

Received September 18, 2019, accepted September 29, 2019, date of publication October 11, 2019, date of current version October 24, 2019.

Digital Object Identifier 10.1109/ACCESS.2019.2947069

Classification of Degraded Traffic Signs Using Flexible Mixture Model and Transfer Learning

ABDUL MANNAN^{1,2}, KASHIF JAVED¹, ATTA UR REHMAN¹, HAROON A. BABRI¹, AND SEROSH K. NOON²

¹Department of Electrical Engineering, University of Engineering and Technology, Lahore 54890, Pakistan

²Department of Electrical Engineering, NFC Institute of Engineering and Technology, Multan 59080, Pakistan

Corresponding author: Abdul Mannan (mannan@nfciet.edu.pk)

ABSTRACT Automatic detection and recognition of traffic signs is a topic of research for various applications like driver assistance, inventory management and autonomous driving. Poorly maintained traffic signs degrade by losing their colors or some part is weird due to aging and hence making the task more challenging. The problem is mainly related to the developing world and has gained less attention in the literature on automatic traffic sign detection and recognition. To handle the degradation issue, we present a novel flexible Gaussian mixture model based technique with automatic split and merge strategy. This adaptive scheme works as a preprocessing step which facilitates locating traffic signs under a certain degree of degradation in a real world scenario. A multiscale convolutional neural network augmented with dimensionality reduction layer is proposed to recognize contents of the sign. Since, there is no available benchmark dataset for this purpose, we collected a number of images containing partially degraded signs from the famous German Traffic Sign Detection Benchmark (GTSDB) and augmented it with manually and naturally degraded traffic sign images taken from the longest highway in the country of authors' residence. Experimental results show that our proposed technique outperforms many state of the art and recent methods for detection and recognition of degraded traffic signs.

INDEX TERMS Degraded traffic sign, Gaussian mixture model, Transfer learning, Convolutional neural networks, dimensionality reduction.

I. INTRODUCTION

Traffic signs are placed to convey important road and road-side information to drivers e.g., the recommended speed or warning about a pedestrian or school etc. In almost all European and Asian countries, traffic signs are mainly found with either red rim or filled with blue color. However, in the country of authors' residence, there are mostly the former category of signs, therefore this work is focused on red bordered traffic signs only. These signs are either circular (mandatory) or triangular (danger/warning). In order to facilitate drivers and reduce potential accidents, computer vision and image processing techniques are in place for detection and recognition of traffic signs since early 90s [1]. Detection is the process of segmenting out a region in a captured image which contains a traffic sign whereas recognition is meant for reading its contents. Broadly, we find four different types

of methods used for segmenting a traffic sign in a complex outdoor scene i.e., (1) thresholding in various color spaces like RGB, HSI, etc. [1]–[5] (2) using statistical parameter estimation for colors of interest [6]–[8] (3) using saliency and/or dictionary learning to segment windows containing a traffic sign with a certain confidence level [9], [10] and (4) using convolutional neural networks (CNN) to indicate that a real world image contains a traffic sign [11]–[13]. CNN based methods are very famous but most of the times post processing is required to exactly crop the portion containing pictogram of the traffic sign [11]. The second step i.e., recognition refers to the identification of internal contents that contain instructions for the driver. There are mainly two types of methods i.e. (1) techniques that use hand crafted feature extraction e.g., histogram of oriented gradients [14], local binary patterns [15], scale invariant feature transform and spectral methods like discrete wavelet transform and discrete cosine transform [1] etc. (2) feature learning based methods where features are learned from the training data like

The associate editor coordinating the review of this manuscript and approving it for publication was Wei Liu.

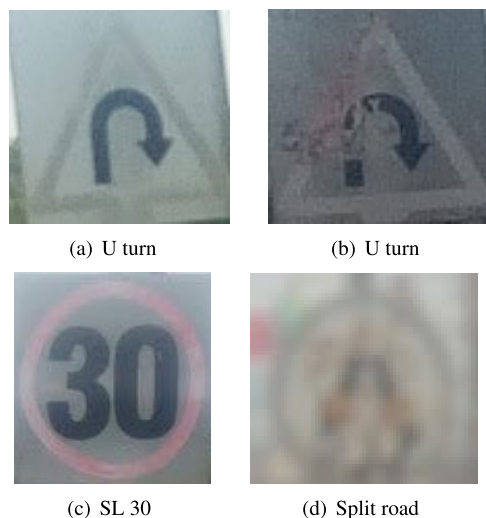


FIGURE 1. Examples of degradations found in real world traffic signs.

convolutional neural network (CNN) [6], [16]. CNN based methods are useful even for recognition purposes but since a lot of parameters are to be learned, they require huge training data and need specialized computing hardware [17], [18].

Traditionally, traffic signs are made in such a way that they provide large amount of contrast between the outer red rim and the background objects (e.g., buildings or trees etc.) and at the same time between the foreground and background of their internal contents. However, in some cases traffic sign detection and recognition becomes more challenging due to aging and/or poor maintenance. Traffic signs are degraded in one of the following two ways (especially in the developing world):

- Outer red rim loses its color and
- Contents are weired

Examples of some degraded signs are shown in Fig. 1. It can be seen that either the outer red rim has lost its color by becoming pink or the internal contents are not properly visible due to aging. These complications make the tasks of detection and recognition more challenging. Since existing color based detection techniques are not designed to tackle weired/degraded signs, they tend to miss traffic signs and produce higher false negative rate. To handle these issues, the proposed method presents a novel scheme especially designed for detection and recognition of degraded traffic signs. Our main contribution is summarized as under:

- In order to detect degraded traffic signs in natural scenes, a unique flexible automatic split and merge strategy based on Gaussian mixture model is proposed.
- For recognition of degraded traffic signs, a pretrained convolutional neural network (CNN) is employed to extract multiscale features from natural and synthetically occluded images.
- By replacing the fully connected layer of the pretrained CNN with newly added dimensionality reduction and classification layers, the system is able to be benefited

of CNN's recognition accuracy with reduced computational complexity.

Rest of this paper is organized as follows: Section II describes work related to degraded traffic sign and convolutional neural networks, Section III explains at length our proposed method for detection and recognition of degraded traffic signs, in Section IV, we provide experimental settings and results and comparison with other state of the art and recent methods, finally, Section V demonstrates conclusion and future work.

II. RELATED WORK

As far as this work is concerned, we concentrate on degraded traffic signs only. For state of the art and recent work on standard traffic sign detection and recognition, we refer to [1] and [19]. This section is dedicated to related work (though small) regarding detection and recognition of degraded red bordered traffic signs (both prohibitory and warning).

In the plethora of literature on traffic sign detection and recognition, only a few efforts can be found for degraded signs [20]–[22]. Floros *et al.* [22] presented a method to detect degraded signs in images with the help of thresholding in famous RGB domain. Due to the unavailability of benchmark datasets, they collected a data of 300 images in two cities of Greece but the method performed poorly due to lesser number of images available for training and testing purposes. Li and Ma [21] and Li *et al.* [23] proposed a combination of blurring, affine invariants and probabilistic neural network to recognize degraded and occluded traffic signs on synthetic images. Their work was, however, focused mainly on tuning parameters of a probabilistic neural network. Another related work on degraded traffic signs was presented by Ishida *et al.* [20]. They used statistical feature extraction in combination with multiclass support vector machine classifier to recognize degraded traffic signs. For experiments, they used a synthetically generated dataset of standard and degraded traffic signs. Further, we find a few efforts for detection of occluded traffic signs i.e., a related issue. In some cases, a traffic sign can be occluded by a tree or a pole etc. which partially hide contents. Rehman *et al.* [24] use tone mapping and color channel histograms to detect partially occluded traffic signs for a synthetic dataset. Color Cubic local binary pattern technique along with adaboost classifier was used in [25] for partially occluded traffic signs detection. Another technique based on HOG-LBP combination was present by Hou *et al.* [26] for occluded signs detection and was found partially successful for their recognition too.

