Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

Deep Learning for Large-Scale Traffic-Sign Detection and Recognition

Naila Fathima1, Imreena Ali (Ph.D)2 , Prof. P.V.Sudha3

1PG Scholar, Department of CSE, ISL Engineering College, Hyderabad, India.

2Assistant Professor; Department of CSE. ISL Engineering College, Hyderabad, India.

3Professor,, Dept. of Computer Science & Engineering, UCE, Osmania University, Hyderabad, India.

(Received: 04 August 2023 Revised: 12 September Accepted: 06 October)

KEYWORDS

Deep learning, Traffic-sign detection and recognition, Traffic-sign dataset, Mask R-CNN, Traffic-sign inventory management

ABSTRACT:

Automatic detection and recognition of traffic signs plays a crucial role in management of the traffic-sign inventory. It provides accurate and timely way to manage traffic-sign inventory with a minimal human effort. In the computer vision community the recognition and detection of traffic signs is a well-researched problem. A vast majority of existing approaches perform well on traffic signs needed for advanced drivers assistance and autonomous systems. However, this represents a relatively small number of all traffic signs (around 50 categories out of several hundred) and performance on the remaining set of traffic signs, which are required to eliminate the manual labor in traffic-sign inventory management, remains an open question. In this paper, we address the issue of detecting and recognizing a large number of traffic-sign categories suitable for automating traffic-sign inventory management. We adopt a convolutional neural network (CNN) approach, the Mask R-CNN, to address the full pipeline of detection and recognition with automatic end-to-end learning. We propose several improvements that are evaluated on the detection of traffic signs and result in an improved overall performance. This approach is applied to detection of 200 traffic-sign categories represented in our novel dataset. Results are reported on highly challenging traffic sign categories that have not yet been considered in previous works. We provide comprehensive analysis of the deep learning method for the detection of traffic signs with large intra-category appearance variation and show below 3% error rates with the proposed approach, which is sufficient for deployment inpractical applications of traffic-sign inventory management.

I. INTRODUCTION

Effective traffic-sign inventory management is crucial to public safety and smooth traffic flow [1, 2]. This is often done by hand. Vehicle-mounted cameras record traffic signs, which are then manually localized and recognized in an offline process to ensure they match the database. However, when applied to hundreds of kilometers of roads, such physical labor may be immensely time consuming. Having this process automated would greatly minimize the amount of human labor required and increase safety by allowing for the faster identification of broken or missing traffic signs [3].

Automating this process requires first eliminating the need for human intervention in the location and identification of traffic signs. Already, good detection and identification algorithms have been suggested [4, 5, 6] for the challenge of traffic-sign recognition in the computer-vision community. However, these methods have only been developed for a limited number of classes, mostly for ADAS-related traffic signs [7] and autonomous vehicle-related traffic signs [8].

There is still some uncertainty about how to detect and recognize a wide variety of traffic-sign types. The problem of traffic-sign identification and detection has been addressed by a number of prior benchmarks [9, [10], [11], [12], [13]. Several of these studies, however, only looked at traffic-sign recognition (TSR) and not the far more difficult issue of traffic-sign detection

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

(TSD), in which the precise position of a traffic sign must be determined.

Other benchmarks that do use TSD often only cover a subset of traffic-sign categories, typically those relevant to ADAS and autonomous vehicle applications. Most benchmark categories may be easily spotted by handmade detectors and classifiers due to their unique look and minimal inter-category variation. Signs of this kind include the obligatory circular sign and the forbidden triangle symbol. While the current benchmarks are useful for identifying the most common types of traffic signs, there are numerous additional classes that may be far more challenging to detect due to their great degree of variance in appearance. The size, shape, and color of these objects, as well as the presence or absence of certain text and symbols (like arrows), might vary widely even across instances of the same class. Because of the similarities in appearance between items of various categories, this usually results in a high degree of intra-category (within-category) appearance variance but a low degree of inter-category (between-category) appearance variation. However, this would be a time-consuming effort, especially when considering that many traffic-sign looks are not constant between nations, therefore it is likely that current approaches would need to be modified with hand-crafted features and classifiers to handle such categories. Using feature learning based on actual instances is a far better strategy. It's easy to see how accommodate and record a wide range of visual differences across several traffic signs. Recent deep learning advancements have showed encouraging outcomes in the identification and recognition of common items. While deep learning methods have been applied for traffic-sign identification and recognition in the past [6], previous works' assessment had only included a small fraction of traffic-sign categories. The absence of a comprehensive dataset with hundreds of categories and enough cases for each category is a major roadblock to the widespread use of deep learning to analyze traffic signs. To avoid overfitting, enormous amounts of samples are required in deep learning, where models might include tens of millions of learnable parameters.

For the purpose of managing the stock of traffic signs along roads, we tackle the problem of learning and identifying a large number of categories in this study. Our major contribution is a system based on deep learning and convolutional neural networks for training a large number of traffic sign classifications. Our method is based on the cutting-edge detector Mask RCNN [14], which has shown impressive precision and speed in object identification applications. Since the TSR and the region proposal network share the same network design, the whole learning process is streamlined. The convolutional approach, in contrast to the more traditional methods that rely on carefully crafted features, is applied across a wide range of categories, including those in which the appearance of individual traffic-sign instances can vary significantly both within and between categories. We also suggest enhancements to Mask R-CNN that are particularly important in the field of traffic signs. To improve the recall rate, we suggest several adjustments and offer a new augmentation approach tailored to traffic-sign categories, especially for smaller signs[1].

II. RELATED WORK

There is a vast body of work on TSR and TSD, with many review articles accessible [11, 15]. Several recent research [15, 16] highlight the absence of a uniform publicly accessible benchmark dataset that would comprise a large number of different traffic-sign categories, making it impossible to determine which strategy offers the best overall results. Most writers test their methods using one of the several publicly available datasets that provide a modest amount of data.

limited number of traffic-sign categories:

- Among these benchmarks is the German Traffic-Sign Detection Benchmark (GTSDB) [10], which consists of three main categories designed for detection.
- The German Traffic-Sign Recognition Benchmark (GTSRB) [9] has 43 distinct categories for the exclusive purpose of traffic-sign recognition.
- Detection and recognition data from 62 categories in the Belgium Traffic Signs (BTS) collection [17].
- A road maintenance evaluation service in Croatia was recently obtained using data from Mapping and Assessing the State of Traffic Infrastructure (MASTIF) [18], which initially

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

included 9 categories but has now been expanded to 31 categories [19].

