Handwritten Digits Classification Through Multi-Classifier Bag Of Visual Words

R.Hariharan ,M.Dhilsath Fathima, Vibek Jyoti Roy, Yasir Ahmad Khan

Abstract: ***Today our world moving towards the smart technology in many ways. In this smart world we are making everything easy. Instead of typing with hand we can convert our hand-written letters to text format. There are many technology's available to recognize handwriting and convert into text format, but still many cases are getting flaw in accurate prediction. Many machine learning classifiers available for recognize and classifies hand written digit. Bag of visual words is one of the simple classification method. Bag of features is detecting the surface and finding features of image and creating a vocabulary with visual words. This paper propose a multi-classifier bag of features methodology to identify hand-written digits.***

Index Terms**: Bag of visual words ,Bag of features ,HDR, multi classifiers,**

I. INTRODUCTION

Handwriting recognition is the ability of a computer to recognize the natural hand-written characters and other symbol by pattern recognize process. It involves the automatic conversion of handwriting text in an image into letter codes and vice versa which we can do process in text processing application within computer, so that it helps to increase the accuracy of our handwriting. It's very common to raise a question why we should use handwriting recognition instead of typing? When we write something by hand, we can spread our strong conceptual thinking which we implement in paper that makes the paper unique, understandable and gives a clear concepts to reader than by typing. As a input, handwriting data is converted to digital form by scanning the writing paper. Conventional techniques center on spreading each characters for recognition but in modern techniques focus on cognizing all character in a line of text. Online handwriting recognition implicate the automatic transformation of text but in offline recognition at first the system scan the text which is written on a paper then converted into a thin line and broken down by a system, allocating letters and their position in a words and also look for split, where the line changes from a letter into an attachment. The algorithm works to replicate handwriting style by allusion scanned examples. For same letter the system will search the best one from which can get best result for the word. For an example ' vibek yasir khan ' have been selected, the system puzzle out the exact gap in between each letter and also identify the height and position also.

Revised Manuscript Received on May 10 ,2019

HARIHARAN.R, Department of Information Technology , VelTech Rangarajan Dr Sagunthala R&D Institute of Science and Technology, Avadi ,Chennai, India

DHILSATH FATHIMA.M, Department of Information Technology , VelTech Rangarajan Dr Sagunthala R&D Institute of Science and Technology,Avadi ,Chennai, India

VIBEK JYOTI ROY, Department of Information Technology , VelTech Rangarajan Dr Sagunthala R&D Institute of Science and Technology, Avadi ,Chennai, India.

YASIR AHMAD KHAN, Department of CSE, VelTech Rangarajan Dr Sagunthala R&D Institute of Science and Technology, Avadi ,Chennai, India

II.ORGANIZATION OF PAPER

This paper discussed about introduction about HDR and multi image classifiers and usage. How other classifiers are used in various papers and what's pros and cons. Data set collection process. And how to evaluating design model and what all the metrics and its description. Discussion about two Machine learning classifiers and working methodology. Proposed bag of visual words working methodology and architecture of designed model. Detailed result analysis.

III.LITERATURE SURVEY

[4]For Kannada characters a standard database was made available. We have to give our own Kannada and English characters as present characters from different persons like school, colleges, and common people. After segmentation by using connected component techniques, databases stored into bmp format. All segmented characters are binaries by using morphological opening and closing operations and broken characters are connected by using morphological dilation operation. All binaries characters are normalized into 32x32 dimensions. Then feature extraction method is applied where all preprocessed characters images are extracted. Then they are divided into 64 zones or none overlapping. Then computation of pixel is performed for each zone. Based on their experiments 5100 character images of Kannada and English were classified with a recognition accuracy of 83.02%.[3] Nowadays the offline character recognition is an active research area. The work by the researcher in the handwritten recognition is very limited. In[9] 2002, Kundu & Chen used HMM to recognize 100 postal words and reported 88.2% recognition accuracy. In 2007, Tomoyuki et al. used 1646 city names of European countries in the recognition experiment and accuracy of 80.2% is achieved. In 2006,[10] Gatos et al. used K-NN classifier to recognize 3799 word from IAM database and reported 81% accuracy. The process of scanning the character image is known as image acquisition. Then all the handwritten character image are converted into '.bmp' or '.jpg' for processing next character samples of 10 different people age of 15-50 years are collected where each contributor provide 5 samples of the English alphabet. In this way 10*5*26=1300 characters image are collected.