Deep learning models such as convolutional neural networks (CNNs) have been designed especially for images and have shown excellent performance in many areas of computer vision and image processing [27]. Previously, CNNs of varying architectures have been widely used to recognize contents of standard traffic signs [1], [6] but their training requires a lot of data and specialized computational resources like Graphical Processing Units (GPUs). For many applications, pretrained CNN models like googlenet [28] or Alexnet [29] have

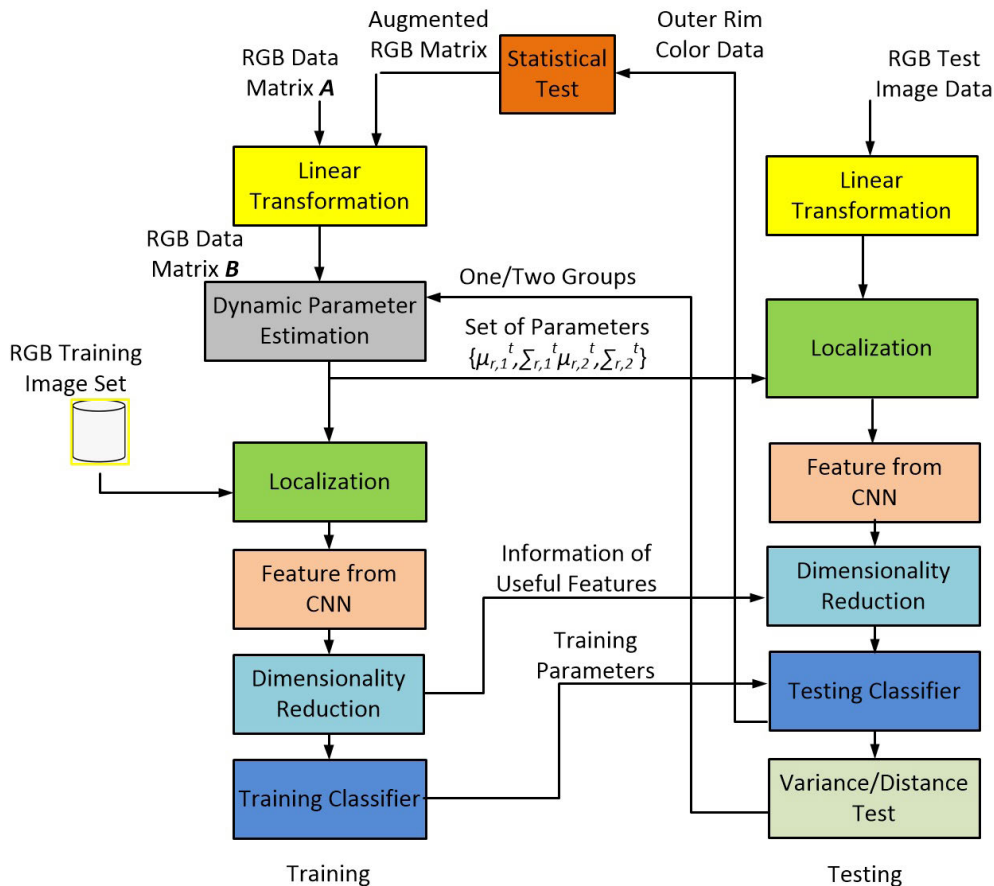


FIGURE 2. Block diagram of our proposed method for detection and recognition of degraded traffic signs. The left branch shows steps performed for training whereas the right branch describes testing. Arrows extending between "Training" and "Testing" branches show exchange of information as the system runs on test data.

been used to utilize features already learnt by the network i.e. these already trained networks are mainly used for the purpose of feature extraction with minimum computational burden. In literature, we find many efforts to detect and recognize standard traffic signs using CNNs [1], [19] but their use for degraded signs is yet to be explored.

To summarize, we find a number of methods for standard traffic sign detection and recognition but very few for signs degraded or partially occluded by other objects. Moreover, convolutional neural networks (CNNs) are commonly used to recognize contents of standard traffic signs but their use for degraded and partially occluded signs is limited.

III. OUR PROPOSED METHOD

In this section, we provide detailed description of our proposed detection and recognition method for degraded signs. Block diagram of the proposed method is shown in Fig. 2. The left branch is for training whereas the right branch shows steps for testing. Details of blocks, forward and feedback links are discussed at length below.

A. DEALING WITH DEGRADATIONS

Detection of a degraded sign is challenging for color based segmentation because either colors used to draw the traffic

signs are faded or some contents have been removed due to aging. There are other methods capable of picking signs on the basis of their shapes only [30], [31] but they miss important information contained in color(s). To locate a degraded sign in an image, we present a technique that utilizes both color and shape information. To tackle the issue of faded outer rim color, a novel flexible linear mapping technique has been proposed.

1) LINEAR TRANSFORMATION

RGB data is collected from training images set and is manually divided in two classes i.e., (1) outer rim class and (2) the background class. As shown in Fig. 3, this data is then placed in the form of a matrix **A** whose each row contains RGB information of a pixel. The total rows are equal to the number of pixels in each image (p) times the total number of training images (n) making its size equal to $(n \times p) \times 3$. A linear transform is then invoked on matrix **A** ensuring that (1) Euclidean distance between pixels corresponding to the outer rim and the background pixels can be maximized (2) the dimensionality of the input RGB data be reduced. Mathematically, for input RGB data matrix **B** obtained by collected sample data from training images, the transformed

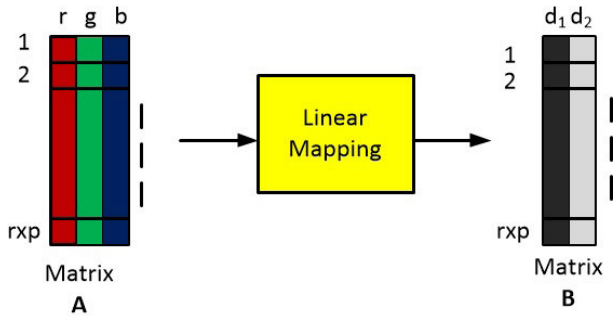


FIGURE 3. Mapping RGB data to transformed domain using a linear transformation.

matrix A is obtained with the help of Equation 1.

$$b_k^T = E \times a_k^T \quad (1)$$

Here a_k^T is the transpose of RGB data picked from k^{th} row in data matrix A and b_k^T is the same data mapped to a two dimensional space. Previous studies have shown the effectiveness of Ohta transform based two dimensional space for segmentation of traffic signs in natural outdoor scenes [7], [8]. In Equation 1, elements of matrix E are the coefficients of Ohta transform [32] ensuring that the higher dimensional data is converted to a lower dimension by maximizing distance among pixels of different classes and simultaneously minimizing distance among pixels of the same class [6], [7], [33]. As mentioned earlier, we have two classes to handle i.e., the background class and the outer rim class. Further, distribution of pixels in the outer rim class is assumed to be Gaussian with a certain mean and covariance.

2) DYNAMIC PARAMETER ESTIMATION

In the transformed domain matrix (B), the two dimensional data corresponding to the outer rim class is assumed to have Gaussian shape. In order to estimate its parameters, we initially use maximum likelihood estimation (MLE) technique [34]. But as the system runs, parameters are updated in a flexible style according to the following rules:

- Parameters are updated at run time whose instantaneous values are shown with superscript t . For example, $\mu_{r,i}$ and $\Sigma_{r,i}$ are mean and variance of i^{th} outer rim group at any times instant t .
- As the system runs, the group containing pixels corresponding to the outer rim splits into two if the variance exceeds a certain threshold. Parameters set for the outer rim class has a provision of the following four parameters i.e.,

$$\Theta^t = \{\mu_{r,1}^t, \mu_{r,2}^t, \Sigma_{r,1}^t, \Sigma_{r,2}^t\}$$

$\mu_{r,1}^t$ and $\Sigma_{r,1}^t$ are dynamic parameters for outer rim class when the training data consists of visible and legible fine traffic signs and we use only one group for the purpose of segmentation. The shape of the group is given in Fig. 4(a). As the system runs on a mixture of fine and degraded signs,

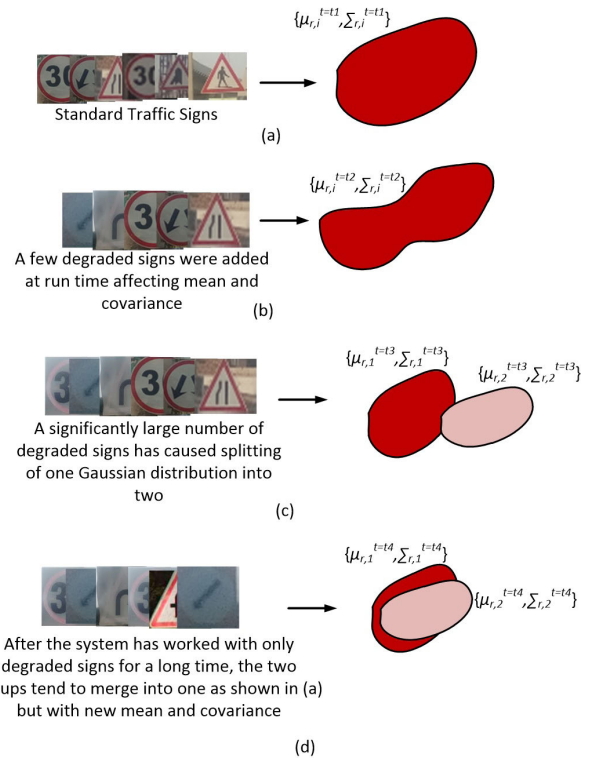


FIGURE 4. Process of splitting and merging the assumed distribution for outer rim colors indexed by time (best seen in color) (a) shows the case when the training data contains only one red outer rim color and hence only one group (b) when the dynamically added data corresponds to degraded images includes pink outer rim colors and the group variance is increased (c) After the variance has exceeded a certain threshold (th_1), the group splits into two and the parameters are now estimated with the help of expectation maximization algorithm on Gaussian mixture models (d) After some time, if the recent history suggests that the distance between the two groups reduces below th_2 , the two groups are merged into one.

the class variance increases as shown in Fig. 4(b). On fulfillment of the following condition, the group splits in two:

$$trace(\Sigma_{r,i}^t) > th_1$$

Here $trace(\Sigma_{r,i}^t)$ is the trace of covariance matrix and th_1 is an empirically chosen threshold.

