Paper Review 03

A Model Based on Convolutional Neural Network (CNN) for Vehicle Classification

Md. Alahi Almin Tansen (190321028)

Department of Computer Science and Engineering, European University of Bangladesh

CSE – 459: Internet of Things

Date: 26-03-2023

Purpose of This Research

The purpose of the research is to propose a convolutional neural network (CNN) for classifying common types of vehicles, with the goal of eliminating traffic-related road accidents. The research uses three different learning methods to identify the vehicles - MobileNetV2, DenseNet, and VGG 19 - and demonstrates their detection accuracy on a real-time standard dataset, which includes 4800 photographs of vehicles from Nepal. The best accuracy is achieved by the MobileNetV2 architecture, with a training accuracy of 97.01% and validation accuracy of 98.10%.

Previous Research Gap

Many researchers have already applied pattern analysis, image processing, and machine learning to the area of vehicle identification, with promising findings that have aided fundamental science and engineering applications

- Ng et al. proposed a method that uses HOG-SVM-dependent handcrafted features to train SVM classifiers with HOG features and Gaussian kernel functions. The classifier was tested on a surveillance camera dataset containing 2800 images and achieved a 92.3% accuracy rate in correctly categorizing motorcycles, cars, and lorries.
- In a separate study, Chen et al. proposed a classification system for vehicles that extracts texture and HOG features and uses a fuzzy influenced SVM classifier for classification. The system was tested on a dataset of 2000 photographs and achieved a 92.6% accuracy rate in correctly classifying vehicles, trucks, and buses.
- Matos et al. proposed a combined approach that uses two neural networks to embed characteristics of cars such as height, distance, and bounding boundaries for classification. The classifier was tested on a dataset of 100 images and achieved a score of 69%.
- Cui et al. used SVM to characterize a dataset consisting of 340 photographs of vehicles, minibusses, and trucks. They proposed a combined model for feature extraction using two Scale Invariant Feature Transform (SIFT) descriptors and a Bag of Words (BoW) approach. The classifier achieved 90.2% accuracy on the given dataset. The method was also used to separate the data into vehicle and non-vehicle types.
- Wen et al. proposed an AdaBoost-based quick learning vehicle classifier and an algorithm for extracting Haar-like features for rapid classifier learning. The classifier was tested on the public Caltech dataset and achieved 92.89%

accuracy. R. Sindoori et al. proposed a technique for determining vehicle disclosure structure from groveling images in 2013 using pixel clever grouping. Alpatov et al. proposed a method for real-time vehicle tracking and counting using photographs from a stationary camera to ensure road safety in 2018.

- Vishwanath P. Baligar and Mallikarjun Anandhalli suggested a color-based algorithm for detecting and tracking vehicles.
- Seda Kul proposed an analysis paper on vehicle identification and classification in 2017
- In 2019, Watcharin Maungmai introduced a method for categorizing vehicles based on vehicle model and color using CNN as a deep learning algorithm. They achieved accuracy rates of 81.62% and 70.09% for the two cases respectively. The analysis was conducted using 763 seconds of video data converted into photographs.
- In 2018, Bensedik Hicham and colleagues presented a CNN-based method for vehicle classification. They used a dataset of 2400 photographs and achieved an overall precision of 89% across four different styles of vehicles.

In this paper the researcher uses the most recent advances in the field of vehicle detection to gain the best possible result using DenseNet, VGG19, MobileNetV2.

Proposed System

The paper proposes using three different CNN architectures, including DenseNet, VGG 19, and MobileNetV2, to detect different types of vehicles. The dataset used for training and validation has undergone preprocessing by converting the images to grayscale, resizing them to 128x128 pixels, setting labels, normalizing the images, and converting them to numpy arrays. The performance of the models is measured based on several metrics, including accuracy of the training and validation datasets, recall, precision, and F1 score. These metrics provide a measure of how well the models are able to accurately classify vehicles in real-world scenarios. The results of the study can be used to develop better vehicle detection algorithms and improve traffic control systems.

Architecture

In this research, CNN architecture is used for image processing and classification. It includes DenseNet, VGG 19, and MobileNetV2, to detect different types of vehicles. DenseNet is a CNN-based transfer learning architecture that requires fewer parameters for high accuracy and less data for training. The architecture includes five convolution layers, pooling layers, three transition layers, one classification layer, and two DenseBlocks. DenseNet uses the weight of Imagenet and has a powerful feature where individual layers can access the gradient provided by the loss function. VGG19 is a variant of the VGGNet architecture, which is a CNN-based architecture for image classification using deep learning. It has been trained on millions of image samples and consists of 19 layers, including 16 convolutional layers, 3 fully connected layers, 5 MaxPool layers, and 1 SoftMax layer. VGG19 is known for its fully connected layers with 4096 channels, each followed by another fully connected layer with 1000 channels to predict 1000 labels. The VGG model has other variants, such as VGG11, VGG16, and others. Another architecture is MobileNetV2 which is a CNN-based architecture that uses 28 deep convolutional neural network layers and an inverted residual structure with residual connections between bottleneck layers. It employs depthwise separable convolutions to create lightweight deep models suitable for mobile and embedded vision applications. The architecture includes an initial fully convolutional layer with 32 filters, followed by 19 residual bottleneck layers. MobileNets are designed to be small, low-latency, and low-power models parameterized to meet various resource constraints in different use cases. Three of those architecture used to detect different types of vehicles.

Experimental Procedure

Researcher used numpy arrays for all images and trained the models on 80% of the data with a random state of 42. The trainable layer included average pooling with a pool size of (2,2) and Relu activation in the first convolution layer with 128 neurons. The output layer included four neurons divided into four types with SoftMax activation functions. Stochastic Gradient Descent was used for backpropagation and learning, and the loss function was measured using categorical cross-entropy. The results showed that MobileNetV2 achieved the highest accuracy of 98.10%, followed by DenseNet at 95.37% and VGG19 at 92.68%. The recall, accuracy, F1-score, and precision were calculated for each model, with MobileNetV2 outperforming the others in terms of precision (96.39%), accuracy (96.46%), F1-score (95.23), and recall (96.63%). The MobileNetV2 model successfully identified categorical vehicles in tested pictures accurately.

Future Plan

The researcher used DenseNet 121, MobileNetV2, VGG-19 model to detect the vehicles. MobileNetV2 performed very well and gain the highest accuracy of 98.10 % on the MobileNetV2 architecture. Besides, DenseNet achieves 95.37% accuracy and 92.68% gain on VGG 19. In the future, researcher will work with a broader dataset and integrate IoT-cloud with this model for centralized monetization.

References

1. A Model Based on Convolutional Neural Network (CNN) for Vehicle Classification by FM Javed Mehedi Shamrat, Sovon Chakraborty, Saima Afrin, Md Shakil, Mahdia Amina Moharram, Tonmoy Roy