Pre-processing is done in offline by selecting more and accurate handwritten characters. In pre-processing input image format is converted into grayscale format .Binarization is an important image processing step by which we are able to minimize the unwanted information and details present in the image. Here they prepared a training sample and extracted features. The binary image is resized and reshaped .In this experiment, 26 characters image is of size 180*1, where 180 neurons are used is the input layer and 26 neurons are used in the output layer neural network classifies. For better results 80 neurons were kept in hidden layer by trial and errors method. This method gives an outstanding accuracy of 85.62%.[5]Some important steps closely connected in offline handwritten digit recognition (HDR). 'Image scanning' is the process of scanning an image as n input image through digital input device in JPEG, BMP format. After that in 'pre-processing' input image format is converted into grayscale format and also smoothing, filtering, resizing occurs so that it can be made more accurate. In segmentation part the digital image is divided into multiple segments to simplify to make it more meaningful and easier to analyses. Image is extracted to reduce its dimension to avoid confusion about lot of features, so that in image processing algorithm is used to extract more relevant features from the image. After extracting the image neural network is used for classification of the digit. Initially create the neural network, then using training data train the network and then bring to consciousness the network with the input data. [1]OCR allows converting mechanical or electronic image based text into machine en-codable text through an optical mechanism. The main feature of OCR is that it can stimulate the human reading capabilities which in turn makes understandable for the computer to read and understand and even edit it. The authors experimented by two types of characters. One is isolated image character and the other is sentential. In isolated character recognition they experimented on 62 English characters with four types of font. They found it very interesting with a huge margin of success rate up to 91.53%. In sentential case character recognition, they used a sentence with again 4 different fonts where they were able to get a success rate of 80.65%. They came to the conclusion that success rate for sentence fall down because of image segmentation, the character edge is cut from the real shape of character. A human eye is able to see object through eye but image characterization by the machine is a herculean task which was made possible through them. The accuracy can be increased by the design of optimal neural network architecture using number of neurons.

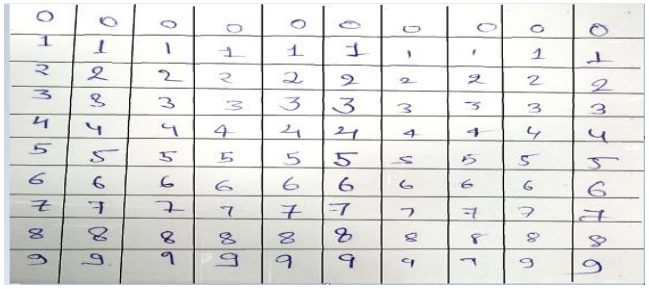


FIG 1: SAMPLE DATASET

IV.DATA SET

Real time Data set has created and data collected from 100 different people. 20 samples of each digits from each people in different situation. That's person A given 20 samples of 0, 20 samples of 1 till 20 samples of 9 from different timing. The same has repeated for different 100 people. Total data set size is 20000 with 2000 samples in each digits. Fig 1.Shown sample data set

V. METRICS

Designing a new model is a biggest and innovate task. More than design our model we have to clime like in which aspect is the best. For that we need to evaluate our model with some important metrics as Discussed in table 1.

VI. PROPOSED MODEL

Today we are analyzing everything in a scientific manner using images. We having many algorithm to find and split various portion of digital image that's particularly we calling character recognition. There are many method available for character recognition like edge detection, corner detection, blob detection, ridge detection, scale invariant feature transform. Based on these methods we can recognition the character. In image classification feature extraction is main phase. Feature extraction is start from beaning of the image processing to extract the information [6] from the image and selecting feature is the process of dimensionality reduction process it means the data store as a raw variable is cluster into more reasonable group for further processing. When processing the data it won't change in original source. If the data should be more that case they should process each and every time as algorithm has processed each data but if we see step by step some same process will repeating till the ends of batch process. The algorithm repeating for same set of data again and again for same data to avoid the repetitions we can minimize the feature using the feature vector selection when selecting the feature we should not us lose important features that desired task performed by many ways but principal component analysis (PCA) is giving best performance. In this paper we are proposing a method creating the visual word using multi-classifier First we will discuss about classifier.Classifier is use to perform classification problem. Classification is nothing but predicting the target with the help of training data. There is main difference between regression and classification. Regression is nothing but working with continuous data but in classification the target is categorical. There are many classifier available like a logistic classifier KNN, SVM, decision tree, random forest and so on. Every classifier having some its own pros and cons but in this visual word creation we are using SVM and KNN.