When the outer rim group of pixels (taken from matrix B) splits into two, parameters must be estimated by considering it a mixture of 2 two-dimensional Gaussian distributions. It should be noted here that setting maximum number of groups equal to two and minimum equal to one was found to be the best option empirically. Parameters of the mixture are estimated using expectation maximization algorithm [35] with an objective to find mean and variance of the two probability distributions related to the two groups present in the data. The probability associated with each data point (b) is given by:

$$P(b) = \sum_{i=1}^2 \pi_{r,i}^t P(b|\mu_{r,i}^t, \Sigma_{r,i}^t)$$

where $P(b|\mu_{r,i}^t, \Sigma_{r,i}^t)$ is the likelihood of data b belonging to distribution $\mathcal{N}(\mu_{r,i}^t, \Sigma_{r,i}^t)$ and $\pi_{r,i}^t$ is the prior probability of group i such that at any time instant t , $\sum_{i=1}^2 \pi_{r,i}^t = 1$.

Expectation maximization (EM) process for the proposed flexible mixture model starts with a random guess of means and variances for two Gaussian distributions and iterates between expectation (E) and maximization (M) steps. During expectation (E) step, belongingness of each data item to the two distributions is determined as follows:

$$P(\mu_{r,i}^t, \Sigma_{r,i}^t | b) = \frac{\pi_{r,i}^t P(b|\mu_{r,i}^t, \Sigma_{r,i}^t)}{\sum_{l=1}^2 \pi_{r,l}^t P(b|\mu_{r,l}^t, \Sigma_{r,l}^t)}$$

After we have determined the probability with which each data point in RGB matrix B was generated by the two Gaussian distributions, parameters $\mu_{r,i}^t$ and $\Sigma_{r,i}^t$ are updated using maximization (M) step as follows:

$$\mu_{r,i}^t = \frac{\sum_{m=1}^{n \times p} b^m P(\mu_{r,i}^t, \Sigma_{r,i}^t | b^m)}{\sum_{m=1}^{n \times p} P(\mu_{r,i}^t, \Sigma_{r,i}^t | b^m)}$$

$$\Sigma_{r,i}^t = \frac{\sum_{m=1}^{n \times p} [b^m - \mu_{r,i}^t][b^m - \mu_{r,i}^t]^T P(\mu_{r,i}^t, \Sigma_{r,i}^t | b^m)}{\sum_{m=1}^{n \times p} P(\mu_{r,i}^t, \Sigma_{r,i}^t | b^m)}$$

Prior probabilities equivalent to the proportional presence of each class are updated using the following formula:

$$\pi_{r,i}^t = \frac{\sum_{m=1}^{n \times p} P(\mu_{r,i}^t, \Sigma_{r,i}^t | b^m)}{n \times p}$$

Finally, if the system is physically transported to another place where it is presented consistently either with all standard or all degraded signs, these two groups gradually merge into one as depicted in Fig. 4(d), mathematically:

$$\|\mu_{r,1}^t - \mu_{r,2}^t\| \leq th_2$$

Realizing that the distribution of pixels corresponding to the outer rim can consist of one or two groups, we define a piecewise likelihood function $p(b|c_r)$ given as:

$$p(b|c_r) = \begin{cases} \max[p(b|c_{r,1}), p(b|c_{r,2})], & \text{if } \mu_{r,1}^t \text{ \& } \mu_{r,2}^t \neq NULL \\ p(b|c_{r,i}), & \text{if } \mu_{r,i}^t \neq NULL \text{ \& } \mu_{r,j}^t = NULL \quad \forall i \neq j \end{cases} \quad (2)$$

The first case for $p(b|c_r)$ corresponds to the situation depicted in Fig. 4(c), where there are two groups with different parameters. In this case, we pick the greater value of the likelihood generated from the two distributions. The second case, however, represents the situation given in Fig. 4(a) in which there is only one group (hence one distribution) and the parameters of the other distribution are set to null.

Finally the likelihood $p(b|c_r)$ is plugged in Equation 3 to find the posterior probability of the data $p(c_r|b)$.

$$P(c_r|b) = \frac{P(b|c_r)P(c_r)}{P(b)} \quad (3)$$

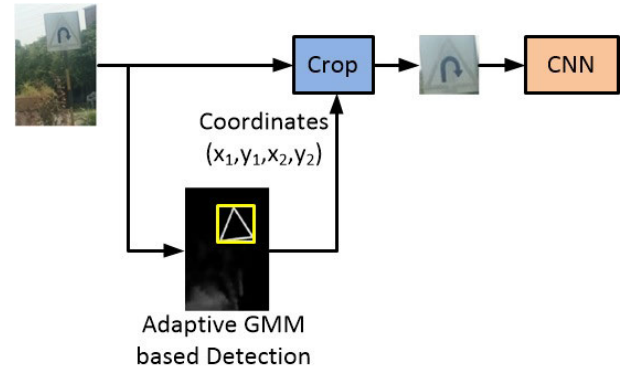


FIGURE 5. Cropping detected traffic sign for recognition.

$P(c_r)$ is the prior term based on proportional representation of class c_r computed as

$$P(c_r) = \frac{\sum_{k=1}^{n \times p} I(b_k \in c_r)}{n \times p}$$

where $I(b_k \in c_r)$ is an indicator function that produces a one if a row in B matrix corresponds to class c_r . However, total number of rows in matrix B are $n \times p$.

3) SEGMENTATION

A rim color boosted image (L) is obtained with the help of the posterior computed using Equation 3. Every pixel in $l(x, y)$ in image L is obtained by scaling linearly the probability $P(c_r|b)$ so that the pixel values are between 0 and 255.

$$l(x, y) \propto P(c_r|b) \quad \text{s.t.} \quad 0 \leq l(x, y) \leq 255$$

L is a grayscale image with foreground pixels taking higher grayvalues than the background pixels. In order to threshold it to a black and white image, L is subjected to maximally stable extremal region (MSER) technique [2], [36]. Using this method, the input grayscale image is thresholded repeatedly at different threshold values and the regions found to be the most stable at majority of the threshold values are retained. The resulting blobs are cropped and used for further processing. The process can mathematically be written as,

$$BW = Th_{mser}[L]$$

$$I = crop_{dim}^{blob}[IRGB]$$

where BW is the black and white (Th stands for thresholding) image obtained using maximally stable extremal region (MSER) operation and I is the corresponding color portion cropped out of the original image $IRGB$. The proposed adaptive GMM based technique served to segment traffic sign with or without degraded rim. Once segmented, the color version of the extracted blob is used for the purpose of recognition. The complete process is given in self explanatory Figure 5.

B. MULTISCALE CONVOLUTIONAL NEURAL NETWORK

Features (preferably invariant) can be extracted from images either using fixed/hand crafted techniques such as histogram

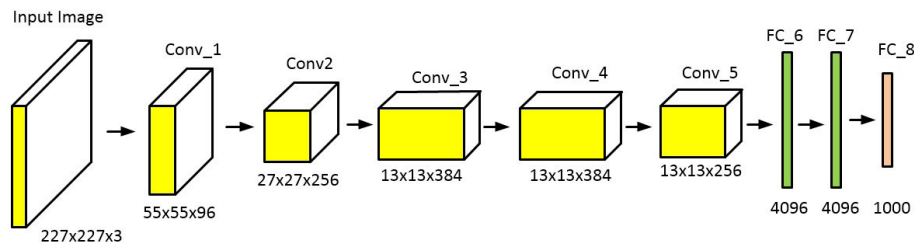


FIGURE 6. Architecture of Alexnet convolutional neural network [29].