- 10 classes, for detection, in the Swedish Traffic Sign Dataset (STSD) [20].
- The LISA Dataset [11]: 49 types of traffic signs collected from American roadways by the Laboratory for Intelligent and Safe Automobiles.
- The Tsinghua-Tencent 100K dataset [13]: a massive dataset with tens of thousands of photographs, ninety percent of which include traffic signs.

Some methods [21], [22] sample photos from several datasets to do the assessment, which helps to expand the pool of potential traffic signs to evaluate. But many more writers rely on their own personal datasets [4, [23], [24], [25].

According to our best estimates, [24]'s private dataset, which differentiates between 131 kinds of non-text traffic signs from the roadways of the United Kingdom, has the biggest collection of categories ever studied.

There are a lot of traffic-sign datasets, but it's still hard to compare detectors across many different categories. Our comprehensive dataset contains 200 traffic-sign categories, including a large number of categories with significant intra category variability, whereas existing benchmarks tend to focus on small numbers of super categories (GTSDB [10]) or on small numbers of simple traffic signs (BTS [17], MASTIF [18], STSD [20], LISA [11]). Tsinghua-Tencent's 100K dataset is the closest massive one, but even so, they only evaluate on 45 basic traffic signs. In contrast, our data collection allows for a deep dive into detectors as they pertain to traffic-sign stock-taking. TSR and TSD have used several different approaches. The histogram of oriented gradients (HOG) [12, 24], [26], [16], [5], [19], [10], the scale invariant feature transform (SIFT) [5], the local binary patterns (LBP) [16], and the integral channel features [26] are all examples of features that have traditionally been created by hand.

Support vector machine (SVM) [24], [16], [27], logistic regression [28], and random forests [16], [27], as well as extreme learning machine (ELM) [19]-style artificial neural networks, have also been used.

TSR and TSD, along with the rest of computer vision, have recently benefited from the rebirth in CNN technology. Using a contemporary CNN method, [29] was able to automatically extract multi-scale features for TSD. Automatic feature representation learning and final classification in TSR have both been accomplished using convolutional neural networks [30, [31], [32], [33]. Combining a convolutional neural network (CNN) with a multilayer perceptron (MLP) was used in [34] to boost recognition accuracy, while [30], [32] advocated using an ensemble classifier made up of many CNNs to do the same. CNN-based feature learning followed by ELM classification is used in [35], whereas a deep network with spatial transformer layers and a tweaked inception module is used in [36].

According to [37], CNNs are superior than humans at GTSRB in terms of recognition accuracy. In recent research [6, 13], CNNs were used to solve both the TSR and TSD issues at once. For the latter, they utilize an OverFeat [38] network with some significant tweaks, while for the former, a fully convolutional network was used to generate an image heat map, which was then detected using a region proposal approach. Finally, the collected areas were classified using a dedicated CNN.

Our deep learning-based method is unique in comparison to similar studies. We propose comprehensive feature learning with end-to-end learning as an alternative to more conventional techniques that rely on hand-crafted features and machine learning [12, 24]. In addition, our strategy is distinct from others that use deep learning to detect traffic signs. Instead of a dedicated technique for producing region recommendations, like in [6] and [13], we leverage deeper networks based on the VGG16 [39] and ResNet-50 [40] architectures in our Mask R-CNN-based approach. We also use a network that has already been trained on ImageNet, which drastically decreases the number of samples needed for training compared to both [6] and [13]. Furthermore, we have incorporated a number of enhancements that have resulted in increased efficiency.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

Lee and Kim (2018) present a unique convolutional neural network (CNN) traffic-sign identification system that provides accurate predictions of both the sign's position and boundaries. Even though the precision was outstanding, Lee's crew had to work with very detailed photographs. Hu et al. (2016) narrowed their attention on traffic signs, automobiles, and bicycles. Their proposed framework used a single learning technique to identify all three categories. The traffic sign detector in their model required the least amount of time since less individual sub-detectors were utilised. Time required to complete the detection process rose considerably once further characteristics were introduced to facilitate the identification of other items. Regardless of the weather, the approach provided in (Greenhalgh and Mirmehdi, 2012) is able to reliably recognize the picture areas having signals as maximal stable extremal regions. Support vector machine (SVM) classifiers that were trained using Histogram of Oriented Gradient (HOG) features are used for sign recognition. The recall and accuracy rates, however, were only 86% and 80%. Later, Greenhalgh and Mirmehdi (2015) used a scene structure to zero down on picture search locations where the presence of a traffic sign was very likely. False positives have resulted in significant losses to the accuracy parameter, and the frame rate has decreased from 14 frames/s to 6 frames/s as a result of the removal of structural information.

The Extreme Learning Machine (ELM) technique was used to train a two-module solution for traffic sign identification (Huang et al., 2017) that consists of a HOG extraction feature and a single classifier. More than half of the photos were incorrectly classified, and the model's performance was dependent on the tuning parameter. In lieu of the standard CNN method, Huang et al. (2020) developed a visually-based automated recognition system. However, the attained accuracy was much lower on average than the suggested RMR-CNN approach. While Yang et al. (2016) detail a lightning-fast technique for traffic sign identification that combines SVM and CNN, its accuracy falls short of that of state-of-the-art systems such Mask R-CNN. A method for detecting traffic signs was proposed by Chen and Lu (2016); it combines Adaptive Boosting (Adaboost) with Support Vector Recognition (SVR). Liu et al. (2016) reported a TSR strategy that makes use of high contrast area extraction, an extended sparse representation, a color enhancement technique, and voting of nearby features. The negative of their methodology is that the colors of other objects, such as those on traffic signals, are boosted, which causes a delay in TSR.

Temel et al. (2020) developed a model that could recognize traffic signs regardless of environmental factors such as rain or a dirty camera lens. Only 80% were correct, however. For TSDR, Kamal et al. (2019) propose SegU-Net, a hybrid of SegNet (a state-of-the-art segmentation network) and U-Net. The model's 95.29% accuracy on the German Traffic Sign Detection Benchmark dataset is lower than that of more traditional approaches such as Deep Neural Networks with Convolutional layers and Spatial Transformer Networks. MicroNet is a small neural network architecture developed by Wong et al. (2018) specifically for TSR. In terms of computational speed, modern neural network designs such as Mask R-CNN and Faster R-CNN excel. In Avramovi et al. (2020), the authors propose a CNN-based TSDR with a YOLO (You Only Look Once) architecture. The primary goal of this work is to enhance the speed and accuracy of detection based on high-definition photographs by zeroing in on specific areas of interest within those images. However, the method used in the aforementioned research might result in the selection of areas in a picture that may not have any traffic signs. The difficulties in seeing and understanding Chinese characters on road signs were detailed by Guo et al. (2020). The downside of this approach was that occlusion affected the precision with which Chinese characters were detected, leading to certain characters being unrecognized and others being recognized wrongly.