VI.1 MODEL DESCRIPTION

FIG 2 . the modules is listed 1.preprocessor 2.clustering to same class 3.Dividing testing and training data 4.creating bag of visual words 5.traing the classifiers.6. Evaluating metrics. With the reference of figure step 1. Is prepossessing the image our input data is collected from multiple

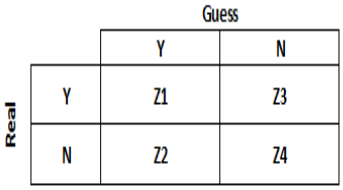
1	Confusion Matrix		Confusion matrix is a matrix formatted output which detail describe the performance of the system. Consider Binary classification we have class Y and N and with our model train and test our dataset and the final result value is segregated as four different category Z1,Z2,Z3,Z4 based on this classes we can measure our results. Z1- if real data belong Y class and model guessed as Y class Z2- if real data belong to N class and model guessed as Y class Z3- if real data belong to Y class and model guessed as N class Z4- if real data belong to N class and model guessed as N class
2	Accuracy	$\frac{Z1+Z4}{Z1+Z2+Z3+Z4}$	Accuracy is defined as how many samples predicted correctly from total number of samples.
3	Sensitive /Recall	$\frac{Z1}{Z1+Z2}$	Sensitivity otherwise called as true positive rate defined as How much guessed as Y and how much is really y
4	Specificity	$\frac{Z3}{Z3+Z4}$	Is also called as false positive defined as How much guessed as N and how much is really N
5	Precision	$\frac{Z1}{Z1+Z3}$	Number of true positive with true positive and false positive
6	F1Score	$\frac{2*Precision*Recall}{Precision+Recall}$	Calculation for F1 formulation given

TABLE 1

multiple person and it was in different size and different color so we resizes the images to 27 x 27 x 1 that is column and rows with pixel size of 27 and the colure scale in black and white. Now we having pre-processed data. But all 20,000 have in same label that we have to cluster based on their classes. In our case we have 10 class of data which contain 0 to 9.every class have 2,000 data which have same digital value with different person handwriting. All 2000 data has classified based on class. Now the data has to divide in two parts that's is testing data and training data, generally testing and training ratio can be 30:70. In our case we tried with 30:70 as well as 20:80. We have achieved better accuracy in 20:80 comparing to 30:70. So we have to randomly divide our data set 20% for testing and 80% for training. Now we have to train our data set with bag of visual words with its label. The figure (3) shows how the bag of visual words are created.

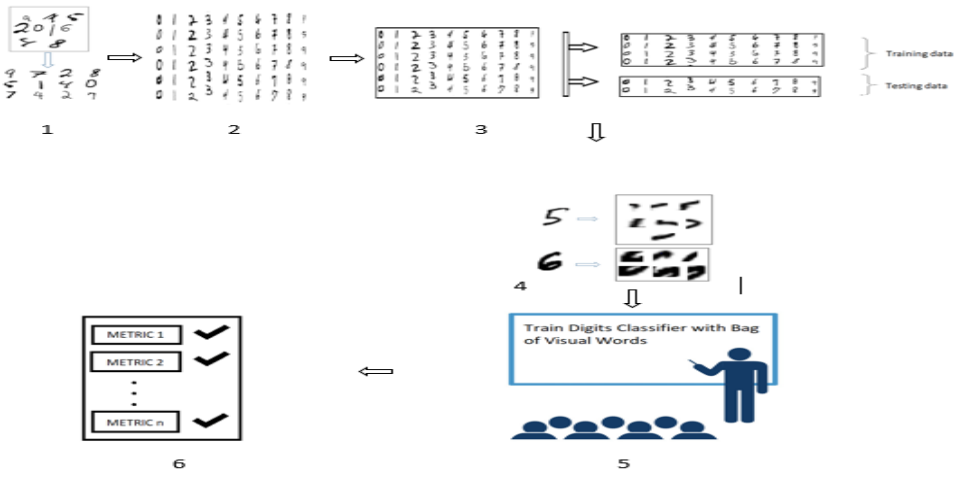


FIG 2: MODEL DESIGN

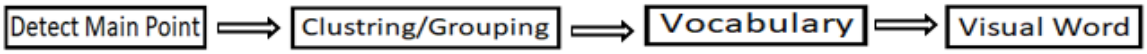


FIG:3: CREATING BAG OF VISUAL WORDS

VI.2.CREATING BAG OF FEATURES

The input image is taken and it detects the image surface and extracting the features and its key point [2][7][8]. The key point is nothing the important features are changing in the pattern. In our case we extracted 2700 features, among that keeping approximately 80% of strongest features from each category that is 2160 features has detected and kept in a temporary memory. The surface features detection is one of the best features detection method. The given image is divided in grid format and each detected features has noted. All the features has keep in an array format and store in a vector format that's also called as features descriptor. Based on the homogenous data cluster centroid is initialized and homogeneous features move towards it the centroid of homogenous characters. Now we have a 530 cluster. Cluster has done in 9/100 iteration (0.02 sec/iteration). Now the vocabulary has created similar features and based on vocabulary visual words has created.

