. When images are given as input, CNN extracts several features from it which can be used for classification. So, it is desirable to have sufficient set of images in dataset. The precisions of deep convolutional neural networks with such dataset on handwritten characters recognition with modifiers have reached close to perfection.

These systems approaching to human classification ability of Handwritten Character recognition but not completely. As we know Gujarati handwritten character and modifier recognition (GHCMR) system can play an important role in daily routine work for example while we are doing data entry we can avoid retyping, postal address reading by scanning handwritten address directly, conversion of handwritten documents into softcopy etc. As per National Education Policy 2020 of India, student's mother tongue or regional language focused at least up to Grade 5. With the use of technology this will be yardstick in education too. Similarly, e-commerce platforms are also showing interest to work in regional language.

Handwriting specially in Gujarati can be varied by person to person due to inducing factors like temperament, state of mind, time, work pressure etc. which enhances the error rates of the HCNR system. To deal with such complications CNN is probable to be a suitable choice for handwritten image classification. Feature extraction by CNN is a related key task for successful recognition. In CNN feature extraction perform edit self from provided image set. Specially for Roman character and digit recognition substantial achievements have been made due to good availability of datasets like CENPARMI, CEDAR, and MNIST. However, for offline GHCMR, significant work on CNN and Dataset is reported in proposed system.

The present work reports the problem of offline GHCMR from the potential of deep learning and the efficacy of CNN model for the same. The outcome of the study is a model

✉ Priyank D. Doshi
doshipriyank76@gmail.com

Pratik A. Vanjara
pavanjara@gmail.com

[1] Atmiya University, Rajkot, Gujarat 360005, India

[2] Shree M.P. Shah Commerce College, Surendra Nagar, Gujarat 363002, India

Published online: 31 May 2024

relied on the Deep Conventional Neural Network (DCNN) and Image Set (Doshi and Vanjara 2022a, b) for GHCMR.

In the next of this paper, we are presenting Domain Knowledge of Gujarati Language in second section, Literature Review in third section, Proposed work in fourth section, Results in fifth section, conclusions in sixth section and ended with Acknowledgements.

Domain knowledge of Gujarati language

In the model GHCMR some facts are taken into consideration about Gujarati Language which are necessary to solve the problem. Some of them are required to avoid exceptionally from solution.

Gujarati language characteristics

- There is no upper line as it is in Devanagari language.
- Many half characters are joined with other displayed differently. Vowels are attached with consonants to form a word (Doshi and Vanjara 2022a, b).
- Skewed Characters and modifiers are also giving variety to data set. This type of characters varied very frequently person to person when written by their hands so it become very difficult to recognize by machine (Patel and Desai 2011).
- In handwritten character recognition size of hand written characters are also varying from person to person so it should be standardized for recognition.
  ‘શા’ - ‘રા’ ‘જ્ઞા’ - ‘રા’
- Some characters with vowels like and etc. can be written confusingly.
- All half characters are not used in real life like half characters of જ્ઞ, , ક્ષ etc.
- Among different Indian scripts if we compare Devnagari, Gujarati and Odia, we can observe the difference of connected components. We can find more connected components in Devnagari compared to others. In Devnagari, Bangla and Odia character recognition research has been presented by Ragunath Dey, Rakesh Chandra Balabantaray and Sidharth Samanta using DEEP CNN VGG-16 networks for noisy samples (Dey et al. 2022).
- Handwritten Connected and disconnected words are possible as per given below.
- Above problem indicates we need to extra provision for each type of variety of connected as well as disconnected components for recognition process. That affects size of data set (Eq. 1) and number of output probabilities (Table 3). In our solution vowels are included in the data set and prediction set. So, they can be separately detected as well as combined character. Both the way the can results same output.

Literature review

In detection of brain tumors from MRI images, Hashemzehi et al. proposed hybrid model using CNN and NADE. This model consisted two separate networks. CNN used to extract features from raw high dimensional data automatically and exploit patterns that lead the model to correct output. A ConvNADE is an autoregressive model which uses convolutional layers instead of fully connected hidden layers (Hashemzehi et al. 2020).

For remote sensing image recognition with fused deep learning architectures, Özyurt proposed model which is using per-trained architectures for feature extraction based on Alexnet, VGG16, VGG19, GoogleNet, ResNet and SqueezeNet. Obtained features applies to the Relief feature selection algorithm to obtain efficient features. Then, selected features are given to the support vector machine classifier with the CNN-learned features to obtain excellent results (Özyurt 2020).

In solution of web personalization recommendations to provide information quickly and efficiently to user, the work of Singh Rajawat and Jain proposed intelligent map reducer model is based on machine learning concept. In this training and testing of big datasets for Map Reduce Fusion Deep Learning Based on Back Propagation Neural Network algorithm established classification problems. In this fusion recommender System using deep learning they used hybrid filtering and back propagation Neural Network to give solution (Singh Rajawat and Jain 2020).

Deep learning architecture proposed by Ting et al. to capture the hidden features required to predict the travel time for the vehicle traveling on the freeway which includes the GRU neural network model, the XGBoost model, and the Hybrid model that combines the GRU and XGBoost through linear regression (Ting et al. 2020).

Basiri et al. proposed system by which patient-written medical and health-care reviews and useful textual content on social propose can be processed. It uses two deep fusion models based on three-way decision theory to analyze the drug reviews. “The first fusion model, 3-way fusion of one deep model with a traditional learning algorithm (3W1DT) developed using a deep learning method as a primary classifier and a traditional learning method as the secondary method that is used when the confidence of the deep method during classification of test samples is low. In the second proposed deep fusion model, 3-way fusion of three deep models with a traditional model (3W3DT), three deep and one traditional models are trained on the entire training data and each classifies the test sample individually" (Basiri et al. 2020).

In the era of the covid-19, to detect face mask Loey et al. proposed a hybrid deep transfer learning model with machine learning methods. For feature extraction Resnet50 model used, while in the second part to detect face mask, classical machine learning algorithms Support Vector Machine (SVM), decision trees have been used (Loey et al. 2021).

Based on deep learning and machine learning a fusion models proposed by Liu et al. to classify benign, malignant, and intermediate bone tumors based on patient clinical characteristics and conventional radiographs of the lesion (Liu et al. 2022).

