PROJECT REPORT

TRAFFIC SIGN RECOGNITION using ARTIFICIAL NEURAL NETWORK

Submitted to

IIT JODHPUR

In Partial Fulfilment of the Requirement for the Award of

BACHELOR'S DEGREE IN COMPUTER SCIENCE AND ENGINEERING

BY

MAKARAND GOMASHE UG201310019

NIKHIL TAJI UG201310021

UNDER THE GUIDANCE OF

Dr. CHIRANJOY CHATTOPADHYAY

|| त्वं ज्ञानमयो विज्ञानमयोऽसि ||

DEPARTMENT of COMPUTER SCIENCE and ENGINEERING
IIT JODHPUR

Abstract

To understand Neural Network and applying the concepts to model the dataset of Traffic Signs, and henceforth recognizing the test images using the trained model.
The approach follows an efficient image recognition using properly trained artificial neural network implementation.

Problem Statement

To identify and recognize varied Traffic signs affected by impact of nature (dust or distortions) or illumination using Artificial Neural Network model to improve recognition capabilities and analyzing the accuracy and metrics on different parameters.

Introduction

Traffic Sign Recognition gains its importance and motivation as it is still a challenging task to recognize Traffic signs in various natural background viewing conditions of weather and illumination. Moreover, traffic sign recognition has computational difficulties with the performance of system in real time.

The project aims at efficient recognition of traffic signs using properly trained implementation of neural network. Initial phase of understanding involved exploring machine learning and its several techniques. As the widely stated, **Machine learning** is a subfield of computer science that evolved from the study of pattern recognition and computational **learning** theory in artificial intelligence. It can be applied in several ways such as regression, classification or prediction. This project directly deals with classification and identification of a test image from one of the trained Traffic sign images. It involves extracting the unique features from the data of interest (Set of traffic signs pre-classified, also known as supervised data) which could be used in mapping the dataset trough the neural network model and henceforth identifying the test images.

The detailed procedure followed is mentioned in the methodology section.

Brief Literature Survey

S.NO	TITLE OF PAPER	AUTHOR	JOURNAL/ CONF.	YEAR OF PUBLICATION
1	Robust Method for Road Sign Detection and Recognition	Piccioli et al.	*Image and Vision Computing 14*, pp.208-223.	1996
2	Using Color to Detect, Localize and Identify Informational Signs	Yuille et al.	*Proc. International Conference on Computer Vision ICCV98*, Bombay, India, pp. 628-633.	1998
3	Road Sign Classification Using Laplace Kernel Classifier	Novovicova et al.	*Pattern Recognition Letters 21*, pp. 1165-1173.	2000
4	The German Traffic Sign Recognition Benchmark	Johannes Stallkamp et al.	*International Joint Conference on Neural Networks*	2011

Methodology

The project was worked out in following phases :

I. **Data Set Identification**

A good Dataset for developing a Machine Learning model is a primary but most important task. A good dataset is one with varied possible training images such that the ML model is able to learn the complete possible variations and henceforth correctly map the data and increasing the accuracy.

The Datasets used in this project are :

1. German Traffic Sign Recognition Benchmark, GTSRB dataset (39,209 images in 43 classes)
2. Belgium Traffic Sign Dataset (7125 images in 63 classes)

	Training	Testing	Classes	Diversity
GTSC	26640+12569	12630	43	~30 images/sign
BTSC	4591	2534	62	~3 images/sign

Output of this phase : Images for Training the model

II. **Feature Extraction**

Raw Images are matrices of RGB components in the pixels. All of which do not segregate of uniquely identify a particular image. Hence, identifying and extracting those particular features and representing them quantitatively is important process. This phase involved processing the images obtained from phase one. This dataset was processed using Caffe for feature extraction.

Caffe is an open-source neural network library developed in Berkeley, with a focus on image recognition. It can be used to load one of the pretrained models.

BVLC GoogLeNet model will be loaded into caffe to get required image feature vectors.

Also, three datasets of pre extracted feature vectors(Hue, HOG and Haar) of GTSC dataset will also be considered for analysis.