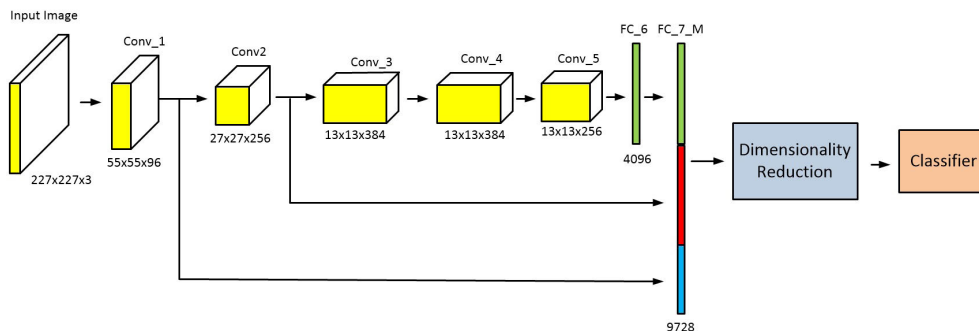


FIGURE 7. Block diagram of our proposed traffic sign recognition method employing a pretrained CNN for feature extraction followed by dimensionality reduction layer and a classifier trained on our training data.

of oriented gradients (HOG) [14] or can be learned during training like in artificial neural networks (ANN) [16], [37]. With the wide spread of computing capabilities, use of deep neural networks has become very popular [38]. These networks are termed as deep because of a number of hidden layers between input and output layers. This makes training computationally expensive and hence fast and parallel computing architectures like GPUs are common. Convolutional Neural Networks [29] are specific type of deep networks used for image processing and computer vision applications. A convolutional neural network contains a number of convolutional and pooling layers followed by fully connected layer(s). There is nothing predefined; coefficients of filter masks used in convolutional layers and weights in fully connected layers are adjusted with the help of training data. CNNs have shown excellent recognition performance for images but like any other deep network, they require a lot of training data to avoid over fitting and take very long to train due to their complex architecture. Fortunately, there are a number of CNNs already trained on millions of images are available for free. They contain a number of convolutional, max pooling, normalization, Rectified Linear Units (ReLU) and fully connected layers and are able to classify images in thousands of categories e.g., Alexnet [29], Googlenet [28] etc. To save training time, these pretrained networks can either be used as such or with the addition of extra layers. The objective, however, is to customize the network according to a given application.

In this work, we propose a convolutional neural network based traffic sign recognition scheme. A multiscale CNN is

augmented with a dimensionality reduction layer so that only highly relevant and least redundant features may be used for multi-class classification. In the following subsections, we will explain the procedure to learn multiscale features and the information theory based dimensionality reduction criterion.

1) THE PROPOSED CNN MODEL

We have used pretrained Alexnet convolutional neural network to extract multiscale features from our training set which contains both standard and degraded traffic sign images. Architecture of Alexnet is given in Fig. 6; it consists of five convolutional and three fully connected layers. Since the original network was trained on two GPUs in parallel, the final feature set is available at fully connected layer FC_7 i.e., the second last layer. The last layer, however, uses a softmax classifier [35]. We, in this work, employ a pretrained Alexnet for the purpose of multiscale feature extraction as shown in Fig. 7. In order to collect multiscale features, the network is tapped at three different convolutional layers and finally the results are concatenated. Outputs taken from different convolution layers are passed through one or more max pooling layers (purposely not shown in Fig. 7). Details of these layers including the information regarding layer of the original network tapped, max pooling parameters and the output size are given in Table 1. The output after Conv_1 layer is passed through three max pooling layers whose output size is $4 \times 4 \times 96$ which gives 1536 dimensional feature vector. Similarly Conv_2 and Conv_5 layers are processed to obtain 4096 dimensional descriptor each and hence the

TABLE 1. Composition of 9728 dimensional multiscale feature vector extracted using a pretrained CNN according to the scheme shown in Fig. 7.

Layer	Tapped	Max Pooling			Output Size	Features Extracted
	Input size	Size	Stride	Padding		
Conv_1	27 × 27 × 96	3 × 3	2	0	13 × 13 × 96	1536
	13 × 13 × 96	3 × 3	2	0	6 × 6 × 96	
	6 × 6 × 96	3 × 3	1	0	4 × 4 × 96	
Conv_2	13 × 13 × 256	3 × 3	2	0	6 × 6 × 256	4096
	6 × 6 × 256	3 × 3	1	0	4 × 4 × 256	
Conv_5	6 × 6 × 256	3 × 3	1	0	4 × 4 × 256	4096
Total Features						9728

total 9728 dimensional feature vector is obtained through concatenation. It must also be noted that the only extra burden put by our proposed feature extraction method is the use of three additional max pooling layers which is not significant.

Mathematically, feature descriptor d_i computed on a traffic sign image (I_i) using our proposed multiscale technique is:

$$d_i = T_{multiscale}(I_i) = cat\{T_{cnn}(I_i^{conv-1}), T_{cnn}(I_i^{conv-2}), T_{cnn}(I_i^{conv-5})\}$$

where T_{cnn} refers to the feature extraction operation performed using pruned pretrained Alexnet CNN. Finally, the data matrix D (shown in Fig. 8) is computed for n number of training images as:

$$D = [d_1 \quad d_2 \quad d_3 \quad \dots \quad d_n]^T$$

2) DIMENSIONALITY REDUCTION LAYER

In order to collect features at various spatial scales, our proposed feature descriptor (d) taps the pretrained Alexnet CNN at different places in the chain. This results in a highly representative descriptor but containing irrelevant and redundant features too. Therefore, before training a multiclass classifier on this data, we selected, using a dimensionality reduction scheme, a reduced set of highly relevant and least redundant features. Previous studies have shown that this reduced set helps in achieving better recognition accuracy with higher execution speed [39].

As shown in Fig. 8, different pre-trained Alexnet layers were used on training images (I_i) as features and finally a data matrix D is populated. In order to obtain a highly descriptive subset of features (S) we apply feature interaction based dimensionality reduction method on matrix D [40], [41].

$$S = T_{dim_red}[D] \tag{4}$$

Mutual information [42] is a popular measure of similarity computed on two random variables, feature interaction [42], however, involves three variables and is a bit more computationally expensive. The advantage, however, is a measure of simultaneous similarity among three variables. Following dimensionality reduction algorithm was used to compute matrix S from matrix D :

- 1) To start with, features are picked from matrix D one by one. Class relevance is computed between a feature

(f) and the class label (c) using mutual information i.e., $MI(f; c)$. Once the metric is computed for all features, they are ranked in descending order of their class relevance.

- 2) The feature f that scores the highest class relevance is put in the empty subset S as the first feature and is removed from the matrix D .
- 3) To add more highly relevant and least redundant features to the subset S , feature interaction is computed among every remaining feature in D , the feature(s) present in S and the class label c , i.e., $I(f_D, S; c)$. The feature that scores the highest is put in S and is removed from D . This process is repeated until we get the required number of features in matrix S .

This subset of the most important features (i.e., S) is used for the subsequent classification operation. However, computational complexity of the computed feature interaction operation $I(f_D, S; c)$ can be resolved by assuming features present in matrix S independent of each other i.e.,

$$f_S^i \perp f_S^j \quad \forall i \neq j \tag{5}$$

This assumption greatly simplifies the feature interaction operation i.e., the features are taken from set S one by one and the following operation is executed:

$$I(f_D, S; c) \approx \sum_{s_s \in S} I(f_D, s_s; c) \tag{6}$$

Since mutual information and feature interaction can be defined in terms of joint and/or marginal entropy terms, the simplified operation in Equation 6 can be expanded as follows:

$$I(f_D, s_s; c) = H(f_D, s_s) + H(c) - H(f_D, s_s, c) \tag{7}$$

One obvious advantage of writing the feature interaction operation as in Equation 7 is that the entropy of class label i.e. $H(c)$ remains the same for all features. Therefore, we can drop this term and can safely use the approximate formula for feature interaction given in Equation 8. Finally, we get the subset S filled with required number of features using this formula.

$$I(f^+, s_s; c) \approx H(f^+, s_s) - H(f^+, s_s, c) \tag{8}$$

Algorithm 1 gives details of the dimensionality reduction algorithm used to reduce the dimensions of the features

