Research Article

Identification of Fake VS Original logos using Deep Learning

Ranjith K C[1], Sharath Kumar Y H[2]

[1]Assistant Professor, Department of Computer Science & Engineering,Mandya, (Karnataka), INDIA
[2]Professor, Department of Information Science & Engineering,Mandya, (Karnataka), INDIA

Article History: Received: 11 January 2021; Revised: 12 February 2021; Accepted: 27 March 2021; Published online: 23 May 2021

ABSTRACT: Logo or trademark conveys the significance of an organization or a product. In view of determining of the authenticity of a logo, a precise extraction of features is crucially important for any automated recognition system to succeed. Here in this proposed work, identification of fake logos are done by fusion based approach of features from SIFT and SURF which considerably does better compared to the existing methods.

Keywords: CNN, Fake Logo, PCA, SIFT, SURF

Abbreviations: SIFT, scale invariant feature transform; SURF, Speeded up robust features; PCA, principle component analysis.

I. INTRODUCTION

Logos in some cases also called as a trademark have high significance in the present advertising world. Logo or trademark conveys this significance since it conveys the identity of the organization or a product. Logo coordinates and elements are essential to determine whether the logo is original or duplicate. Testing image set consists of images with various aspects like scale, pivot, relative twisting, enlightenment clamor, exceptionally blocked commotion. Filter descriptor, surf descriptor are excellent features to use among the current strategies to perceive the logo pictures from all such troubles more precisely.

Precise extraction of features is crucially important for any automated recognition system to succeed. Now with technological advancement, there are many improvements in image processing, pattern recognition and object detection.

Logo detection and classification has been extensively researched in the literature, but a very few researchers have implemented or addressed this fake logo detection.

The proposed system helps in recognition of the Genuine and counterfeited logos of Brands namely Renault, Volkswagon, Rolls Royce, Ferrari and Honda. We have created our own logo dataset by using Image processing techniques such that any image, low or high resolution will be further scaled and segmented as per the pixel data. Once done, we will use Convolutional layer with vectored images for further training. We will then prepare a test dataset to measure the accuracy with split, train and test module. Once complete, we will have a web application from where we will upload the test photos to understand, if the brand and uniqueness of the logo is identified and correctly classify for any fake logo.

In this work, Features of the logo images are extracted by using the SIFT and SURF techniques and then the features are fused to obtain the significant features to detect the query image. From the literature, we are able to understand that the authors are proposed works are limited to the dataset, Feature extraction fusion and classification. In this work we have used our own dataset and effective feature extraction techniques likes SIFT and Surf with CNN as classifier.

II. LITERATURE SURVEY

Zhang Nan et.al [1], has utilized the element extraction strategy PCA and the Kernel PCA. They use back propagation feed forwarding neural network classifier, KNN and SVM, for comparative study and analysis. The overall results indicated accuracy up to 97.5%.

Bailing Zhang et.al [2], has found that the framework for the most part comprises of two stages, logo detection and the logo classification. The first stage in the process is helped by the two pre-logo location modules, which are vehicle locale recognition and the little Region of Interest (ROI) division. Return on initial capital investment that spreads logos divided relative with labels, which can be absolutely restricted from frontal vehicle pictures. A two-organize course classifier continues with the divided ROI, utilizing a Support Vector Machine (SVM) with Adaboost mix, bringing about exact logo situating. Results show a promising rush of 95% exactness.

Chun Pan et.al [3], he stated that the classifier uses subsampling procedure for identification of logo and then the same is given as an input to CNN for further classification. But they separate both approaches, where they use CNN alone and then SIFT with SVM as comparative study. Normal identification precision pace of the methodology as for CNN is 8.61% more noteworthy than consequences of the methodology dependent on SIFT. So, they have preferred CNN as it gave them the accuracy of 99.23%.

N Vinay Kumar et.al [4], proposed the classification model which makes use of the global features of logo images for the classification. Texture of the image, color of the image, and the shape of the logo derives the global features of the logo. Here they have limited very limited number of the features for the fusion process which has led their model efficiency moderate level.

Changbo Hu et.al [5], has designed a model to handle the brand identification using the multimodal fusion procedure which integrates the image based on logo identification using the convolutional neural networks and context feature. Shreyansh Gandhi et.al [6], has presented the computer vision based offensive and the non-compliant image identification system for data sets of large size.

Linghua Zhou et.al [7], has derived a novel procedure to detect the vehicles logo when it is with motion blur and the combinations of the Filter-DeblurGAN and the VL-YOLO. Here they have made use of the traditional way of detecting the logo instead of any advanced methods which could have optimized the accuracy of the logo detected.

Qin Guet et.al [8], has presented a work on logo detection which uses the SIFT method for feature extraction as discussed in any pattern recognition forums. Here, identification before-area structure is proposed for the vehicle logo identification, where SIFT feature extractor is utilized to separate thick SIFT descriptors of the multi scale standard vehicle logo. Afterward, vehicle logo identification before confinement is set up dependent on thick SIFT coordinating vitality and the SIFT flow consistency.