Output of this phase : Five datasets of corresponding Feature Vector of images with their Classification labels

III. **Training the Model**

This phase involved creating the trained Neural Network model using the image vectors obtained in the previous phase. The training set with the classified labels was given for mapping to the Neural Network model.

Scikit-learn package in python was used to apply the Multi-Layer Perceptron. For this MLP, L-BFGS algorithm was used for implementation.

Output of this phase : Trained Neural Network

IV. **Testing the Model**

This is the important phase where the developed Neural Network is tested for its classification. The test set is given to model and corresponding classification of the image is stored as vector of test outputs. This is also termed as validation.

Output of the phase : Validated Neural Network and data for metric analysis

V. **Analysis of the Model**

This phase analyses the model for its accuracy. The test set is given to the Model and the outputs obtained are verified. Accuracy can be calculated as :

Accuracy = True Positives / Total Data

The analysis is covered in the next section of Results and Conclusion.

Metrics Used

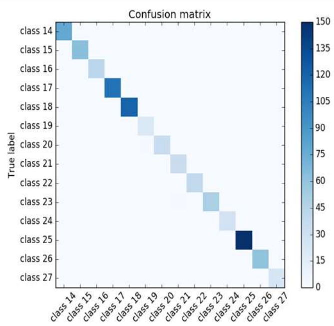
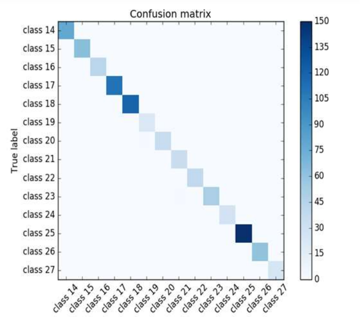
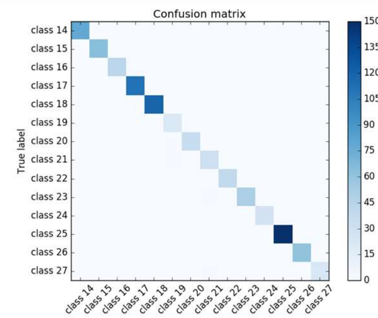
1. Accuracy
2. Precision
3. Recall
4. F1-score
5. Support
6. Confusion Matrix

Results and Conclusion

<u>Observation 1 :</u>

Dataset 1 : GTSC caffe extracted

Training : (3135,4097)
Testing : (855,4097)
Parameters : {}

Hidden Neurons = 1500	Hidden Neurons=2500	Hidden Neurons=4000
 Fig. Confusion Matrix with Predicted outputs on horizontal and True outputs on vertical axes with 14 different classes.	 Fig. Confusion Matrix with Predicted outputs on horizontal and True outputs on vertical axes with 14 different classes.	 Fig. Confusion Matrix with Predicted outputs on horizontal and True outputs on vertical axes with 14 different classes.
Precision = 1 Recall = 1 F1-score = 1 Support = 855	Precision = 1 Recall = 1 F1-score = 1 Support = 855	Precision = 1 Recall = 1 F1-score = 1 Support = 855
Accuracy = 99.87 %	Accuracy = 99.766 %	Accuracy = 99.53 %

<u>Conclusion 1 :</u>

This Germany dataset has feature vector of length 4096 as given by GoooogleNet model through Caffe. This instance of training has 3135 images in the training set and 855 images in the testing set(475 from the training and rest half new unknown vectors). This data belongs to 14 different Traffic signs (14 Classes).

