

Transfer Learning Models for Traffic Sign Recognition System

Sowjanya Palavanchu

KoneruLakshmaiah Education Foundation, Vaddeswaram,AP,India
Email: 195031008@kluniversity.in

ABSTRACT

Computer vision has become the most happening technology in Artificial Intelligence in recent decades, and the continued advancement of Intelligent Transportation has been a major consequence of this success. Traffic signs detection, recognition and classification is an important functionality of the Advanced Driver Assistance System (ADAS) which helps the Smart transportation system to identify the traffic, signals and signs on road and perform necessary actions. Extensive research shows that artificial neural networks are an ideal choice for image classification and recognition challenges. Transfer learning has taken Deep learning to another level by enabling the models trained on a task to be re-used for another task which in result reduces the development and training time of the model. Five efficient transfer learning models are explored which are available in Keras libraries - Xceptionnetwork, InceptionV3 Networks, Residual Networks ResNet50, VGG-16 and EfficientNetB0 models for the detection, recognition, and classification of GTSRB traffic signs. The main focus of this paper is to apply and as well as compare five recent and most successful deep learning strategies to verify which model can stand-out in feature extraction and classification of the traffic sign data available. Accuracy, loss, training time and model parameters are considered in grading these models. Xception network has been proven to be highly successful in terms of accuracy (95.04%), minimum lossvalue (0.2311) and affordable speed and training time, whereas ResNet50 and EfficientNetB0 obtained good accuracy with fewer model parameters for traffic signs detection, recognition and classification.

Keywords

EfficientNetB0, Inception-v3, ResNet-50, Traffic sign recognition systems (TSRS), VGG16, Xception

Introduction

Time is crucial in intelligent transportation systems as the autonomous car, or the driver should be cautious and aware of the surroundings and road traffic signs to take necessary steps for safe driving. Safety of the passengers and the driver is of paramount in intelligent transportation system. Traffic sign recognition system is the main element in ADAS (Advanced Driving Assistance Systems) in autonomous and automated vehicles to recognize traffic signs that are encountered on the road in real-time situations. Therefore, to identify and classify traffic signs encountered on the roadway with great accuracy, the traffic-sign recognition model should work efficiently. Traffic-sign recognition model comprises of two steps: 1. Object Detection 2. Traffic sign classification from the images. The images were collected from a front-facing camera that scans the roads for any traffic signs continuously and the images captured are provided as input to the neural network for classification to intimate or alert the driver or the car to take necessary steps according to the sign. When the driver or the autonomous car is not aware of the traffic signs, accidents can happen which will jeopardize the driver or passenger's safety. The input

image quality plays a vital role in detecting and classifying a traffic sign and poor lighting, blur caused by motion or bent traffic signs could be a reason for poor image quality. Gaussian blur is used to the input data images to increase the quality of an image by reducing the noise in the image which helps in recognition of the image with better accuracy. Multi-task Conventional Neural Networks [1] work efficiently for Image classification and object detection needs to be performed before the CNN can be trained on the images for classification.

Object detection and feature extraction process with classification using Recurrent Neural networks [2] were developed with great recognition accuracy. In this paper, efficient transfer learning models are explored for feature extraction as dimensionality reduction and image classification to boost the performance of the model with minimum training. Transfer Learning is efficacious and easy as the models are already trained on a big dataset and the well pre-trained network is bootstrapped and customized according to the input dataset. Out of all the pretrained models available, five most efficient models are chosen in this experiment to observe the individual performance of these models. Deep Transfer learning is very popular in computer vision era and these models are utilized as a backbone for several research models to elevate model performance. This is achieved by modifying the pretrained model architecture by tweaking the outer layer according to the input dataset classification classes.

