Handwritten Text Recognition

Project Reference Number: 46S_BE_4772

Nitte Meenakshi Institute Of Technology, ECE Department

Alka Ashok 1NT19EC011

Haripriya Reddy 1NT19EC057

Under the guidance of Dr Parameshachari B D

Keywords—HTR(Handwritten Text Recognition), CNN(Convolutional Neural Networks), SVM(Support Vector Machine).

Introduction

Human beings are different from computers in terms of intelligence. Humans can do various tasks that machines are not able to perform without supervision. Handwritten text recognition is one such example. Over the past few years, there has been a dramatic growth in the number of handwriting methods developed to identify words in written content. However, the accuracy in recognizing characters and its efficiency are still not up to the mark for commercial use. Due to the various and unique styles of handwriting recognition algorithms don't perform well on various types of handwriting samples. With the help of artificial intelligence (AI), humans and computers are becoming more and more similar. Computer vision is one such discipline. Making computers behave and perceive like people is the main goal of this field. Handwritten text recognition (HTR) remains one of the most popular research areas due to the many benefits that can be derived from recognizing text accurately. The introduction of Deep Learning (DL) has brought substantial strides in the science of handwritten text recognition. A plethora of neural networks are proposed for the recognition but none of them seem to have the accuracy required for commercial use cases. Only a handful of research is performed in the recognition of South Asian languages. This project uses Convolutional Neural Networks and Support Vector Machines because it was indicated from the previous research that they are the most powerful learning algorithms that exist today and give over 90% accuracy.

Objectives

A comparison of CNN and SVM models for recognizing Kannada numbers is proposed, which includes automatic feature generation using CNN and output classification and prediction using the SVM classifier. The proposed model can be improved in the future to recognize handwritten characters in different languages such as Tamil, Kannada, Punjabi, Bengali, Gujarati, and so on. Later many letters, words, and sentences can also be predicted with high accuracy in various languages. The

objective is to recognize handwritten documents, which include characters, words, lines paragraphs, etc. There is extensive work in the field of handwritten recognition and a number of reviews exist. In this approach, it is required to recognize handwritten Kannada numbers.

Methodology

The suggested approach in this paper is to build a Convolutional Neural Network and an SVM classifier and compare them using the same dataset. CNN can learn invariant local features very effectively and functions in a similar way to how humans do. From unprocessed digit pictures, it can extract the most discriminating information. In an area where data items from various classes are divided by a hyperplane, Support Vector Machine (SVM) represents multi-dimensional datasets. On untried data, the SVM classifier has the capacity to reduce the generalization error. The term "optimal hyperplane" also refers to the separating hyperplane. SVM is deemed to be terrible for noisy data and proven to be good for binary classification. Learning deep features can be challenging due to SVM's thin architecture. There is a split between "training data" and "testing data" in the collection of data. The photos in this collection have been normalized and resized to a standard 28x28 pixels before being included. Once the dataset is split, grayscale normalization is performed to reduce the effect of illumination differences. In this project 28*28*1 image dimensions and kernel size of 3*3 are used. For CNN the kernel size of 3*3 and activation function ReLU is used. CNN will automatically extract the features and forms feature maps. A 2*2 pooling layer will be added for downsampling.

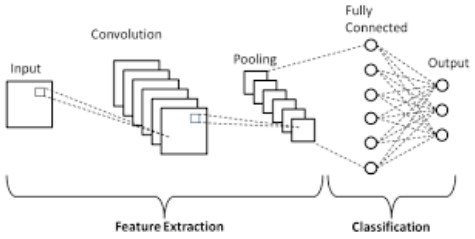

Fig. 1. CNN Basic Architecture

Issues with classification are where SVM really shines. The support vector machine (SVM) method aims to locate the best hyperplane (or decision boundary) for classifying the information at hand. The best feasible choice boundary is a hyperplane. The SVM kernel is a function that takes an input space with few dimensions and maps it to a larger-dimensional output space, so turning a non-separable problem into a separable one. Nonlinear separation problems are where it really shines. To put it plainly, the kernel undergoes incredibly sophisticated data alterations before deciding whether to partition the information depending on the labels or outputs provided.