III. TRAFFIC-SIGN DETECTION WITH MASK R-CNN

Here we introduce our traffic-sign detection system, which makes use of a modified version of the Mask R-CNN detector. To begin, we introduce the Mask R-CNN detector and then proceed to describe our modified version of the algorithm specifically designed to learn traffic sign categories.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

A. Mask R-CNN

For a more in-depth explanation of Mask R-CNN, we recommend reading [14]. Similar to Faster R-CNN [41], the Mask R-CNN network [14] consists of two nodes. The first component is a region proposal network (RPN), a deep fully convolutional network that receives an input picture and outputs a series of rectangular object suggestions, each with an objectness score. The second component is a convolutional neural network (CNN) tailored specifically to the task of categorizing suggested areas into existing ones; it goes by the name Fast R-CNN. Fast R-CNN is very effective because it uses shared convolutions across different proposals. To further improve the quality of the suggested areas, bounding box regression is also carried out. Together, RPN and Fast R-CNN share their convolutional characteristics to form a single unified network. The RPN module directs the attention of the Fast RCNN module, using language popularized by recent discussions of neural networks equipped with a "attention" mechanism. Then, by integrating a Feature Pyramid Network (FPN) into the base network design, Mask R-CNN enhances the original system [42]. Since the FPN collects features from lower layers of the network before the down-sampling destroys essential characteristics in tiny objects, it helps the detector perform better on these items. In Mask R-CNN, a residual network (ResNet) [40] is used in lieu of the VGG16 [39] used in Faster R-CNN's underlying network architecture. Both Faster and Mask R-CNN are taught to perform classification and region suggestion tasks. A stochastic gradient descent is used to do this. Mask R-CNN uses end-to-end learning to concurrently learn both networks. To do this, the first version of Faster R-CNN used a 4-step optimization procedure that switched back and forth between the two jobs. However, the more recent end-to-end learning strategy from Mask R-CNN may be used with Faster R-CNN as well. Before training on the target domain, it is common practice to initialize both networks using an ImageNet-pretrained model. In the testing phase, both approaches allow for rapid detection and identification. The trained model produces a collection of item bounding boxes for each input picture, with each box having a category name and a softmax score in the range [0, 1].

B. Adaptation to traffic-sign detection

Mask R-CNN is a generic approach to object identification and recognition. We created and implemented a number of TSD-specific enhancements to make it suitable for use in this area.

First, we add OHEM (online hard-example mining) to the Fast R-CNN module, which is responsible for classification learning. We update the technique for picking ROIs that are transmitted to the classification learning module, building on the work of Shrivastava et al. [43], who developed OHEM for Faster RCNN. Typically, 256 regions of interest (ROIs) per picture are chosen at random, with some being designated as foreground (traffic signs), and others as background (nontraffic signs). In our method, ROIs are no longer chosen at random but rather depending on the value of their classification loss. Each region's loss is ranked, and only those with a high enough loss are sent on to the module that learns to classify them. This guarantees that the network is learning from the most challenging cases, or the samples on which it made the most mistakes. To guarantee that each gradient descent step has a enough number of positive and negative samples, we execute selection independently for the background and foreground objects. Using the preexisting classification module, we are able to retrieve the classification losses for ROIs, allowing us to build OHEM as a fully-fledged learning system. Keep in mind that the RPN only calculates the top ROIs based on their object ness score when calculating classification loss, which is a criterion for picking ROIs. To get rid of the redundant ROIs, we use a non-maxima suppression (NMS) on a dataset of 2000 regions. This is a common method used in Mask R-CNN to narrow down the pool of potential regions of interest (ROIs) before making a learning decision. We tried training with more than 2000 areas in front of the NMS, but the slower NMS made this approach prohibitively slow with no discernible improvement in performance.

b) Selected training sample distribution: The suggested method further enhances the technique used to pick training samples for the region proposal network. Mask R-CNN used to choose regions of interest (ROIs) at random. The foreground and background are processed independently. However, random selection causes imbalance into the learning

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

process when both tiny and big items are present in the picture at the same time. The disparity arises because there are many more ROIs covering big objects than covering little ones. The learning process would be skewed if samples were selected according to this distribution, since bigger items would be noticed more often and preferred considerably more than smaller ones. We address this problem by adjusting the proportion of training samples from each object size. To do this, we assign each item in the picture a uniform number of region of interest (ROI) selections.

c) Sample weighting: We use sample weighting to make the learning process more efficient. Our testing revealed that Mask R-CNN falls short of a perfect recall score on occasion because of missing area recommendations. We solve this problem by giving various training zones equal weight. Both foreground and background regions are chosen during training, while the latter tends to be favored due to the rarity of good region proposal candidates for tiny traffic signs. Without any kind of weighting, the learning process will tend to pay more attention to, and so learn more about, items in the background. We solve this issue by giving less weight to the background areas, which makes the network focus on understanding what's in the front. Backgrounds are weighted 0:01 for the region proposal network (RPN) and 0:1 for the classification network (CN) throughout their respective training processes.

Since regions ignored at this stage in the pipeline cannot be recovered by the classification module and would lead to poor overall recall if not addressed, this enhancement is especially critical for the RPN.

d) **Adjusting region pass-through during detection**: Finally, during the detection phase, we modulate the amount of ROIs sent to the classification network via the RPN. Due to the enormous number of relatively tiny objects present in the traffic-sign domain, the total number of areas traversed must be modified. We double it, from 1,000 to 10,000 areas, every FPN level in advance of the NMS. The NMS 2000 regions are kept after combining ROIs from all FPN levels. Enhanced information The size of the training set is a critical component in deep model learning. Millions of trainable parameters render the system indeterminate without a large enough sample size. We suggest an extra data augmentation in addition to a pre-trained model that was trained on 1:2 million photos from ImageNet to help with this problem. Because of how traffic signs work, we can easily create a large number of new samples by applying arbitrary transformations to existing traffic sign instances. By manipulating subsamples of real-world training data, we can generate new synthetic traffic-sign occurrences. The proposed dataset includes pictures of traffic signs tagged with bounding boxes so that they may be extracted from the background of the training photos (see Figure 5). Geometric/shape distortions (including changes in perspective and scale) and appearance distortions (including shifts in brightness and contrast) were carried out. Each instance of a traffic sign was first normalized, and then geometric and visual distortions were applied. We used L*a*b contrast normalization to standardize the appearance normalization, and homography between instance annotation points and a geometric template for a certain traffic-sign class to standardize the geometric normalization. Several classes (such as the railway crossing sign, direction signs with the form of an arrow, etc.) were exceptions, but for the vast majority of classes we manually generated templates. We also created new synthetic instances for these classes, but unlike before, we did not subject them to geometry normalization or geometric distortions.