Ruilong Chen et.al [9], has presented a work on Vehicle logo identification in transportation frameworks. Best in class vehicle logo identification approaches utilize naturally took in highlights from the Convolutional Neural Networks (CNNs). In any case, CNN's don't work well when the pictures are turned in exceptionally boisterous. This paper proposes a picture identification structure with a case organize. This is the situation of social affair of the neurons, whose measure can address the nearness probability of a component or part of a substance. Y H Sharath Kumar et.al [10], has presented a work on logo detection using central moments and PCA. was employed to reduce the extracted features. He has proposed a model for the purpose identify logos from the input document. Here they have eliminated the text in the beginning and later the logos are segmented from the document. Here the results were been matched with the reference to five human experts tracing of the logo from the given document.

III. PROPOSED METHOD

This part describes sufficient detail so that all procedures can be repeated. These can be divided into subsections as several methods are described.

A. SIFT (Scale Invariant Feature Transform)

There will be numerous highlights in any article, interesting spotlights on items which can be removed to give a "include" portrayal of an article. These depictions would then be able to be utilized during endeavoring to discover the thing in an image which contains various articles. Observation is more challenging when bifurcating such highlights and also the way it is recorded. Filter picture highlight gives a lot of highlights of the item which are not subjective by numerous individuals of inconveniences experienced at different strategies, for an example, object scaling and revolution.

However considering the article to be observed in the master plan SIFT technique incorporates moreover mulling over inquiries on various photos of the comparative zone, observed from the various circumstances inside the earth, to be seen. Channel features are in like manner incredibly adaptable with the effect of "disturbance" in picture.

Here in SIFT approach, the image feature extraction, it accepts the image and transform that to the "large collection of local feature vectors". All such component vectors are invariant to the scaling, turn if not the interpretation of the picture. These approaches share several features by responses of neurons in the primate vision. The SIFT calculation uses the 4 phase distinguished approaches for helping the extraction of these highlights:

Scale Spacing Extrema Detection

Here in this phase of differentiating endeavor towards distinguish such areas and the scales that are recognizable by various perspectives on a same kind of article. This can viably practice with a "scale space" work. Additionally it has been showed up that it is established using the Gaussian limit under reasonable suspicions. The following limit is used to describe the Scale space:

$$L(m, n, \sigma) = G(m, n, \sigma) * I(m, n)$$

Here * symbolizes the convolution operator given as, G(m, n, σ) indicates variable-scale Gaussian and the input image is represented by I(m, n).

Many techniques would now have the option to be used to identify stable key point territories in scale space. Qualifications of the Gaussians are a kind of framework, discovering scale space extrema, D (m, n, σ) enrolling that complexity between both images, with one image having the scale value which is k-times the value of the other image. The value of D(m, n, σ) is determined as:

$$D(m, n, \sigma) = L(m, n, k\sigma) - L(m, n, \sigma)$$

When distinguishing closer maximum and minimum of D(m, n, σ) each points are compared and its eight neighbors at a comparable scale and its nine neighbors back and forth in a scale. If the value is the reference point or the breaking point to all of these centers, by then this point will be an extrema.

Localization of Key Point

At this level, it attempts to wipe out different attentions from rundown of key points by selecting those with little complexity on an edge that are ineffectively limited. Area of the extremum, z, is determined as:

$$z = -\frac{\partial^2 D}{\partial \mathbf{x}^2}^{-1} \frac{\partial D}{\partial \mathbf{x}}$$

In breaking point respect Z is under the mark respect, that's the point kept up a key good ways from. With low distinction, this clears extreme. To avoid extreme dependent on poor containment, there is a vast standard shape on top of the edge in some cases yet a little repetitive pattern in opposite way as in qualification of the Gaussian breakpoint. If such qualifications are below the limit of the largest to the smallest own vector, the key point will be expelled from the 2x2 Hessian structures at the zone and key point level.

Assigning Orientation

These progressions focus to dole out reliable direction to the key point dependent on nearby picture features. The key point descriptor, shown below, can be expressed in contrast to the direction, rendering the revolution invariance. The methodology for discovering the direction is:

- Use the key focus scales to pick the smooth Gaussian image L from the top
- Register slope greatness,
- Compute orientation, θ
- Structure the direction histograms from the inclination directions of test highlights
- Choose the top most elevated in the histogram. Use these zeniths and others nearby top within 80% of this top's stature to make a key point with this bearing.
- A couple centers will by then be designated various bearings.
- Fit the parabola to three histogram regards next to each of them top to interpose the zenith locale.

Key point Descriptor

In addition, key point descriptors can be done by close-by-edge info, as illustrated above. Angle data pivoted to align with the key point direction and subsequently weighted with the Gaussian by value of 1.5* key point scale fluctuation. Then the obtained data value is utilized to make a lot of histograms over a key point-focused window.

A key point descriptor usually uses 16 histograms, mounted in a 4x4 frame, each with eight direction boxes, for all the basic compass bearings and 1 for all of mid-purposes of these headings.