Five transfer learning models considered in this experiment are Xception, InceptionV3, ResNet50, EfficientNetB0 and VGG-16. These models are trained with ImageNet dataset weights and are implemented and available in Keras libraries. These pretrained model weights are retained by freezing the lower layers which was used as a base model for the project and then a small neural network is bootstrapped to the base model to customize according to the output features of traffic sign dataset. Xception model performs better for this dataset in terms of accuracy, loss values. EfficientB0 is better in model parameters area as it consists of only 5.3 million trainable parameters whereas the Xception model has 22.8 million.

In this paper, German Traffic Sign Recognition Benchmark (GTSRB)[3] is considered as it is a standard dataset in traffic sign recognition systems with almost 50,000 real-time photos of traffic signs with annotations for training data. The images in the dataset are real-life like frames in terms of lighting, image blur and position or angle of the traffic sign changes in real-time driving hinged on the time and unsuitable weather conditions. Data transformation is performed on the dataset and then fed to pre-trained neural network model to accomplish the classification task. Main tasks of this paper are as follows

- a. Data pre-processing is implemented and data augmentation is done to increase the size of the dataset to improve efficiency of the transfer learning model as it is trained on ImageNet dataset which is a large dataset.
- b. The Outer layer of the network/model is modified to fit the total number of classes and all the lower layers are frozen to not update the pre-trained weights and model training is performed with training images
- c. Traffic sign detection, efficient dimensionality reduction such as feature extraction and classification of the pictures is obtained with higher accuracy and performance for validation data. A deep neural network is implemented as a base model to compare the performance or efficiency of the transfer learning models.

The paper is organized as follows - in section 2, related research work is discussed, section 3 comprises of the brief introduction to transfer learning models used in this paper, section 4 explains the methodology of traffic sign recognition using the pretrained model, section 5 consists of the analysis of the results achieved from the five models with comparison to a deep neural network and comparison with models implemented until now and finally section 6 concludes the article with findings and analysis.

Literature Review

H.Lou and Y.Yang had proposed the usage of Multi-task convolutional neural networks (2018)[1] for traffic sign recognition system in which the Region of Interest (ROIs) are extracted, refined and classified. F. Chollet has introduced Xception networks [4] that are used in this evaluation which is a deep neural network with depth wise separable convolutions in 2017, these networks are trained on ImageNet dataset that contains 17,000 classes. According to F.Chollet, the gain in the DNN performance is due to decrease in model parameters rather than the increase in model capacity.

Simonyan, Karen, and Andrew Zisserman introduced VGG-16[5] models for image recognition and classification in 2014. Visual Geometry Group (VGG) from Oxford developed a convolutional neural network architecture in which the depth of the neural network is extended by very small 3x3 convolutional filters to the 16- and 19-layer network. In 2016, Inception-v3[6] network is developed by Szegedy, et.al. in which several convolutional layers are stacked upon each other with max pooling and inception layers in between to increase the performance metric of image classification.

K.He, X. Zhang, S. Ren and J. Sun implemented residual networks called ResNets[7] in 2016 where the network connections are skipped which will in effect take care of the vanishing gradients problem which is the result of deep neural networks. ResNet50 is the model used in this paper. Tan, M. and Le, Q.V proposed EfficientNet[8] in which model scaling is performed with balancing the depth,width and resolution for better performance. Pan and Yang [11] discussed several transfer learning approaches depending on the source and the target domains and tasks and whether the source and target data is labeled or not.All the strategies are discussed to help in selecting a transfer learning strategy for a user specific task.

Several techniques were developed before deep neural networks to improve the object detection by extracting the important features by using Extreme machine learning with adaboost and achieved great performance [2,14,15].Spatial transformers, HoG features [13] and the stochastic gradients [16,18] were used in traffic sign recognition. RCNN, Fast RCNN, Faster RCNN is proposed/implemented in later research for object detection [16,17] later ruled the world as it has of two networks: Region proposal network to get the ROI proposals of the object and another network to recognize the object/sign from proposals. These proposals are fed into the deep artificial neural networks and a classifier to train and classify the images.