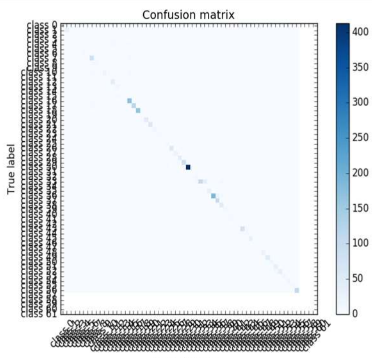
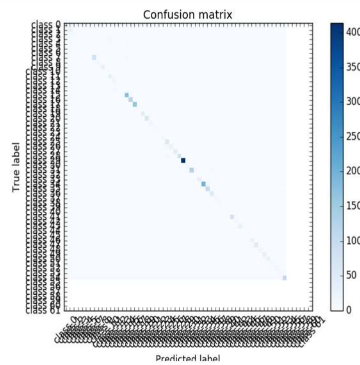
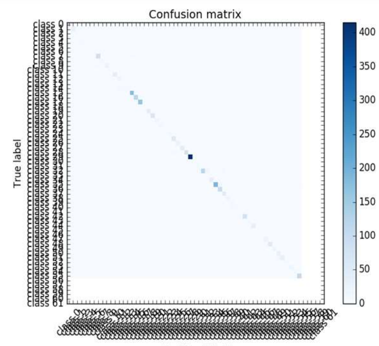
When such image vectors are provided as input to Neural Network, the model performs good with high accuracy > 99% henceforth high precision of 1. A minute increase is seen on reducing number of hidden layer nodes.

This high accuracy can be attributed to good dataset selection, proper feature extraction and high precision mapping done by the neural network during the training.

Observation 2:

Dataset 2 : BTSC Caffe Extraced

Training : (4575,4097)
Testing : (2520,4097)
Parameter []

Hidden Neurons = 500	Hidden Neurons = 2000	Hidden Neurons = 4000
 Fig. Confusion Matrix with Predicted outputs on horizontal and True outputs on vertical axes with 62 different classes for predictions.	 Fig. Confusion Matrix with Predicted outputs on horizontal and True outputs on vertical axes with 62 different classes for prediction.	 Fig. Confusion Matrix with Predicted outputs on horizontal and True outputs on vertical axes with 62 different classes for predictions.
Precision = 0.93 Recall = 0.91 F1-score = 0.91 Support = 2520	Precision = 0.93 Recall = 0.92 F1-score = 0.92 Support = 2520	Precision = 0.94 Recall =0.92 F1-score=0.92 Support = 2520
Accuracy = 91.03	Accuracy = 91.94	Accuracy = 92.063

Conclusion 2 :

This Belgium dataset has feature vector of length 4096 as given by GooogleNet model through Caffe. This instance of training has 4575 images in the training set and 2520 images in the testing set(475 from the training and rest half new unknown vectors). This data belongs to 62 different Traffic signs (14 Classes).

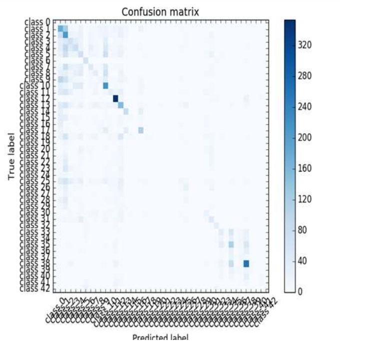
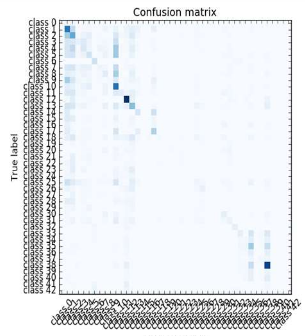
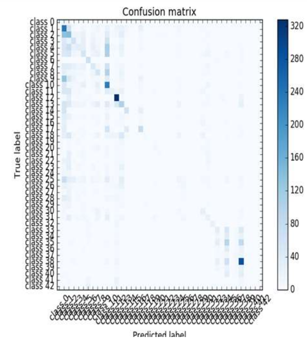
On providing such image vectors as input, the model sows a significant good result of accuracy>90%. It performs even better on increasing the number of hidden layer nodes. Increase is seen in recall and f1-score as wel

This accuracy could be increased slightly more by providing a larger training data and maintain a ratio of 1:3 for validating a test dataset.