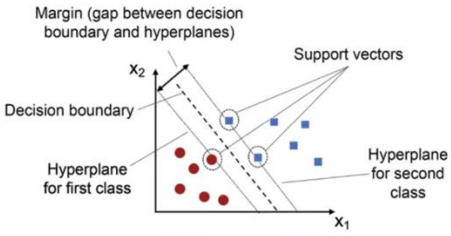

Fig. 2. Support Vector Machine

Results and Conclusions

In this paper, models of CNN and SVM model are proposed for recognizing Kannada numbers The proposed model compares the dataset using CNN and SVM. On performing the experiment, it is found that CNN has a higher accuracy than SVM. The dataset that is being used is from Kaggle. The research on CNN and SVM models is improving and growing very popular.

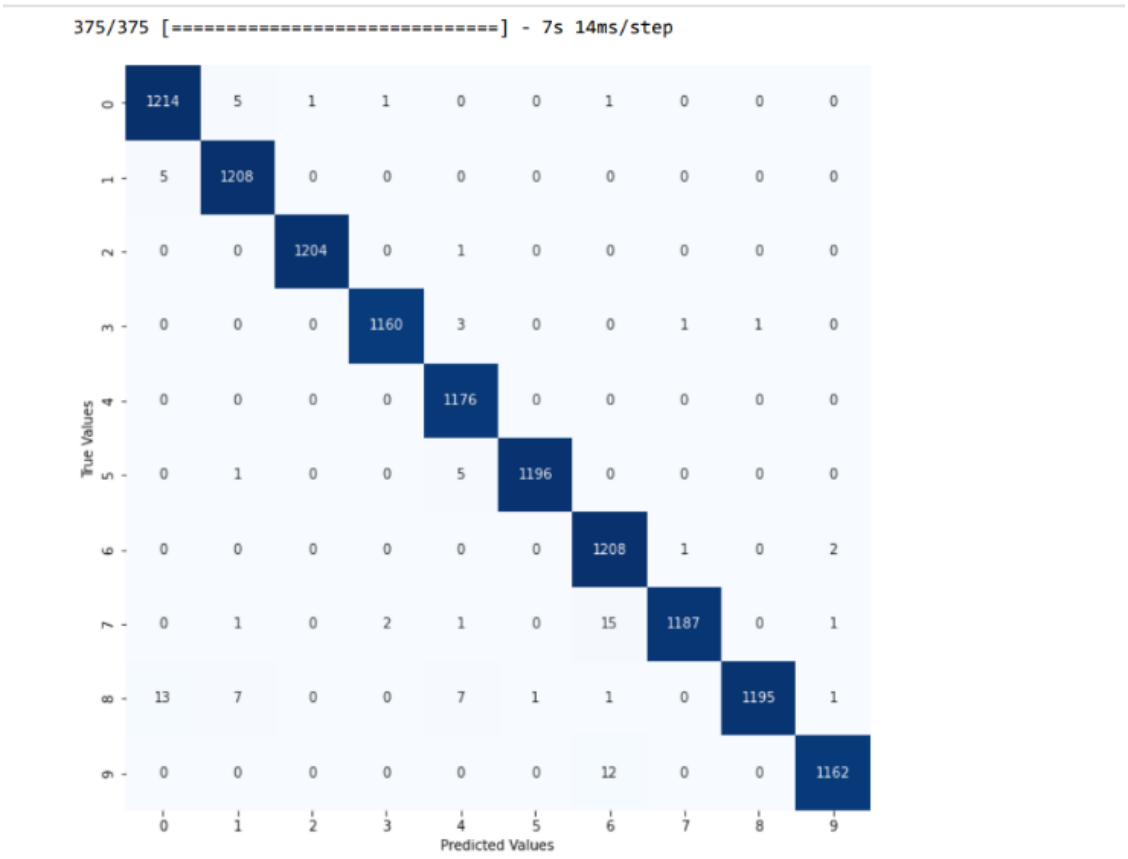

Fig. 3. Confusion Matrix for CNN to evaluate model performance on ten different target values (0-9)