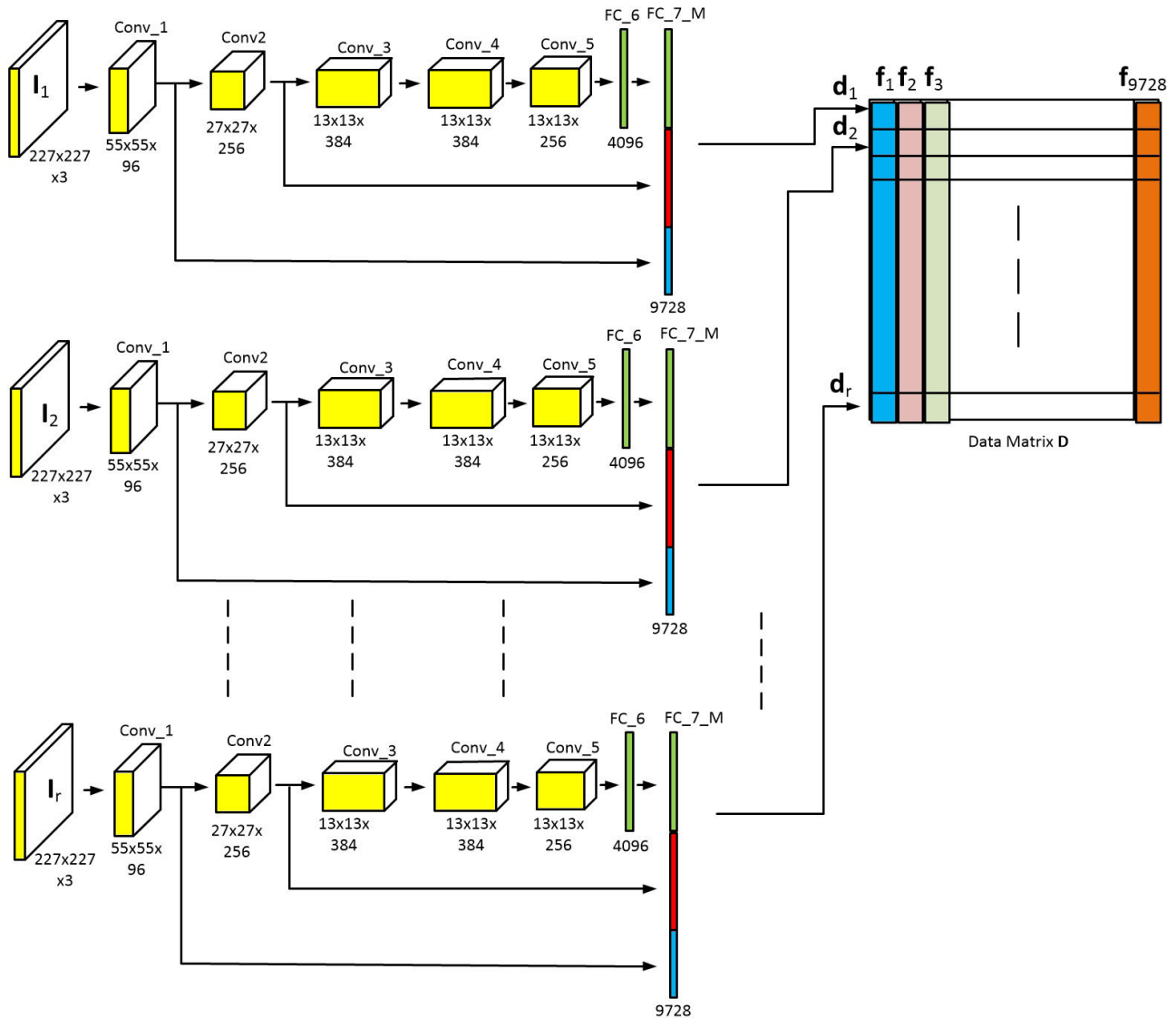


FIGURE 8. Proposed multiscale CNN based architecture for feature learning. The matrix D is then subject to dimensionality reduction so that the computational complexity may be reduced in subsequent layers.

extracted using pretrained CNN. The input is data matrix D , input class labels c and the number of desired features n_d . Features in data matrix D are sorted according to the decreasing order of their class relevance computed on the basis of mutual information. Then those n_d features are added to the subset S which produce the highest joint mutual information computed using Equation 6.

3) CLASSIFIER LAYER

As shown in Fig. 7, classification is the second layer we have added to the pre-trained Alexnet CNN for the purpose of recognizing degraded traffic signs. This layer is next to the dimensionality reduction layer which supplies highly descriptive subset (S) computed from the data matrix (D). We have used a multiclass support vector machine classifier with linear kernel. The output of the layer classifies red

bordered traffic signs in different classes belonging to either prohibitory or danger types.

IV. EXPERIMENTS AND RESULTS

All experiments were conducted on core i7 computer with 8 GB RAM. Implementation of pretrained Alexnet model available at Mathworks was used for feature extraction and LibSVM implementation of support vector machine classifier (originally written in C language [43]) was employed for multiclass training and testing.

A. DATASET

A dataset was prepared using samples from famous German traffic sign detection benchmark (GTSDB) [44], our self collected and synthetically prepared images. The dataset contains more than 28,684 images divided in 32 different classes

Algorithm 1 Joint Mutual Information Based Dimensionality Reduction

Input: $D = [d_1 \ d_2 \ d_3 \ \dots \ d_{9728}]$ ▷ Input Data matrix
Input: c ▷ Vector containing class labels
Input: n_d ▷ Desired number of features
Output: $S = [s_1 \ s_2 \ s_3 \ \dots \ s_{n_d}]$ ▷ Subset of selected features

$E = [e_1 \ e_2 \ e_3 \ \dots \ e_{9728}] \leftarrow \text{sort}_{\text{rank}}(D)$
 $s_1 \leftarrow e_1$
for $i \leftarrow 1 : n_d$ **do** ▷ Select desired number of features
 for $j \leftarrow 2 : 9728$ **do**
 $\text{arr}_j \leftarrow MI(e_j, S; c)$ ▷ Equation 6
 end for
 $k \leftarrow \text{ind}_{\text{max}}(\text{arr})$
 $S \leftarrow [S \ e_k]$ ▷ Add winning feature to S
end for



FIGURE 9. Samples of standard traffic sign whose degraded versions were used for experiments.

plus one negative class (33 in total). Since the set of actually collected images is not balanced (participation of each class of signs is not the same), we added different permutations on traffic sign proposals like rotation, scaling, skewing etc. to make the resultant dataset balanced.

Fig. 9 shows screen shots of some standard prohibitory and warning traffic signs taken either from the GTSDDB or our self collected dataset. In this work, we have used degraded versions of these samples for experimentations. For degraded traffic signs, it is mentioned in Table 2 that either the traffic signs were found degraded at the time of data collection or samples from GTSDDB and field surveys were synthetically degraded.

B. STATISTICAL DYNAMIC PARAMETER ESTIMATION

For the only dataset available for experimentation, we used 60% data for training and reserved remaining 40% for testing purposes. As mentioned in Section III, the right hand side branch of Fig. 2 shows the test procedure. In this branch, the block named ‘Statistical Test’ performs two sample t-test between RGB data used for training (A) and the data of recently detected traffic signs (A_{aug}). This arrangement facilitates adaptation between standard and degraded traffic

signs by employing flexible Gaussian mixture model technique discussed earlier. Set A_{aug} is of size equal to only 10% of the data matrix A and its contents are updated with every new detection.

In order to know whether the contents of data matrix A are to be updated with A_{aug} , we need to perform a statistical test with the following null and alternate hypotheses at 95% confidence interval [45]:

$$H_0 : \mu_A = \mu_{A_{aug}}$$

$$H_1 : \mu_A \neq \mu_{A_{aug}}$$

Consequently, the new RGB data matrix (A_{aug}) is augmented to the originally available data matrix (A) only if $\mu_A = \mu_{A_{aug}}$, statistically. The t-statistic computed from the two data matrices A and A_{aug} is computed using the following Hotelling’s t -squared formula [46]:

$$t^2 = \frac{n_A n_{A_{aug}}}{n_A + n_{A_{aug}}} (\mu_A - \mu_{A_{aug}})^T \Sigma_p^{-1} (\mu_A - \mu_{A_{aug}})$$

Number of samples given in matrices A and A_{aug} are given by n_A and $n_{A_{aug}}$ respectively and Σ_p is the estimated pooled covariance:

$$\Sigma_p = (n_A - 1)\Sigma_A + (n_{A_{aug}} - 1)\Sigma_{A_{aug}}$$

where Σ_A and $\Sigma_{A_{aug}}$ are covariance matrices of A and A_{aug} matrices.