We followed the distribution of the training set's geometry and appearance variations to develop synthetic training examples that are as realistic as feasible. Both the distribution of Euler rotation angles (along the X, Y, and Z axes) and the distribution of averaged intensity values were determined for the training instances used in the geometry transformation. Using the sizes of instances with their geometry corrected, we also approximated the distribution of scales. We utilized a Gaussian mixture model to account for all shifts, but with just one mixture component (K=1) to account for geometry and appearance and two (K=2) to account for scale. Figure 2 displays a number of instances of authentic, standard, and synthetically manufactured samples.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

Figure 1: Sample traffic signs and their corresponding annotation masks.

A histogram and its corresponding distributions for different distortions are depicted in Figure 3. When generating synthetic distortions we sampled random values from the corresponding distributions. However, variance that is twice as large as the variance in the observed distribution was used to increase the likelihood of generating larger distortions. In the appearance distortion the distributions were not generic for all classes, but instead, we used different distribution for each classes. We used class specific mean instead of mean over all categories but we still applied common variance calculated from all the categories. This guarded us from generating invalid contrast values for very dark/bright categories, such as gray or white direction signs. To emulate the real-world settings, the newly generated traffic-sign instances were inserted into the street-environment like background images. Background images were acquired from the subset of the BTS dataset [17], which contains no other traffic signs. At least two, and at most five, traffic signs were placed in a non-overlapping manner in random locations of each background image, avoiding the bottom central part where only the road is usually seen. With the whole augmentation process, we generated enough new instances to ensure each category has at least 200 instances. This resulted in around 30;000 new traffic-sign instances spread over 8775 new training images.

IV. THE TRAFFIC-SIGN DATASET FROM THE DFG

Our dataset was acquired by the DFG Consulting d.o.o. company for the purpose of maintaining inventory of traffic signs on Slovenian roads. The RGB images were acquired with a camera mounted on a vehicle that was driven through several different Slovenian municipalities. The image data was acquired in rural as well as in urban areas. Only images containing at least one traffic sign were selected from the vast corpus of collected data. Moreover, the selection was performed in such a way that there is usually a significant scene change between any pair of selected consecutive images. Since images were acquired for the purpose of maintaining traffic-sign inventory, this allowed the image acquisition to be performed in the day-time avoiding bad weather conditions

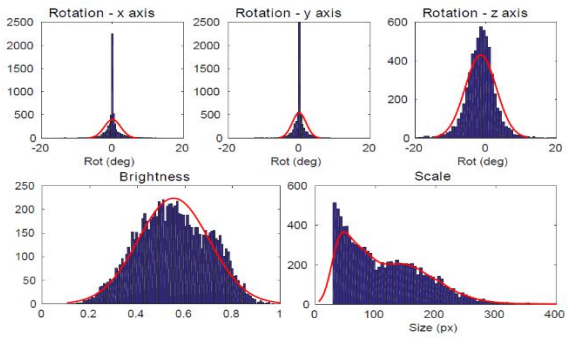


Figure 2: Distributions of traffic-sign distortions computed for rotation in the top row, appearance (i.e. brightness) in the bottom left side and scale in the bottom right side. Red lines represent the Gaussian distributions, which are sampled when generating new examples.

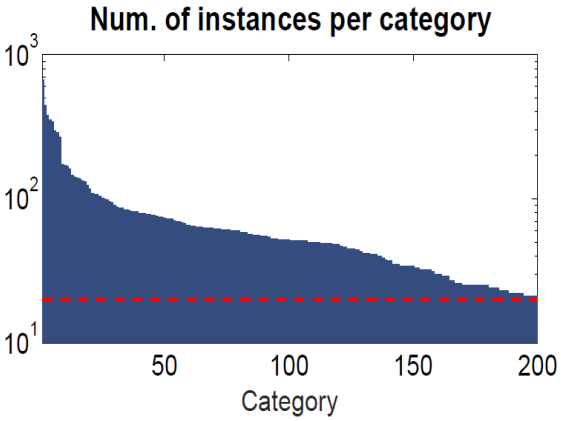
such as rain, snow and fog. Nevertheless, the dataset does include other difficult variations in the weather and the environment that are present in the real-world

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

environment such as: rural and city/urban landscape, different levels of natural occlusions and shadows, and various ranges of a cloudy sky and direct sunlight. Images taken under winter conditions with snow cover were also included. The dataset, termed the DFG traffic-sign dataset1, contains a total of 6957 images with 13;239 tightly annotated traffic-sign instances corresponding to 200 categories. The total number of instances is different for each category (see Figure 1). Each image contains annotations of all traffic signs larger than 25 pixels for any of the 200 categories in a tightly annotated polygon (see Figure 4). Categories in the dataset represent a subset of all categories from the corpus of raw images provided by the company; however, some categories in the corpus did not meet the necessary criteria to create a quality dataset. In particular, all categories in the public dataset now meet the following three criteria: (a) each category has a sufficient number of instances (at least 20 instances with a minimal bounding box size of 30 pixels), (b) each category represents a planar object and (c) each category contains traffic signs that have at least some visual consistencies. Among all categories in the DFG traffic-sign dataset roughly 70% of them correspond to traffic signs with low appearance changes, while a significantly larger appearance variability is present in the remaining 30%. Latter signs can be of variable aspect ratio or color and can contain various text and numbers. See 200 categories of traffic signs depicted in Figure 1. Note that the dataset contains annotations as small as 25 pixels. However, annotations smaller than 30 pixels are flagged as difficult and are not considered neither for the training nor for the testing. We selected 30 pixels as a minimal size based on down-sampling of features in Faster and Mask R-CNN,



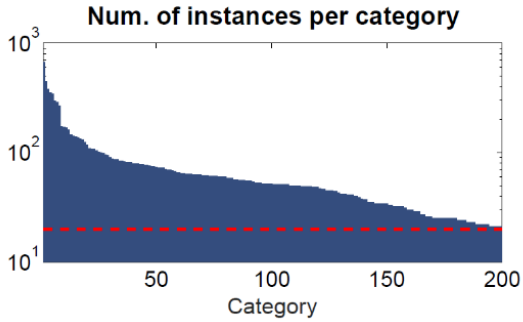


Fig. 3: Distribution of number of instances over categories in the DFG traffic-sign dataset. Horizontal red dashed line represents 20 instances per category, which we use as a cutoff point. Note, the distribution is shown in the logarithmic scale.

which is performed 5-times and results in 32x32 pixels being represented by 1x1 feature pixel. A suitable train-test split was generated to provide a sufficient number of samples for both the training and the test set. A restriction was set that 25% of traffic-sign instances for each category have to appear in the test set. For the smallest categories with only 20 instances, this ensured a minimum number of 15 samples for the training set and a minimum number of 5 samples for the test set. Images were assigned randomly to either the training or the test set. However, additional constraint mechanism was employed to ensure all images of the same physical object are always present either in the test set or in the training set but never in both of them at the same time. This was ensured by clustering images within 50 meter distance and assigning whole clusters to the training or the test set. In this way, we generated a training set with 5254 images and a test set with 1703 images.