B. SURF (Speeded up Robust Features)

The SURF incorporates identifier actuates by performing a derived Gaussian second backup spread to an image at various scales. Since the segment identifier applies cover along each center point on 45 degree to center point is better turned than Harris corner. The philosophy is fast an eventual outcome of the utilization of a significant picture during the estimation of a pixel (x,y) is obtained as aggregate of all the qualities in square shape depicted by the cause and the pixel (x,y). The entire pixels of such image inside the square state of any size in the source picture can be found as the outcome of four undertakings. It allows the rectangular front of several sizes to be applied from nearby no enrolling period.

For detection of the features, a Hessian matrix is assembled here is the convolution for the 2nd subordinate of Gaussian using the picture at that point. Covers utilized is unrefined estimate and are appeared in Fig 1. Hessian determinate qualities to a similar picture as Figure 1 are appeared in Fig 2 for the scope of identifier windows that were utilized in this work. Legitimate highlights is identified as nearby maximum over a 3x3x3 territory in which the third measurement are locator window sized, so a component should locally one of the kind over spatial range and scope of scales. The SURF creators utilized quick find calculation for the non-most extreme concealment; we have not completed this yet.

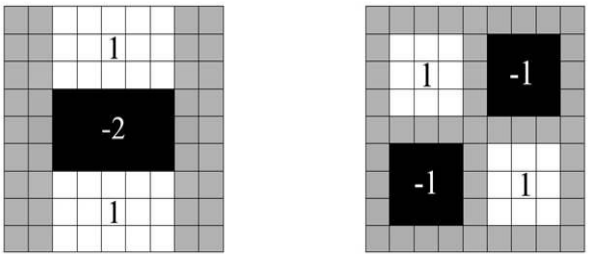

Fig. 1. Approximate Gaussian second derivatives utilized for SURF detector.

SURF descriptors are necessary for invariant scale and also invariant rotation. To order to overlook the scale, these descriptor is evaluated on the window that corresponds to size of the window to which it was identified, as well as the scaled representation of the component is another image the descriptor is used to evaluate the component along the comparative region.

Upset is dealt with finding the overall heading of component and rotating the looking at window to agree with that point. At the point when the turned neighborhoods gained are part into sixteen subs (Ai) squares and then again every sub square is isolated to four squares. Auxiliaries in the x and y headings are considered in such last squares. The sub square (Ai) descriptor is the aggregate of x subordinates over its 4 sections, total of through estimations of x backups and furthermore for y value. The hard and fast descriptor is four characteristics for each (Ai) for an aggregate of sixty-four characteristics. These vectors are institutionalized to length 1 and are the component descriptor. The procedures are condensed on Fig 2.

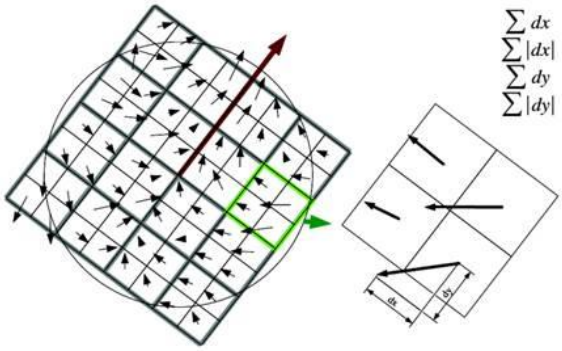

Fig. 2. SURF descriptor Graphical representation

C. PCA

Principal Component Analyses (PCA) is the performance straight change strategy which is comprehensively enabled over different field, most recognizably for the feature extraction and dimensionality decline. Other well-known uses of the PCA make use wide information examinations and the de-noising sign on securities exchange and in the field of bioinformatics.

The symmetrical tomahawks (head segments) of new subspace should be translated as headings of the most extreme difference provided the requirement that new element tomahawks is symmetrical for one another, is represented at the accompanying Fig 3.

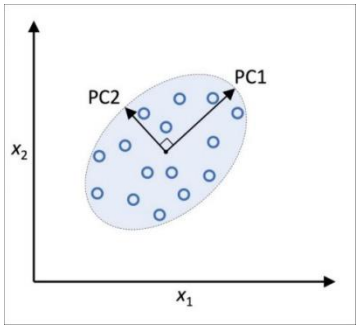

Fig. 3. Principal components v/s original features

In the former figure, x1 and x2 are the first element tomahawks, and PC1 and PC2 are the main parts.

On the off chance that we use PCA for dimensionality decrease, we develop a d x k–dimensional change network W which enables us to outline test vector x onto another k–dimensional component subspace that has less measurements than the first d–dimensional element space:

$$\boldsymbol{x} = [x_1, x_2, \ldots, x_d], \qquad \boldsymbol{x} \in \mathbb{R}^d$$
$$\downarrow \boldsymbol{x}\boldsymbol{W}, \quad \boldsymbol{W} \in \mathbb{R}^{d \times k}$$
$$\boldsymbol{z} = [z_1, z_2, \ldots, z_k], \qquad \boldsymbol{z} \in \mathbb{R}^k$$

Because of changing first d-dimensional information on this new k-dimensional subspace (ordinarily k ≪ d), primary head part should have biggest conceivable difference, and all ensuing head segments will have the biggest fluctuation given imperative that segments will be symmetrical to the next head segments regardless of whether the info highlights are connected, the subsequent head segments will be commonly symmetrical.