Observation 3 :

Dataset 3 : BTSC Hue Feature Set

Training : (30061,256)
Testing : (7842,256)
Parameters {}

Hidden Neurons = 100	Hidden Neurons = 250	Hidden Neurons = 500
 Fig. Confusion Matrix with Predicted outputs on horizontal and True outputs on vertical axes with 43 classes.	 Fig. Confusion Matrix with Predicted outputs on horizontal and True outputs on vertical axes with 43 classes.	 Fig. Confusion Matrix with Predicted outputs on horizontal and True outputs on vertical axes with 43 classes.
Precision = 0.26 Recall = 0.29 F1-score = 0.25 Support = 7842	Precision = 0.23 Recall = 0.26 F1-score = 0.21 Support = 7842	Precision = 0.24 Recall = 0.26 F1-score = 0.22 Support = 7842
Accuracy = 29.09	Accuracy = 26.09	Accuracy = 25.97

Conclusion 3 :

This Germany based dataset with hue feature extraction is very important case. This data was downloaded readily from the internet, with hue feature pre extracted. This data has 30061 training image vectors and 7842 test image vectors. Since hue feature was looked upon. Every image is mapped to 255 columns corresponding to pixels, and each columns having a value in between 0 to 255.

This feature training model doesn't work good and has shown very less accuracy <30 %. This fact is also evident from the confusion matrices of varied instances. Even after reducing number of hidden layer nodes there is not much significant increase in accuracy.

This can be attributed to less uniqueness and weak mapping of the hue feature set and henceforth resulting into redundant and incorrect classification. This helps us know, hue alone cannot be used for unique feature extraction.

Observation 4 :

Dataset 4 : BTSC HOG Feature Set

Training : (35288,1568)
Testing : (7842,1568) }
Parameters ()

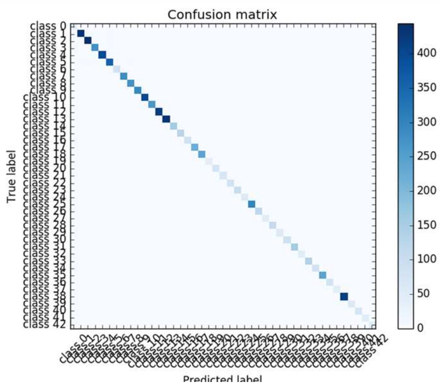
<table>
<tr><td>Hidden Neurons = 100</td><td>Hidden Neurons = 250</td></tr>
<tr><td>

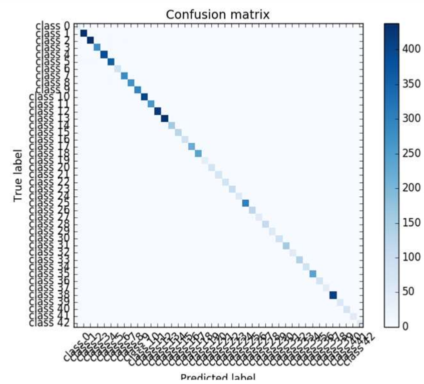
Fig. Confusion Matrix with Predicted outputs on horizontal and True outputs on vertical axes with 43 different classes.</td><td>

Fig. Confusion Matrix with Predicted outputs on horizontal and True outputs on vertical axes with 43 different classes.</td></tr>
<tr><td>Precision = 0.99
Recall = 0.99
F1-score = 0.99
Support = 7842</td><td>Precision = 0.99
Recall = 0.99
F1-score = 0.99
Support = 7842</td></tr>
<tr><td>Accuracy = 99.05</td><td>Accuracy = 99.05</td></tr>
</table>

Conclusion 4 :

This instance considers a pre exrtracted HOG feature set for Germany dataset. It has feature vectors of size 1568 with 35288 in training set and 7842 images in the test set. This is huge training dataset covering a large variety. This dataset classifies image into 43 different classes.

