VISVESVARAYA TECHNOLOGICAL UNIVERSITY

"JNANA SANGAMA", BELAGAVI-590018.

2020-2021

Mini Project Report on

Learning Semantic Features for Classifying Very Large Image Datasets Using HOG and Machine Learning Techniques

Submitted in partial fulfillment for the award of the degree of

BACHELOR OF ENGINEERING

IN

ELECTRONICS AND COMMUNICATION ENGINEERING

Submitted By

Pranesh B	**[1JB18EC068]**
Nitin T	**[1JB18EC064]**
Shree Charan	**[1JB18EC087]**
D P Tejash	**[1JB18EC102]**

UNDER THE GUIDANCE OF

Dr Mahantesh K
Associate Professor
Dept. of ECE, SJBIT

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

SJB INSTITUTE OF TECHNOLOGY

B G S HEALTH AND EDUCATION CITY

Kengeri, Bangalore-560060

|| JAI SRI GURUDEV||

Sri Adichunchanagiri Shikshana Trust ®

SJB INSTITUTE OF TECHNOLOGY

BGS Health & Education City, Kengeri, Bangalore-560060.

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

SJBIT

CERTIFICATE

Certified that the Mini Project work entitled "**Learning Semantic Features for Classifying Very Large Image Datasets Using HOG and Machine Learning Techniques"** carried out by **Pranesh B [1JB18EC068], Nitin T [1JB18EC064], Shree Charan [1JB18EC087], D P Tejash [1JB18EC102]** are Bonafede students of **SJB Institute of Technology** in partial fulfillment for the award of "**BACHELOR OF ENGINEERING**" in **ELECTRONICS AND COMMUNICATION ENGINEERING** as prescribed by **VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI** during the academic year **2020-2021**. It is certified that all corrections/suggestions indicated for internal assessment have been incorporated in the report deposited in the department library. The report has been approved as it satisfies the academic requirements in respect of Mini Project work prescribed the said degree.

Dr Mahantesh.K	**Dr K.V. Mahendra Prashanth**	**Dr. Ajai Chandran C.K**
Guide & Associate Professor	**Professor & Head**	**Principal**
Dept. of ECE, SJBIT	**Dept. of ECE, SJBIT**	**SJBIT**

EXTERNAL VIVA-VOCE

Name of the Examiners	**Signature with date**
1. _____	_____
2. _____	_____

ACKNOWLEDGEMENT

SJBIT

We would like to express our profound grateful thanks to **His Divine Soul Jagadguru Sri Sri Sri Padmabhushana Dr. Balagangadharanatha Mahaswamiji** and His Holiness **Jagadguru Sri Sri Sri Dr. Nirmalanandanatha MahaSwamiji** for providing us an opportunity to be a part of this esteemed institution.

We would also like to express our profound thanks to **Reverend Sri Sri Dr. Prakashnath Swamiji,** Managing Director, SJB Institute of Technology, for his continuous support in providing amenities to carry out this Mini project in this admired institution.

We express our gratitude to **Dr. Ajai Chandran C K,** Principal**,** SJB Institute of Technology, for providing us excellent facilities and academic ambience, which helped us in satisfactory completion of Mini project work.

We extend our sincere thanks to **Dr. K V Mahendra Prashanth**, Professor & Head, Department of ECE, for providing us invaluable support throughout the period of our Mini project work.

We wish to express our heartfelt gratitude to our guide, **Dr.Mahantesh K**, for his/her valuable guidance, suggestions and cheerful encouragement during the entire period of our Mini project work.

We express our truthful thanks to **Mrs. Latha S & Mrs. S Nithya** Mini project coordinators, Dept. of Electronics and Communication Engineering, for their valuable support.

Finally, We take this opportunity to extend our earnest gratitude and respect to our parents, teaching & technical staff of the department, the library staff and all our friends, who have directly or indirectly supported us during the period of our Mini project work.

Regards,

PRANESH B [1JB18EC068]
NITIN T [1JB18EC064]
SHREE CHARAN [1JB18EC087]
DP TEJASH [1JB18EC102]

DECLARATION

We hereby declare that the entire work embodied in this Mini project report has been carried out under the supervision of **Dr. Mahantesh.K, Associate Professor** in partial fulfilment for the award of "BACHELOR OF ENGINEERING" in ELECTRONICS AND COMMUNICATION ENGINEERING as prescribed by VISVESVARAYA TECHNOLOGICAL UNIVERSITY, BELAGAVI during the academic year 2020 - 2021.

Pranesh B [1JB18EC068]

Nitin T [1JB18EC064]

Shree charan[1JB18EC087]

D P Tejash [1JB18EC102]

ABSTRACT

The intention behind this project is to propose a transform technique that is derived by applying 2D wavelets in the DCT domain to categorize objects in large multi-class image datasets. Transform techniques such as Coslets are created by using wavelets and DCT sub-bands to analyze the pixel distribution of segmented regions in a multi-resolution framework. The obtained coefficients are projected onto a lower-dimensional feature space using PCA to produce compressed and de-correlated feature vectors. Later, DCT brings out low frequency components expressing the image's visual features and further wavelets decompose these coefficients into multi-resolution sub-bands giving an advantage of spectral analysis to develop robust and geometrically invariant structural object visual features.

A set of observed data is mapped onto a lower-dimensional feature space with a transformation matrix using PCA along with a histogram of oriented gradients as its feature vectors. Finally, different distance measure techniques and machine learning algorithm analysis are used for classification to obtain an average recognition rate for object categorization. To obtain the highest classification rates in comparison to several benchmarking techniques investigated in this project.

Content-Based Image Retrieval methods are emerging as influential next-generation tools for accurate image retrieval from massive digital image databases, with a wide range of applications in fields such as shape recognition, medical diagnosis, remote sensing, digital forensic, radar engineering, and robotics. The main idea behind this work is to use segmented images to represent smoothness along edges in Covid-19 X-ray radiography database to map point singularities on to objects that can have the best match in the compressed domain and Convolution neural network with accuracy of 97% of validation data.

Real-time object detection is the task of doing object detection in real-time with fast inference while maintaining a base level of accuracy. It is a challenging computer vision task that requires both successful object localization in order to locate and draw a bounding box around each object in an image, and object classification to predict the correct class of object that was localized. YOLO works by applying a single neural network to the full image input. The network divides each input image into an S-by-S grid and each grid cell predicts a predetermined number of bounding boxes

The scope of this project can be extended in object detection in detecting real-time trained instances. This project also identifies the main qualitative research traditions that can be applied to Radiography image data.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF ABBREVIATIONS

SL No	Particulars	Acronym
1	Discrete Cosine Transform	DCT
2	Histogram Of Oriented Gradients	HOG
3	Principal Component Analysis	PCA
4	Daubechies Wavelet	db Wavelets
5	Machine Learning	ML
6	Kullback-Leibler Divergence	KL Divergence
7	Jensen-Shannon Divergence	JS Divergence
8	Euclidean and Manhattan Distance	-
9	Convolutional Neural Network	CNN
10	You Only Look Once	YOLO

CHAPTER 1

1.1 Introduction

The innovation and research and development sectors are growing rapidly due to the digital collection of data. The development of the internet and digital media techniques increased the demand for data. so, classification is an essential and required process for searching and efficient indexing of data. This leads to the existence of data classification and detection techniques and significant interest in the research community.

The traditional method for extracting the image from huge datasets is by interpreting the image through words and later using the keywords for retrieval from the databases, but manually annotating the images is very time consuming. Moreover, due to the consistent natural failure of keywords to link the semantic gap, the accuracy of methodology is frequently questioned. However, manually annotating images for a wide spectrum of images is obviously a cumbersome and expensive task for large image datasets which is often subjective to human perception, context sensitive and incomplete. Since text-based methods failed to support a variety of task-dependent queries, content-based image retrieval (CBIR) was introduced as an effective alternative in the early 1980's. In CBIR, images are indexed by their visual content, such as color, texture & shapes. The problem of extracting/matching images has remained primarily statistical in nature, image retrieval systems employ pattern recognition methods to define the visual content with partial semantics.

Many sophisticated algorithms designed to describe color, shape, and texture features, cannot adequately model image semantics and portray limitations while dealing with broad content image databases. In this project

we are also using some trending machine learning algorithms. Machine Learning (ML) is constructing computer programs that develop solutions and improve with experience to solve problems which cannot be solved by enumerative methods or calculus-based techniques. Intuition is to model human way of solving some problems which require experience.

1.2 Motivation

- Classification is an essential and required process for searching and efficient indexing of data.
- Digitalization of data and evolution of internet increased the demand for data.
- Image classification plays an important role in remote sensing images and is used for various applications such as environmental change, agriculture, surveillance, geographic mapping, disaster control, and object detection. This process involves extracting the required features from image data.
- The Content Based Image Retrieval (CBIR) method is emerging as an influential next generation tool for accurate image retrieval.

1.3 Literature survey

For image classification tasks, various methods are used, including methods based on low-level image feature representation, which considers an image as a set of low-level characteristics such as texture, form, scale, colour, and so on, and methods based on mid-level visual feature constructions. Image processing with different ensemble methods has a large number of publications. The benefits and applications of image classification in the fashion domain are numerous, and numerous research papers have been presented on the topic. For image classification, [13] suggested HOG techniques. Turk and Pentland [14] were the first to use PCA to extract significant features from a function space. For face recognition, Pankaj and Wilscy [15] proposed a method for comparing PCA, and features.

We developed a subspace-based algorithm for improving the correctness rate on large image datasets in linear and nonlinear subspaces, taking into account the advantages of real space as described in [16] and the properties of PCA as described in [17]. The following is how the paper is organized: In Sects. 2 and 3,the suggested approach and classification have been discussed respectively. Nowadays, usage of HOG and Principle component analysis (PCA) the use of Principal Component Analysis with HOG, as discussed in [12], allows us to obtain features from successful regions. PCA is a statistical procedure for converting a large number of correlated data into a smaller number of uncorrelated data, known as principal components [8]. HOG feature vectors are fed into PCA, which rewrites HOG features in terms of new variables using feature vectors derived from HOG. Feature extraction techniques examine images inorder to remove the most influential features that reflect different object classes.

Classifiers use features as inputs to allocate them to the class that they represent. Feature extraction allows you to minimize the amount of data you have by analyzing certain properties of images that contain relevant data, or features, that differentiate

one pattern from another. There are various types of features, such as shape-based, color-based, and texture-based features, among others. The texture is one of the most important characteristics for identifying objects or regions of interest in an image. The following are some of the texture features used in this method.

Following that, DCT extracts low-frequency components that convey image visual features, and wavelets decompose these coefficients into multi-resolution sub-bands, allowing spectral analysis to create geometrically invariant structural object visual features. PCA is used to map a set of observed data (i.e. transformed coefficients) onto a lower-dimensional feature space using a transformation matrix. Finally, various distance measurement techniques are used for object categorization to achieve an average correctness score. On two extremely difficult datasets, we demonstrated the proposed work's approach and obtained the highest classification rates in comparison with several benchmarking techniques explored in the literature.

Nowadays, the usage of Logistic regression is to obtain an image model that is trending. In several image and medical data classification tasks, logistic regression is the model of choice. From a technical standpoint, we summarise the differences and similarities of these models and compare them to other machine learning algorithms in this analysis. We provide considerations for evaluating the consistency of models and the findings based on these models objectively.

1.4 Objective of Work

An efficient image representation and extracting discriminative features in compressed domain is attracting researchers in computer vision and pattern recognition to develop efficient algorithms to classify and annotate images in large datasets. There are numerous methods of classifying images nowadays the intent of the classification process is to categorize all pixels in a digital image.

The main objective of this Case study is to provide awareness and to identify the research areas related to COVID-19. It may help improve the understanding of this disease and describe the psychological impacts of this pandemic and how these could change as the disease spreads.

Since the recent sudden surge of COVID-19 infections across the world, many alternative screening approaches have been developed to identify suspected cases of COVID-19.

The intention of our next case study in on real time object detection is to detect all instances of objects from a known class, such as people, cars or faces in a multiple video input device/feed.