Note that PCA bearings is exceptionally delicate for information scaling, and also we have to institutionalize the highlights before PCA if the highlights were estimated on various scales and we need to dole out equivalent significance to all highlights.

Before taking a gander at PCA calculation to dimensionality decrease for more detail, we should outline the methodology in a couple of basic advances:

- Institutionalize the d-dimensional data set.
- Create the lattice of covariance.
- Deteriorate the network of covariance into its own vectors and values.
- Sort eigen values by diminishing request to rank the relating eigen vectors.
- Select the k eigen vectors which are used to relate the k largest eigen values. The value k indicates the dimensionality of new component subspace where k ≤ d.
- Construct a projection lattice W using the "top" k eigen vectors.
- Alter d-dimensional dataset X utilizing the projection lattice W to get new k-dimensional element subspaces.

D. CNN

The idea of Convolutional Neural Networks is motivated from working principles of the human brain. The theory of hierarchy plays the significant role in the human brain. The data is kept using the structures of the

patterns, following the successive order. The brain consists of an outermost layer called neocortex, which stores information in hierarchical structure. In the neocortex, data is arranged as cortical columns, or organized evenly as groups of the neurons. Fukushima (1980), a researcher developed the hierarchical neural network model and this model is named as the neocognitron by him. The idea of the Simple and Complex cells is the motivation to develop the model. The neocognitron model can recognize patterns using the knowledge learned from the shapes of objects. The concept of Convolutional Neural Network was introduced by Le Cun, Haffner, Bottou and Bengio in 1998. LeNet-5, is a Convolutional Neural Network created by them is capable of classifying the numbers from the hand-written digits. The architecture of Convolutional Neural Networks is different from the architecture of regular Neural Networks. In case of Regular Neural Networks, the input is transformed by passing it over a number of hidden layers. Every layer comprises a set of neurons and neurons in every layer have been fully connected to entire neurons present in previous layer. Lastly, there is an output layer, which is fully-connected layer which represents predictions. In Convolutional Neural Network, the layers are structured through three dimensions: height, depth and width. The neurons of every layer are not connected to all neurons of following layer however it connects to small part of it. Lastly, the output obtained is converted to single vector with probability scores and structured with depth dimension.

CNN has two parts:

- Feature extraction part (Hidden layers): The sequences of convolutions and the pooling tasks are accomplished by the network during which the features are identified in this part.
- Classification part: On top of the extracted features in the feature extraction part, a set of fully connected layers are served as a classifier in this part. The object on the image predicted by the algorithm is assigned the probability.

The Convolutional Neural Network mainly carries out four operations as shown in the Fig 4 below:

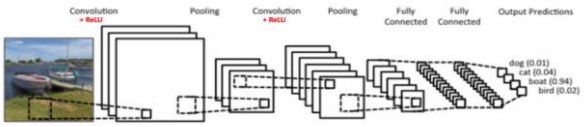

Fig. 4. Layers of Convolutional Neural Networks

1. Convolution
2. Non-Linearity (ReLU)
3. Pooling or Sub Sampling
4. Classification (Fully Connected Layer)

These four processes serve as the basic building blocks in every Convolutional Neural Networks.

1. Convolution

The key purpose of the Convolution is to extract the features from the given input image in case of a ConvNet. The spatial relationship between pixels is conserved through the convolution by learning the features of image by means of small squares of the input data.

Every image is represented by matrix in terms of the pixel values. For instance, consider the 5 x 5 image as shown in the Fig 5, the pixel values for the image are zero (0) and one (1).

1	1	1	0	0
0	1	1	1	0
0	0	1	1	1
0	0	1	1	0
0	1	1	0	0

Fig. 5. 5x5 Matrixes for Convolution

Consider the 3 x 3 matrix presented in Fig 6 below:

1	0	1
0	1	0
1	0	1

Fig. 6. 3x3 Matrixes for Convolution

Next, the Convolution function of 5 x 5 matrixes and 3 x 3 matrixes is calculated as presented in the Fig 7:

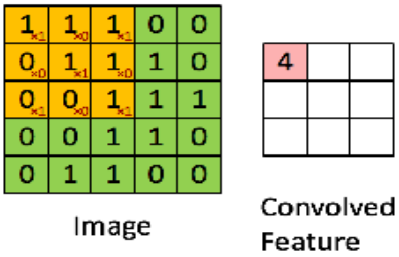

Fig. 7. Computation of 5x5 and 3x3 Matrix

The orange matrix slides over the original image colored with green by one pixel, it is also termed as 'stride'. Then for each position, an element wise multiplication between 2 matrices is computed and result of multiplication is added to obtain final value which gives the single element for output matrix colored with pink. The 3×3 matrix observes the portion of input image in every stride.