V. EXPERIMENTAL EVALUATION

In this section, we perform extensive evaluation of deep learning methods that are appropriate for the traffic-sign detection and recognition. We focus on evaluating two state-of-theart, region-proposal-based methods: Faster R-CNN and Mask R-CNN. We first perform evaluation on the existing public traffic-sign dataset to establish a baseline comparison with the related work. Swedish traffic-sign dataset (STSD) is used for this purpose. Then, an extensive evaluation on newly proposed DFG traffic-sign dataset is performed with a comprehensive analysis of the proposed improvements.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

A. Implementation details

A publicly available Caffe2-based, Python implementation of the Detectron [44] is used for both Faster and Mask RCNN2. For the Faster R-CNN, we employ the VGG16 [39] network with 13 convolutional layers and 3 fully-connected layers, while for the Mask R-CNN, we employ a residual

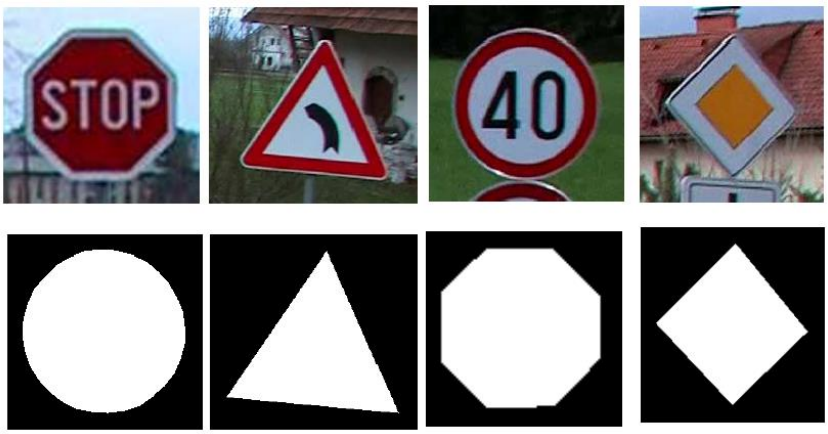


Fig. 4: Several examples of traffic signs in the DFG traffic-sign dataset with their corresponding annotation asks showing the precision of the annotation mask.

network [40] with 50 convolutional layers (ResNet-50). The ResNet-50 architecture consists of 16 convolutional filters with kernel sizes of 3x3 or larger. Mask R-CNN also implements Feature Pyramid Network (FPN) [42], which collects features from different layers of the network to capture the information from small objects, which may be removed in higher layers due to down-sampling. Both networks are initialized with a model pre-trained on ImageNet as provided by [44]. We also experimented with larger variant of the residual network using

101 layers (ResNet-101), but performance did not improve compared to ResNet-50. We therefore focused only on the ResNet-50, which at the same time is faster with half the layers of ResNet-101. Both methods use similar learning hyper-parameters. A learning rate of 0:001 is used for Faster R-CNN with a weight decay of 0:0005, while a learning rate of 0:0025 and a weight decay of 0:0001 is used for Mask R-CNN. Both approaches also use momentum of 0:9. The same hyperparameters are used in all experiments. Note that the same hyper-parameters are used in [44] to pre-train the model on ImageNet dataset. Both methods are trained end-to-end with simultaneous learning of both the region proposal network and the classification network. We learn both methods for 95 epochs and reduce the learning rate by a factor of 10 at the 50th and 75th epoch. We use two images per batch per GPU and train on STSD with 2 GPUs and on DFG dataset with 4 GPUs. This resulted in effectively using 4 images per batch on the STSD and 8 images per batch on the DFG dataset.

B. Performance metrics

Several different metrics are used in this study to evaluate the proposed approach. As a primary metric, we report mean average precision (mAP), which is commonly used in the evaluation of visual object detectors. We use two variants of the mAP: (i) mAP50, based on the PASCAL visual object challenge [45], and (ii) mAP50:95, based on the COCO challenge [46]. Both metrics define a minimal intersectionover- union (IoU) overlap with the groundtruth region for a detection to be considered as a true positive, and both compute

average precision (AP) as the area under the precision-recall curve to accurately capture the trade-off between the miss rate and the false-positive rate.

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

Table 1. Results of Indian Traffic Sign detection.

Traffic sign	**R-CNN**	
	Precision (%)	**Recall (%)**
Speed 30	100	100
Speed 40	100	98.4
Speed 60	99.7	97.5
Speed 70	100	100
Speed 80	99	91.2
No horn	100	90.5
Warning	98.9	98.1
Average	**99.6**	**96.75**

Table 2. Precision and Recall % comparison for Indian Traffic Sign detection.

Traffic sign	**Fast R-CNN**		**Mask RCNN**	**R-CNN**		
	Prec (%)	**Rec (%)**	**Prec (%)**	**Rec (%)**	**Prec (%)**	**Rec (%)**
Sign 40	92.1	92.5	98	96.5	98.7	97.6
Sign 60	91.3	97.1	94.5	97.3	99.5	98.1
Sign 70	92.4	92.4	81.4	99.1	88.7	98.5
Go left	99.1	94.2	99.5	92.4	98.6	93.6
Go right	98.9	95.1	94.6	95.5	96.3	95.1
warning	81.2	91.5	96.5	96.5	98.7	98.5
Speed 80	96.4	95.2	96.8	96.8	99.1	95.9
Average	93.05	94	94.4	96.7	97.08	96.75

Note: Prec – Precision, Rec – Recall.