If the two mean values are found statistically the same at time $t = t_0$, the new data matrix A to be used at a later point in time called $t = t_1$ is given by:

$$A^{t=t_1} = [A^{t=t_0} \ A_{aug}^T]^T$$

The proposed system has the ability to update itself continuously based on the number of traffic signs detected with a certain degree of confidence. This property makes our proposed technique adaptive and it is able to cope into a new situation. As mentioned in Section III, if the covariance of data matrix ($A^{t=t_1}$) exceeds a certain threshold, parameter estimation module splits the group into two (see Fig. 4) and the new parameters are estimated using the proposed flexible mixture model.

As the system is tested and traffic signs are detected, the size of RGB data matrix A keeps on increasing which may slow down parameter estimation in future. To cater for this problem and to ensure that the system is able to train itself continuously, we only use latest Z entries (where Z is equal to the size of matrix A at time $t = 0$) and neglect any older data.

C. EVALUATION MATRICES

For experimental results on two datasets using our proposed techniques, following evaluation metrics were used:

$$\text{Precision} = \frac{TP}{TP + FP}$$

$$\text{Recall} = \frac{TP}{TP + FN}$$

TABLE 2. Description of the dataset used for experimentation.

Instances	Classes	Size Train(Test)	Composition
28,684	33	500×450×3 (227×227×3)	Standard samples taken from GTSDDB and field surveys Synthetically degraded GTSDDB samples Degraded samples from field survey Synthetically degraded samples from field survey

$$Accuracy = \frac{TP + TN}{TP + FP + TN + FN}$$

$$Error_rate = \frac{FP}{FP + TN}$$

where TP: true positive, FP: false positive, TN: true negative and FN: false negative.

D. RESULTS AND DISCUSSIONS

In this section, we present results obtained for detection and recognition of degraded traffic signs and compare those with other state of the art and recent methods. Since degraded signs have been dealt with only by a few already published research works, the performance of the proposed method is also compared with the techniques meant for standard signs.

1) DETECTION

As mentioned in Section III, the proposed dynamic detection method continuously updates itself i.e., it is trained in the beginning once and after some predictions, a small amount of recent data after passing a statistical test is added to the original data matrix A for further parameter estimation. This allows to adapt to new situations at run time by adjusting number of groups for outer rim color and their parameters. This unique feature enables the proposed algorithm to perform well in a combination of standard and degraded traffic signs. Figure 10 shows a number of signs containing degradations at certain place on the outer rim (where color is flushed out). These signs were correctly detected using our proposed flexible Gaussian mixture model based segmentation technique.

Table 3 shows the performance of various detection techniques in comparison with the proposed method on our dataset containing both standard and degraded traffic sign images. As obvious, color thresholding based methods [3], [4] fail because the thresholds are set tight enough to suppress background objects. This makes them unsuitable for signs with large variation in colors due to aging etc. Optimization based method proposed in [13] uses color edge data to localize traffic sign in a real world scene and use Markov random field to obtain a tight box around a detected sign. Color probability maps (CPM) [6] based method uses statistical parameter estimation using Ohta transform [32]. The technique performs poorly for degraded signs because the variance of the Gaussian distribution increases causing increased number of false alarms. Pretrained versions of Googlenet [29] and Alexnet [47] are also not able to detect many degraded traffic signs especially in cluttered background. Our proposed



FIGURE 10. Samples containing standard and with degradations present on the outer rim of the traffic sign detected with our proposed flexible mixture model based segmentation technique.

method achieves the highest recognition accuracy for both prohibition (circular) and danger (triangular) traffic signs with degradations.

E. MULTISCALE CONVOLUTIONAL NEURAL NETWORK

As mentioned earlier in the text, recognition refers to reading internal contents of a traffic sign. Interior of a standard traffic sign is composed of either digits (e.g. speed limits) or graphical objects (e.g. school ahead, left turn, round about etc.). For

TABLE 3. Detection accuracy of various methods to detect degraded traffic signs.

	Prohibition	Danger
RGB Thresholding [3]	0.51	0.43
HSI Thresholding [4]	0.55	0.45
Optimization [13]	0.71	0.65
CPM [6]	0.69	0.74
Googlenet [29]	0.65	0.69
Alexnet [48]	0.70	0.68
Our Proposed Method	0.83	0.85

degraded signs with weird contents or with poor contrast between foreground and background of the interior of the sign, a specialized feature extraction technique was devised. Use of a multiscale pretrained Alexnet neural network for feature extraction followed by a powerful dimensionality reduction layer produces a reduced set of features demonstrating great intraclass similarity and interclass separation.

1) HAND CRAFTED VS AUTOMATIC FEATURES

In this section we present a comparison between hand crafted features like HOG, LBP etc. with the features we learn from a pretrained Alexnet CNN. Near optimal performance on GTSDDB dataset has been achieved using convolutional neural networks [27], [44] but the degraded signs present in cluttered background are far more challenging. Table 4 gives precision and recall values of different hand crafted and automatic feature extraction methods in comparison with our proposed method. HOG [2], LBP [1], their combination [33], Eigen [3] and Energy [48] based methods correspond to hand crafted class and have been reported to extract invariant features for standard traffic signs in recent literature but our results show that they are not successful for degraded signs. Eigen based method, however, is the best among all hand crafted feature extraction techniques. In a recent work, Yang *et al.* [6] trained a convolutional neural network from scratch whose architecture is very close to the Alexnet CNN. Both precision and recall values obtained using their techniques are close to our proposed method. However, since our proposed method uses a pretrained Alexnet for the purpose of feature extraction (followed by dimensionality reduction and classification layers), the time for training is saved and the chance of over fitting is reduced because the initial layers of the pretrained CNN used in this work are already trained on millions of images. Finally, the proposed method was compared to very recent work where hybrid features are computed using a combination of steered discrete cosine transform and circular histogram of oriented gradients [33]. Both precision and recall of the proposed system are very close to the hybrid method.

TABLE 4. Comparison of "Precision" and "Recall" among different method.

	HOG [2]	LBP [50]	HOG+LBP	Eigen [3]	Energy [49]	CNN [6]	Hybrid [33]	Proposed
Precision	0.61	0.60	0.65	0.67	0.59	0.72	0.78	0.81
Recall	0.63	0.59	0.52	0.49	0.78	0.69	0.74	0.79

2) EFFECT OF DIMENSIONALITY REDUCTION

Dimensionality reduction is supposed to decrease the size of shape descriptor and increase the recognition accuracy by removing irrelevant and redundant features [33], [39]. We have tested feature interaction based dimensionality reduction technique on matrix of shape descriptors i.e. D and obtain a subset S . Table 5 shows how error rate varies as the number of features are increases using our proposed dimensionality reduction techniques. The error rate is high for lesser number of features, though these are the most important. As the number of features increase, the error rate decreases and saturates to its approximate least value. This table also shows that the execution speed decreases as we use more and more number of features. Table 5 shows that with only approximately 5,000 features, we are able to obtain the optimum recognition error and hence can save up to 40% computational cost. The rule of dimensionality reduction mentioned in Section III, however, is extracted only once and its computational cost during testing is not mentioned here.

To investigate the effect of dimensionality reduction on overall recognition accuracy and execution speed, we compared our proposed method with a number of state of the art and recent methods. Results are shown in Figure 11 that compares four different methods with our proposed scheme. As the difference among all techniques is not significantly different with very small number of features, it can be seen that with about fifty percent of the features, our proposed method outperforms all other feature extraction schemes. In order to judge any benefit obtained for execution speed, the same methods were compared again versus increasing number of features as shown in Figure 12. For small number of features (around 1,000) our proposed method doesn't perform well but with almost fifty percent of the features, it obtained slightly higher execution speed than HOG based method but the advantage in recognition accuracy is significantly higher. It should also be noted that grid sizes were adjusted for all competing methods so that equal length of shape descriptor be obtained to fulfill the requirement of a fair competition [33].

3) OVERALL PERFORMANCE

In order to compare various methods for both detection and recognition tasks of degraded traffic signs, we selected a number of state of the art and recent methods tabulated in Table 6. Signs with degradations on the outer rim and on the interior were manually separated and the performance of different methods was compared. The method reported in [2] uses color thresholding by employing maximally stable extremal region for detection and histogram of oriented gradients for

TABLE 5. Effect of dimensionality reduction on error rate and execution speed using our proposed method.