We have many classifiers for training our data like KNN, SVM, Decision Tree and Random Forest. In our model we have proposed multi classifier that is the combination of SVM and KNN are using.

Support vector machines is linear classifier. It's a discriminate classifier (discriminate means its classifies different type of data) support vector is potent at classification, pattern recognition task & numeral prediction.

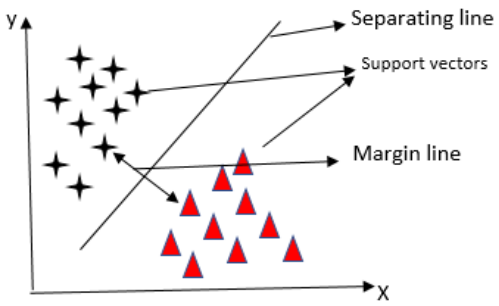


FIG 4: SVM

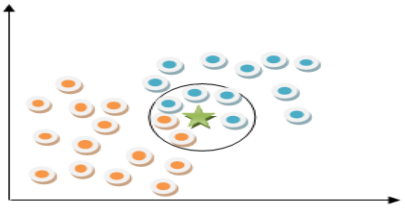


FIG 5:KNN

In above figure FIG 4.it can be seems , here are two type of support vector are given one is in form of star (class-1) and another is in form of delta (class-2), to separate these here separating line is using to separate both the class, distance between either side the nearest point is maximized. The maximizing distance is called margin, the line between two points is called as margin line.

KNN is the way of searching nearest possibilities. It works based on nearest distance from K. It's one of the simplest classification algorithm. It will predict the possibilities which part is the majority carrier within a certain distance from k. Here, the star is the unknown input element 'k'. If we consider k=5 then it will predict the given element by comparing with nearest 5 element of 'k'. In this graph we have marked the nearest 5 element with a circle. In this circle the majority carrier is blue .Then it will predict that the given element is blue.

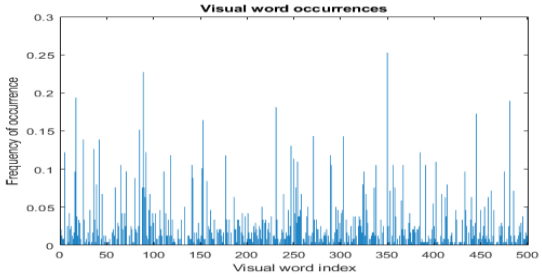


FIG: 6 VISUAL WORD OCCURRENCE

The features are creating a frame work with SVM classifier to train data. In SVM all the homogeneous data clustering based on same features. After that using KNN all the homogeneous group of training data its move towards its adjacent neighbor and it will create the feature histogram of word count VS visual word index. In this feature histogram (Fig 6) we are having the cluster of visual word and its numbers of occurrence. Its feature has detector and stored in array vector format. Now repeat that previous step for n-number of times (n-number of training data) and train our classifier with that digit data and its class label [2][7][8]. After that test our data set with 20% of random testing data. In our model is giving 98% of accuracy and our model is undergone with few metrics which is discussed in result

VII. RESULTS

With reference of confusion matrix(fig.7) the prediction level of our model is detail given, all '0' are predicted as '0',all given '1' as predicted as '1',97% of '2'is predicted as '2', 3% of '2' is predicted as '3',97% of '3' is predicted as'3',3% of '3' is predicted as '5', all '4' is predicted as '4',87% of '5' is predicted as '5',13% of '5' is predicted as '3',all '6' is predicted as '6',all '7' is predicted as'7',all '8' is predicted as'8',all '9' is predicted as '9'. 0,1,4,6,7,8,9 is predicting 100% and rest of the numbers are predicting maximum level that is 2,3 with 97% and 5 with 87%. From all of the prediction we are getting average accuracy level of 98% it is

PREDICTED

KNOWN	0	1	2	3	4	5	6	7	8	9
0	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0.00	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
2	0.00	0.00	0.97	0.03	0.00	0.00	0.00	0.00	0.00	0.00
3	0.00	0.00	0.00	0.97	0.00	0.03	0.00	0.00	0.00	0.00
4	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00	0.00	0.00
5	0.00	0.00	0.00	0.13	0.00	0.87	0.00	0.00	0.00	0.00
6	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00	0.00
7	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00	0.00
8	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00	0.00
9	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	1.00

FIG 7: CONFUSSION MATRIX

more than user expectation level (95%). From the confusion matrix we are understanding very minimum error level that is 0.2% we getting through our model. The rate of sensitivity is 98.2% we are getting. Our satisfying other metrics also. Comparatively our model is satisfying minimum expectation level with previous available model.