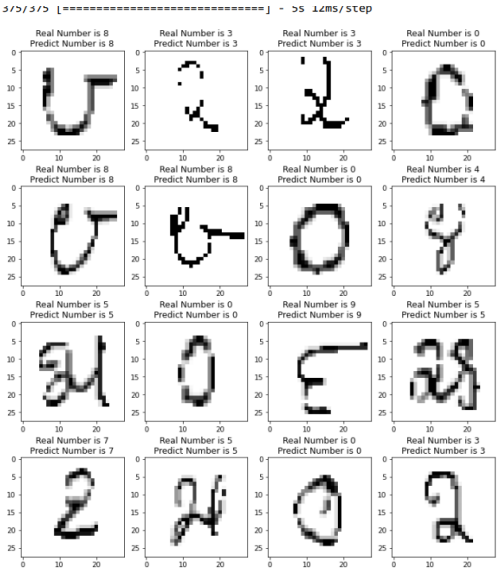

Fig. 4. The final results.

	Precision		Recall		f1-score		Support	
	CNN	SVM	CNN	SVM	CNN	SVM	CNN	SVM
0	0.995122	0.990338	0.974522	0.979299	0.984714	0.984788	628.0000	628.0000
1	0.972359	0.979032	1.000000	0.998355	0.987013	0.988599	608.0000	608.0000
2	1.000000	1.000000	0.991817	0.993453	0.995892	0.996716	611.0000	611.0000
3	0.994690	0.977273	0.984238	0.978984	0.989437	0.978128	571.0000	571.0000
4	0.996575	0.988115	0.994872	0.994872	0.995723	0.991482	585.0000	585.0000
5	1.000000	0.993344	0.996656	0.998328	0.998325	0.995830	598.0000	598.0000
6	0.990244	0.963434	0.988636	0.983766	0.989439	0.973494	616.0000	616.0000
7	0.984000	0.977049	0.990338	0.959742	0.987159	0.968318	621.0000	621.0000
8	0.998305	0.998299	0.998305	0.994915	0.998305	0.996604	590.0000	590.0000
9	0.982759	0.996448	0.996503	0.980769	0.989583	0.988546	572.0000	572.0000
accuracy	0.991500	0.986167	0.991500	0.986167	0.991500	0.986167	0.9915	0.986167
Macro avg	0.991605	0.986333	0.991589	0.986248	0.991559	0.986251	6000.0000	6000.0000
Weighted avg	0.991584	0.986244	0.991500	0.986167	0.991503	0.986164	6000.0000	6000.0000

TABLE 1. Comparison between CNN and SVM

There is need for improvement in the suggested framework for the identification of handwritten characters in a variety of spoken languages, including but not limited to Tamil, Kannada, Punjabi, Bengali, Gujarati, etc. To further improve categorization results, certain optimization methods may be explored.

Scope for Future Work

In this paper, models of CNN and SVM model are proposed for recognizing Kannada numbers The proposed model compares the dataset using CNN and SVM. On performing the experiment, it is found that CNN has a higher accuracy than SVM. The dataset that is being used is from Kaggle. The dataset was found to be a clean dataset. The research on CNN and SVM models is improving and growing and

is very popular. There is a need for improvement in the suggested framework for the identification of handwritten characters in a variety of spoken languages, including but not limited to Tamil, Kannada, Punjabi, Bengali, Gujarati, etc. In future work, we intend to improve the work by making use of hybrid datasets and experimenting with different activation functions, also increasing the number of neural network layers. To further improve categorization results, certain optimization methods may be explored.