VI. RESULT ANALYSIS

In this section, we demonstrate the performance of our approach on traffic-sign detection with additional qualitative analysis. We focus only on the best performing model, namely Mask R-CNN using ResNet-50 with our adaptations and data augmentation. All results in this section are reported on the test set of the DFG traffic-sign dataset. A per-class distribution of AP50 is depicted in Figure 2. This graph clearly shows that a large number of traffic-sign classes (108) are detected and recognized with average precision of 100%, i.e. with no errors. For the remaining categories our approach still achieved AP of above 90% on 60 of them, and above 80% on 23 of them. Figure 3 further

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

shows the traffic-sign classes with their corresponding AP50 sorted by their AP50 in descending order. The best performing categories at the top of the list are mostly traffic signs with low intra-category variations, i.e. with fixed sizes and fixed appearance. This includes various triangular danger signs, circular prohibitory signs, speed limit signs, rectangular information signs, etc. On the other hand, the worst performing signs at the bottom are traffic signs with a large variation of their sizes/aspect ratios as well as with a large intra-category variations, i.e., their content significantly varies from instance to instance.

Fig. 5: DFG traffic-sign categories sorted by average precision (AP50) calculated when using Mask R-CNN ResNet-50 with our adaptations and data augmentation.

This includes particularly complex class of mirrors (both rectangular and round mirrors), speed feedback signs, various direction signs and signs marking the start or the end of the towns. Traffic signs with high intra-category variations and good performance: Figure 6 reveals several traffic signs with extremely good detection rate despite having large intra-category variations in their appearance. Samples for three such traffic sign categories are depicted in Figure 6, namely they are: (i) large-direction-with-separate-lanes, (ii) left-arrow-shaped direction and (iii) right-gray-direction. Each row in this figure depicts one category with eight instances. For clarity we display only the relevant part of the image. True detections are shown in green, false detections in red and missing detections in magenta. Examples are also sorted by their descending detection score from left-to-right. Therefore if true (green) and false (red) positive detections can be successfully separated with a threshold then false detections can be trivially eliminated by setting an appropriate detection threshold. Note that this is important when looking at false detections as many of them are not problematic at all.

Figure 6: Examples of complex traffic signs with variable content and good detection on the test set of the DFG traffic-signdataset. True positives are depicted in green, false positives in red, and missing detections (false negatives) in magenta. (*)

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

Note, the last detection in the first row is not false since actual traffic sign was not annotated due to high occlusion.

When focusing on the large-direction-with-separate-lanes traffic-sign category in the first row in Figure 6, an extremely good performance is clearly shown for the traffic signs that have quite significant variation in their content as well as large variation in their sizes and aspect ratios. The first image in the top row depicts a good example of this as the traffic sign was detected with a high score despite having completely different color combination than other instances of the same class. Several detected instances are also quite small, yet our approach successfully detects them. Moreover, the last image in the first row shows a false detection of a small instance; however, a close inspection reveals that it is a correct detection. This instance was not annotated in the dataset due to small size and high occlusion of the tree. The second row in Figure 6 depicts detections of a left arrow- shaped-direction traffic sign. This category is fairly difficult to detect as aspect ratios vary quite significantly from instance to instance, mostly due to wide viewing angles, yet the detector did not have significant issues finding them. The second-to-last example in the second row is also significantly cropped; however, the detector is still able to correctly find it. Finally, detections for the right-gray-direction traffic sign are shown in the last row in Figure 6. Detection of this category is difficult mostly due to significant variation of the content. Those traffic signs also often appear side-by side in multiple rows which makes it difficult to generate the correct region proposal. Nevertheless, most instances have been correctly found.

Figure 7: Examples of traffic signs with fixed content but poor detection on the test set of the DFG traffic-sign dataset. Truepositive detections are marked in green, false positives in red and missing detections (false negatives) in magenta. (*) Note that false detections in the first row occur due to two almost identical traffic-sign categories in the dataset (one with distance label below and one without). True detections with the other category detector are shown in dashed green line.

Traffic signs with poor performance and low intra category variations: Next, we focus on three worst performing traffic signs despite having low appearance variation within a category, namely: (i) left-into-right-lane-merger, (ii) train crossing and (iii) work-in-progress. Samples are depicted in Figure 7 and are organized in a similar manner as in Figure 6, with eight examples per category in a row, sorted by their descending detection score. The worst results are achieved for the left-into-right-lanemerger traffic sign with the AP50 of 57%. Mask R-CNN correctly detects four out of five test instances, but appears to detect four

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

false traffic signs as well, as can be seen in the top row. However, those false detections should not be considered problematic as the traffic sign is identical to the left-into right- lane-merger sign with the only difference in the distance value printed below the sign. Since the correct category is also detected (shown with the dashed green line), those false detections would be eliminated by the across-category non maxima suppression, meaning that even in this case the issue is not as bad as it might seem. Still, such extremely minor differences between those two categories appear to pose a challenge for deep learning and point to a existing limitations of deep learning methods.

CONCLUSION

In this work, we have addressed the problem of detecting and recognizing a large number of traffic-sign categories for the main purpose of automating traffic-sign inventory management. Due to a large number of categories with small interclass but high intra-class variability, we proposed detection and recognition utilizing an approach based on the Mask RCNN [14] detector. The system provides an efficient deep network for learning a large number of categories with an efficient and fast detection. We proposed several adaptations to Mask R-CNN that improve the learning capability on the domain of traffic signs. Furthermore, we proposed a novel data augmentation technique based on the distribution of geometric and appearance distortions. As an important contribution, we also present a novel dataset, termed the DFG traffic-sign dataset, with a large number of traffic-sign categories that have low inter-class and high intra-class variability. This datasetas been made publicly available together with the codefor our improvements, allowing the research community to make further progress on this problem and enabling reliable and fair comparison of different methods on a large-scale
traffic-sign detection problem. We also extensively evaluated our proposed improvements and compared them against the original Faster and Mask R-CNN. Our evaluation on the DFG and the Swedish traffic-sign datasets showed that the proposed adaptations improve the performance of Mask R-CNN in several metrics. This includes improvement in the miss rate of the RPN network for smaller objects, improvement in the overall recall of the full pipeline for both small and large objects, as well as improvement in the overall performance in the mean average precision. Despite excellent performance of the proposed approach there is still room for improvement. Our analysis revealed that the ideal performance is still not achieved, mostly due to several missed detections that are being lost by the classification network. Future improvements should focus on improving this part of the system.