Posture and human health status are somehow related to each other. Liaqat et al. proposed a system to monitor remotely the posture such as standing, sitting and walking. "This proposes a novel hybrid approach based on machine learning classifiers (i.e. Support vector machine (SVM), logistic regression (KNN), decision tree, Naive Bayes, random forest, Linear discrete analysis and Quadratic discrete analysis) and deep learning classifiers(i.e., 1D-convolutionalneural network (1D-CNN), 2D-convolutional neural network (2D-CNN), LSTM and bidirectional LSTM) to identify posture detection" (Liaqat et al. 2021).

Feature extraction of connected or disconnected component, Number of end point, number of close loop and similar concept like adjacent pixel connectivity and curvature based pattern matching and classification used by Madushanka et al. (2017) in their HCR method for Sinhala Language.

Ben Atitallah et al. proposed a novel randomly initialized CNN (RND-CNN) architecture for the recognition of COVID-19. RND-CNN model yields 94% accuracy for the COVIDx dataset and 99% accuracy on the enhanced COVID-19 dataset. The proposed RND-CNN consists of an input layer and four hidden blocks for feature learning and extraction, followed by two fully connected layers and a SoftMax layer for case classification (Ben Atitallah et al. 2022). As mentioned in their paper there is a prone of overfitting by initializing random weights and higher and lower weights initialization leads to slower optimization. In our model learning is based on first six layers of feature extraction as shown in Fig. 4.

Raghunath et al. presented Handwritten Digit Recognition System Based on Sliding window with Edit distance. They have validated the proposed system on different datasets. Also, there was a problem of poor handwriting written by writers and that affects prediction. Here in my case the same problem experienced. Similarly in our case pooling layer 2,4,6 as shown in Fig. 4 works as sliding layer and its results feed to next CNN layer (Raghunath and Jayashree 2020).

By reviewing above proposed model shows that machine learning and deep learning can be used according to their need so in the model Gujarati handwritten character and modifier recognition (GHCMR) system proposed in this paper DNN used for feature extraction with sufficient data set and hybrid way achieved with CNN to use as a classifier for achieving good result.

LSTM (Long Short-Term Memory networks are a type of recurrent neural network (RNN) and it is used to model sequential data. LSTM is useful in solving problem based on NLP (natural language processing) such as LM (Language Modeling) in which various probabilistic techniques used to determine or analyze bodies of text data like words, phrases etc. occurring in a sentence for word predictions. It can be used in sentiment analysis, translation, time series prediction and speech recognition. These models are more computationally expensive. It can be applied in video recognition using temporal dependencies but here in our problem we used CNN which is more suitable for character recognition.

Proposed work

In proposed model which can be considered in three phases. Segmentation of Image Acquisition results are given in 28×28-pixel images. This is achieved by segmenting image into line, word and character segmentation gradually. In second phase feature extraction also carried out by extra CNN layers so deep learning and machine learning both are applied here together so this proposed model is called hybrid and Classifier will give prediction output and recognize a character that required to be mapped with any Guajarati character or modifier or combination of both. So, in post processing, model is converting output prediction into a Gujarati character with modifier in required font. This mapping is possible by equivalent bytecode of that font character.

Data set description

In the proposed model data set prepared is based on cross product of two sets. Since the span or size of Data Set depends on all possible combination outcome of each character with all vowels or modifiers. The set of Gujarati Character let's say A and another is set of vowels (modifiers) say set B. The set C is defined as follows which includes cross product or all possible combinations of these two sets with all elements of A and B with one extra element ''ક્ષ''.

C =A U B U{A X B} U {'ક્ષ'} (1)

Set A and B given as under. Where Set A = {અ, ક, ખ, ગ, ઘ, ચ, છ, જ, ઝ, ટ, ઠ, ડ, ઢ, ણ, ત,

થ, દ, ધ, ન, પ, ફ, બ, ભ, મ, ય, ર, લ, વ, શ, સ, હ, ળ,ક્ષ, જ્ઞ, ષ} and Set B = {'◌ા', '◌િ', '◌ી', '◌ુ', '◌ૂ', '◌ે', '◌ૈ', '◌ો', '◌ૌ', '◌ં', '◌ઃ', '◌્', '◌્ર' }.

Let Total number of members in Set C = Mc, Total number of Members in Set A = Ma and Total number of members in Set B = Mb then Mc can be calculated as under.

Here, Ma = 35 and Mb = 13 then

$$\mathrm{Mc} = \mathrm{Ma} + \mathrm{Mb} + \mathrm{Ma} \times \mathrm{Mb} + 1 \tag{2}$$

So, value of Mc = 35 + 13 + 3 × 13 + 1 = 504. Members of Mc are given as under.