CHAPTER 2

2.1 Problem Statement

- To bridge semantic gap between high level and low-level features.
- Image representation and multi resolution analysis in compressed domain for image classification.
- Covid-19 prediction using Chest Xray image data.
- Real time object detection from multiple camara feed.

2.2 Proposed Methodology

In this section, the frequency domain (DCT) and combined time-frequency domain techniques are described in this section (Wavelets). By isolating edge discontinuities and capturing the maximum energy of the given image, DCT (Discrete Cosine Transforms) is used to convert 2-D signals into elementary frequency components of lower frequency and wavelets representing point singularities. In the following pages, we'll go through the specifics of combining DCT and wavelets.

2.2.1 Data pre-processing

Image recognition is perhaps one of the most well-known computer vision activities. It allows a given image to be classified into one of a set of predefined categories. It's a set of operations on a picture that you can perform. These operations are often required to convert an image into a format suitable for training. Converting to grayscale, rotating, cropping, and detecting edges are the operations available. Basic statistical features can be used to classify image data. The first, second, and third moments. It is possible to locate the moments since images can be interpreted in terms of pixel values.

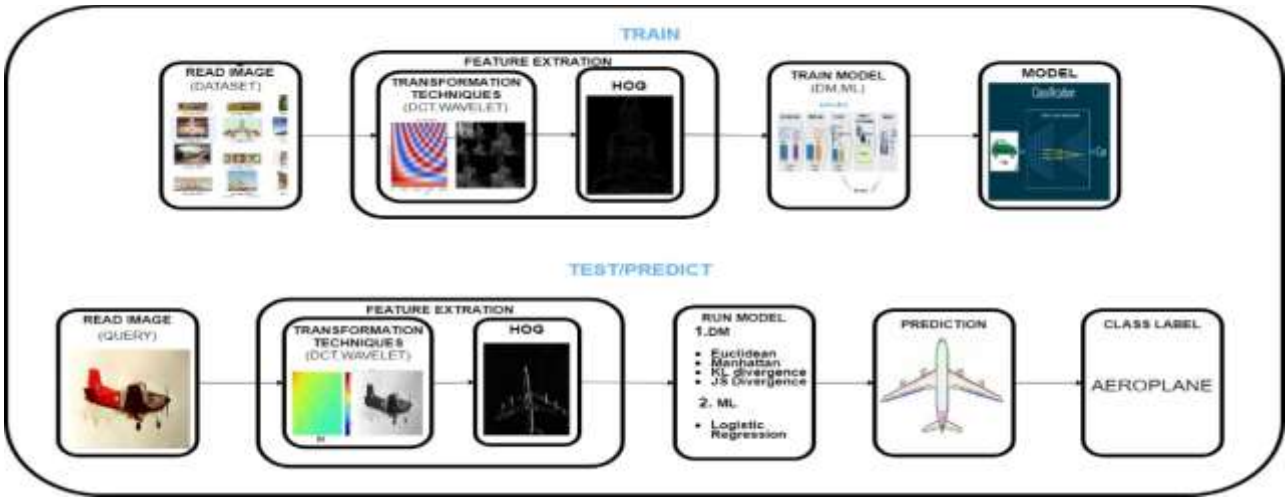


Fig1: Proposed Methodology

2.3 Transformation Techniques

Transformation is a function or operator that takes an image as its input and produces an image as its output. This section follows some transformation techniques like DCT, PCA and wavelets. Discrete Cosine Transform, used in image and video compression. Wavelet Transform, used to perform discrete wavelet analysis, denoise, and fuse images. PCA is a process of computing the principal components and using them to perform a change of basis on the data, sometimes using only the first few principal components and ignoring the rest.

2.3.1 Principal component analysis (PCA) on image Data

As mentioned in [12], the use of Principal Component Analysis with HOG enables us to obtain features from effective regions. Since feature extraction is so important, the investigation of subspace-based feature extraction begins with principal component analysis. It reduces the amount of correlated image data available by converting it into a smaller number of linearly uncorrelated variables known as score features [14]. The set of images Ii of size s*s is considered, with I = 1 to n (n varying depending on the dataset used, e.g., Caltech n = 101/256). The average of both of these images is calculated using this Eq .

$$m = \frac{1}{n} I_i \tag{1}$$

The images are mean centered as $I_i' = I_i - m$. Further covariance matrix is obtained as follows:

$$C = \frac{1}{n} \sum_{i=1}^{n} I_i' I_i'^T \tag{2}$$

The covariance matrix is used to obtain the eigenvectors e_i and corresponding eigenvalues λ_i. After projecting the image into the $S_i' = e_k^T I_i'$, the e_k eigenvectors of the largest eigenvalues are considered for feature extraction by PCA.

We get its score features S_i' from the eigenspace, which is used for classification. PCA is found to have weaknesses in terms of lacking class separability and localization, so MPCA is used to address this limitation.

Steps involved in principal component analysis.

- Subtraction of the mean of each variable.
- Covariance Matrix Calculation.
- Computing the Eigenvalues and Eigenvectors.
- Sort the Eigenvalues in order of decreasing Eigenvalues.
- Selecting a subset from the rearranged Eigenvalue matrix.
- Data transformation.

(a) Base image **(b) BAND – 1** **(c) BAND – 2** **(d) BAND – 3**

Fig2: PCA Images based on different bands

2.3.2 Discrete Cosine Transforms

Since PCA is found to have weaknesses in terms of lacking class separability and localization The DCT transform is widely used in image compression and face recognition [12, 13]. The signal intensity of most images is concentrated at low frequencies, and DCT effectively divides image spectral sub-bands based on visual features. It converts a signal from one frequency domain to another.

Let N be the length of the series f(x),x = 0,1,2,...,N1. The following equation describes 1D DCT:

$$F(u) = \left(\frac{2}{N}\right)^{\frac{1}{2}} \sum_{x=0}^{N-1} \Lambda(x) cos \left[\frac{(2x+1)u\pi}{2N}\right] f(x) \quad (3)$$

The DCT of a m x n image f(x,y) is given by:

$$C(u,v) = \frac{2}{\sqrt{mn}} \alpha(u)\alpha(v) \sum_{x=1}^{m}\sum_{y=1}^{n} f(x,y) cos \left[\frac{(2x+1)u\pi}{2m}\right] cos \left[\frac{(2y+1)v\pi}{2n}\right] \quad (4)$$

where u = 1,2,...,m and v = 1,2,...,n are scaling factors, as opposed to a 1D signal.

Fig.3(a) depicts a color aero plane image with a scale of 256 X 256 pixels, Fig.3(b) depicts the color map of its DCT's log magnitude, indicating high energy compaction at the origin, and Fig.3(c) depicts the scanning strategy of DCT coefficients for an image using the traditional zigzag technique. Low-frequency coefficients are preserved by DCT coefficients in order to distinguish local knowledge that is invariant to lighting, occlusion, and clutter, effectively nullifying the effect of high-frequency coefficients.

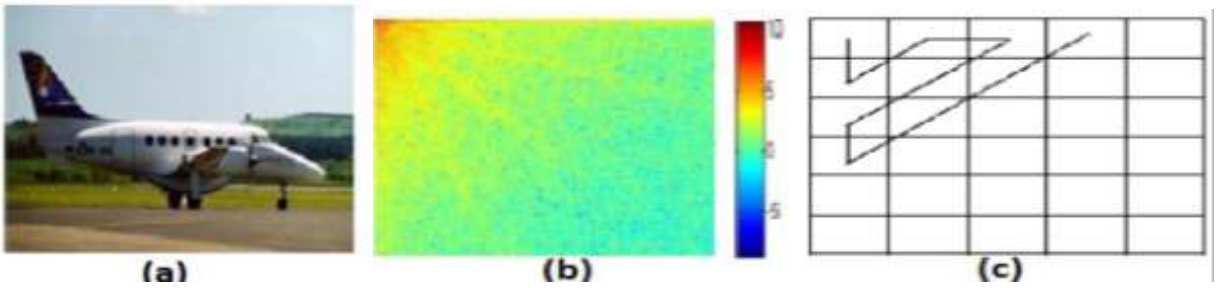


Fig 3 (a) original image; (b) color map of quantized magnitude of DCT;(c) scanning strategy of DCT coefficients

2.3.3 Wavelets

The objective of image compression is to use the fewest number of bits feasible to display an image while ensuring the integrity of the recovered image to meet the needs of specific implementations. The Daubechies' extremal phase wavelets are the dbN wavelets. The number N denotes the number of vanishing moments. Wavelet is a mathematical method used in a variety of fields that has recently seen rapid development and use in image processing and object detection.

The Daubechies feature extraction is described in the same way as the Haar wavelet transform by calculating running averages and differences using scalar products with scaling signals and wavelets. The only difference is how these scaling signals and wavelets are defined.

The scaling signals and wavelets have slightly longer supported for the Daubechies wavelet transformations, i.e. Using just a few such values from the signal, they generate correlations and differences. This minor update, on the other hand, vastly improves the capabilities of these new transforms. They offer us a collection of powerful tools for performing fundamental image processing tasks. These tasks include neural network compression and noise reduction, as well as image enhancement and optimization. Daubechies combined the principles of multi-resolution analysis and pyramid coding techniques, then outlined their similarities and differences, and initially incorporated wavelet transform with the conventional definition of filter banks. Currently, the Daubechies wavelet is a typically used orthogonal wavelet.

Even if a source image is decomposed, four equal-sized output images are produced:

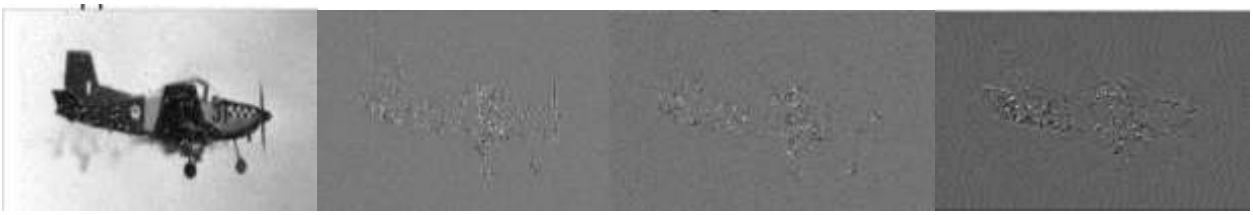


Fig 4: (a) (b) (c) (d)

Fig 4 (a): An approximation image**:** Vertical and horizontal lowpass filtering is used to create the 'Approximation' image.

Fig 4 (b): is a horizontal detail image**:** Vertical high pass and horizontal lowpass filtering are used to create the 'Horizontal detail' image

Fig 4 (c): is a vertical detail image: Vertical low pass and horizontal high pass filtering are used to create the 'Vertical detail' image.

Fig 4 (d): is a diagonal detail image**:** Vertical and horizontal high pass filtering is used to create the 'Diagonal detail' image.

2.4 Feature extraction

Feature extraction is a process of dimensionality reduction by which an initial set of raw data is reduced to more manageable groups for processing. A characteristic of these large data sets is a large number of variables that require a lot of computing resources to process.

2.4.1 HOG

HOG, or Histogram of Oriented Gradients, is a feature descriptor that is often used to extract features from image data.

(a) Original image

(b) HOG representation

Fig 5: HOG extraction general view – graphical description of all the extraction step

It is widely used in computer vision tasks for object detection. The configuration used in this work differs slightly from the best one found in Dalal and Triggs' work. The selected configuration consists of linear gradient voting into 9 orientation bins in 0◦ to 180◦; 16x16 pixel blocks of four 8x8 pixel cells.