Deep Neural networks (DNN)[19] became really popular in image classification problems by analyzing the images and giving highest probability of the object belonging to a target/output class. In DNNs, data flows from input layers to output layers and therefor they are feed forward networks with no loop back from output layer to input layer. The problem with deep neural

networks can be over-fitting of the data, huge computational time due to high training parameters and vanishing gradients as the network goes deep.

Transfer Learning

Transfer Learning is an important machine learning process in which a model already trained with a dataset is used as a baseline point for another model which has different input dataset. Transfer Learning is possible when input datasets of both the models are similar and is mostly used for image classification, object detection and feature extraction tasks. There are three transfer learning strategies [10] available such as Transductive, Inductive and Unsupervised. The Inductive transfer learning is when source and target operations are different but have similar domain and the labeled data is available in target domain. This is possible in two cases when there is unlabeled data in source side and labeled data is present in source domain. Transductive Learning is when source and target operations are same, but their respective domains are different, and the labeled data is there in source domain. Also, Unsupervised transfer learning is mostly same as inductive transfer learning where the source and target activities are different but have similar domain but there is no labeled data present in both the domains.

The transfer learning models that will be explored in this paper are EfficientNetB0, Xception, ResNet Inception-v3 and VGG16. Therefore, Transfer Learning is developed for various reasons such as model performance improvement, reducing network's training iterations/time and the ability to train a neural network with small amount of data.

A. Xception

Xception - Depth wise separable convolutional neural networks is proposed by Francois Chollet (Google) in 2017 and it is extreme extension of Inceptionv3 network which was developed by Google in 2015. Xception is a 71-layer deep convolutional neural network that contains modified depth-wise separable convolutions with point wise convolutions. Xception differs from Inception-v3 in two minor ways, first is the order of channel-wise spatial convolutions and 1x1 convolutions are reversed and the second difference is that intermediateReLU non-linearity is not present after the first operation in Xception network.

B. VGG16

VGG-16 is a deep convolutional neural network model proposed by K. Simonyan and A. Zisserman from the University of Oxford which takes a 224x224 image as an input and has multiple 3x3 kernel-sized filters sequentially with max pooling and fully connected layers with ReLU activation and a softmax classifier in the last layer with 1000 classes output. VGG16 is very slow to train, the network weights are too large, and it is over 533MB in memory due to its number of fully connected layers and also the depth of those layers.

C. ResNet50

ResNet50- is a 50-layer classic deep neural network that is proposed by K. He, X. Zhang, S. Ren and J. Sun of Microsoft in 2016. ResNet models skip connections in the network layers and overcome the exploding or vanishing gradients issue that all deep learning models suffer with because of the network convergence when more layers are added to the network. This convergence results in degradation of performance of the model as accuracy does not improve and rapidly deteriorates as training continues. Residual connections can be made in two ways,

First one is when the input and output are of equal dimensions, the identity shortcuts x can be used directly and second way is when the dimensions are different, identity mapping will be done by increasing the dimension with extra padding of zeroes.

D.EfficientNetB0

EfficientNetB0 is one of the family of neural network models proposed by Tan and Le in 2019 which are created by randomly choosing the scaling factor of the models. Scaling factor is modified by restricting the choice of resolution, width and depth of the network which effectively increases the accuracy and model performance. Computational resources are wasted if the resolutions are not divisible by 8 or 16 as they result in zero padding at the limits, therefore resolution chosen for B0 is 224. Width and depth of the EfficientNet blocks should have channel size in multiples of 8. Resolution can be kept constant in the case of memory limitation and depth/width can still increase for improving the model's performance.