Resultant Set C = '◌ા', '◌િ', '◌ી', '◌ુ', '◌ૂ', '◌ે', '◌ૈ', '◌ો', '◌ૌ', '◌ં', '◌ઃ', '◌્', '◌્ર', 'અ', 'અ◌ા', 'ઇ',
'ઈ', 'ઉ', 'ઊ', 'અ◌ે', 'અ◌ૈ', 'અ◌ો', 'અ◌ૌ', 'અં', 'અઃ', 'અ્', 'અ્ર', 'ક', 'કા', 'કિ', 'કી', 'કુ', 'કૂ', 'કે', 'કૈ', 'કો', 'કૌ',
'કં', 'કઃ', 'ક્', 'ક્ર', 'ખ', 'ખા', 'ખિ', 'ખી', 'ખુ', 'ખૂ', 'ખે', 'ખૈ', 'ખો', 'ખૌ', 'ખં', 'ખઃ', 'ખ્', 'ખ્ર', 'ગ', 'ગા', 'ગિ',
'ગી', 'ગુ', 'ગૂ', 'ગે', 'ગૈ', 'ગો', 'ગૌ', 'ગં', 'ગઃ', 'ગ્', 'ગ્ર', 'ઘ', 'ઘા', 'ઘિ', 'ઘી', 'ઘુ', 'ઘૂ', 'ઘે', 'ઘૈ', 'ઘો', 'ઘૌ', 'ઘં',
'ઘઃ', 'ઘ્', 'ઘ્ર', 'ચ', 'ચા', 'ચિ', 'ચી', 'ચુ', 'ચૂ', 'ચે', 'ચૈ', 'ચો', 'ચૌ', 'ચં', 'ચઃ', 'ચ્', 'ચ્ર', 'છ', 'છા', 'છિ', 'છી', 'છુ',
'છૂ', 'છે', 'છૈ', 'છો', 'છૌ', 'છં', 'છઃ', 'છ્', 'છ્ર', 'જ', 'જા', 'જિ', 'જી', 'જુ', 'જૂ', 'જે', 'જૈ', 'જો', 'જૌ', 'જં',
'જઃ', 'જ્', 'જ્ર', 'ઝ', 'ઝા', 'ઝિ', 'ઝી', 'ઝુ', 'ઝૂ', 'ઝે', 'ઝૈ', 'ઝો', 'ઝૌ', 'ઝં', 'ઝઃ', 'ઝ્', 'ઝ્ર', 'ટ', 'ટા', 'ટિ', 'ટી',
'ટુ', 'ટૂ', 'ટે', 'ટૈ', 'ટો', 'ટૌ', 'ટં', 'ટઃ', 'ટ્', 'ટ્ર', 'ઠ', 'ઠા', 'ઠિ', 'ઠી', 'ઠુ', 'ઠૂ', 'ઠે', 'ઠૈ', 'ઠો', 'ઠૌ', 'ઠં', 'ઠઃ', 'ઠ્', 'ઠ્ર', 'ડ',
'ડા', 'ડિ', 'ડી', 'ડુ', 'ડૂ', 'ડે', 'ડૈ', 'ડો', 'ડૌ', 'ડં', 'ડઃ', 'ડ્', 'ડ્ર', 'ઢ', 'ઢા', 'ઢિ', 'ઢી', 'ઢુ', 'ઢૂ', 'ઢે', 'ઢૈ', 'ઢો', 'ઢૌ', 'ઢં',
'ઢઃ', 'ઢ્', 'ઢ્ર', 'ણ', 'ણા', 'ણિ', 'ણી', 'ણુ', 'ણૂ', 'ણે', 'ણૈ', 'ણો', 'ણૌ', 'ણં', 'ણઃ', 'ણ્', 'ણ્ર', 'ત', 'તા', 'તિ', 'તી',
'તુ', 'તૂ', 'તે', 'તૈ', 'તો', 'તૌ', 'તં', 'તઃ', 'ત્', 'ત્ર', 'થ', 'થા', 'થિ', 'થી', 'થુ', 'થૂ', 'થે', 'થૈ', 'થો', 'થૌ', 'થં', 'થઃ', 'થ્',
'થ્ર', 'દ', 'દા', 'દિ', 'દી', 'દુ', 'દૂ', 'દે', 'દૈ', 'દો', 'દૌ', 'દં', 'દઃ', 'દ્', 'દ્ર', 'ધ', 'ધા', 'ધિ', 'ધી', 'ધુ', 'ધૂ', 'ધે', 'ધૈ', 'ધો',
'ધૌ', 'ધં', 'ધઃ', 'ધ્', 'ધ્ર', 'ન', 'ના', 'નિ', 'ની', 'નુ', 'નૂ', 'ને', 'નૈ', 'નો', 'નૌ', 'નં', 'નઃ', 'ન્', 'ન્ર', 'પ', 'પા', 'પિ',
'પી', 'પુ', 'પૂ', 'પે', 'પૈ', 'પો', 'પૌ', 'પં', 'પઃ', 'પ્', 'પ્ર', 'ફ', 'ફા', 'ફિ', 'ફી', 'ફુ', 'ફૂ', 'ફે', 'ફૈ', 'ફો', 'ફૌ', 'ફં', 'ફઃ',
'ફ્', 'ફ્ર', 'બ', 'બા', 'બિ', 'બી', 'બુ', 'બૂ', 'બે', 'બૈ', 'બો', 'બૌ', 'બં', 'બઃ', 'બ્', 'બ્ર', 'ભ', 'ભા', 'ભિ', 'ભી', 'ભુ',
'ભૂ', 'ભે', 'ભૈ', 'ભો', 'ભૌ', 'ભં', 'ભઃ', 'ભ્', 'ભ્ર', 'મ', 'મા', 'મિ', 'મી', 'મુ', 'મૂ', 'મે', 'મૈ', 'મો', 'મૌ', 'મં', 'મઃ',
'મ્', 'મ્ર', 'ય', 'યા', 'યિ', 'યી', 'યુ', 'યૂ', 'યે', 'યૈ', 'યો', 'યૌ', 'યં', 'યઃ', 'ય્', 'ય્ર', 'ર', 'રા', 'રિ', 'રી', 'રુ', 'રૂ', 'રે', 'રૈ',
'રો', 'રૌ', 'રં', 'રઃ', 'ર્', 'ર્ર', 'લ', 'લા', 'લિ', 'લી', 'લુ', 'લૂ', 'લે', 'લૈ', 'લો', 'લૌ', 'લં', 'લઃ', 'લ્', 'લ્ર', 'વ', 'વા', 'વિ',
'વી', 'વુ', 'વૂ', 'વે', 'વૈ', 'વો', 'વૌ', 'વં', 'વઃ', 'વ્', 'વ્ર', 'શ', 'શા', 'શિ', 'શી', 'શુ', 'શૂ', 'શે', 'શૈ', 'શો', 'શૌ', 'શં',
'શઃ', 'શ્', 'શ્ર', 'સ', 'સા', 'સિ', 'સી', 'સુ', 'સૂ', 'સે', 'સૈ', 'સો', 'સૌ', 'સં', 'સઃ', 'સ્', 'સ્ર', 'હ', 'હા', 'હિ', 'હી',
'હુ', 'હૂ', 'હે', 'હૈ', 'હો', 'હૌ', 'હં', 'હઃ', 'હ્', 'હ્ર', 'ળ', 'ળા', 'ળિ', 'ળી', 'ળુ', 'ળૂ', 'ળે', 'ળૈ', 'ળો', 'ળૌ', 'ળં', 'ળઃ',
'ળ્', 'ળ્ર', 'ક્ષ', 'ક્ષા', 'ક્ષિ', 'ક્ષી', 'ક્ષુ', 'ક્ષૂ', 'ક્ષે', 'ક્ષૈ', 'ક્ષો', 'ક્ષૌ', 'ક્ષં', 'ક્ષઃ', 'ક્ષ્', 'ક્ષ્ર', 'જ્ઞ', 'જ્ઞા', 'જ્ઞિ', 'જ્ઞી', 'જ્ઞુ', 'જ્ઞૂ',
'જ્ઞે', 'જ્ઞૈ', 'જ્ઞો', 'જ્ઞૌ', 'જ્ઞં', 'જ્ઞઃ', 'જ્ઞ્', 'જ્ઞ્ર', 'ષ', 'ષા', 'ષિ', 'ષી', 'ષુ', 'ષૂ', 'ષે', 'ષૈ', 'ષો', 'ષૌ', 'ષં', 'ષઃ', 'ષ્', 'ષ્ર', 'ઋ'
}