The size of the descriptor can be calculated as:

$$\boldsymbol{Descriptor\ size = b_s^2\, n_b\, (v_{cells} - b_s + \sigma))(h_{cells} - b_s + \sigma)} \quad (5)$$

Descriptor size equation

The detection window has a size of 64 128. The first step in creating the HOG descriptor is to compute the 1-D point derivatives Gx and Gy in the x and y directions by converging the gradient masks [−1 0 1] and [−1 0 1]T with the raw image I

The equations defining the gradients are, respectively (being I the image and (i, j) the pixel coordinates):

$$G_x(i,j) = \frac{\partial I}{\partial x}(i,j) \qquad G_y(i,j) = \frac{\partial I}{\partial y}(i,j) \quad (6)$$

Gradient equations

The gradient degree |G(x,y)| and direction angle (x,y) for each pixel are then calculated using the derivatives Gx and Gy. The gradient degree indicates the intensity of the gradient at a given pixel:

$$M(i,j) = \sqrt{G_x^2(i,j) + G_y^2(i,j)} \quad (7)$$

Gradient magnitude equation

The gradient degree is simply employed as a weighting factor for the direction histogram. The gradient direction angle can be easily calculated using the following formula:

$$\theta(i,j) = arctan\left(\frac{G_y(i,j)}{G_x(i,j)}\right) \quad (8)$$

Gradient angle equation

Orientation / Spatial Binning introduces nonlinearity into the descriptor and implies the creation of a histogram for a collection of local spatial areas, which we refer to as cells. A weighted vote is issued for each pixel within a cell for the bin corresponding to the angle of its gradient, and the issued votes are then accumulated to form the final histogram for that cell. This histogram represents angles that are evenly spaced between 0 and 180 or between 0 and 360, depending on whether the angle is signed or unsigned. As a result, the resulting equation for computing the kth bin of the histogram is:

$$h_k = \sum_{i,j} = M(i,j)1|\Phi(i,j) = k| \quad (9)$$

Bin k equation

In our specific case, k ranges from 1 to 9. Given the previous angle division in bins, probably given an angle it will fall between to bin centers, so the vote is calculated via a bilinear interpolation between the two neighboring bin centers

$$f(x|x_1,x_2) = f(x_1) + \frac{f(f(x_2)-f(x_1))}{x_2-x_1}\,(x-x_1) \quad (10)$$

Linear interpolation equation

For the sake of a robust descriptor, some kind of illumination normalization must be done, this is evidenced when paying attention to the variations in the

gradient's strength. Actually, this step is essential to achieve good results. Let's define first v as the vector containing all the histograms for a given block, ||v||k the k-norm of v, with k ∈ 1, 2, and let's ∈ be a small constant.

$$L1 - norm: v \rightarrow \frac{v}{||v||_2 + \epsilon} \quad (11)$$

L1-normalization for the v descriptor vector of a block

$$L1 - squared\ norm: v \rightarrow \sqrt{\frac{v}{||v||_2 + \epsilon}} \quad (12)$$

L1-squared-normalization for the v descriptor vector of a block

$$L2 - norm: v \rightarrow \frac{v}{\sqrt{||v||_2^2 + \epsilon^2}} \quad (13)$$

L2-normalization for the v descriptor vector of a block

An L2-norm followed by a clipping, limiting the values of v to 0.2 and re-normalizing again. This can be achieved by getting the result of the L2-norm and cutting it and normalizing again. The experiments showed that either L1-norm-squared, L2-norm, or L2-norm-Hys performs similarly and achieves good results, but L1-norm decreases performance in a 5%. Not normalizing penalizes enormously the performance in around a 27%.

2.5 Classification Models

This section briefly describes the classifiers used in the proposed method calculating the visual similarities between test image and the database images.

2.5.1 Distance Classifiers

We first look at minimum distance classifiers that use feature variance. In multi-feature space, the distance classifier is used to classify unknown image data into classes with the shortest distance between the image data and the class.

Euclidean Distance method

It is used in situations where the population groups' variation differs from one another. The distance of the Euclidean is a distance of two points. The two points P and Q in Euclidean two-dimensional spaces, and P with coordinates (p1, p3) (q1, q2). A hypotenuse of a right angles triangle will be the line segment with the endpoint P and Q. The distance between 2 points p and q is defined as the square root of the sum of the different squares between the coordinates of the respective points. The Euclidean geometry of two dimensions, the Euclidean distance between a = (ax, ay) and b = (bx, by) of two points, are defined as:

$$d(a,b) = \sqrt{(bx - ax)^2 + (by - ay)^2} \tag{14}$$

The Euclidean distance and the similarity index are technically equivalent. Column x of the column and a collection of column vectors in a codebook matrix cb are computed by the Euclidean distance algorithm. The algorithm calculates the minimum distance from x and identifies the closest column vector cb to x.

$$d(a,b) = |p - q| \tag{15}$$

$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2 + \cdots + (p_n - q_n)^2} = \sqrt{\sum_{i=1}^{n}(p_i - q_i)^2} \tag{16}$$

The distance between x1 and x2 in a line in one dimension is simply the absolute value of the difference between the two points, as follows:

$$\sqrt{(X_2 - X_1)^2} = |X_2 - X_1| \tag{17}$$

In two dimensions, the distance between P = (p1, p2) and q = (q1, q2) as:

$$\sqrt{(p_1 - q_1)^2 + (p_2 - q_2)^2} \tag{18}$$

Manhattan distance method

The Manhattan distance is used to determine how far two real-valued vectors are apart. The sum of the absolute differences between the two vectors is used to measure Manhattan distance. The distance in the grid between two points is based on comparison with the diagonal on a strictly horizontal and/or vertical path. The distance of Manhattan is simply the sum of the vertical and horizontal components, whereas it is possible to calculate the diagonal distance using the Pythagoras theorem. The distance from L1 is also called and of u = (x1,y1) and v = (x2,y2) and two points, then the distance from Manhattan between u and v given by

M_{dist} = Manhattan distance

$$MH(u, v) = |x_1 - x_2| + |y_1 - y_2| \tag{19}$$

Instead of two dimensions, if the points have n-dimensions, such as a = (x1, x2,..., xn) and b = (y1, y2,..., yn) then, equation (1) can be generalized by defining the Manhattan distance between a and b as

$$MH(u, v) = |x_1 - y_1| + |x_2 - y_2| + \cdots \ldots + |x_n - y_n| = \sum_{i=1}^{n} |x_i - y_i| \tag{20}$$

Kullback-Leibler Divergence

The Kullback-Leibler Divergence index, also known as the KL divergence score, measures how much one probability distribution varies from another.
The KL divergence between two distributions Q and P is commonly expressed as:

$$KL\ (P \parallel Q)$$

Where the "||" operator indicates "divergence" or P's-divergence in relation to Q. KL-divergence may be calculated and multiplied by the log of the event in P over the event-probability in Q and as a sum of each event in P.

$$KL\ (P \parallel Q) = \text{sum x in X } P(x) * \log(P(x) / Q(x)) \quad (21)$$

For zero-valued odds, an alarm "log of zero" will be thrown. Take care of it as you see fit. It seems fair to add a slight additional value to all probabilities. (If applicable, renormalize)

The KL divergence score's intuition is that there is a broad divergence when the likelihood of an event from P is high but the same event in Q is low. There is also a large divergence when the probability from P is small and the probability from Q is large but not as large as in the first case. It can be used to determine the difference between discrete and continuous probability distributions, where the integral of the events is determined rather than the sum of the discrete events' probabilities in the latter case.

The KL divergence score is not symmetrical, as seen in the following example:

$$KL\ (P \parallel Q)! = KL(Q \parallel P) \quad (22)$$

As a result, for image classification, we use an asymmetric version of the Kullback-Leibler divergence, denoted by

$$KL(P \parallel Q) + KL(Q \parallel P)/\ 2 \quad (23)$$

Jensen-Shannon Divergence

Another way to measure the difference (or similarity) between two probability distributions is the Jensen-Shannon divergence or JS divergence.

It calculates a symmetrical normalized score by using the KL divergence. This implies P's divergence from Q is the same as Q's divergence from P, or, to put it another way, P's divergence from Q is the same as Q's divergence from P.

$$JS(P \parallel Q) == JS(Q \parallel P) \tag{24}$$

The following formula can be used to measure the JS divergence:

$$JS(P \parallel Q) = 1/2 * KL(P \parallel M) + 1/2 * KL(Q \parallel M) \tag{25}$$

M is determined as follows:

$$M = 1/2 * (P + Q) \tag{26}$$

And the KL divergence mentioned in the previous section is used to measure KL().

By using the base-2 logarithm, it provides a smoothed and normalized version of KL divergence, with scores ranging from 0 (identical) to 1 (maximally different), making it more useful as a metric.

The Jensen-Shannon distance, or JS distance for short, is calculated by taking the square root of the score

.

2.5.2 Implementation Of Machine learning algorithm

In this section, we will discuss about how machine learning algorithms works by exploring data and identifying patterns, and involves minimal human intervention

Logistic Regression

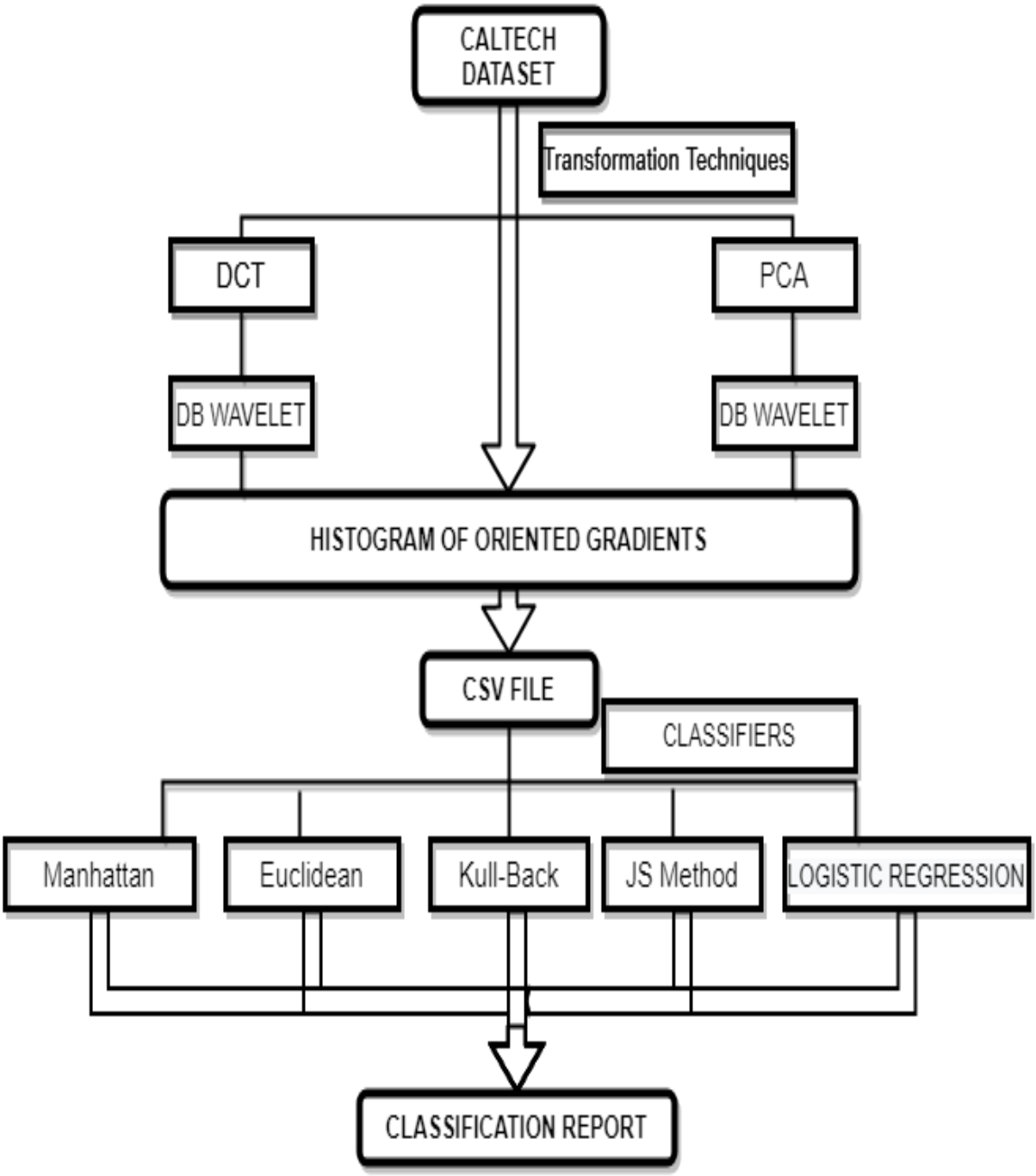
Logistic regression is widely used in machine learning and statistics. In both binary and multiclass grouping, it performs admirably. The dependent variable, in simple terms, is binary with either one (success/yes) or 0 (failure/non) coding of results. A logistic regression model mathematically predicts P(Y=1) as an X-function .