No. of Features	Error Rate	Execution Time (ms)
1000	0.71	90
2000	0.7	95
3000	0.63	100
4000	0.45	123
5000	0.21	130
6000	0.22	155
7000	0.21	169
8000	0.24	170
9000	0.30	187
10000	0.25	191

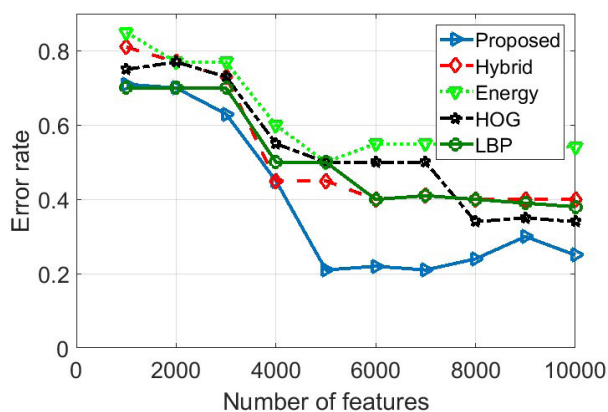


FIGURE 11. Comparison of error rate among different feature extraction methods subject to dimensionality reduction.

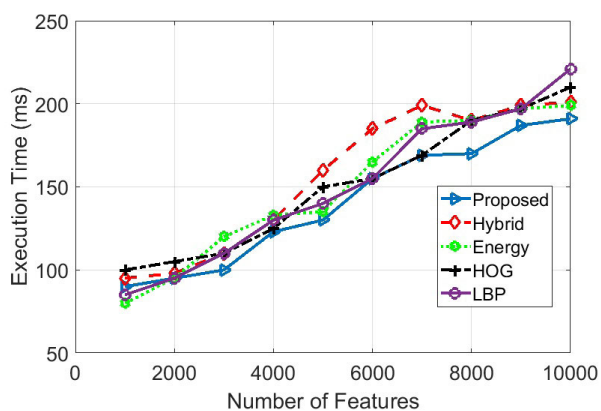


FIGURE 12. Comparison of execution speed among different feature extraction methods subject to dimensionality reduction.

recognition. It achieves higher error rate for degradations present on both rim and on the interior of the signs. Fleyeh and Davami [3] presented color thresholding and principal component analysis but the approach is not effective for degraded examples. The method in [6] uses statistical parameter estimation using Ohta transform for detection and convolutional neural network for recognition. A recently proposed random forest based technique [4] using HSI thresholding and HOG features performed poorly for degradations present on the

TABLE 6. Comparing error rate of various methods on degradations present at different places on a traffic sign.

	HOG [2]	PCA [3]	CPM [6]	RM [4]	MT [36]	Proposed
Rim	0.35	0.49	0.34	0.6	0.3	0.16
Interior	0.4	0.30	0.27	0.3	0.4	0.22



FIGURE 13. A number of correctly detected and recognized degraded traffic signs using our proposed technique.



FIGURE 14. A number of correctly detected but incorrectly recognized degraded traffic signs using our proposed technique.

outer rim. The last method [36] uses multiple thresholding technique similar to maximally stable extremal region for detection and log polar transforms for recognition. Our proposed method obtained the best detection rate for degradations present on rims. However, the error rate is very close to the neural network based methods for degradations found on the interior of the sign.

Finally, keeping in view the natural and synthetically added occlusions, Fig. 13 shows a number of correctly detected and recognized degraded traffic signs whereas Fig. 14 shows the signs correctly detected but could not be recognized correctly. It is evident that the proposed method can reliably detect traffic signs under degradations i.e., either with lost rim color or weird contents.

V. CONCLUSION

Detection and recognition of degraded traffic signs in real world cluttered backgrounds is a challenging task. Very few efforts are found in literature for aged/degraded signs whereas both detection and recognition of standard traffic signs has been an active field of research. We restrict our work to red bordered prohibitory and warning signs only with degradations present on the outer rim or on the interior. We, in this work, present a dynamically updating split and merge scheme to detect traffic signs catering for lost rim color. Group of pixels corresponding to the rim color is split to two if the variance exceeds a certain threshold. The two groups, however, merge back to one if their overlap achieves another threshold.

This property makes the system adaptable to perform in varying outdoor situations. For recognition, we make use of a pretrained convolutional neural network to obtain invariance for degradations present on the interior of a sign. In order to capture features at varying spatial scales, the pretrained network is tapped at different places and a long feature vector is formed. In order to make efficient and correct predictions, a joint mutual information based dimensionality reduction layer was added before classification layer. Using this technique, we are able to avail the functionality of convolutional neural networks without training a network on a lot of images from scratch.

The results have shown that using our dynamically updating detection scheme, we are able to detect degraded traffic signs in difficult background conditions compared to many state of the art and recent techniques including end to end deep learning methods. This is because our proposed system takes care of possible degradations in advance which make the CNN based detections more accurate. A comparison of hand crafted features versus our proposed feature extraction technique demonstrates the superiority of our proposed approach to recognize contents of a degraded traffic sign. Finally, experiments validate that the dimensionality reduction method proposed in this work helps achieving better recognition accuracy with reduced computational complexity.