E.Inception-v3

Inception-v3 is the third version of deep convolutional neural-networks proposed by C.Szegedy, Google in 2016 and was trained on the ImageNet dataset which contains more than a million images with 1000 output classes and the network is 48 layers deep. The model consists of convolutions, avg pooling, max pooling, dropouts, concat and fully connected layers with symmetric and asymmetric building blocks. Batch normalization is also adequately used wherever needed for the model input activation. Softmax performs the classification in the last/output layer to classify the signs.

Methodology

Transfer Learning in Traffic sign recognition systems is led by performing data pre-processing on the training images from the GTSRB dataset and augmentation of data is implemented using torchvision transforms APIs from PyTorch libraries to crop, rotate and resize all the training set images and after that these images are loaded into data loaders to train the pre-trained network in batches to tune the hyper parameters of the model. The training data is then split into train and validation data in 80-20% with shuffling and the data is fed into the pre-trained network which is already trained on the ImageNet dataset. Model architecture is modified by freezing the lower layers, changing the number of filters in every layer and the output features are modified to suit the needs of the input dataset. The categorical cross entropy loss function, optimizer such as Adam or SGD with momentum of 0.9, is used with adaptive learning rate depending on the model to train the pre-trained network. This paper intends to show that transfer learning is really beneficial in increasing the performance of the model performance and which in turn reduces the time taken to build and train a neural network from in certain image recognition and classification tasks.

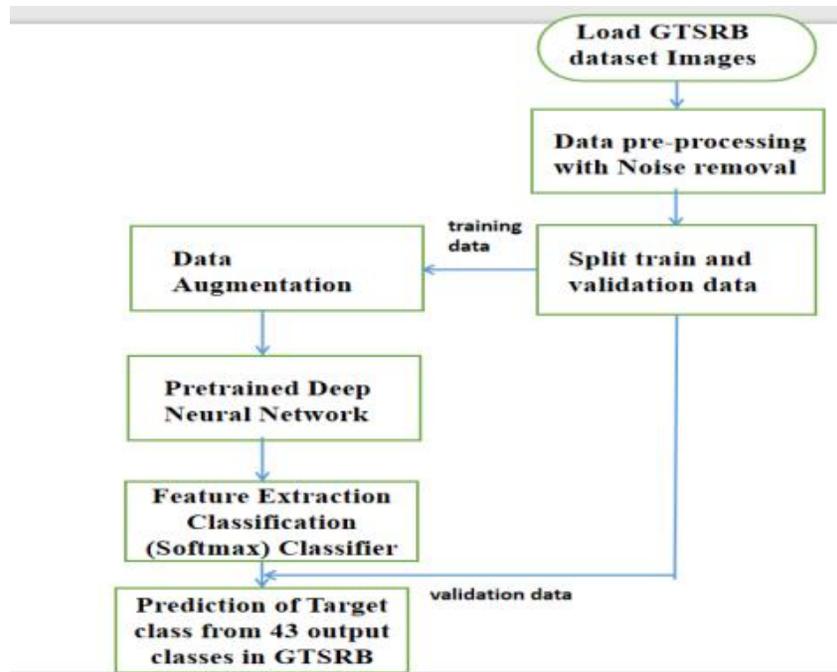


Figure1. process flow for TSRS with transfer learning

Figure 1 shows the flow of the process that is being followed to experiment with transfer learning models to achieve the desired result of traffic sign classification.

A. Input Dataset

The GTSRB [35] dataset is a single image multi-class classification dataset for which a competition was held at IJCNN 2011. The GTSRB dataset includes more than 50,000 images for training/validation and testing with corresponding labels and annotations for the training data. This dataset has an imbalance in the number of images in each class, some classes have around 2200 images and some of them have only 210. The training pictures can be split into training: validation in 80:20 ratios respectively in this implemented model. There are forty-three (43) target classes that each of the images can belong to train and testing set images needs to be later classified into a specific target class.



Figure 2. total classes in GTSRB dataset

Figure 2 shows the total number of classes that are present in GTSRB dataset. The input images containing the traffic sign is classified into one of the forty-three classes. The dataset contains real-life like images with great variations in lighting/illumination, weather-conditions, and visual appearance of signs from the point of distance, partial occlusions, and rotations.