It is a simple ML algorithm that can be used to detect spam, diabetes prediction, cancer detection, etc for several classification problems. In our project, we are using the Caltech 101 and 256 data sets. We're taking 30 images for training and the rest for testing the model. The training images are first transformed using techniques such as DCT (Discrete Cosine Transformation, PCA (Principal Component Analysis), wavelets, and so on. These Transformation techniques aid in noise removal, dimensionality reduction, smoothening, and other tasks. The HOG feature descriptor receives the processed image.

It returns the HOG feature parameters of the trained images, which are saved in a file for later use. Now we're importing the file, which will be used to build and train a logistic regression model that can take extracted parameter values and determine whether or not an image belongs to that specific class. We must be conscious of the following assumptions in relation to the same. The target variables in binary logistic regression must always be binary, and the desired outcome is represented by factor level 1.The model should be free of multicollinearity , which means the independent variables must be independent

of one another. we must include relevant variables for logistic regression, we can use a broad sample size.

2.6 Flow Chart



CHAPTER 3

3.1 Dataset

A dataset is a collection of data. It consists ideally of all the information neatly normalized and uniformly formatted each has a description. This section focuses on the description of the dataset used and the results obtained from the proposed methodologies.

3.1.1 Caltech 101 Dataset

Caltech-101 is made up of 9144 images divided into 101 different objects (such as faces, watches, ants, pianos, and so on) with the varying pose, color, and lighting intensity, as well as a context category (for a total of 102 categories) with 31 to 800 images per category collected by Fei-Fei et al.[13]. Figure 3. (a) depicts three different types of the dataset Caltech-101 has a high classification score. Demonstrates that the proposed Gabor wavelet model is superior to the most commonly used techniques [26, 27]. & 28] based on related datasets and experimental procedures found in the literature.

3.1.2 Caltech 256 Dataset

Griffin et al. [17] downloaded Google images and manually screened out various object classes to construct a daunting collection of 256 object categories containing 30,607 images in total. Images have a lot of variety in terms of intensity, clutter, object size, position, and pose, and there are at least 80 images in each category. We used the standard experimental procedure described in [17] to evaluate the output of the proposed process, labeling the first 15 & 30 images per category for generating the train feature vector and the remaining images as a test.

Fig 6 : Caltech 101 Dataset

Fig 7: Caltech 256 Dataset

3.2 Results and Discussion

This section discusses the results and analysis performed on the Caltech-101 and Caltech-256 datasets. Caltech-256 is an extension of Caltech-101 and consists of images from 256 object categories. It has images ranging from 80 to 827 in each category, for a total of 30608. The database has more variability between classes and within classes than Caltech-101. Figures 6 and 7 depict a few object categories from two datasets with the highest and lowest classification, respectively.

Table 1:Analysis of Distance Classifier

Distance metrics	Caltech 101 Dataset 15 Train	Caltech 101 Dataset 30 Train	Caltech 256 Dataset 15 Train	Caltech 256 Dataset 30 Train
HOG	35.1316	43.3317	11.1489	15.1768
DCT+HOG	40.5596	46.9413	13.6233	15.2541
Wavelet+HOG	39.6711	46.1189	12.8623	15.9632
DCT+Wavelet+HOG	48.3920	53.0854	16.6244	19.0058
PCA+Wavelet+HOG	44.0699	48.0862	14.0952	17.2614

This section discusses the results and analysis performed on the Caltech-101 and Caltech-256 datasets. Caltech-256 is an extension of Caltech-101 and consists of images from 256 object categories. It has images ranging from 80 to 827 in each category, for a total of 30608. The database has more variability between classes and within classes than Caltech-101. Figures 6 and 7 depict a few object categories from two datasets with the highest and lowest classification, respectively

Table 2 : Analysis of Machine Learning Algorithm

Distance metrics	Caltech 101 Dataset 15 Train	Caltech 101 Dataset 30 Train	Caltech 256 Dataset 15 Train	Caltech 256 Dataset 30 Train
Euclidean Distance	35.2246	38.1712	12.3379	14.5542
Manhattan Distance	40.5264	44.2474	17.5962	20.5483
Kullback-Leibler Method	23.8639	25.5652	-	-
Jensen-Shannon Divergence Method	28.1980	32.6593	-	-

The dataset was split into train and test images using random 15 and 30 images per category, for a total of 1515 and 3030 images as training and testing, respectively. In the Caltech-101 and Caltech-256 datasets, the categories with the highest accuracy in terms of color and shape information are harmonium, currency, cartoon character and bus, piano, earth, and so on. As a result, the performance of the HOG descriptor in conjunction with the given classification algorithms outperforms other descriptor techniques, and the overall performance of the descriptor is improved.

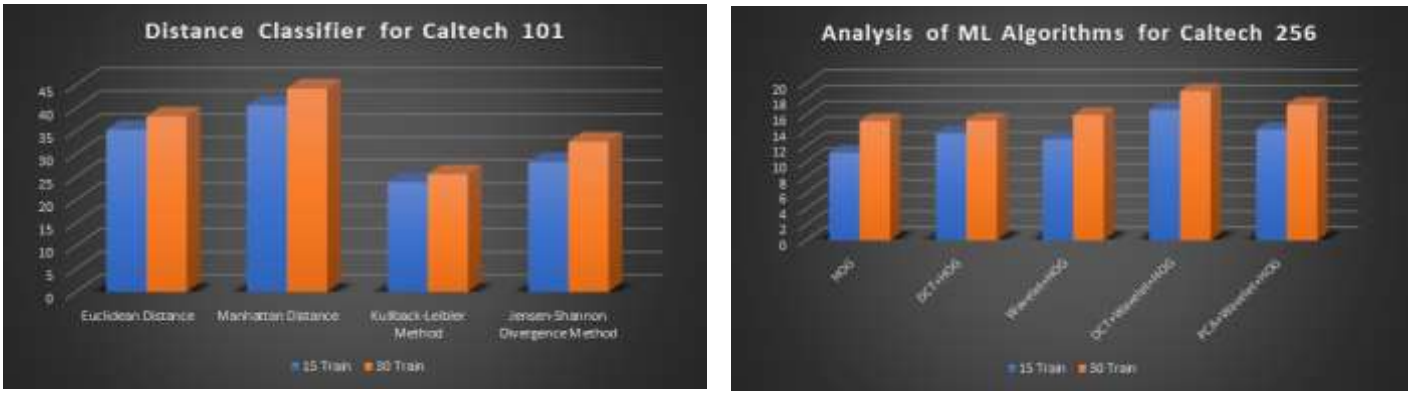


Fig 8: Analysis of Distance Classifier on Caltech 101 & Caltech

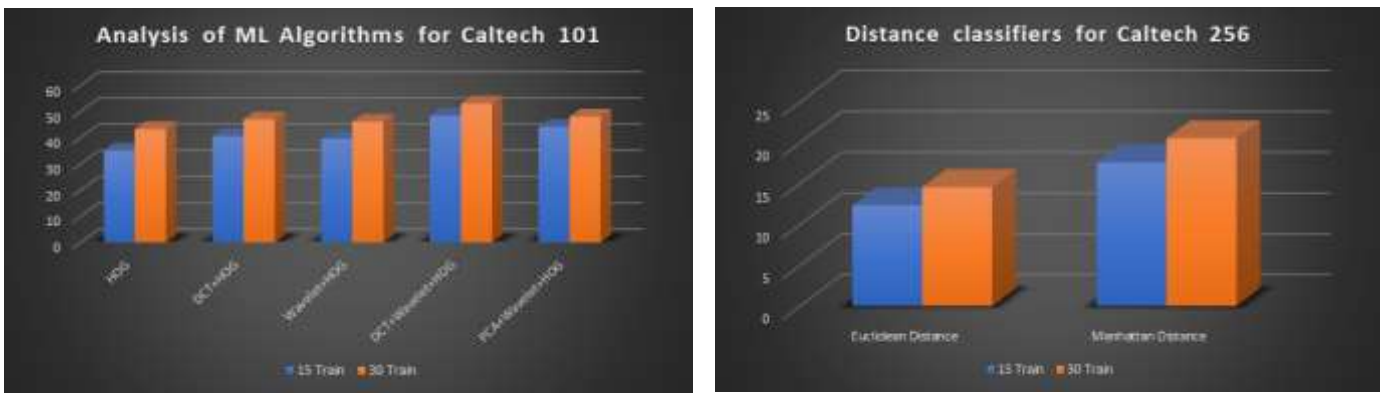


Fig 9 Analysis of Machine Learning Algorithm on Caltech 101 & Caltech 256

Table 1 and Table 2 exhibits the performance analysis of the proposed model over few state of art techniques mentioned in the related works. From the table it is noticed that the HOG mixture model along with distance measure (DM) and machine learning algorithms (ML) as a classifier outperforms the mixture models. In general performance of mixture models is improved by applying the transformation techniques , especially for mixture models with HOG proved to be very efficient and progressive when compared to the mixture models with DM and ML as classifiers. Fig.8 and Fig 9 exhibits the performance evaluation of proposed methods across recent dictionary learning methods on manifolds. The experiments were done as considered in [24],for Caltech-10/1 dataset, 15 and 30 images per category are considered for training with 20 and 30 images per category for testing respectively.

In Caltech-256 dataset 15 and 30 images per category are considered for training with 15 images per category as testing. The proposed method results in an impressive recognition rate just by using a single visual descriptor when compared to [21],[22], [24], [25] where grids of HOG features are found to be highly primitive and significant in outperforming the manifold techniques.

CHAPTER 4

4.1 CASE STUDY 1: COVID-19 Prediction using Chest X-rays Images

4.1.1 Introduction

The proposed system considered input of the X-ray images to identify COVID-19. First of all, this system converted images from RGB to grayscale and identified the region of interest (ROI) by removing the unwanted regions. Furthermore, the system considered two feature extractors: histogram-oriented gradient (HOG) and CNN. First, the HOG technique was used to extract a feature vector from the X-ray COVID-19 dataset. Then the CNN method was used to extract another feature vector from the same images. These two features were fused and used as the input to train the classification model.

The number of features extracted by one technique was not large enough to accurately identify COVID-19. However, the fusion approach of extracting features by two different techniques could provide a large number of features for accurate identification. Fusion was considered as a concatenation between the two individual vectors in this context. Speckle-affected and low-quality X-ray images along with

good quality images were used in our experiment for conducting tests. If training and testing are performed with only selected good quality X-ray images in an ideal situation, the output accuracy may be found higher. However, this does not represent a real-life scenario, where the image database would be a mix of both good- and poor-quality images. Therefore, this approach of using different quality images would test how well the system can react to such real-life situations.

A modified anisotropic diffusion filtering technique was employed to remove multiplicative speckle noise from the test images. The application of these techniques could effectively overcome the limitations in input image quality. Next, the feature extraction was carried out on the test images. Finally, the CNN classifier performed a classification of X-ray images to identify whether it was COVID-19 or not.