In the CNN terminology, a 3×3 matrix is known as the 'feature detector' or 'filter' or 'kernel'. The process of sliding filter through the image and then computing dot product results in a matrix, which is termed as the "Feature Map" or "Convolved Feature" or "Activation Map". The filter serves as feature detectors from original input image.

It is obvious that dissimilar values in a filter matrix create diverse Feature Map values for identical input image. The values of these filters are learned by the Convolution neural network on its own through training process but values of parameters like the architecture of network, number of filters, filter size etc. are to be specified earlier the training process. If the number of filters used are more, then the number of image features which can be extracted are more and hence the network turn out to be better in detecting patterns in the unseen images.

The size of Convolved Feature (Feature Map) is controlled with the 3 parameters, which should be decided before performing convolution process:

i. Depth: The depth parameter is related to the quantity of filters used for the convolution process.

The convolution of the original boat image is performed utilizing 3 separate filters as shown in the network and three different feature maps produced are as shown in Fig 8. The 3 feature maps are viewed as stacked 2-dimensional matrices, thus the value of 'depth' parameter of feature map is 3.

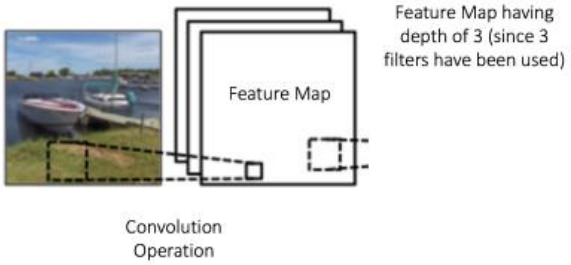

Fig. 8. Convolution Operation

ii. Stride: The stride parameter represents the amount of pixels used to slide the filter matrix through input matrix. If the value of stride is one, then filters are moved by 1 pixel at a time. If the value of stride is two, then filters are jumped by two pixels at a time during the sliding process. With the larger stride, smaller feature maps will be produced.

iii. Zero-padding: The filter is applied on the bordering elements in the input image matrix by padding the input matrix with 0's (zero-padding) nearby the border. One of the good features of zero-padding is that it permits the users to control the size of feature maps. The process of using zero-padding is termed as the wide convolution, and the not using zero-padding process is called as the narrow convolution.

2. Non-Linearity (ReLU)

After each Convolution process, an additional action called ReLU (Rectified Linear Unit) is used. ReLU is a non-linear process and the output of ReLU operation is shown in Fig 9.

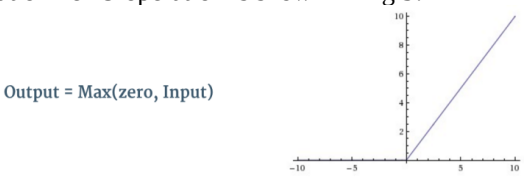

Fig. 9. The ReLU operation

The ReLU operation is performed element wise (applied per pixel). It substitutes all the pixel values which are negative in the feature map with a value of zero. The use of ReLU in ConvNet is to introduce the non-linearity, as most of the real-world data used by the ConvNet to learn is non-linear.

3. Pooling Step

It is also called as down sampling or subsampling or Spatial Pooling. The dimensionality of every feature map is reduced by pooling but preserves the most significant information. There are many types of Spatial Pooling: Sum, Average, Max etc.

In the Max Pooling, a spatial neighborhood (for instance, 2×2 windows) is defined and then the biggest element in rectified feature map is taken within that window. As an alternative to the biggest element, if the average value of elements is considered, then it is called Average Pooling, if sum of all elements is considered in that window then it is called Sum Pooling. The Max Pooling works better compared to other types in practice.

A sample of Max Pooling process in the Rectified Feature map by using 2×2 window after convolution and ReLU operation is shown in Fig 10.

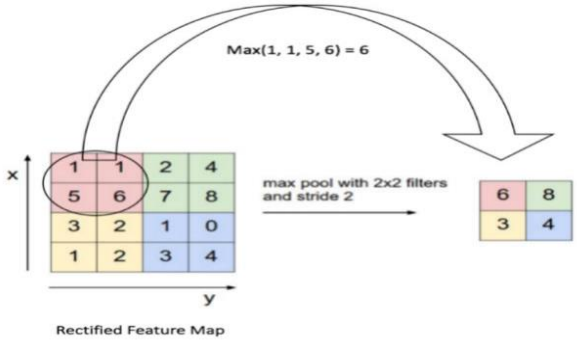

Fig. 10. Max Pooling Operation

The 2 x 2 window slides over through two cells, the process is termed as 'stride and the extreme value in every region is used. It decreases the dimensionality of feature map as shown in Fig 10.

The pooling operation is applied on every feature map separately in the network shown in the Fig 11. Due to this, 3 output maps are obtained from 3 input maps.