If we collect images between the range of 300 to 500 images for each member of set C then huge image set has been generated. We know that deep learning extracts features itself using DNN.

Hence necessary condition is to include 504 members and sufficient condition would be that having enough images required for deep learning, then we can justify the solution of the problem of Gujarati handwritten character and modifier recognition (GHCMR). Here in this model enough images included within specified range for all 504 members of set C.

Feature extraction

As far as machine learning is concerned it is very obvious that it is well proven technology in solving wide range of problems. With the use of deep learning for feature extraction the efficiency became even better which in turn increases the applicability in other new problem domains with existing Gujarati handwritten character and modifier recognition (GHCMR).

Stacked convolutional neural network are intentionally included in the deep neural network (DNN) for auto encoding or feature extraction. The layers for classification are combined into DNN for the purpose of classification and efficient learning. There are various ways for image normalization. Image normalization is required before providing images as input to DNN. The input layer and middle

Table 1 Connected components in different India scripts

Devnagari	Gujarati	Odia
आम	આમ	ଓଣ୍ଣୁ
काशी	કાશી	କାଶୀ
नीर	નીર	ପାଣି

Table 2 Variation in Gujarati handwritten language

ક + ઓ = કો	કો	કો
ક+ઈ =કી	કી	કી

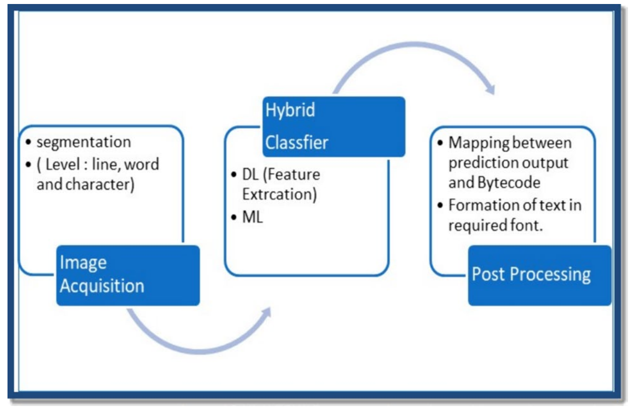

Fig. 1 Proposed model Gujarati handwritten character and modifier recognition (GHCMR) with post processing in bytecode

layers are used for feature extraction. Then the layers near to output layers called high-level layer are used for classification depending on the previously extracted features. The output layer which is the last layer uses the SoftMax function. Accordingly, these structured layers calling framework, delivers enhanced accuracy as compared to training of classification and feature extraction independently.

Classifier (CNN)

In the presented model mainly two tasks are addressed feature extraction and classification. The architecture of CNN contains layers for different types having different operations namely, convolutional, pooling and fully connected (FC) layers. We have one input and one output layer. The total number of convolutional, pooling and fully connected layers and number of nodes in each layer in architecture of CNN are varies depending on the image size and number of output classes of classifier. Underfitting and overfitting problem depends on number of epochs less or more respectively. Training on dataset is allowed to continue till the point from where the performance of system starts to degrade. Pooling layer down samples on the input map. The input and output feature maps are same in this layer but dimension is reduced.

The proposed CNN architectures for Gujarati handwritten character and modifier recognition (GHCMR) system are shown in Table 1. The architecture of CNN comprises of an input layer, twice repeated convolutional and max-pooling layer sequence for feature extraction task, two fully connected layers for classification task and one output layer. The neurons of input layer are connected to each input image and each layer of the network connected with the output of the previous layer as input and passes the output to the next layer as next input.

In final layer, total number of nodes are equal to the number of predicted classes. To overcome the problem of overfitting "Dropout" regularization method used. In this regularization some neurons removed randomly which makes insensitive the learned weight of some neurons.

The parameters of concerning layers for Gujarati handwritten character and modifier recognition (GHCMR) are given in Table 2. In this work architecture using repeated sequence of convolution and pooling layers for feature extraction and fully connected layers for classification. Convolutional layer here consists 1024 mappings. The 28×28 input image is fed to first input layer. The input feature maps (N), dimension of filter (F) and stride (S) used to calculate output feature map dimension by the formula $(((N-F)/S+1)$. For a (28×28) dimensional input image, if filter size is 3×3 and stride is 2 for convolution layer the output for convolutional layer output will be $(28-3)/1+1=26$. The output of input layer is 26×26. Additionally, learnable parameters in Ith convolution layer calculated by given formula or simply we can say ((shape of width of the filter × shape of height of the filter × number of filters in the previous layer + 1) × number of filters). In our case we can have number of parameters for convolution are $((3 \times 3 \times 1024) + 1) \times 1024 = 9{,}438{,}208$. Based on size of filter (F), feature map of previous layer ($FMi-1$), feature map of current layer (FM_i) as $P_i = (((F \times F \times FM_{i-1}) + 1) \times FM_i$. Here, 1 is the bias for each feature map.

Experimental results and discussions

In this paper, Gujarati handwritten character and modifier recognition (GHCMR) system have been used for the extraction of features using deep learning and recognition process is attempted through CNN.

Output using bytecode in Gujarati font

Now the objective to recognize all Gujarati characters and their maximal possible conjunction with modifiers (vowels/"Barkshari") can be fulfilled by generating output in sequence of figures given as under (Fig. 1).