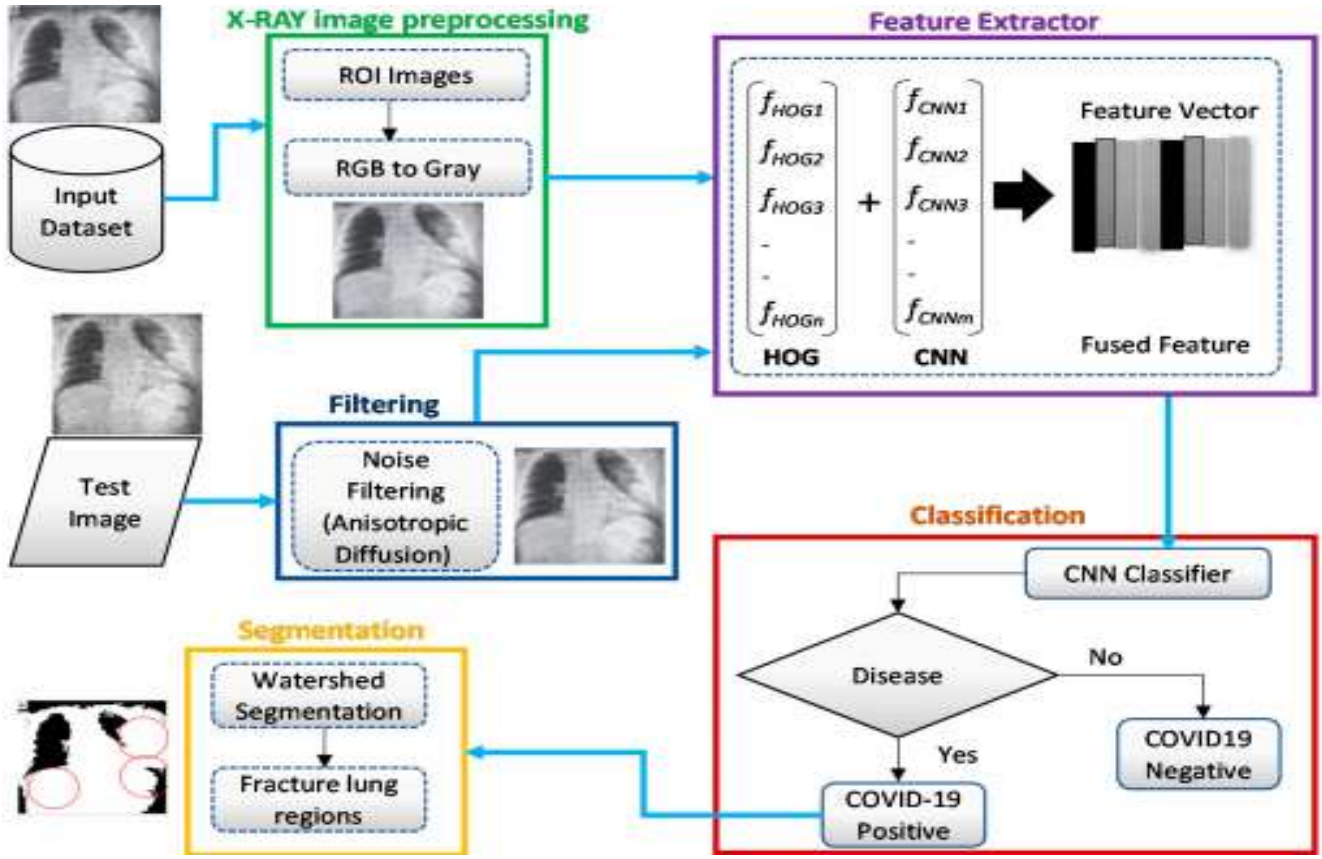


Fig10: Overview of the proposed intelligent system architecture for identifying COVID-19 from chest X-ray images

4.1.2 Dataset and Algorithmic Description

A team of researchers from Qatar University, Doha, Qatar, and the University of Dhaka, Bangladesh along with their collaborators from Malaysia in collaboration with medical doctors have created a database of chest X-ray images for COVID-19 positive cases along with Normal and Viral Pneumonia images.

Algorithm Proposed Algorithm for COVID-19 Detection

Input: COVID-19 Chest X-ray image dataset (D) with resize image (M)
Extraction: Extract Feature Matrix (f).
CNN Feature Vector (Fc).
Step 1: Initialize $F_c \geq M_i.i = 1$
Step 2: Extract each image feature $\mathbf{D}(i,1,570)$.
Step 3: $F_c\ (i,1) = \mathbf{M}(x,1) + F_c\ (i,1)$.
Step 4: F_c = overall CNN features.
Histogram Oriented Gradient (HOG).
Step 1: Initialize. $H_0 = Low\ pass\ output, H_1 = Band\ pass\ output$.
Step 2: HOG $(i,1) = H_0\ (i,1) + H_1\ (i,1)$.
Step 3: HOG = overall Histogram Oriented Gradient
Fusion of features in Vector (V).
$Training\ feature\ (V) = [F_c\ , HOG]$.
$test\_image = imread(img)$.
Extract test feature (T) = repeat step 1, 2 from test_image.
$result\ (i) = classify\ (training\ feature, T)$.
Output: $result\ (i) =$ **COVID19 Positive or Normal**.

This COVID-19, normal and other lung infection dataset is released in stages. In the first release it was 219 COVID-19, 1341 normal and 1345 viral pneumonia chest X-ray (CXR) images. In the 2nd update, the database increased to 3616 COVID-19 positive cases along with 10,192 Normal, 6012 Lung Opacity (Non-COVID lung infection) and 1345 Viral Pneumonia images. We used two classes form this dataset to train our proposed model. Basically, we intended to perform binary classification between the class positive and negative.

4.1.3 Comparison of Chest Xray Images

This Figure illustrates a comparison of the COVID-19 and normal chest X-ray images. In general, similar to pneumonia, the density of the lungs is increased in the case of the COVID-19, which causes whiteness in the lungs on radiography. An experienced radiologist can confirm the disease by the appearance of a ground-glass pattern (ground glass opacity) due to the increased whiteness

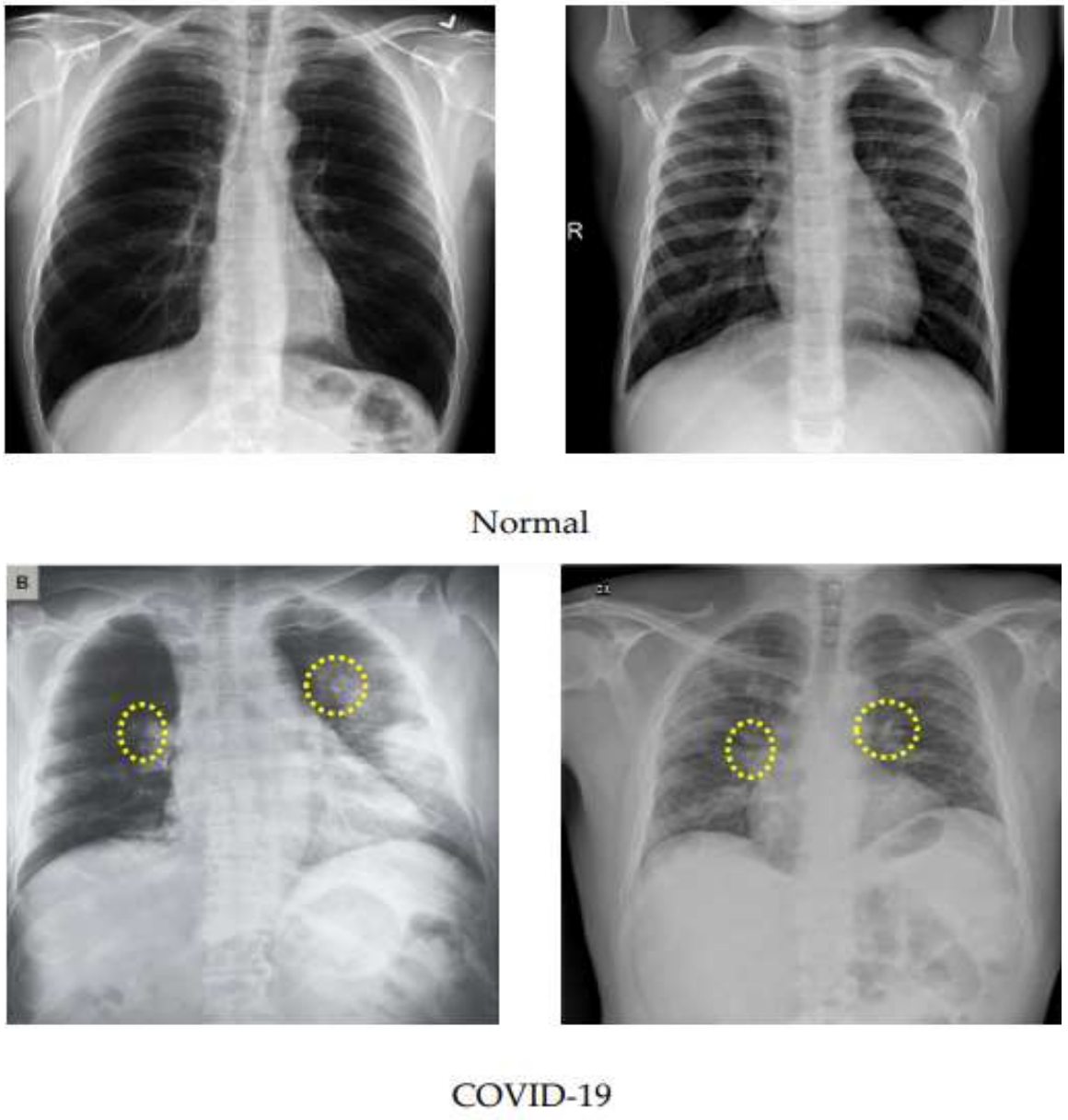


Fig11: Comparison of the COVID-19 and normal X-ray images.

4.1.4 Methodology and Implementation

Data fusion was applied in several machine learning and computer vision applications [. Particularly, feature fusion can combine more than one feature vector. Two feature extractors provide a feature vector of 1 × 4096 and 1 × 3780. The feature selection process was mathematically explained by Equations (27) – (31) [54]. Equations (27) and (28) represent features extracted by HOG and CNN, respectively. The extracted feature vectors are combined by concatenation and represented by Equation (29).

$$f_{HOG\ 1\times n} = \{ HOG_{1\times 1}, HOG_{1\times 2}, HOG_{1\times 3} - - - - - - - - - HOG_{1\times n}\} \quad (27)$$

$$f_{VGG19\ 1\times m} = \{ VGG19_{1\times 1}, VGG19_{1\times 2}, VGG19_{1\times 3} - - - - - - VGG19_{1\times m}\} \quad (28)$$

$$Fused\ (features\ vector)_{1\times q}^{cat} = \{f_{HOG\ 1\times n}, f_{VGG19\ 1\times m}\} \quad (29)$$

Then the features extracted by HOG and CNN are fused with 7876 features. 1186 score-based features were selected out of 7876 features based on maximum entropy. When the value of i = 1, it recalls HOG features and when i = 2, it recalls VGG19 features and finally adds them together. For the purpose of selecting optimal features, entropy was employed considering score values. The probability of features and entropy is defined by the following Equations. The final selected features were fed to the classifiers in order to identify COVID-19 images.

$$B_{He} = -NHe_b \sum_{i=1}^{n} p(f_i) \quad (30)$$

$$F_{select} = B_{He}(\max(f_i, 1186)) \quad (31)$$

where f is fused vector (1 × 1186), p denotes features probability, and He represents entropy. This entropy-based feature selection method selects similar or highly related features from the fused vector. Unrelated features were removed, and appropriate features were considered for classification. In this proposed system, the appropriate feature vector is 1 × 1186. CNN classifier uses the selected features where the

process of feature extraction and selection is shown in Figure 8. Furthermore, the feature vectors obtained by HOG and deep learning were fused to validate the proposed approach in this work.