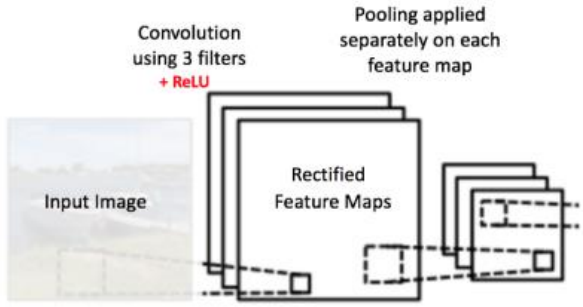

Fig. 11. Pooling Operation on Separate feature map

The purpose of pooling operation is to gradually decrease the spatial size in input representation. The functions of pooling are

- It makes the feature dimension (input representation) smaller and manageable
- It controls over fitting by decreasing the number of parameters and the calculations in a network.
- The network is made invariant to smaller transformations, translations and distortions in input image. The smaller distortion in input does not alter the result of Pooling as the average or maximum value in the local neighborhood is taken.
- It aids in arriving at nearly scale invariant representation of the image. It is called the "equi-variant".

Fully Connected Layer

This layer is the Multi-Layer Perceptron and the activation function, softmax is used in the output layer. Each neuron in the earlier layer is connected to all neurons in following layer, so it is called "Fully Connected".

The convolutional and the pooling layers produce output which represents high-level features obtained from input image. The function of Fully Connected layer is to utilize these features to categorize input image into several classes depending on the training dataset. Consider the image classification task as an illustration which has 4 potential outputs as presented in Fig 12.

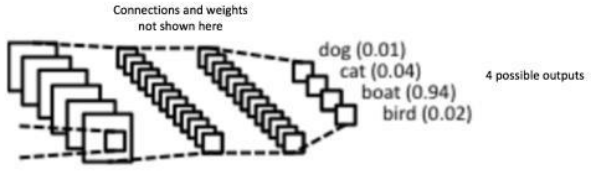

Fig. 12. Fully Connected Layer

The use of fully-connected layer, apart from classification also helps to learn the non-linear combination of these features. Many features obtained after the convolution process and pooling steps are good to perform the classification job, however the combining those features are even well.

4. Classification

The classification part after the convolution and the pooling layers comprises of few fully connected layers. The final layers in Convolutional neural networks are fully connected. The fully connected layers consist of neurons which contains full connections to all activations in earlier layer. The steps used to train the Convolution neural Networks can be summarized as follows:

1) The parameters / weights and filters are initialized with random values

2) The training image is given as input to network, and then passed into the convolution process, ReLU layer and the pooling process along with the forward propagation process in Fully Connected layers. It determines output probabilities for every class. Consider the sample image used in the above step with the rough output probabilities as [0.2, 0.4, 0.1, 0.3]. As the random values are assigned for the weights in the 1st training instance, the output probability values are also random.

3) The total error is computed as summation of all four classes at the output layer.

Total Error = Σ ½ (target probability–output probability) ^{2}

4) To determine the gradients of the error considering all the weights in the network, Back propagation is used. To minimize the output error, all filter values / weights and values of parameter are updated using the gradient descent. The weights are tuned proportionally considering their contribution for total error. If an identical image is used as input again, then values of output probabilities are [0.1, 0.1, 0.7, 0.1] now, those are nearer to values in target vector [0, 0, 1, 0]. It implies that the network has learned to sort the specific image properly by altering the filters/weights so that the value of output error is decreased. The factors like architecture of the network, number of filters, size of filters etc. are fixed before starting the training process and they are not changed during the process of training – only filter matrix values and the connection weights are updated.

5) The steps from 2 to 4 are repeated for all images in the training set.

These 5 steps are used to train the ConvNet, where all parameters and the weights of the ConvNet are optimized to classify the images correctly present in the training set.

The network passes over the forward propagation process and then outputs the probability value for every class, when the new (unseen) image is passed as input to a ConvNet. The output probabilities for a new image are computed utilizing the value of weights that are optimized to properly categorize all earlier training instances. Larger the training set, network generalizes better to the new images to classify them into the right classes.

IV. RESULTS AND DISCUSSION

In this work for the purpose of Experimentation we have created our own dataset of 4 different categories of general classes like Renault, Volkswagon, Rolls Royce, and Honda brand logos as listed in Table I which is shown at Figure 13.

Table I. Shows the Dataset of different brand logos

Here, extended the experiment over created and real datasets with the intention of reveal the capability of suggested method. The work has been implemented in a Matlab R2015a using an Intel Pentium 4 processor, 2.99 GHz Windows PC with 2 GB of RAM. In this work, the experimentation is conducted individual and fusion feature of SIFT and SURF are carried out. The features are reduced using PCA, here we fix the reduction percentage of PCA to 50. Precisely, pick logos randomly from the databases and experimentation is conducted more than five iterations. The effort outcomes are emphasizing to witness through maximum accuracy obtained in all cases. Moreover, the results obtained for reduced PCA with 50 percent over logos dataset with the appropriate training such as 30, 50 and 70 percentage plots are shown in Figures 13, 14 and 15. In addition,

experimentation is conducted without applying the reduction method PCA is shown in 16, 17 and 18. However, analysis of graphical representation can be noticed that the fusion-reduced features with reduction method achieve relatively higher accuracy in all cases. In addition, outcomes of Fusion approach have maximum accuracy when compared to individual feature due to its computational excellence.