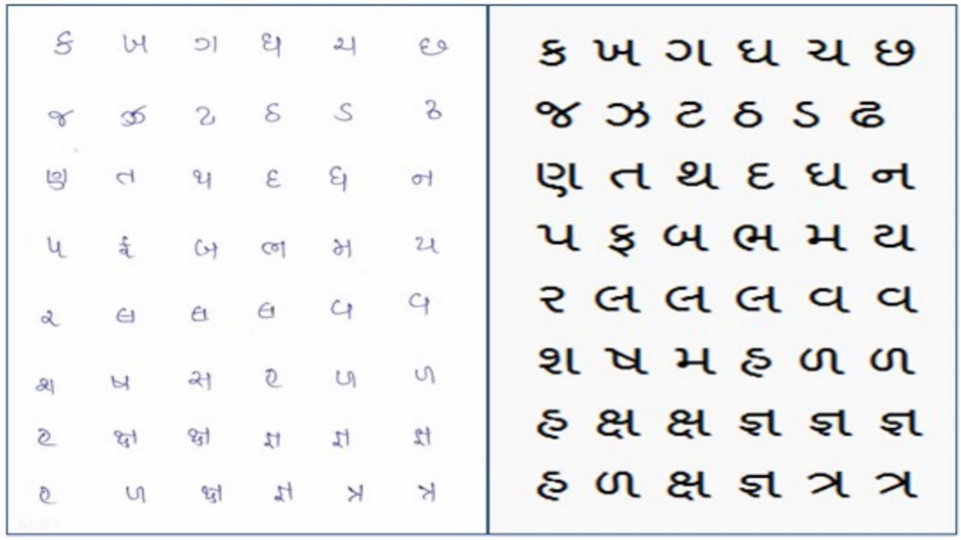

Fig. 2 Gujarati characters—scanned and final output (digitally converted Gujarati characters using Bytecode)

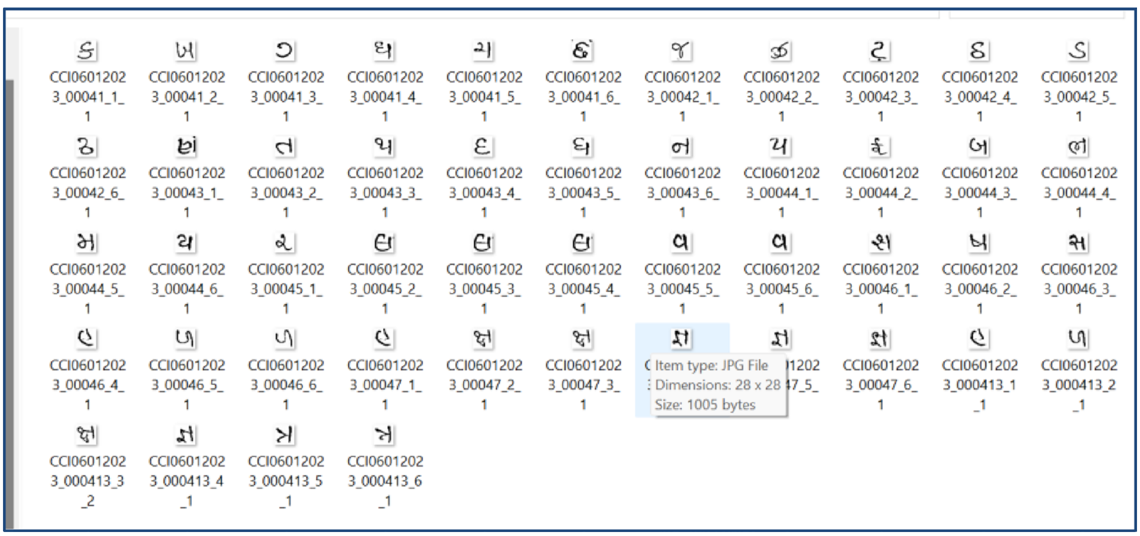

Fig. 3 Segmented Gujarati characters

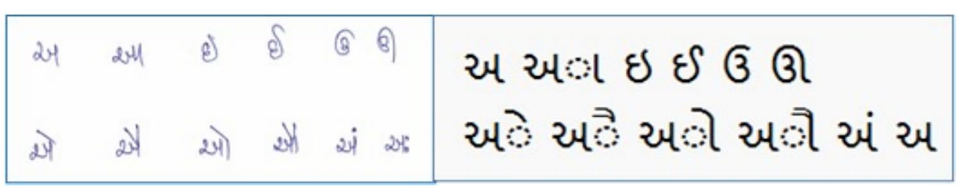

Fig. 4 Handwritten Gujarati vowels and vowel output of classifier using bytecode in shruti.ttf

Figure 2 shows scanned input of Gujarati Characters and final output by classifier and Fig. 3 gives images of segmented characters. Figure 4 shows Gujarati scanned vowels and final output by classifier. Segmentation of vowels given in Fig. 5.

In Fig. 8, scanned image of a character 'ક્' with modifiers shown and successively it is converted into segmented images (given in Fig. 9) and at last it is converted into digital form using bytecode in Fig. 10. In Fig. 10 we can observe four words are given that corresponds to Fig. 8 and 9. The additional thing in Fig. 10 that character combined with modifiers and output is given by model in combination with modifiers (Figs. 6, 7).

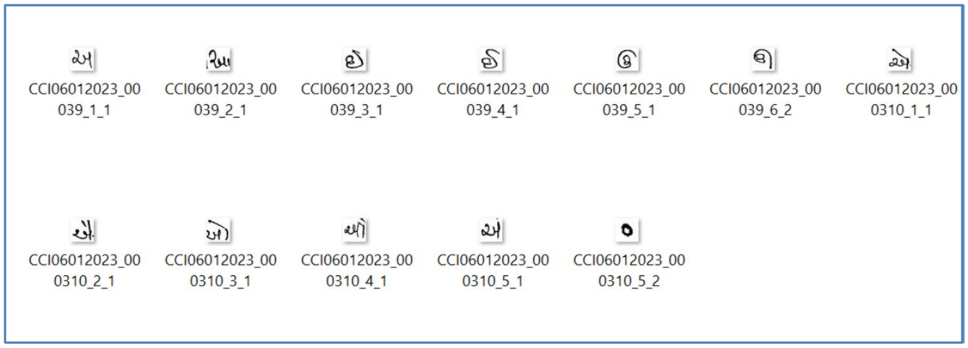

Fig. 5 Segmented Gujarati vowels

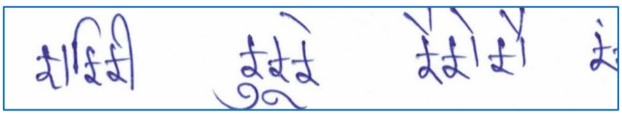

Fig. 6 Gujarati character ‘ક’ with modifiers

Second approach for segmented characters with vowels

As per figure given below, we have characters with combination of vowels (Fig. 8). Sometimes characters segmented along with modifiers or vowels so Fig. 9 is the result of segmentation process. It is also converted using classifier into digital format using bytecode as shown in Fig. 10 (Doshi and Vanjara 2022a, b).