VIII.CONCLUSION

As we have discussed many machine learning algorithm for handwritten digit recognition. Many classifier are predicting handwritten digit, even they not able to achieve the expected accuracy (95%). This paper we have used bag of visual words using multi classifier. When we are using single classifier we have received average prediction level, but when we used with multi classifier with bag of feature, given more accuracy what we have expected that we have already discussed in result part. With the reference of Result discussion we can assure that our model is getting low error rate and high accuracy rate and give best performance in few metrics that is discussed. This multi classified satisfying the user expectation.

REFERENCE

1. Md. Abdullah-al-mamun,Tanjina Alam ,"An approach to empirical Optical Character recognition paradigm using Multi-Layer Perceptorn Neural Network ",18th internatrional conference on computer and information technology (ICCIT),21-31 December,2015,PP 132-137.
2. Neetika Singhal ; Nishank Singhal ; V. Kalaichelvi,"Image classification using bag of visual words model with FAST and FREAK" 2017 Second International Conference on Electrical, Computer and Communication Technologies (ICECCT),DOI: 10.1109/ICECCT.2017.8117861
3. Amit Choudhary,Rahul Rishi, Savita Ahlawat,"Off-Line Handwritten Character Recognition using Features Extracted from Binarization Technique "AASRI Procedia volume 4 ,P.no 306 – 312,2013.
4. Gururaj Mukarambi, Dhandra B. V, and Mallikarjun hangarge (2012), "A zone Based Character Recognition Engine for Kannada and English Scripts", Procedia Engineering, Vol. 38, pp. 3292-3299.
5. J. Pradeep, E. Srinivasan and S.Himavathi,"International Journal of Computer Applications (0975 – 8887) Volume 8– No.9, October 2010 17 Diagonal Feature Extraction Based Handwritten Character System Using Neural Network", vol.8-no.9,pp 17-22, 2010.
6. Gaurav Kumar ; Pradeep Kumar Bhatia," A Detailed Review of Feature Extraction in Image Processing Systems",2014 Fourth International Conference on Advanced Computing & Communication Technologies,DOI: 10.1109/ACCT.2014.74
7. K. Kesorn et al., "An Enhanced Bag-of-Visual Word Vector Space Model to Represent Visual Content in Athletics Images", IEEE Transaction on Multimedia, Vol. 14, No. 1, 2012, pp. 211- 222.
8. F. Jie et al., "Bag-of-Visual-Words Based on Clonal Selection Algorithm for SAR Image Classification" , IEEE Geoscience and Remote Sensing Letters, Vol. 8, No. 4, 2011, pp. 691-695.
9. Kundu, Y. H., Chen, M., 2002. "Alternatives to variable duration HMM in handwriting recognition", IEEE Trans Pattern Anal Mach Intell, 20(11), pp. 1275–1280.
10. Gatos, B., Pratikakis, I. & Perantonis, S. J., 2006. "Hybrid off-line cursive handwriting word recognition", in proceedings of 18th international conference on pattern recognition (ICPR'06), 2, pp. 998–1002.

AUTHORS PROFILE

Mr.R.Hariharan working as Assistant professor in Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology with 5 years of experience. He is a member in IEEE and CSI. .His research area is image processing and deep learning. He has published 13 papers in various international journal and conference.

Ms.Dhilsath Fathima.M has been working as an assistant professor in Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology with 8 years of professional experience in teaching & learning. Some of her thrust areas are Data Mining, Machine learning, deep Learning, and image processing.

Mr.Vibek jyoti roy doing is bachelor of technology in Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, his area of interest is image processing and green computing. He has presented 2 papers in national level conference.

Mr.Yasir ahmed khan doing is bachelor of technology **in** Vel Tech Rangarajan Dr.Sagunthala R&D Institute of Science and Technology, his area of interest is image processing and deep learning. He has presented 1 paper in national level conference.