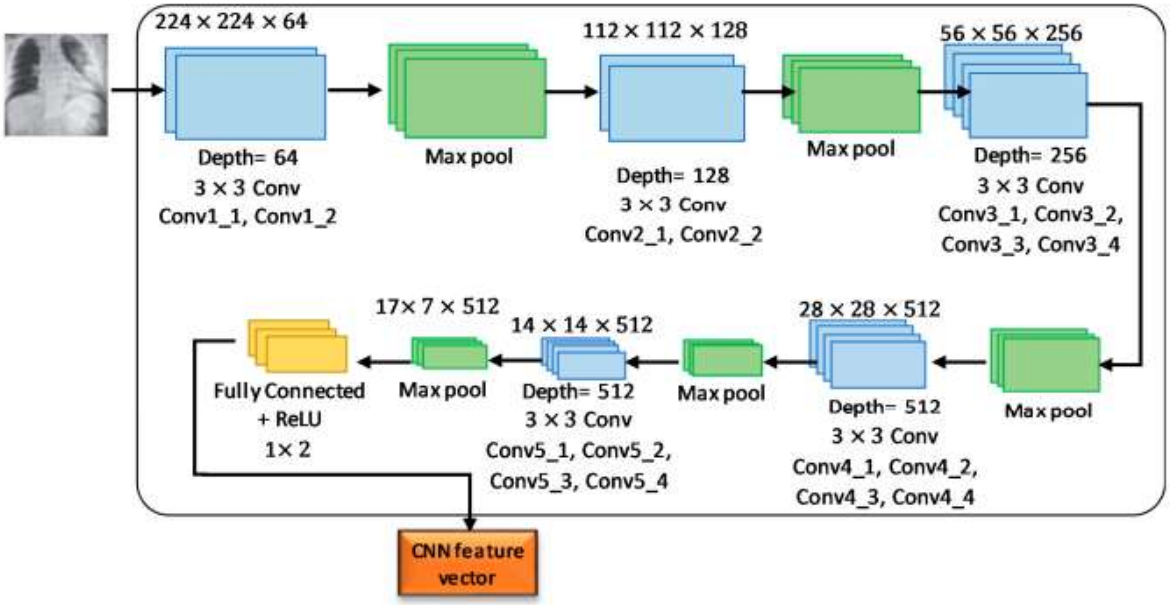


Fig12 Illustration of image feature extraction by VGG19 pre-trained convolutional neural network (CNN) model

The VGG19 architecture consisted of 16 layers of CNN, 3 connected layers with 1 final output layer for conducting SoftMax function. No changes in the number of connected layers and final layer are required to build the network architectures. In addition, 2 × 2-pixel windows with stride 2 were used for the max-pooling layer. The first two layers and third layer of the three fully connected layers provided 4096 features and 1000 channels, respectively. The final layer represents the output layer with two neurons (COVID-19 and normal).

4.1.5 Results & Classification Performance

This work proposed a fusion of feature vectors obtained by a combination of HOG and CNN techniques. This fusion vector was deliberated as the final input for the training and test datasets. Figure presents a comparative study of different feature extraction approaches. The performances of different individual feature extraction techniques were less satisfactory than the fusion approach. This demonstrated that the proposed approach could classify COVID-19 cases more accurately than the single feature extraction approaches.

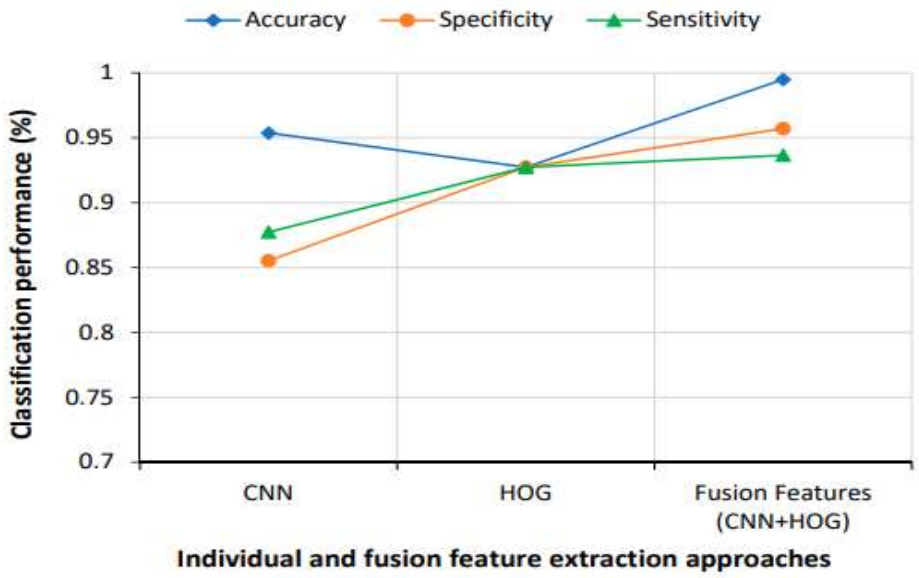


Fig13: Comparative results of individual and fusion features.

CNN classifier was trained by 5-fold cross-validation after feature extraction where feature vector was randomly distributed into 5 sub folds. Four sub folds were selected for the training dataset, and one single sub fold was selected for the testing dataset. This Table presents the different CNN classification results using individual and combined feature extraction methods. Single feature extraction methods provided less accuracy compared to the proposed fusion vector.

Feature Extraction Methods	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Mean Accuracy
HOG	0.8732	0.8789	0.8741	0.8675	0.8730	0.8734
CNN	0.9378	0.9367	0.9387	0.9367	0.9321	0.9364
Proposed fusion (HOG+CNN)	0.9856	0.9847	0.9813	0.9827	0.9833	0.9836

Table 3: Overall classification accuracy measured using 5-fold cross-validation

One of the limitations of this work was the imbalance of data in the datasets used for training and testing. In general, balanced data set with an equal number of normal and COVID-19 X-ray images makes the model building more comfortable, and the developed model can provide better prediction accuracy. Furthermore, the classification algorithm finds it easier to learn from a balanced dataset. Naturally, in any opensource database, the number of normal images would be higher than the COVID-19-positive images. As the images used in this study were taken from open-source databases, the imbalance in the training and testing data sets was obvious. However, the ratio between the number of normal and COVID-19 images was maintained at 1.57 in both the training and testing data sets in order to alleviate the data imbalance problem to some extent.

4.2 CASE STUDY 2: Real Time Object Detection from Multiple Feed

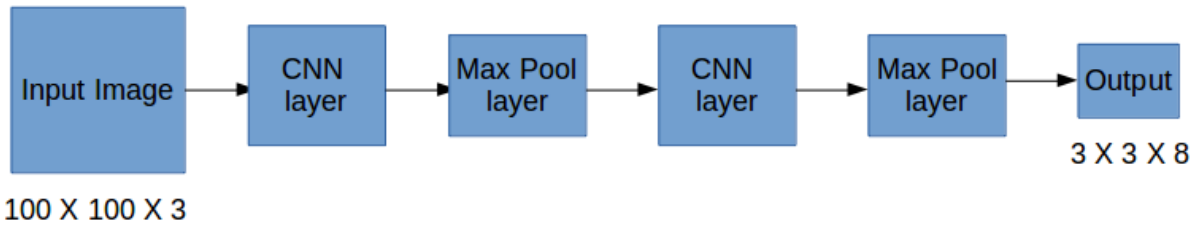
4.2.1 Introduction

Object detection is a technology that detects the semantic objects of a class in digital images and videos. One of its real-time applications is self-driving cars. In this, our task is to detect multiple objects from an image. The most common object to detect in this application is the car, motorcycle, and pedestrian. For locating the objects in the image, we use Object Localization and have to locate more than one object in real-time systems.

There are various techniques for object detection, they can be split up into two categories, first is the algorithms based on Classifications. CNN and RNN come under this category. In this, we have to select the interested regions from the image and have to classify them using Convolutional Neural Network. This method is very slow because we have to run a prediction for every selected region. The second category is the algorithms based on Regressions. YOLO method comes under this category. In this, we won't select the interested regions from the image. Instead, we predict the classes and bounding boxes of the whole image at a single run of the algorithm and detect multiple objects using a single neural network. YOLO algorithm is fast as compared to other classification algorithms. YOLO algorithm makes localization errors but predicts less false positives in the background.

4.2.2 Implementation of YOLO Algorithm

First, an image is taken and YOLO algorithm is applied. In our example, the image is divided as grids of 3x3 matrixes. We can divide the image into any number grids, depending on the complexity of the image. Once the image is divided, each grid undergoes classification and localization of the object. The confidence score of each grid is found. If there is no proper object found in the grid, then the confidence score and bounding box value of the grid will be zero or if there found an object in the grid then the confidence score will be 1 and the bounding box value will be its corresponding bounding values of the found object.



4.2.3 Bounding Box Predictions

YOLO algorithm is used for predicting the accurate bounding boxes from the image. The image divides into S x S grids by predicting the bounding boxes for each grid and class probabilities. Both image classification and object localization techniques are applied for each grid of the image and each grid is assigned with a label. Then the algorithm checks each grid separately and marks the label which has an object in it and also marks its bounding boxes. The labels of the gird without object are marked as zero.

Fig14: Example image with 3x3 Grids

Consider the above example, an image is taken and it is divided in the form of 3x3 matrixes. Each grid is labelled and each grid undergoes both image classification and objects localization techniques. The label is considered as Y. Y consists of 8 values.

$$y = \begin{bmatrix} pc \\ bx \\ by \\ bh \\ bw \\ c1 \\ c2 \\ c3 \end{bmatrix}$$

Pc – Represents whether an object is present in the grid or not. If present pc=1 else 0.

bx, by, bh, bw – are the bounding boxes of the objects (if present).

c1, c2, c3 – are the classes. If the object is a car, then c1 and c3 will be 0 and c2 will be 1.

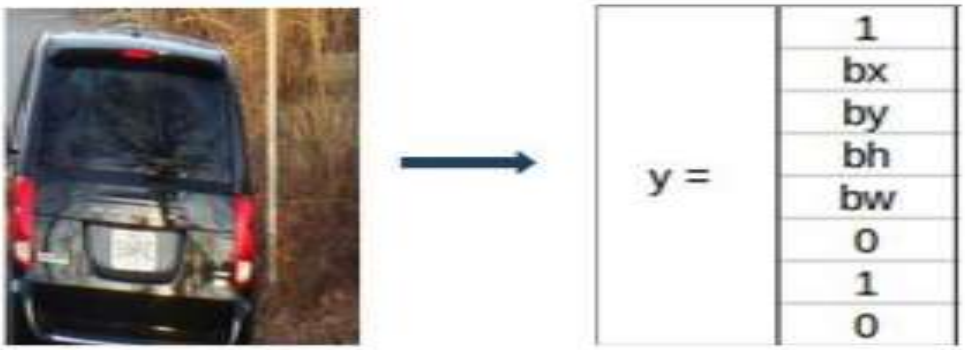


Fig15: Example from base image for Bounding Box Calculations

4.2.4 Class Prediction

Each box predicts the classes the bounding box may contain using multilabel classification. We do not use a softmax as we have found it is unnecessary for good performance, instead we simply use independent logistic classifiers. During training we use binary cross-entropy loss for the class predictions.

This formulation helps when we move to more complex domains like the Open Images Dataset. In this dataset there are many overlapping. Using a softmax imposes the assumption that each box has exactly one class which is often not the case. A multilabel approach better models the data.

YOLOv3 predicts boxes at 3 different scales. Our system extracts feature from those scales using a similar concept to feature pyramid networks. From our base feature extractor, we add several convolutional layers. The last of these predicts a 3-d tensor encoding bounding box, confidence score, and class predictions. In our experiments with COCO [10] we predict 3 boxes at each scale so the tensor is $N \times N \times [3 * (4 + 1 + 80)]$ for the 4 bounding box offsets, 1 confidence score prediction, and 80 class predictions.

Next, we take the feature map from 2 layers previous and up sample it by 2×. We also take a feature map from earlier in the network and merge it with our up sampled features using concatenation. This method allows us to get more meaningful semantic information from the up sampled features and finer-grained information from the earlier feature map. We then add a few more convolutional layers to process this combined feature map, and eventually predict a similar tensor, although now twice the size. We perform the same design one more time to predict boxes for the final scale. Thus, our predictions for the 3rd scale benefit from all the prior computation as well as fine-grained features from early on in the network.