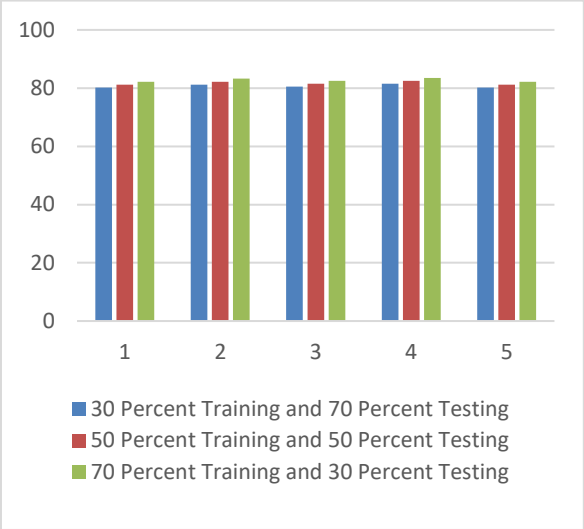

Fig. 13. Shows the Accuracy of SIFT features

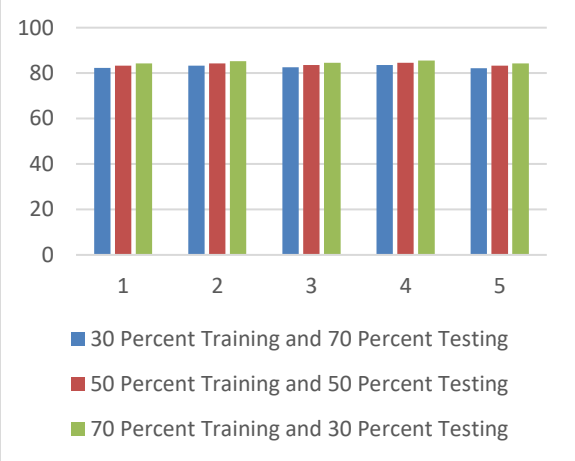

Fig. 14. Shows the Accuracy of SURF features

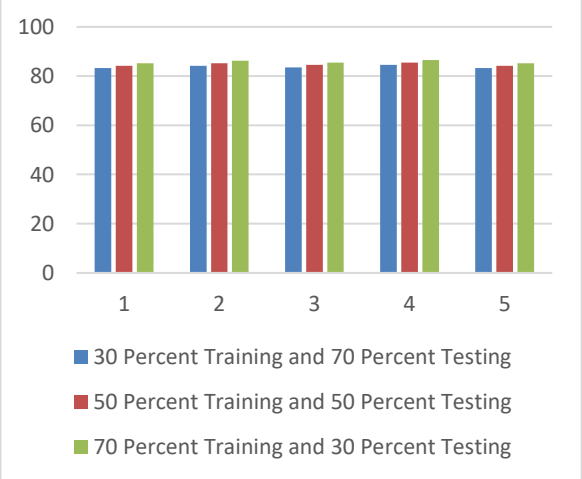

Fig. 15. Shows the Accuracy of Fusion features

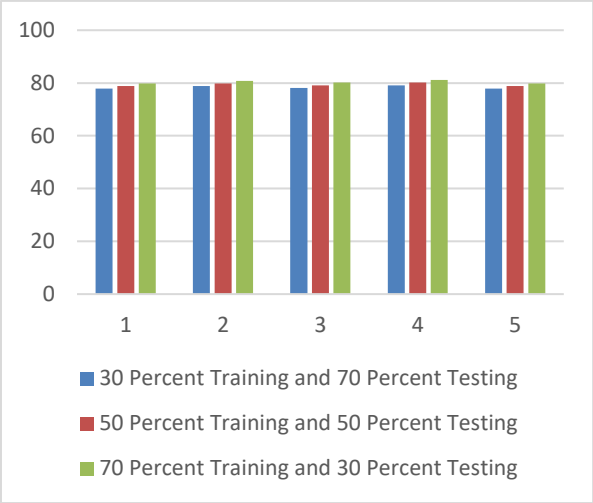

Fig. 16 Shows the Accuracy of SIFT features

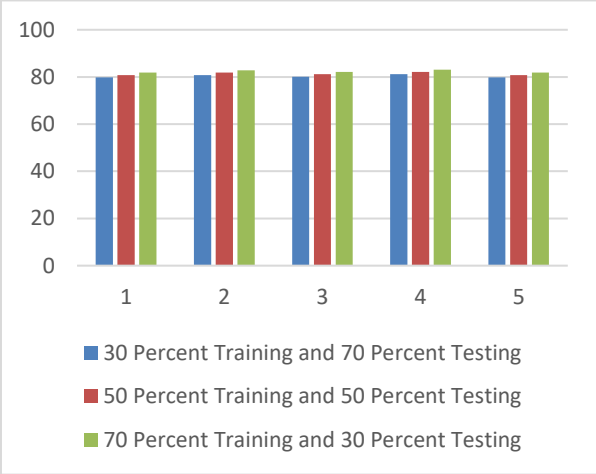

Fig. 17. Shows the Accuracy of SURF features

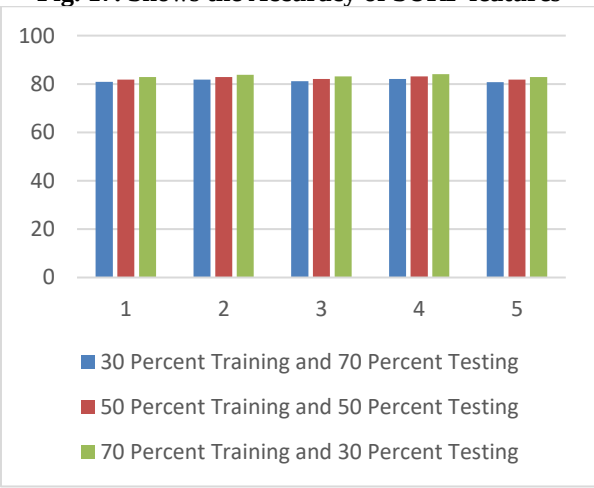

Fig. 18. Shows the Accuracy of Fusion features

V. CONCLUSION

In this paper, we have presented a method for Fake logo detection using fusion approaches of SIFT and SURF features. Features of the logo images are extracted using the SIFT and SURF techniques and the features are fused to obtain the significant features to detect the query image. Further the features were reduced with the use of PCA. The obtained features are stored for the database for querying the input image to determine whether it is a valid or fake logo. The experimentation is done with the own dataset of 144 image of six different brand car logos were considered out of which 20 images were of fake logos.

VI. FUTURE SCOPE

The originality of the logo can be detected more precisely using different optimal features extracted. The work can be enhanced to multi classification of logos for documents to classify the sub classes. Further the tree indexing method can be applied for effective retrieval system.

REFERENCES

[1]. ZhangNan., (2012). Vehicle Logo Recognition Based on Classifier Combinations.IEEE. 10.1109/ICSSEM.2012.6340863.

[2]. Bailing Zhang., Hao Pan., (2013). Classification of Vehicle Logos by an Improved Local mean based Classifier. IEEE.

[3]. Chun Pan., Zhiguo Yan., XiaomingXu., Mingxia Sun., Jie Shao., Di Wu., (2013). Vehicle Logo Recognition Based On Deep Learning Architecture In Video Survelliance For Intelligent Traffic System. ICSSC.