Figure 3. sample images from GTSRB dataset

Figure 3 shows the sample pictures from each class of the GTSRB dataset, and the images vary from 15x15 mages to 250x250 resolution with varied lighting conditions, distance of the traffic signs from the mounted camera and the position of the traffic sign in the images.

B. Data Pre-processing and Augmentation

Data pre-processing is performed by loading all training images into a data-loader and re-sizing them into 224x224 images with 3 channels as RGB and applying Gaussian Blur from opencv libraries to intentionally blur an image with Gaussian function to reduce the noise in the input image. Images with poor quality can be processed, trained and classified as a result of this function.

Data Augmentation is performed to increase the size of the dataset to train on pre-trained networks to increase the accuracy as the network and to provide larger training data. Each Image is stored as several images as transformations such as rotate, horizontal flip, crop and normalizes to enlarge the dataset size and get all angles of the pictures. The training data increases seven times the original data because of the data-augmentation.

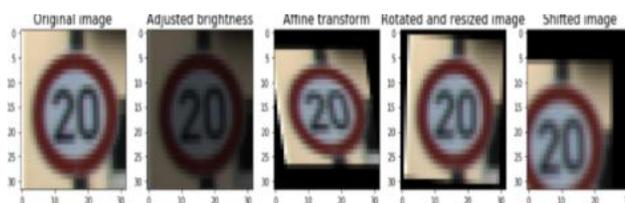


Figure 4. original image and augmented images

Figure 4 shows the actual dataset image and the augmented images resulted from the transformations applied on the image to increase the size of the dataset and for the pre-trained model to train on different varieties of the same image.

D. Training and Evaluation of models

The pre-processed and augmented data is loaded into data loaders with batch size of 64 and then given as an input to the network architecture for training in batches. Batch processing improves the accuracy of the neural network and its one of the important hyper parameters in a neural network to improve performance of the network. Learning rate is also a hyper parameter which could be adjusted when the model is being trained and the model is coded with gradual adjusting learning rate with each epoch. Required code is written to save the best model with least loss value as a bench-mark model in training to prevent overfitting or under fitting and to reach the optimal iterations for training.

Image Recognition and Classification is better performed by convolutional neural networks than any other machine learning models that have been developed and the research to better these convolutional Neural Networks has been intense. The transfer learning models in this particular experiment are enhanced by adding dense and drop-out layers in outer layers and freezing the lower layers of the pre-trained model which avoids the updating the weights for those layers and decreases the time taken to train the network with increased accuracy. Features from each picture are extracted by the pre-trained model and then the input images are classified into one of the 43 classes of GTSRB dataset.

A deep convolutional neural network is built for baseline with 11 hidden layers that include convolutional layers, max-pooling layers, dense and drop-out layers with more than 44 million trainable parameters for image classification tasks. The main purpose of DNN is to show how easier it is to use a pre-trained transfer learning model with efficient performance and accuracy rather than building a neural network from scratch. Memory usage for the pre-trained model is also significantly lower than the deep CNN (convolutional neural network) for each image recognition depending on the batch size.

The testing set images that are unknown to the network are fed into the trained model in batches with data loaders and evaluated by calculating the accuracy score from metrics libraries of python sklearn packages. The predictions are printed with the probable target class names and also several unseen pictures are tested with the model to see if the network predicts to the right target class.

Data visualization is performed to better analyze the performance by plotting the accuracy and loss metrics for training and validation data. The evaluation plays a vital role in finding the best model and hyper parameters for a particular model and the dataset to see if any overfitting is done. Loss and Accuracy attributes are the basic metrics evaluated in any model and an optimal model is a model with less loss value, high top-5 accuracy, and less training time.