4.2.5 Network Design

We implement this model as a convolutional neural network and evaluate it on the COCO detection dataset. The initial convolutional layers of the network extract features from the image while the fully connected layers predict the output probabilities and coordinates. Our network architecture is inspired by the GoogLeNet model for image classification. Our network has 24 convolutional layers followed by 2 fully connected layers. Instead of the inception modules used by GoogLeNet, we simply use 1 × 1 reduction layers followed by 3 × 3 convolutional layers. The full network is shown in Figure. We also train a fast version of YOLO designed to push the boundaries of fast object detection. Fast YOLO uses a neural network with fewer convolutional layers (9 instead of 24) and fewer filters in those layers. Other than the size of the network, all training and testing parameters are the same between YOLO and Fast YOLO.

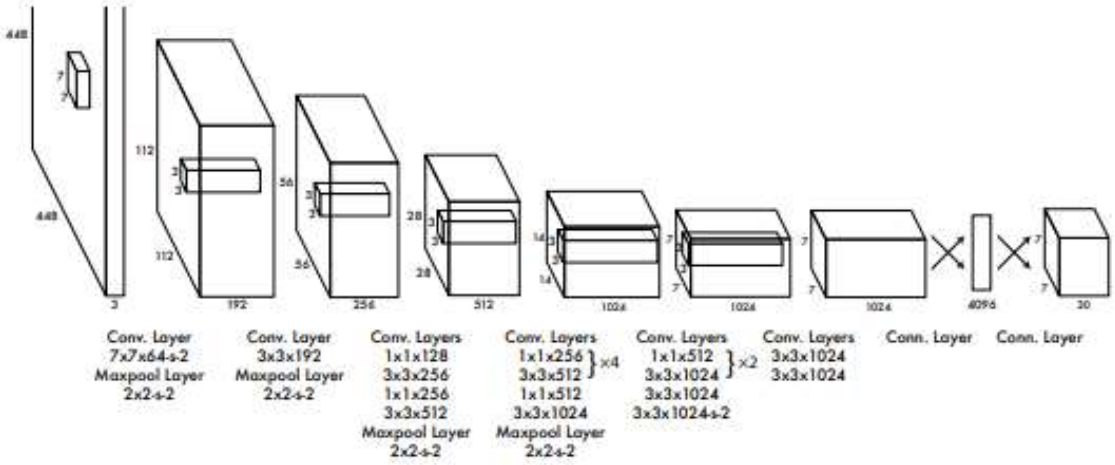


Fig16: Architecture of Yolov3

The Architecture. Our detection network has 24 convolutional layers followed by 2 fully connected layers. Alternating 1 × 1 convolutional layers reduce the features space from preceding layers. We pretrain the convolutional layers on the ImageNet classification task at half the resolution (224 × 224 input image) and then double the resolution for detection.

4.2.6 Limitations of YOLO

YOLO imposes strong spatial constraints on bounding box predictions since each grid cell only predicts two boxes and can only have one class. This spatial constraint limits the number of nearby objects that our model can predict. Our model struggles with small objects that appear in groups.

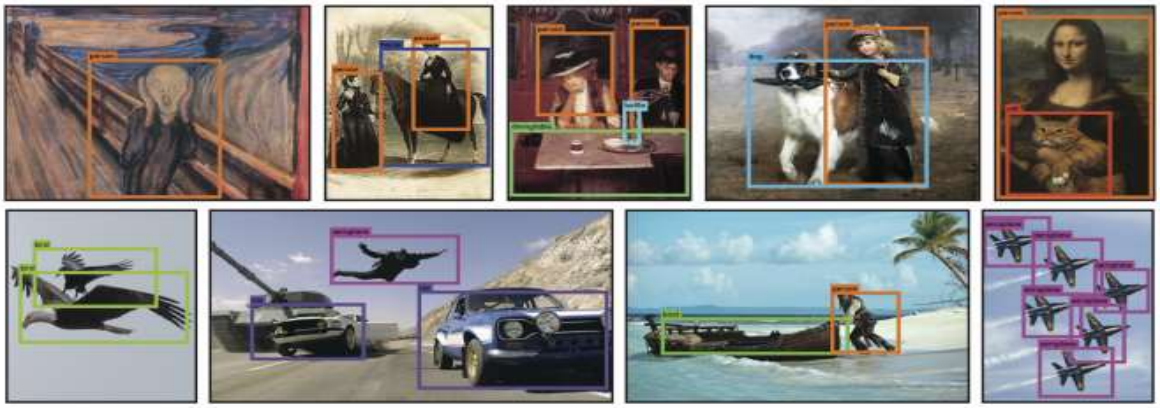


Fig17: Yolo Implementation

Since our model learns to predict bounding boxes from data, it struggles to generalize to objects in new or unusual aspect ratios or configurations. Our model also uses

relatively coarse features for predicting bounding boxes since our architecture has multiple down sampling layers from the input image.

Finally, while we train on a loss function that approximates detection performance, our loss function treats errors the same in small bounding boxes versus large bounding boxes. A small error in a large box is generally benign but a small error in a small box has a much greater effect on IOU. Our main source of error is incorrect localizations.

Live detection of our model

Fig18: Live Results

CHAPTER 5

5.1 Frontend Application Building

5.1.1 Hardware and Software Requirement

- Framework: Django
- IDE: Pycharm, Anaconda (Jupyter notebook)
- Client Side Technologies: HTML, CSS, JavaScript, Bootstrap
- Server Side Technologies: Python
- Data Base Server: SQLite
- Operating System: Microsoft Windows

Image classifications

- Predicts the class of an image along with class label.
- Displays and saves it to the database.
- This application has a great impact in classifying images in large datasets.

Radiography Detection

- Predicts the Covid-19 Positive or Negative from the input Xray image.
- Generates the report with infection rate and stores it to the Database.
- This application will help the user to get the report early compared to RTPCR test.

Real time object Detection

- Accepts feed from multiple video input devices.
- User can monitor from a single base station.

5.1.2 Implementation

Python

Python is a widely used general-purpose, high level programming language. It was initially designed by Guido van Rossum in 1991 and developed by Python Software Foundation. It was mainly developed for emphasis on code readability, and its syntax allows programmers to express concepts in fewer lines of code.

Python is a programming language that lets you work quickly and integrate systems more efficiently.

Python is dynamically typed and garbage-collected. It supports multiple programming paradigms, including procedural, object-oriented, and functional programming. Python is often described as "batteries included" language due to its comprehensive standard library.

HTML

HTML (Hypertext Markup Language) is the set of markup symbols or codes inserted in a file intended for display on a World Wide Web browser page. The markup tells the Web browser how to display a Web page's words and images for the user. Each individual markup code is referred to as an element (but many people also refer to it as a tag). Some elements come in pairs that indicate when some display effect is to begin and when it is to end.

CASCADING STYLE SHEET (CSS)

Cascading Style Sheets (CSS) are a collection of rules we use to define and modify web pages. CSS are similar to styles in Word. CSS allow Web designers to have much more control over their pages look and layout. For instance, you could create a style that defines the body text to be Verdana, 10 points. Later on, you may easily change the body text t Times New Roman, 12 points by just changing the rule in the CSS. Instead of having to change the font on each page of your website, all you need to do is redefine the style on the style sheet, and it will instantly change on all of the pages that the style sheet has been applied to. With HTML styles, the font change would be applied to each instance of that font and have to be changed in each spot.

CSS can control the placement of text and objects on your pages as well as the look of those objects.

HTML information creates the objects (or gives objects meaning), but styles describe how the objects should appear. The HTML gives your page structure, while the CSS creates the "presentation". An external CSS is really just a text file with a .css extension. These files be created with Dreamweaver, a CSS editor, or even Notepad.

JavaScript

JavaScript is a programming language commonly used in web development. It was originally developed by Netscape as a means to add dynamic and interactive elements to websites. While JavaScript is influenced by Java, the syntax is more similar to C and is based on ECMAScript, a scripting language developed by Sun Microsystems.

JavaScript is a client-side scripting language, which means the source code is processed by the client's web browser rather than on the web server. This means JavaScript functions can run after a webpage has loaded without COMMUNICATING with the server. For example, a JavaScript function may check a web form before it is submitted to make sure all the required fields have been filled out. The JavaScript code can produce an error message before any information is actually transmitted to the server.

Like server-side scripting languages, such as PHP and ASP, JavaScript code can be inserted anywhere within the HTML of a webpage. However, only the output of server-side code is displayed in the HTML, while JavaScript code remains fully visible in the source of the webpage. It can also be referenced in a separate .JS file, which may also be viewed in a browser.

Django

Django is a web application framework written in Python programming language. It is based on MVT (Model View Template) design pattern. The Django is very demanding due to its rapid development feature. It takes less time to build application after collecting client requirement.

Bootstrap

Bootstrap is a free and open-source tool collection for creating responsive websites and web applications. It is the most popular HTML, CSS, and JavaScript framework for developing responsive, mobile-first websites. Nowadays, the websites are perfect for all the browsers (IE, Firefox, and Chrome) and for all sizes of screens (Desktop, Tablets, Phablets, and Phones). All thanks to Bootstrap developers – Mark Otto and Jacob Thornton of Twitter, though it was later declared to be an open-source project.

SQLite

SQLite is a self-contained, high-reliability, embedded, full-featured, public-domain, SQL database engine. It is the most used database engine in the world. It is an in-process library and its code is publicly available. It is free for use for any purpose, commercial or private. It is basically an embedded SQL database engine. Ordinary disk files can be easily read and write by SQLite because it does not have any separate server like SQL. The SQLite database file format is cross-platform so that anyone can easily copy a database between 32-bit and 64-bit systems.

5.1.3 System Design

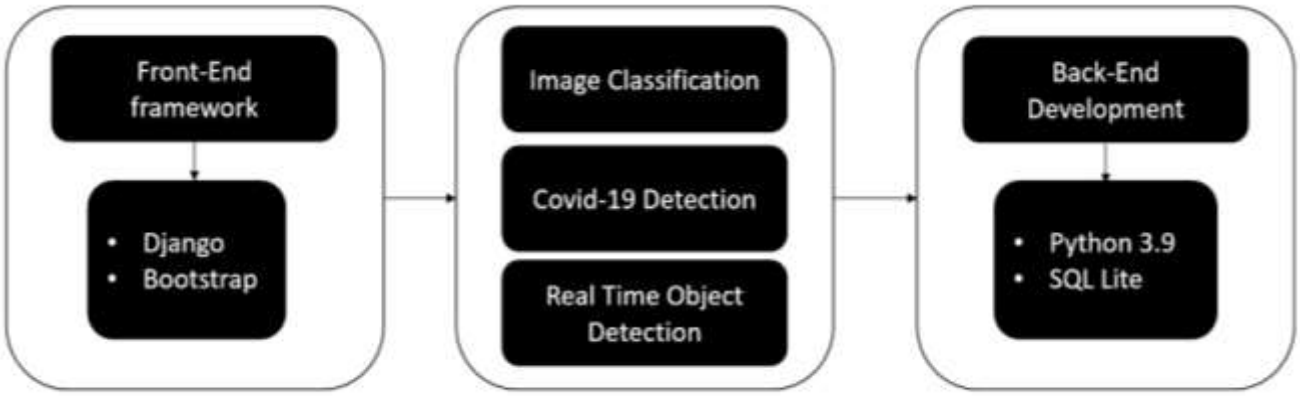
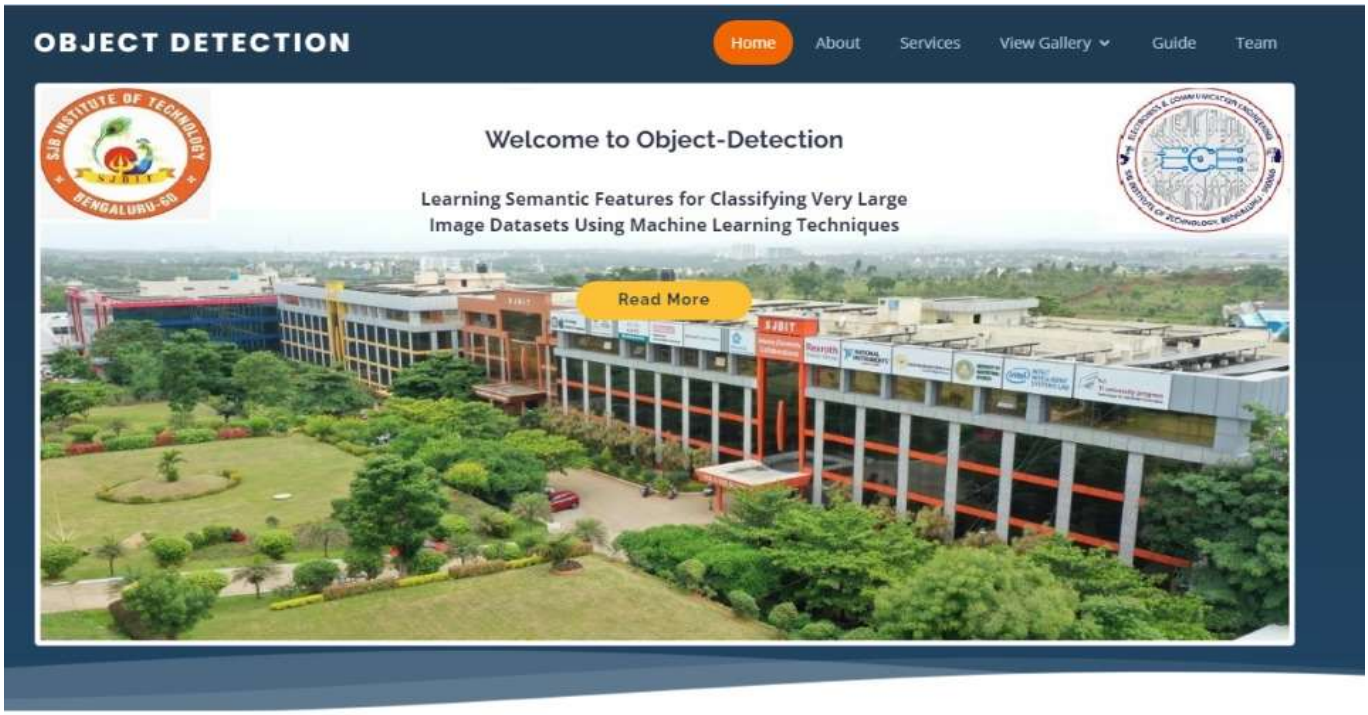