REFERENCES

1. Domen Tabernik and Danijel Skoˇcaj, Deep Learning for Large-Scale Traffic-Sign Detection and Recognition, https://arxiv.org/pdf/1904.00649.
2. RajeshKannan Megalingam, Kondareddy Thanigundala, Sreevatsava Reddy Musani, Hemanth Nidamanuru, Lokesh Gadde, Indian traffic sign detection and recognition using deep learning, International Journal of Transportation Science and Technology
Volume 12, Issue 3, September 2023, Pages 683-699
3. V. Balali, A. Ashouri Rad, and M. Golparvar-Fard, “Detection, classification, and mapping of U.S. traffic signs using google street view images for roadway inventory management,” Visualization in Engineering, vol. 3, no. 1, p. 15, 2015. 1
4. K. C. Wang, Z. Hou, and W. Gong, “Automated road sign inventory system based on stereo vision and tracking,” Computer-Aided Civil and Infrastructure Engineering, vol. 25, no. 6, pp. 468–477, 2010. 1
5. V. Balali and M. Golparvar-Fard, “Evaluation of Multiclass Traffic Sign Detection and Classification Methods for U.S. Roadway Asset Inventory Management,” Journal of Computing in Civil Engineering, vol. 30, no. 2, p. 04015022, 2016. 1
6. J. M. Lillo-Castellano, I. Mora-Jimenez, C. Figuera-Pozuelo, and J. L. Rojo-Alvarez, “Traffic sign segmentation and classification using statistical learning methods,” Neurocomputing, vol. 153, pp. 286–299, 2015. 1, 2

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

7. M. Haloi, “A novel pLSA based Traffic Signs Classification System,” CoRR, vol. abs/1503.0, 2015. 1, 3
8. Y. Zhu, C. Zhang, D. Zhou, X. Wang, X. Bai, and W. Liu, “Traffic sign detection and recognition using fully convolutional network guided proposals,” Neurocomputing, vol. 214, pp. 758–766, 2016. 1, 2, 3, 7, 8
9. R. Timofte, V. A. Prisacariu, L. J. V. Gool, and I. Reid, “Combining Traffic Sign Detection with 3D Tracking Towards Better Driver Assistance,” in Emerging Topics in Computer Vision and its Applications,2011, pp. 425–446. 1
10. Mr. Pathan Ahmed Khan, Dr. M.A Bari,: Impact Of Emergence With Robotics At Educational Institution And Emerging Challenges”, International Journal of Multidisciplinary Engineering in Current Research(IJMEC), ISSN: 2456-4265, Volume 6, Issue 12, December 2021,Page 43-46
11. A. Mogelmose, “Visual Analysis in Traffic & Re-identification,” Ph.D. dissertation, Faculty of Engineering and Science, Aalborg University, 2015. 1
12. Ishaq Bin Mohammed Barabood, Mohd Mohsin Uddin, Mohd Faraz Uddin, Mrs. Asra Sultana, Cellar Ventillation System With Auto Detection And Control, International Journal of Multidisciplinary Engineering in Current Research - IJMEC Volume 8, Issue 4, April-2023, http://ijmec.com/, ISSN: 2456-4265.
13. J. Stallkamp, M. Schlipsing, J. Salmen, and C. Igel, “Man vs. computer: Benchmarking machine learning algorithms for traffic sign recognition,” Neural Networks, vol. 32, pp. 323–332, 2012. 1, 2
14. S. Houben, J. Stallkamp, J. Salmen, M. Schlipsing, and C. Igel, “Detection of traffic signs in real-world images: The German traffic sign detection benchmark,” in IJCNN. Ieee, aug 2013, pp. 1–8. 1, 2, 3
15. A. Mogelmose, M. M. Trivedi, and T. B. Moeslund, “Vision-Based Traffic Sign Detection and Analysis for Intelligent Driver Assistance Systems: Perspectives and Survey,” Transactions on Intelligent Transportation Systems, vol. 13, no. 4, pp. 1484–1497, 2012. 1, 2
16. F. Zaklouta and B. Stanciulescu, “Real-time traffic-sign recognition using tree classifiers,” IEEE Transactions on Intelligent Transportation Systems, vol. 13, no. 4, pp. 1507–1514, 2012. 1, 3
17. Vishnuvardhan Reddy, G. Shivani, J. Sowmya, M. Jyothi, P. Sai Prasad Reddy, B. Vinay Kumar, Design And Analysis Of Auditorium By Using Staad PRO, International Journal of Multidisciplinary Engineering in Current Research - IJMEC Volume 8, Issue 5, May-2023, http://ijmec.com/, ISSN: 2456-4265.
18. Z. Zhu, D. Liang, S. Zhang, X. Huang, B. Li, and S. Hu, “Traffic-Sign Detection and Classification in the Wild,” in CVPR, 2016, pp. 2110– 2118. 1, 2, 3
19. H. Kaiming, G. Gkioxara, P. Dollar, and R. Girshick, “Mask R-CNN,” in International Conference on Computer Vision, 2017, pp. 2961–2969. 2, 3, 12.
20. Mohammed Nadeem Shareef, Junaid Hussain, Mohammed Khaja Adnan Ali Khan,, Dr. Mohammed Abdul Bari .” Crypto Jacking”, Mathematical Statistician and Engineering Applications, ISSN: 2094-0343, 2326-9865, Vol 72 No. 1 (2023), Page Number: 1581 – 1586
21. Mohammed Fahad, Asma Akbar, Saniya Fathima, Dr. Mohammed Abdul Bari ,” Windows Based AI-Voice Assistant System using GTTS”, Mathematical Statistician and Engineering Applications, ISSN: 2094-0343, 2326-9865, Vol 72 No. 1 (2023), Page Number: 1572 – 1580
22. Naif Ismail Ibrahim, Mohd Kamran Khadeer, Mohammed Abdul Kareem, Vikas Kumar Tiwari, Mohammad Rafeeq, Mr.B.Tejavardhan, Design Optimization And Analysis Of Aircraft Landing Gear, International Journal of Multidisciplinary Engineering in Current Research - IJMEC Volume 8, Issue 6, June-2023, http://ijmec.com/, ISSN: 2456-4265.
23. Syed Shehriyar Ali, Mohammed Sarfaraz Shaikh, Syed Safi Uddin, Dr. Mohammed

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

Abdul Bari, “Saas Product Comparison and Reviews Using Nlp”, Journal of Engineering Science (JES), ISSN NO:0377-9254, Vol 13, Issue 05, MAY/2022