Results and Discussion

This paper intends to use different transfer learning models for traffic signs detection, recognition and classification and compare which pretrained network performs the task on GTSRB dataset with best accuracy, minimum loss value, and lower training time.

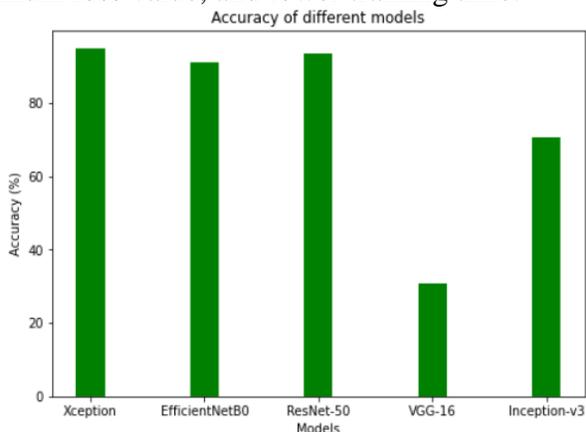


Figure 5. Accuracy of the models

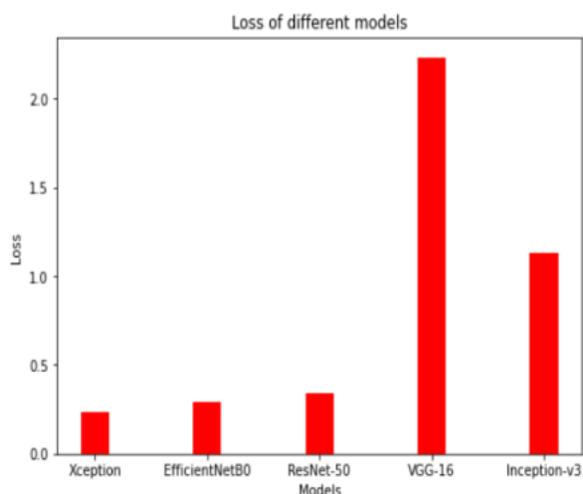


Figure 6. Loss values of the models

Fig 5& 6 shows the accuracy and loss results obtained from different models and from the bar graph it is clearly visible that Xception model performs best compared to all the other models in terms of validationset accuracy. ResNet-50 and EfficientNetB0 are also great models for the task of image classification with better accuracy values and good training times. VGG-16 performs very poorly for this dataset with respect to accuracy and loss results.

From Table 1, it is evident that four models Inception-v3, Xception, EfficientNetB0 and ResNet-50 perform well in traffic sign classification and recognition by observing the metrics such as validation set accuracy and loss functions.

Inception-v3 model performs better in accuracy,model parameters and loss values where as Xception, ResNet-50 and EfficientNetB0 models have comparable performance. Top-1 accuracy is the accuracy of the classifier’s highest probable target class that the image belongs to is the correct class of an image. For example, if the classifier guesses 90% that it is a picture of a cat and the image is a picture of a cat, it is the top-1 accuracy of the classifier.

The deep neural network that is developed for baseline obtained validation accuracy of 74.29% and loss value of 0.8445 with 44 million trainable parameters which is a higher number than the best pre-trained models which will compromise the network performance. Therefore, pre-trained transfer learning models are a better choice for image classification task as there are readily available in keras libraries and with much better performance.

Table 1. GTSRB validation accuracy and loss results for models and a DNN

<i>Model</i>	<i>Xception</i>	<i>EfficientNetB0</i>	<i>ResNet-50</i>	<i>VGG-16</i>	<i>Inception-v3</i>	<i>DNN</i>
Validation Accuracy(%)	95.04	91.0	93.33	30.61	70.74	74.29
Validation Loss	0.2311	0.2910	0.3384	2.2355	1.1348	0.8445
Trainable Parameters(Million)	22	5.3	25	138	24	44
Top-1 Accuracy	78.79%	76.3%	77.15%	74.5%	78.8%	
Training time	1hr55min	1hr44min	1hr44min	1hr8min	58min	1hr19min