[4]. N. Vinay Kumara, Pratheeka, V. Vijaya Kanthaa, K. N. Govindarajua, and D. S. Gurua., (2016). Features Fusion for Classification of Logos. Procedia Computer Science 85 (2016) 370 – 379. Elsevier B.V.

[5]. Changbo Hu., Qun Li., Zhen Zhangy., Keng-hao Chang and Ruofei Zhang., (2020). A Multimodal Fusion Framework For Brand Recognition From Product Image And Context. 978-1-7281-1485-9/20. IEEE.

[6]. Shreyansh Gandhi, Samrat Kokkula, Abon Chaudhuri, Alessandro Magnani, Theban Stanley, Behzad Ahmadi, Venkatesh Kandaswamy, Omer Ovenc., (2020). Scalable Detection of Offensive and Non-compliant Content / Logo in Product Images. IEEE Winter Conference on Applications of Computer Vision (WACV).IEEE.

[7]. Qin Gu., Jianyu Yang., Guolong Cui., Lingjiang Kong., HuakunZheng., ReinhardKlette., (2016). Multi Scale Vehicle Logo Recognition By Directional Dense Sift Flow Parsing. IEEE.

[8]. Ruilong Chen., MdAsif Jalal., Lyudmila Mihaylova., Roger K Moore., (2018). Learning Capsules for Vehicle Logo Recognition. ISIF.

[9]. Y H Sharath Kumar., K C Ranjith., (2016). An Approach for Logo Detection and Retrieval in Documents. RTIP2R, Springer.

[10]. Linghua Zhou., Weidong Min., Deyu Lin., Qing Han., Ruikang Liu., (2020). Detecting Motion Blurred Vehicle Logo in IoV Using Filter-DeblurGAN and VL-YOLO. 10.1109/TVT.2020.2969427, IEEE.

[11]. Chang, S.L., Chen, L.S., Chung, Y.C., Chen, S.W., (2004) Automatic license plate recognition. IEEE Transactions on ITS 5(1), 42–53.

[12]. Ciocca, G., Napoletano, P., Schettini, R., (2015) Iat - image annotation tool: Manual. CoRR.

[13]. Dehghan, A., Masood, S.Z., Shu, G., Ortiz, E.G., (2017). View independent vehicle make, model and color recognition using convolutional neural network. CoRR.