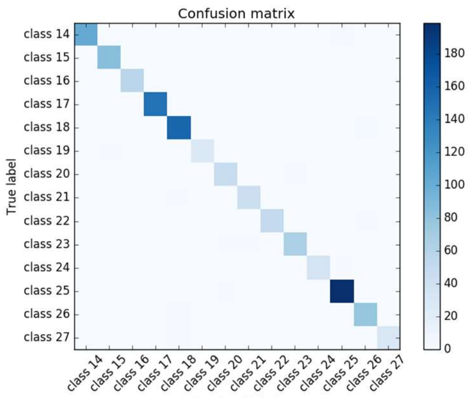
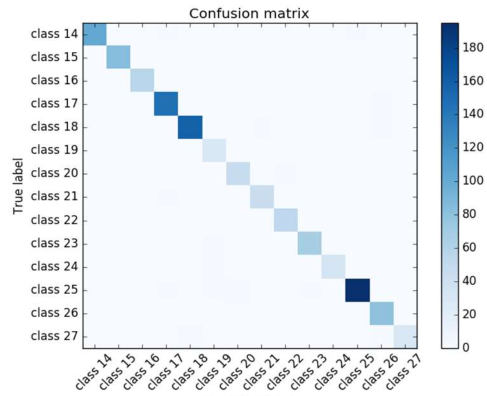
As depicted above this dataset being large has managed to get an accuracy figure > 99.05% and high precision of 0.99. This accuracy is almost maintained and is stable over variations in number of hidden nodes.

This high accuracy can be attributed to good dataset selection, proper feature extraction and high precision mapping done by the neural network during the training.

Observation 5 :

Dataset 5 : BTSC Haar Feature Set

Training : (3135,11584)
Testing : (1140,11584)
Parameter{}

Hidden Neurons = 1000	Hidden Neurons = 2000
 Fig. Confusion Matrix with Predicted outputs on horizontal and True outputs on vertical axes with 14 different classes.	 Fig. Confusion Matrix with Predicted outputs on horizontal and True outputs on vertical axes with 14 different classes.
Precision = 0.98 Recall = 0.98 F1-score = 0.98 Support = 1140	Precision = 0.99 Recall = 0.99 F1-score = 0.99 Support = 1140
Accuracy = 98.42	Accuracy = 98.508

Conclusion 5 :

This data set has pre extracted Haar feature Germany Dataset image vectors. These image vectors are of size 11584 with feature definition entities. Such large vectors help defining a feature more accurately and uniquely. This dataset has 3135 training image vectors and 1140 test image vectors. This instance of analysis classifies the test image into one from the 14 different classes.
When such image vectors are provided as input to Neural Network, the model performs good with high accuracy > 98% henceforth high precision of nearly 1. A minute increase is seen on reducing number of hidden layer nodes.
This high accuracy, here also, can be attributed to good dataset selection, proper feature extraction and high precision mapping done by the neural network during the training.

References

1. Fistrek, T.; Loncaric, S., "Traffic sign detection and recognition using neural networks and histogram based selection of segmentation method," in *ELMAR, 2011 Proceedings* , vol., no., pp.51-54, 14-16 Sept. 2011
2. Yok-Yen Nguwi; Kouzani, A.Z., "Automatic Road Sign Recognition Using Neural Networks," in *Neural Networks, 2006. IJCNN '06. International Joint Conference on Neural Networks* , vol., no., pp.3955-3962, 0-0 0 doi: 10.1109/IJCNN.2006.246916
3. Y.-Y. Nguwi and S.-Y. Cho, "Emergent self-organizing feature map for recognizing road sign images", 2009
4. Mohammad Badrul Alam Miah. Article: A Real Time Road Sign Recognition using Neural Network. *International Journal of Computer Applications* 114(13):1-5, March 2015.
5. C.Y. Fang, C.S. Fuh, P.S. Yen, S. Cherng, S.W. Chen, An automatic road sign recognition system based on a computational model of human recognition processing, *Computer Vision and Image Understanding*, Volume 96, Issue 2, November 2004, Pages 237-268, ISSN 1077-3142
6. J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel. The German Traffic Sign Recognition Benchmark: A multi-class classification competition. In *Proceedings of the IEEE International Joint Conference on Neural Networks*, pages 1453–1460. 2011.