A. Performance analysis of transfer learning models

When comparing with the implemented models with multi-column neural networks[1], there were twenty five deep convolutional neural networks arranged in 5 columns each training on a

dataset that achieved the accuracy of 99.46% and that is higher than the transfer learning models implemented in this paper but the time taken to train, computational resources, model parameters and memory usage for training these models are all lower in the transfer learning models as they are all pre-trained and network's training time is much lower as only the outer layers are trained but still achieved 95.04% and the recognition rate can still be improved by adjusting the hyper parameters like Optimizer with or without momentum and learning rate which will be continued in the future work.

When compared with Real-time traffic sign detection with Faster R-CNN for object detection [11], the transfer learning models obtained higher accuracy (95.04%) even though the model in the experiment was tested for just 300 test images and achieved 91.5% mean average precision(mAP) and the transfer learning models implemented in this experiment tested for more than 7000 images.

Proposed Welm+adaboost [15] has achieved great recognition accuracy 99.12% better than the transfer learning models but training duration of the network is almost 9.5hrs and the transfer learning models highest training period is only ~2hrs.

Table 2. Comparison of other models on GTSRB dataset

<i>Method</i>	<i>Recognition Accuracy(%)</i>	<i>Training Time</i>
<i>Multi-Column DNN[1]</i>	99.46	Several days
<i>WELM+Ada boost [15]</i>	99.12	9.36h
<i>Xception</i>	95.04	1hr 55min
<i>Faster RCNN [11]</i>	91.5	N/A
<i>DNN + Spatial Transformers [14]</i>	99.49	5hr 55min

Conclusion

The comparison of five transfer learning models for traffic sign recognition is experimented in this paper. The analysis of these models is performed by comparing main metrics of these models, such as accuracy, loss, speed, trainable parameters and memory. All of these transfer learning models were pre-trained on ImageNet [ILSVRC] dataset and later fine-tune according to the GTSRB dataset which contains 43 output classes of distinct traffic signs and exactly had been trained on a GPU to maintain the uniformity of the experiment exactly for 10 epochs.

After evaluating the accuracy results of these five models, it is concluded that Xception model obtains the best accuracy(95.04%) with only 22 million trainable parameters and loss value (0.2311) and the second best model is ResNet-50(Residual convolutional neural networks) with the accuracy of 93.33% with only 25million trainable parameters where as EfficientNet-B0 has lowest number of trainable parameters with 5.3million and good enough accuracy of 91.0%

that is a good trade-off for model performance and accuracy to consider for traffic sign classification task. The model VGG-16 and Inception-v3 obtained poor accuracy and high loss values even though the speed of the training and recognition is much higher.

Future work could be done in extending these transfer learning models to train with real-time videos for traffic sign recognition and classification as the GTSRB dataset contains images that were captured on various types of roads with varying visibility conditions

Acknowledgement

Author thanks Dr.Suryakanth.V.Gangashetty, Professor in KoneruLakshmaiah Education Foundation for his immense support and guidance in writing this paper.