Performance chart and confusion matrix

As per figure given below two charts has been plotted, blue line indicates training and yellow line indicates validation you can observe accuracy and are almost unswerving with each other and above 90%. The loss of the proposed model indicates that after each epoch it is reducing.

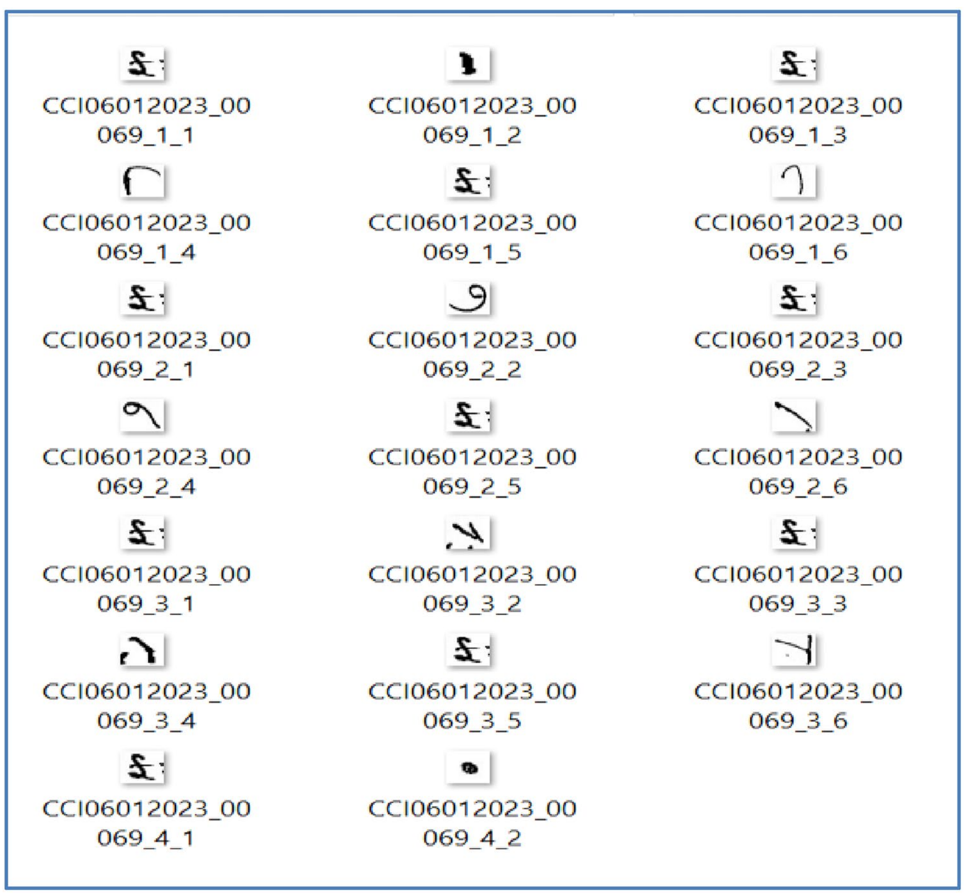

Fig. 7 Segmented images of Fig. 6

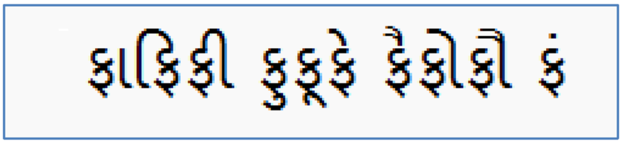

Fig. 8 Digitally converted characters using bytecode of Fig. 7

Here observing the graph, we can see that training results in good fit and validation loss that decreases to stability point having small gap between validation and training loss values. This gap is referred to as the “generalization gap.”

The confusion matrix is a x y matrix, where x = y = number class or output. In confusion Matrix we require to understand terms like True Positive (TP), False Positive (FP),True Negative (TN) and False Negative (FN).

Some writer in Gujarati Language writes vowels which are extremely cursive or artistic in nature. So, it can possibly reduce the rate of accuracy. In effort to improve the result and to rise accuracy we can use concepts of Graphics, Basic Geometry and trigonometry collaborating with a Machine Learning tools or artificial neural network. We can extract or predict shapes in terms of geometrical figures in feature extraction and co-relate it with character so word recognition will be possible. As shown in Fig. 11 basic shapes like arc of circle or ellipse, slop of the line, connected and disjoint components (pixel wise) can be used to map vowels and characters in Gujarati Language. A line used in a zone can be slant or vertically straight which can be determined by its slop and the type of zone where it lies with other shapes, we can possibly predict exact vowel (Table 3).

Tables 4 and 5 indicates latency of prediction function in millisecond per step. We can observe that connected characters as given in in Fig. 9 are more efficiently resulted in single step shown in Table 4 whereas disconnected characters with vowels are differently predicted and their combined output given to text file as shown in Table 5. This latency

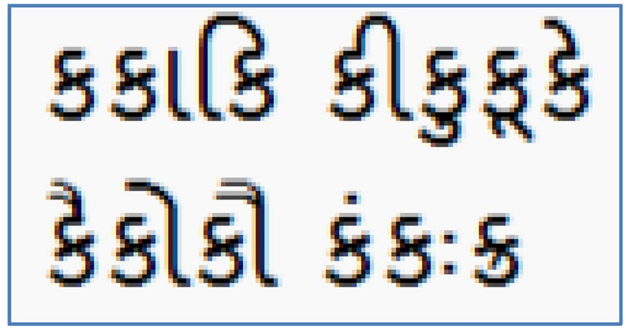

Fig. 10 Output by classifier with bytecode conversion

time varies by complexity of model, data size, image itself and hardware also (Figs. 12, 13).