24. Kassa Mahesh, Male Sathyam Goud, Md. Abdul Khadar, Dasari Sandhya, Mr.Mohd Abdul Hafeez, Mechanical Design And Analysis Of Automatic Pipe Bending Machine, International Journal of Multidisciplinary Engineering in Current Research - IJMEC Volume 8, Issue 6, June-2023, http://ijmec.com/, ISSN: 2456-4265.
25. Hafsa Fatima, Shayesta Nazneen, Maryam Banu, Dr. Mohammed Abdul Bar,” Tensorflow-Based Automatic Personality Recognition Used in Asynchronous Video Interviews”, Journal of Engineering Science (JES), ISSN NO:0377-9254, Vol 13, Issue 05, MAY/2022
26. Mohammed Abdul Bari, Shahanawaj Ahamad, Mohammed Rahmat Ali,” Smartphone Security and Protection Practices”, International Journal of Engineering and Applied Computer Science (IJEACS) ; ISBN: 9798799755577 Volume: 03, Issue: 01, December 2021
27. Shahanawaj Ahamad, Mohammed Abdul Bari, Big Data Processing Model for Smart City Design: A Systematic Review “, VOL 2021: ISSUE 08 IS SN : 0011-9342 ;Design Engineering (Toronto) Elsevier SCI Oct
28. Mohammed Abdul Bari, Shahanawaj Ahamad, Mohammed Rahmat Ali,” Smartphone Security and Protection Practices”, International Journal of Engineering and Applied Computer Science (IJEACS) ; ISBN: 9798799755577 Volume: 03, Issue: 01, December 2021 (International Journal,U K) Pages 1-6
29. Mohammed Shoeb, Mohammed Akram Ali, Mohammed Shadeel, Dr. Mohammed Abdul Bari, “Self-Driving Car: Using Opencv2 and Machine Learning”, The International journal of analytical and experimental modal analysis (IJAEMA), ISSN NO: 0886-9367, Volume XIV, Issue V, May/2022
30. S. B. Wali, M. A. Hannan, A. Hussain, and S. A. Samad, “Comparative Survey on Traffic Sign Detection and Recognition: a Review,” Przeglad Elektrotechniczny, vol. 1, no. 12, pp. 40–44, 2015. 2
31. A. Ellahyani, M. E. Aansari, and I. E. Jaafari, “Traffic Sign Detection and Recognition using Features Combination and Random Forests,” IJACSA, vol. 7, no. 1, pp. 6861–6931, 2016. 2, 3
32. R. Timofte, K. Zimmermann, and L. V. Gool, “Multi-view traffic sign detection, recognition, and 3D localisation,” in WACV, 2009, pp. 1–8. 2, 5
33. S. Segvic and K. Brkic, “A computer vision assisted geoinformation inventory for traffic infrastructure,” in 13th International IEEE Conference on Intelligent Transportation Systems (ITSC), 2010, pp. 66–73. 2
34. Mr. Pathan Ahmed Khan, Dr. M.A Bari,: Impact Of Emergence With Robotics At Educational Institution And Emerging Challenges”, International Journal of Multidisciplinary Engineering in Current Research(IJMEC), ISSN: 2456-4265, Volume 6, Issue 12, December 2021,Page 43-46
35. Mohammed Abdul Bari, Shahanawaj Ahamad, Mohammed Rahmat Ali,” Smartphone Security and Protection Practices”, International Journal of Engineering and Applied Computer Science (IJEACS) ; ISBN: 9798799755577 Volume: 03, Issue: 01, December 2021
36. Shahanawaj Ahamad, Mohammed Abdul Bari, Big Data Processing Model for Smart City Design: A Systematic Review “, VOL 2021: ISSUE 08 IS SN : 0011-9342 ;Design Engineering (Toronto) Elsevier SCI Oct
37. Mohammed Abdul Bari, Shahanawaj Ahamad, Mohammed Rahmat Ali,” Smartphone Security and Protection Practices”, International Journal of Engineering and Applied Computer Science (IJEACS) ; ISBN: 9798799755577 Volume: 03, Issue: 01, December 2021 (International Journal,U K) Pages 1-6

Journal of Chemical Health Risks

www.jchr.org

JCHR (2023) 13(3), 1238-1253 | ISSN:2251-6727

38. Z. Huang, Y. Yu, J. Gu, and H. Liu, "An Efficient Method for Traffic Sign Recognition Based on Extreme Learning Machine," IEEE Transactions on Cybernetics, no. 99, pp. 1–14, 2016. 2, 3
39. F. Larsson and M. Felsberg, "Using fourier descriptors and spatial models for traffic sign recognition," Image Analysis, no. May, pp. 238– 249, 2011. 2
40. H. Li, F. Sun, L. Liu, and L. Wang, "A novel traffic sign detection method via color segmentation and robust shape matching," Neurocomputing, vol. 169, pp. 77–88, 2015. 2
41. X. Yang, Y. Qu, and S. Fang, "Color Fused Multiple Features for Traffic Sign Recognition," in ICIMCS, 2012, pp. 84–87. 2
42. Dr. Abdul Wasay Mudasser, Dr. Pathan Ahmed Khan, "Artificial Intelligence Usage in Wireless Sensor Network: An Overview", International Journal of Multidisciplinary Engineering in Current Research(IJMEC), ISSN: 2456-4265, Volume 7, Issue 10, October 2022,Page 9-14.
43. S. Salti, A. Petrelli, F. Tombari, N. Fioraio, and L. D. Stefano, "Traffic sign detection via interest region extraction," Pattern Recognition, vol. 48, no. 4, pp. 1039–1049, 2015. 2
44. Dr. M.A.Bari, "EffectiveIDS To Mitigate The Packet Dropping Nodes From Manet ", JACE, Vol -6,Issue -6,June 2019
45. M.A.Bari & Shahanawaj Ahamad," Process of Reverse Engineering of Enterprise InformationSystem Architecture" in International Journal of Computer Science Issues (IJCSI), Vol 8, Issue 5, ISSN: 1694-0814, pp:359-365,Mahebourg ,Republic of Mauritius , September 2011
46. M.A.Bari & Shahanawaj Ahamad, "Code Cloning: The Analysis, Detection and Removal", in International Journal of Computer Applications(IJCA),ISSN:0975-887, ISBN:978-93-80749-18-3,Vol:20,No:7,pp:34-38,NewYork,U.S.A.,April 2011
47. Ijteba Sultana, Mohd Abdul Bari and Sanjay," Impact of Intermediate Bottleneck Nodes on the QoS Provision in Wireless Infrastructure less Networks", Journal of Physics: Conference Series, Conf. Ser. 1998 012029 , CONSILIO Aug 2021
48. J. Greenhalgh and M. Mirmehdi, "Real-Time Detection and Recognition of Road Traffic Signs," Transactions on Intelligent Transportation Systems, vol. 13, no. 4, pp. 1498–1506, 2012. 2, 3
49. G. Overett and L. Petersson, "Large scale sign detection using HOG feature variants," in Intelligent Vehicles Symposium, 2011, pp. 326–331. 2