References

- [1] Luo, H., Yang, Y., Tong, B., Wu, F., and Fan, B. (2018). Traffic Sign Recognition Using a Multi-Task Convolutional Neural Network.*IEEE Transactions on Intelligent Transportation Systems*,19(4): 1100-1111.doi: 10.1109/TITS.2017.271469R.
- [2] Han, C., Gao, G. & Zhang, Y. (2019). Real-time small traffic sign detection with revised faster-RCNN. *Multimedia Tools Appl* 78:13263–13278. <https://doi.org/10.1007/s11042-018-6428-0>
- [3] Hou, Y., Hao, X., &Chen, H. (2017). A Cognitively Motivated Method for Classification of Occluded Traffic Signs.*IEEE Transactions on Systems, Man, and Cybernetics: Systems*,47(2): 255-262. doi:10.1109/TSMC.2016.2560126.
- [4] Chollet, F. (2017).Xception: Deep Learning with Depthwise Separable Convolutions. 2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, 2017, pp. 1800-1807, doi: 10.1109/CVPR.2017.195.
- [5] Simonyan, Karen, & Andrew Zisserman (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014)
- [6] Szegedy, C., Vanhoucke, V.,Ioffe, S., Shlens, J. &Wojna, Z. (2016). Rethinking the Inception Architecture for Computer Vision.2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 2016, pp. 2818-2826, doi: 10.1109/CVPR.2016.308.
- [7] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, 2016, pp. 770-778, doi: 10.1109/CVPR.2016.90.
- [8] Tan, M. and Le, Q.V. (2019). Efficientnet: Rethinking model scaling for convolutional neural networks. arXiv preprint arXiv:1905.11946.
- [9] Huang, Z.,Yu, Y., Gu, J., & Liu, H. (2017). An Efficient Method for Traffic Sign Recognition Based on Extreme Learning Machine.*IEEE Transactions on Cybernetics*,47(4): 920-933.doi: 10.1109/TCYB.2016.2533424.
- [10] Liu, C., Chang, F., Chen, Z., & Liu, D. (2016). Fast Traffic Sign Recognition via High-Contrast Region Extraction and Extended Sparse Representation.*IEEE Transactions on Intelligent Transportation Systems*, 17(1):79-92. doi: 10.1109/TITS.2015.2459594.

- [11] Pan, S. J., & Yang, Q. (2010). A Survey on Transfer Learning," *IEEE Transactions on Knowledge and Data Engineering*, 22(10):1345-1359. doi: 10.1109/TKDE.2009.191.
- [12] Stallkamp, M. Schlipping, J. Salmen, C. Igel, Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition, *Neural Networks*, Available online 20 February 2012, ISSN 0893-6080, 10.1016/j.neunet.2012.02.016.
- [13] Yang, H., Luo, H., Xu., & Wu, F. (2016). Towards Real-Time Traffic Sign Detection and Classification. *IEEE Transactions on Intelligent Transportation Systems*, 17(7): 2022-2031. doi: 10.1109/TITS.2015.2482461
- [14] Tian, Y., Gelernter, J., Wang, X., Li, J., & Yu, Y. (2019). Traffic Sign Detection Using a Multi-Scale Recurrent Attention Network," *IEEE Transactions on Intelligent Transportation Systems*, 20(12): 4466-4475. doi:10.1109/TITS.2018.2886283.
- [15] Xu, Y., Wang, Q., Wei, Z., & Ma, S. (2016). Traffic sign recognition based on weighted ELM and AdaBoost. *Electronics Letters*, 52(24): 1988-1990. doi: 10.1049/el.2016.2299.
- [16] Berkaya, Selcan Kaplan, Huseyin Gunduz, Ozgur Ozsen, Cuneyt Akinlar, and Serkan Gunal (2016). On circular traffic sign detection and recognition. *Expert Systems with Applications*. 48: 67-75.
- [17] Arcos-Garcia, A., Alvarez-Garcia, J.A. and Soria-Morillo, L.M., 2018. Deep neural network for traffic sign recognition systems: An analysis of spatial transformers and stochastic optimisation methods. *Neural Networks*, 99, pp.158-165.
- [18] Wu, Linxiu, Houjie Li, Jianjun He, and Xuan Chen (2019). Traffic sign detection method based on Faster R-CNN. *Journal of Physics: Conference Series*, 1176 (3):032045. IOP Publishing, 2019.
- [19] Arcos-Garcia, Alvaro, Juan A. Alvarez-Garcia, and Luis M. Soria-Morillo (2018). Evaluation of deep neural networks for traffic sign detection systems. *Neurocomputing*. 316: 332-344.