Future scope

The future scope will be the applicability of algorithms by putting them online through smartphones and websites. Android and iPhone are widely used in urban and rural areas of Gujarat. Also, we can use Hand Scanner at offices of public or private sectors including government offices like the post, railway, city survey offices, etc. where records and documents are received and stored in regional handwritten Gujarati language. Another stack holder can be schools where primary education will be in regional or mother language according to NEP 2020.

Secondly, reverse of the current research model can be the future scope. That means one can prepare handwritten image dataset of the particular person and create mapping from typed Gujrati characters to his own handwriting.

Just like in English still the problem of creating onto and one to one mapping of hand-written Gujarati characters with vowels recognition. The future scope will be detection and conversion of Gujrati phonetic language to text (Doshi and Vanjara 2022a, b).

Fig. 9 Segmented characters with combination with vowels

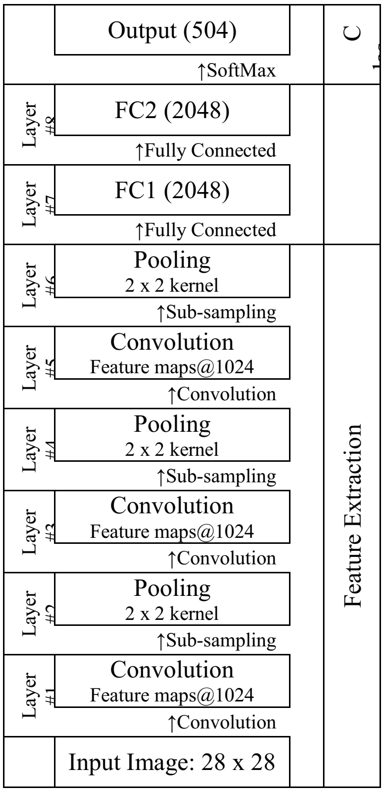

Fig. 11 Architecture of CNN for classification of GHCMR

Table 4 Output of prediction function with latency in millisecond per step for segmented images given in Fig. 9

Connected characters	ms/step	Connected characters	ms/step
ક	113	કૈ	63
કા	49	કો	47
કિ	47	કૌ	63
કી	47	કં	47
કુ	62	કઃ	63
કૂ	47	કૃ	47
કે	48		

Table 5 Output of prediction function with latency in millisecond per step for segmented images given in Fig. 7

ms/step	Character	ms/step	Vowel	Combined characters
63	ક	52	◌ા	કા
63	ક	63	િ◌	કિ
66	ક	63	◌ી	કી
65	ક	63	◌ુ	કુ
63	ક	75	◌ૂ	કૂ
63	ક	64	◌ે	કે
54	ક	63	◌ૈ	કૈ
63	ક	53	◌ો	કો
63	ક	63	◌ૌ	કૌ
78	ક	73	◌ં	કં

Table 3 CNN parameter setup for Gujarati handwritten character and modifier recognition

Layer	Operation in layer	Feature map (In number)	Output shape	No. of parameters	Architecture
Input	Input layer	1024	26×26	10,240	Feature extraction
S1	Max-pooling	1024	13×13	0	
C1	Convolution	1024	11×11	9,438,208	
S2	Max-pooling	1024	5×5	0	
C2	Convolution	1024	3×3	9,438,208	
S3	Max-pooling	1024	1×1	0	
F1	Fully connected	2048	1×1	2,099,200	Classification
F2	Fully connected	2048	1×1	4,196,352	
Output	SoftMax	504	1×1	1,032,696	

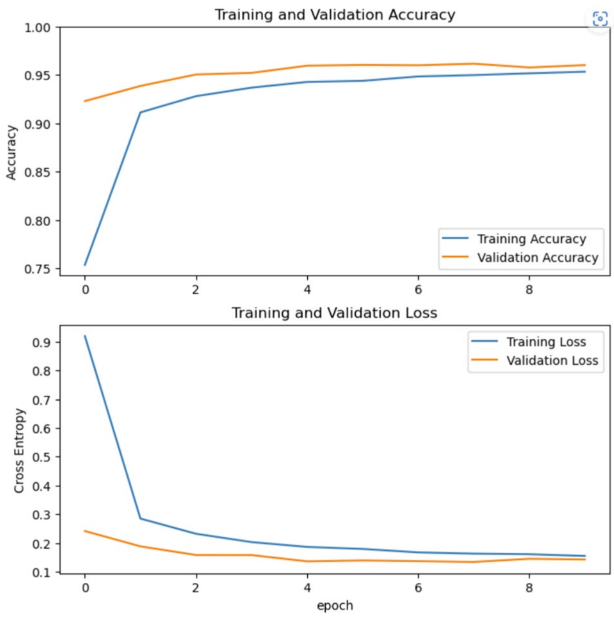

Fig. 12 Performance charts

Conclusions

Handwritten character recognition is a research field in which researchers are continuously contributing with new ideas and techniques to propose systems to show accuracies close to 100%. A hybrid GHCMR system is proposed in this context. The construction and training of this hybrid model has been done and presented. The proposed systems testified outstanding accuracies using sufficient image set. Experimental results concludes that proposed model performs well in classifying Gujarati handwritten characters and modifiers. In this study, a recognition accuracy of 97% for characters and modifiers recognition has been achieved. The proposed method deals with maximally all Gujarati characters with modifiers and tried to work feasibly in real word. The scope is to encompass possibly all characters and all modifiers.