REFERENCES

- [1] A. Gudigar, C. Shreesha, and U. Raghavendra, "A review on automatic detection and recognition of traffic sign," *Multimedia Tools Appl.*, vol. 75, no. 1, pp. 333–364, 2016.
- [2] J. Greenhalgh and M. Mirmehdi, "Real-time detection and recognition of road traffic signs," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 4, pp. 1498–1506, Dec. 2012.
- [3] H. Fleyeh and E. Davami, "Eigen-based traffic sign recognition," *IET Intell. Transp. Syst.*, vol. 5, no. 3, pp. 190–196, Sep. 2011.
- [4] A. Ellahyani, M. El Ansari, and I. El Jaafari, "Traffic sign detection and recognition based on random forests," *Appl. Soft Comput.*, vol. 46, pp. 805–815, Sep. 2016.
- [5] Y. Saadna, A. Behloul, and S. Mezzoudj, "Speed limit sign detection and recognition system using SVM and MNIST datasets," *Neural Comput. Appl.*, vol. 31, no. 9, pp. 5005–5015, 2019.
- [6] Y. Yang, H. Luo, H. Xu, and F. Wu, "Towards real-time traffic sign detection and classification," *IEEE Trans. Intell. Transp. Syst.*, vol. 17, no. 7, pp. 2022–2031, Jul. 2016.
- [7] J. B. Kim, "Detection of traffic signs based on eigen-color model and saliency model in driver assistance systems," *Int. J. Automot. Technol.*, vol. 14, no. 3, pp. 429–439, 2013.
- [8] L. W. Tsai, J. W. Hsieh, C. H. Chuang, Y. J. Tseng, K. C. Fan, and C. C. Lee, "Road sign detection using eigen colour," *IET Comput. Vis.*, vol. 2, no. 3, pp. 164–177, Sep. 2008.
- [9] S. Salti, A. Petrelli, F. Tombari, N. Fioraio, and L. D. Stefano, "Traffic sign detection via interest region extraction," *Pattern Recognit.*, vol. 48, no. 4, pp. 1039–1049, Apr. 2015.
- [10] R. Timofte, K. Zimmermann, and L. Van Gool, "Multi-view traffic sign detection, recognition, and 3D localisation," *Mach. Vis. Appl.*, vol. 25, no. 3, pp. 633–647, Apr. 2014.
- [11] H. S. Lee and K. Kim, "Simultaneous traffic sign detection and boundary estimation using convolutional neural network," *IEEE Trans. Intell. Transp. Syst.*, vol. 19, no. 5, pp. 1652–1663, May 2018.
- [12] V. Balali, A. Jahangiri, and S. G. Machiani, "Multi-class US traffic signs 3D recognition and localization via image-based point cloud model using color candidate extraction and texture-based recognition," *Adv. Eng. Inform.*, vol. 32, pp. 263–274, Apr. 2017.
- [13] Z. Zhu, J. Lu, R. R. Martin, and S. Hu, "An optimization approach for localization refinement of candidate traffic signs," *IEEE Trans. Intell. Transp. Syst.*, vol. 18, no. 11, pp. 3006–3016, Nov. 2017.
- [14] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2005, vol. 1, no. 1, pp. 886–893.
- [15] T. Ojala, M. Pietikäinen, and T. Mäenpää, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Jul. 2002.
- [16] 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, Nov. 2016.
- [17] X. Bangquan and W. X. Xiong, "Real-time embedded traffic sign recognition using efficient convolutional neural network," *IEEE Access*, vol. 7, pp. 53330–53346, 2019.
- [18] Z. Liu, J. Du, F. Tian, and J. Wen, "MR-CNN: A multi-scale region-based convolutional neural network for small traffic sign recognition," *IEEE Access*, vol. 7, pp. 57120–57128, 2019.
- [19] A. Mogelmose, M. M. Trivedi, and T. B. Moeslund, "Vision-based traffic sign detection and analysis for intelligent driver assistance systems: Perspectives and survey," *IEEE Trans. Intell. Transp. Syst.*, vol. 13, no. 4, pp. 1484–1497, Dec. 2012.
- [20] H. Ishida, T. Takahashi, I. Ide, Y. Mekada, and H. Murase, "Identification of degraded traffic sign symbols by a generative learning method," in *Proc. 18th Int. Conf. Pattern Recognit. (ICPR)*, vol. 1, Aug. 2006, pp. 531–534.
- [21] L. Li and G. Ma, "Recognition of degraded traffic sign symbols using PNN and combined blur and affine invariants," in *Proc. 4th Int. Conf. Natural Comput. (ICNC)*, vol. 3, Oct. 2008, pp. 515–520.
- [22] G. Floros, K. Kyritsis, and G. Potamianos, "Database and baseline system for detecting degraded traffic signs in urban environments," in *Proc. 5th Eur. Workshop Vis. Inf. Process. (EUVIP)*, Dec. 2014, pp. 1–5.
- [23] L. Li, G. Ma, and S. Ding, "Identification of degraded traffic sign symbols using multi-class support vector machines," in *Proc. Int. Conf. Mechatronics Automat. (ICMA)*, Aug. 2007, pp. 2467–2471.
- [24] Y. Rehman, I. Riaz, X. Fan, and H. Shin, "D-patches: Effective traffic sign detection with occlusion handling," *IET Comput. Vis.*, vol. 11, no. 5, pp. 368–377, Aug. 2017.
- [25] C. Liu, F. Chang, and C. Liu, "Occlusion-robust traffic sign detection via cascaded colour cubic feature," *IET Intell. Transp. Syst.*, vol. 10, no. 5, pp. 354–360, Jun. 2015.
- [26] Y.-L. Hou, X. Hao, and H. Chen, "A cognitively motivated method for classification of occluded traffic signs," *IEEE Trans. Syst., Man, Cybern., Syst.*, vol. 47, no. 2, pp. 255–262, Feb. 2017.
- [27] Z. Liang, J. Shao, D. Zhang, and L. Gao, "Traffic sign detection and recognition based on pyramidal convolutional networks," *Neural Comput. Appl.*, pp. 1–11, Mar. 2019.
- [28] Z. Zhong, L. Jin, and Z. Xie, "High performance offline handwritten Chinese character recognition using GoogLeNet and directional feature maps," in *Proc. 13th Int. Conf. Document Anal. Recognit. (ICDAR)*, Aug. 2015, pp. 846–850.
- [29] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.
- [30] N. Barnes, A. Zelinsky, and L. S. Fletcher, "Real-time speed sign detection using the radial symmetry detector," *IEEE Trans. Intell. Transp. Syst.*, vol. 9, no. 2, pp. 322–332, Jul. 2008.
- [31] S. Maldonado-Bascon, S. Lafuente-Arroyo, P. Gil-Jimenez, H. Gomez-Moreno, and F. López-Ferreras, "Road-sign detection and recognition based on support vector machines," *IEEE Trans. Intell. Transp. Syst.*, vol. 8, no. 2, pp. 264–278, Jun. 2007.
- [32] Y.-I. Ohta, T. Kanade, and T. Sakai, "Color information for region segmentation," *Comput. Graph. Image Process.*, vol. 13, no. 3, pp. 222–241, 1980.
- [33] A. Mannan, K. Javed, A. Rehman, S. K. Noon, and H. A. Babri, "Optimized segmentation and multiscale emphasized feature extraction for traffic sign detection and recognition," *J. Intell. Fuzzy Syst.*, vol. 36, no. 1, pp. 173–188, 2019.
- [34] E. Alpaydin, *Introduction to Machine Learning*. Cambridge, MA, USA: MIT Press, 2014.
- [35] C. Bishop, *Pattern Recognition and Machine Learning*. New York, NY, USA: Springer, 2006.
- [36] A. Gudigar, S. Chokkadi, U. Raghavendra, and U. R. Acharya, "Multiple thresholding and subspace based approach for detection and recognition of traffic sign," *Multimedia Tools Appl.*, vol. 76, no. 5, pp. 6973–6991, 2017.

[37] S. Haykin, *Neural Networks: A Comprehensive Foundation*. Upper Saddle River, NJ, USA: Prentice-Hall, 1994.

[38] J. Schmidhuber, "Deep learning in neural networks: An overview," *Neural Netw.*, vol. 61, pp. 85–117, Jan. 2015.

[39] I. Guyon and A. Elisseeff, "An introduction to variable and feature selection," *J. Mach. Learn. Res.*, vol. 3, pp. 1157–1182, Jan. 2003.

[40] M. Bannasar, Y. Hicks, and R. Setchi, "Feature selection using joint mutual information maximisation," *Expert Syst. Appl.*, vol. 42, no. 22, pp. 8520–8532, Sep. 2015.

[41] G. Brown, A. Pocock, M.-J. Zhao, and M. Luján, "Conditional likelihood maximisation: A unifying framework for information theoretic feature selection," *J. Mach. Learn. Res.*, vol. 13, pp. 27–66, Jan. 2012.

[42] L. Welch, "Lower bounds on the maximum cross correlation of signals (corresp.)," *IEEE Trans. Inf. Theory*, vol. IT-23, no. 3, pp. 397–399, May 1974.

[43] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011.

[44] 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 *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Aug. 2013, pp. 1–8.

[45] D. C. Montgomery, G. C. Runger, and N. F. Hubele, *Engineering Statistics*. Hoboken, NJ, USA: Wiley, 2009.

[46] K. V. Mardia, "Assessment of multinormality and the robustness of hotelling's T^2 Test," *Appl. Statist.*, vol. 24, no. 2, pp. 163–171, 1975.

[47] S.-G. Lee, Y. Sung, Y.-G. Kim, and E.-Y. Cha, "Variations of AlexNet and GoogLeNet to improve Korean character recognition performance," *J. Inf. Process. Syst.*, vol. 14, no. 1, pp. 205–217, 2018.

[48] A. de la Escalera, J. M. Armingol, J. M. Pastor, and F. J. Rodríguez, "Visual sign information extraction and identification by deformable models for intelligent vehicles," *IEEE Trans. Intell. Transp. Syst.*, vol. 5, no. 2, pp. 57–68, Jun. 2004.

[49] X. Yuan, X. Hao, H. Chen, and X. Wei, "Robust traffic sign recognition based on color global and local oriented edge magnitude patterns," *IEEE Trans. Intell. Transp. Syst.*, vol. 15, no. 4, pp. 1466–1477, Apr. 2014.



KASHIF JAVED received the B.S., M.S., and Ph.D. degrees in electrical engineering from UET, Lahore, Pakistan, where he currently serves as an Associate Professor. He has published a number of articles in reputed international journals hosted by the IEEE, PLoS, Springer, and Elsevier. He is currently acting as a Reviewer for many journals published by the IEEE, Elsevier, and Springer. His research interests include machine learning and intelligent data analysis.



ATTA UR REHMAN received the B.S. degree in electrical engineering from Air university, Islamabad, Pakistan, and the Ph.D. degree in electrical engineering from Loughborough University, U.K. He is currently serving as the Head of electrical engineering with UET, Gujranwala Campus. His research interests include image processing, machine learning, and pattern recognition.



HAROON A. BABRI received the B.S. degree from UET, Lahore, and the M.S. and Ph.D. degrees from Penn State University, USA. He is currently a Professor of electrical and computer engineering with UET. He has worked at numerous universities around the globe and has published more than 80 research articles in reputed journals and conferences.



ABDUL MANNAN received the B.S. and M.S. degrees in electrical engineering from the University of Engineering and Technology, Lahore, Pakistan, where he is currently pursuing the Ph.D. degree in electrical and computer engineering. He is currently a Researcher with the NFC Institute of Engineering and Technology, Multan, Pakistan. His current research interests include image processing for intelligent transport systems, machine learning, and statistical modeling.



SEROSH K. NOON received the B.S. degree from NFC IET, Multan, Pakistan, and the M.S. degree from UET, Lahore. She is currently with NFC IET. Her research interests include image processing in agriculture, bioinformatics, and pattern recognition.

...