	0	1	2	3	4	5	6	7	8	9	10	13	14	15	16	17	18	19	20	21
0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
2	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
3	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
4	0	0	0	0	0.9	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
5	0	0	0	0	0	0.9	0.1	0	0	0	0	0	0	0	0	0	0	0	0	0
6	0	0	0	0	0	0	0.9	0	0	0	0	0	0	0	0	0	0	0	0	0
7	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
8	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
9	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
10	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
13	0	0	0	0	0	0	0	0	0	0	0	0.9	0.1	0	0	0	0	0	0	0
14	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
15	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
17	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
18	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
21	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

	20	21	22	23	27	40	41	54	55	68	69	82	83	96	97	111	125	139	153	167
20	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
21	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
22	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
23	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
27	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
40	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
41	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
54	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
55	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
68	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
69	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
82	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
83	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
96	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
97	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
111	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
125	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
139	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
153	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
167	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

	153	167	181	195	209	222	223	237	250	251	265	278	279	292	293	306	307	320	321	334
153	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
167	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
181	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
195	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
209	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
222	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
223	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
237	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
250	0	0	0	0	0	0	0	0.1	0.9	0	0	0	0	0	0	0	0	0	0	0
251	0	0	0	0	0	0	0	0	0	0.9	0	0	0	0	0	0	0	0	0	0
265	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
278	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0
279	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
292	0	0	0	0	0	0	0	0	0	0	0	0	0	0.9	0	0	0	0	0	0
293	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
306	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
307	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
320	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
321	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
334	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

	335	348	349	362	363	377	391	404	405	418	419	432	433	446	447	461	475	489	502	503
335	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
348	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
349	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
362	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
363	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
377	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0
391	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
404	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0
405	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0
418	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0
419	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0
432	0	0	0	0	0	0	0	0	0	0	0	0.8	0	0	0	0	0	0	0	0
433	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0	0
446	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0	0
447	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0	0
461	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0	0
475	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0	0
489	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	0
502	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0
503	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1

Fig. 13 Confusion matrix (partial)

Acknowledgements I am extremely grateful to Hon'ble Vice Chancellor Dr. Girish Bhimani of Saurashtra University, Rajkot and Dr. Kishor Atkotiya, Professor and Head, Department of Statistics, Saurashtra University, Rajkot. for encouraging, inspiring and mentoring to pursue this research work. This research work could not have been accomplished without the support of students of Atmiya University, Atmiya School, and their parents. They have contributed by writing Gujarati Characters on white papers of the size A4 containing 6 characters per row and 10 rows per page in preparation of Image Set.

Data availability Image set of examples given in this research paper is available with the link provided in the below section. (http://tinyurl.com/mr2vy7rc

References

Basiri, M.E., Abdar, M., Cifci, M.A., Nemati, S., Acharya, U.R.: A novel method for sentiment classification of drug reviews using fusion of deep and machine learning techniques. Knowl. Based Syst. **198**, 105949 (2020)

BenAtitallah, S., Driss, M., Boulila, W., Ben Ghézala, H.: Randomly initialized convolutional neural network for the recognition of COVID-19 using X-ray images. Int. J. Imaging Syst. Technol. **32**, 55–73 (2022)

Dey, R., Balabantaray, R.C., Samanta, S.: Recognition of handwritten characters from Devanagari, Bangla, and Odia languages using transfer-learning-based VGG-16 networks. Imaging Sci. J. **70**(2), 61–74 (2022). https://doi.org/10.1080/13682199.2022.2163348

Doshi, P.D., Vanjara, P.A.: Image set quality optimization for handwritten Gujarati character and its modifier recognition. In: Proc. 2022 9th Int. Conf. Comput. Sustain. Glob. Dev. INDIACom 2022, pp. 693–697 (2022a) https://doi.org/10.23919/INDIACom54597.2022.9763144

Doshi, P.D., Vanjara, P.A. (2022b). A Comprehensive Survey on Handwritten Gujarati Character and Its Modifier Recognition Methods. In: Joshi, A., Mahmud, M., Ragel, R.G., Thakur, N.V. (eds) Information and Communication Technology for Competitive Strategies (ICTCS 2020). Lecture Notes in Networks and Systems, vol 191. Springer, Singapore. https://doi.org/10.1007/978-981-16-0739-4_79

Harish, B.S., Rangan, R.K.: A comprehensive survey on Indian regional language processing. SN Appl. Sci. **2**, 1204 (2020)

Hashemzehi, R., Mahdavi, S.J.S., Kheirabadi, M., Kamel, S.R.: Detection of brain tumors from MRI images base on deep learning using hybrid model CNN and NADE. Biocybern. Biomed. Eng. **40**, 1225–1232 (2020)

Liaqat, S., Dashtipour, K., Arshad, K., Assaleh, K., Ramzan, N.: A hybrid posture detection framework: integrating machine learning and deep neural networks. IEEE Sens. J. **21**, 9515–9522 (2021)

Liu, R., et al.: A deep learning–machine learning fusion approach for the classification of benign, malignant, and intermediate bone tumors. Eur. Radiol. **32**, 1371–1383 (2022)

Loey, M., Manogaran, G., Taha, M.H.N., Khalifa, N.E.M.: A hybrid deep transfer learning model with machine learning methods for face mask detection in the era of the COVID-19 pandemic. Meas. J. Int. Meas. Confed. **167**, 108288 (2021)

Madushanka, P.T.C., Bandara, R., Ranathunga, L.: Sinhala handwritten character recognition by using enhanced thinning and curvature histogram based method. In: 2017 IEEE 2nd Int. Conf. Signal Image Process. ICSIP 2017, January, pp. 46–50 (2017)

Özyurt, F.: Efficient deep feature selection for remote sensing image recognition with fused deep learning architectures. J. Supercomput. **76**, 8413–8431 (2020)

Patel, C., Desai, A.: Zone identification for Gujarati handwritten word. In Proc. - 2nd Int. Conf. Emerg. Appl. Inf. Technol. EAIT 2011, pp. 194–197 (2011) https://doi.org/10.1109/EAIT.2011.47

Raghunath, D., Chandra, B.R., Jayashree, P.: A robust handwritten digit recognition system based on sliding window with edit distance. In: IEEE International Conference on Electronics, Computing and Communication Technologies (CONECCT) (2020). https://doi.org/10.1109/CONECCT50063.2020.9198385

Singh Rajawat, A., Jain, S.: Fusion deep learning based on back propagation neural network for personalization. In: 2nd Int. Conf. Data, Eng. Appl. IDEA 2020 (2020) https://doi.org/10.1109/IDEA49133.2020.9170693

Ting, P., et al.: Freeway travel time prediction using deep hybrid model-taking Sun Yat-Sen freeway as an example. IEEE Trans. Veh. Technol. **69**(8), 8257–8266 (2020)

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.