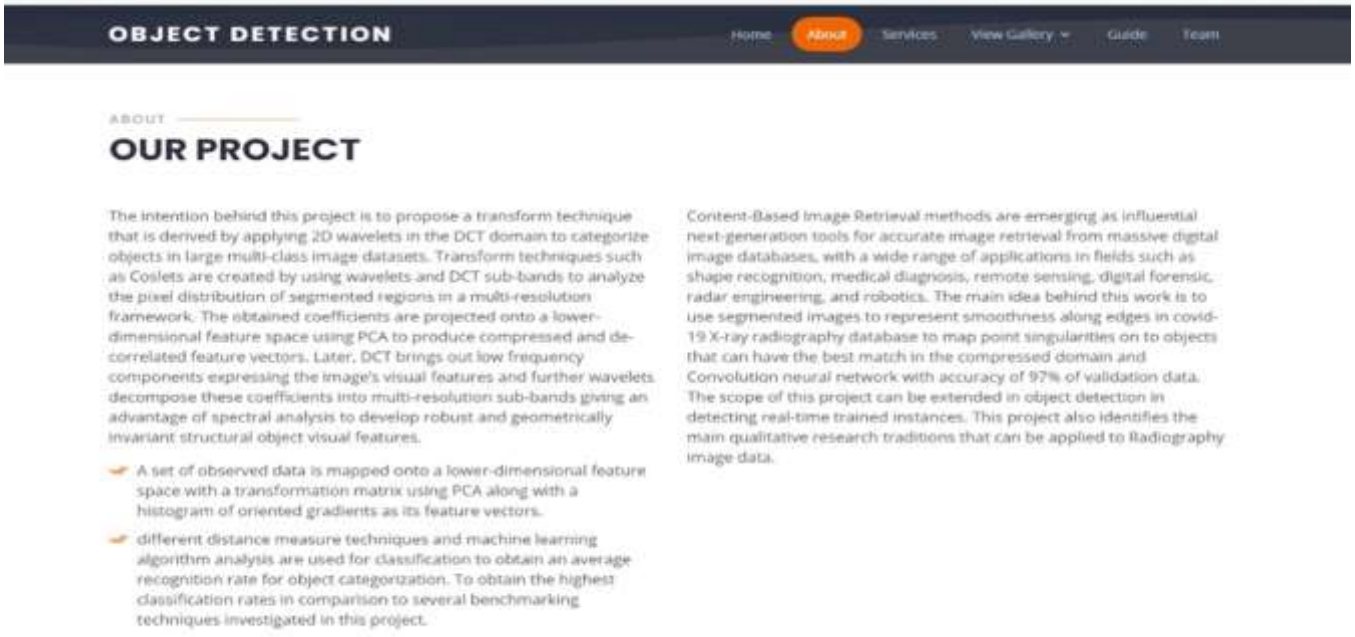
Fig18:System Design

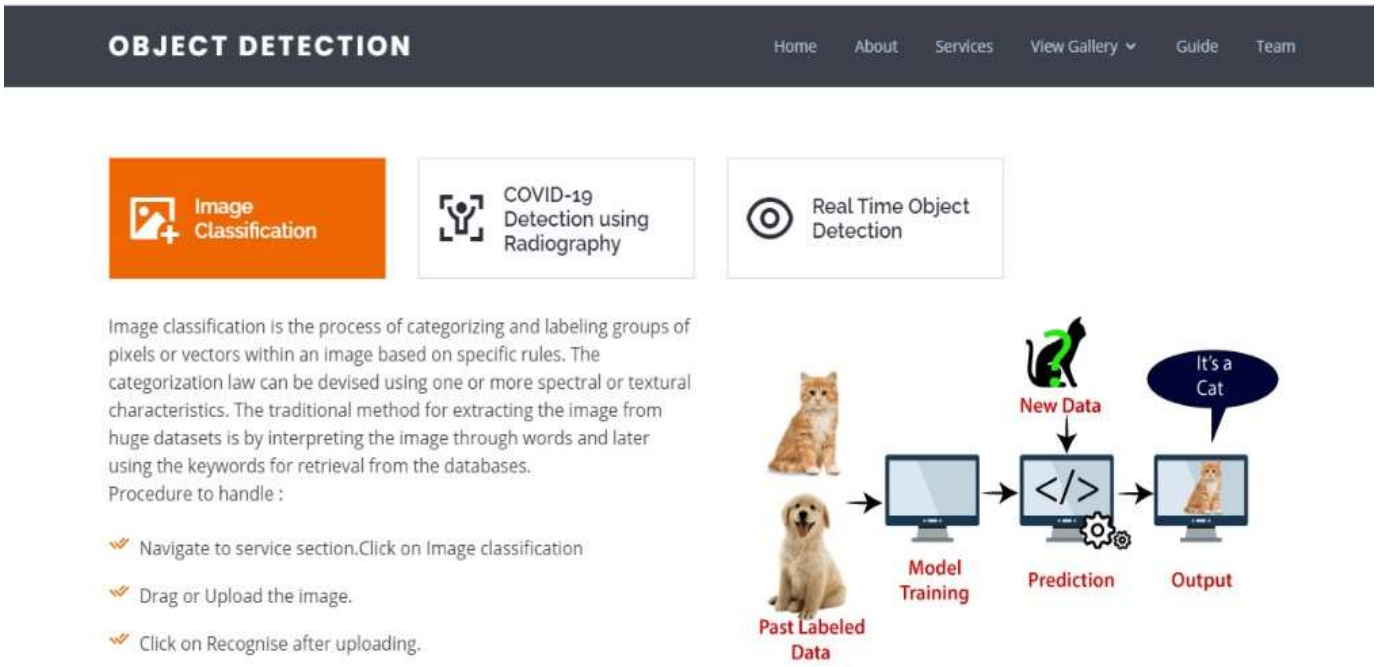
5.1.4 User Screen & Snips from Web Page

Home



About




OBJECT DETECTION
Home
About
Services
View Gallery
Guide
Team
Image Classification
COVID-19 Detection using Radiography
Real Time Object Detection
Image classification is the process of categorizing and labeling groups of pixels or vectors within an image based on specific rules. The categorization law can be devised using one or more spectral or textural characteristics. The traditional method for extracting the image from huge datasets is by interpreting the image through words and later using the keywords for retrieval from the databases.
Procedure to handle :
Navigate to service section.Click on Image classification
Drag or Upload the image.
Click on Recognise after uploading.
You can see the recognised class and the recognition rate.
Images tested are present in the Image DataBase section section.
New Data
It's a Cat
Past Labeled Data
Model Training
Prediction
Output

Services

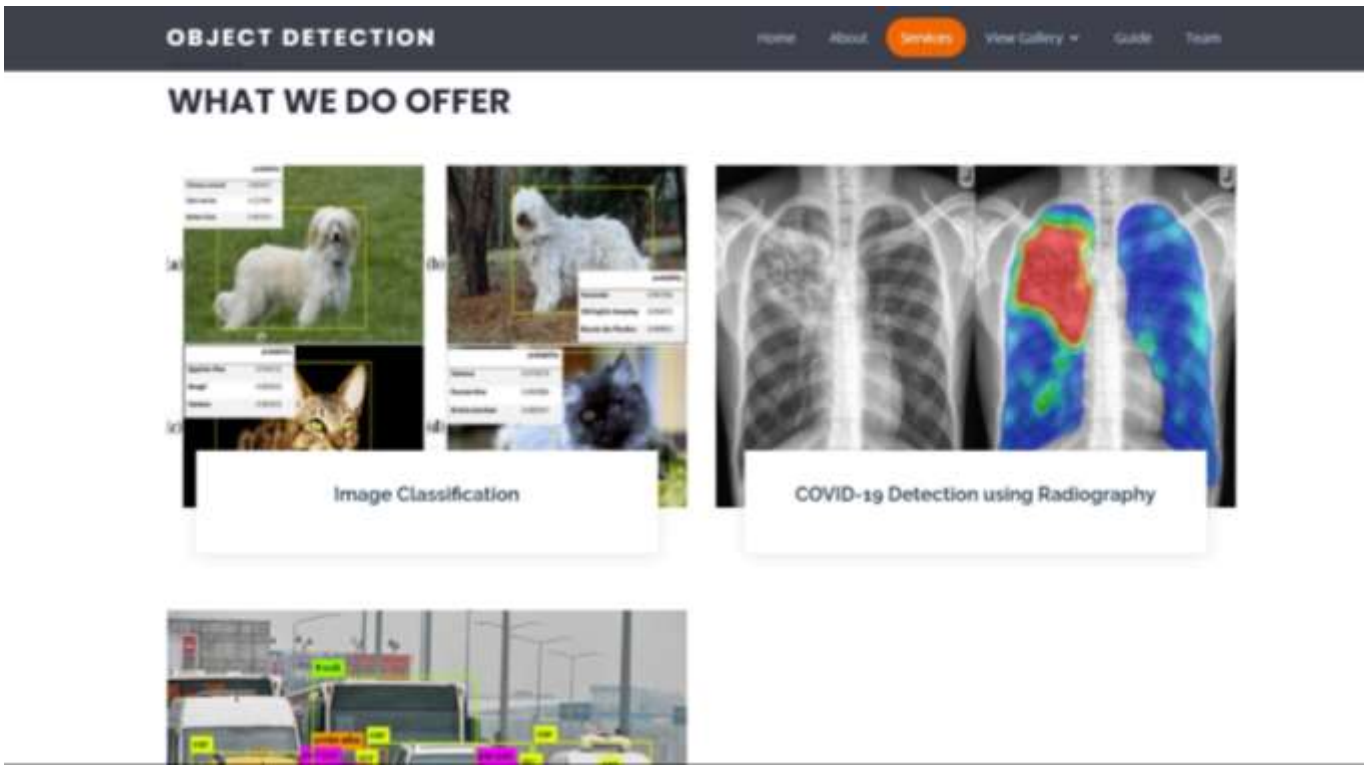
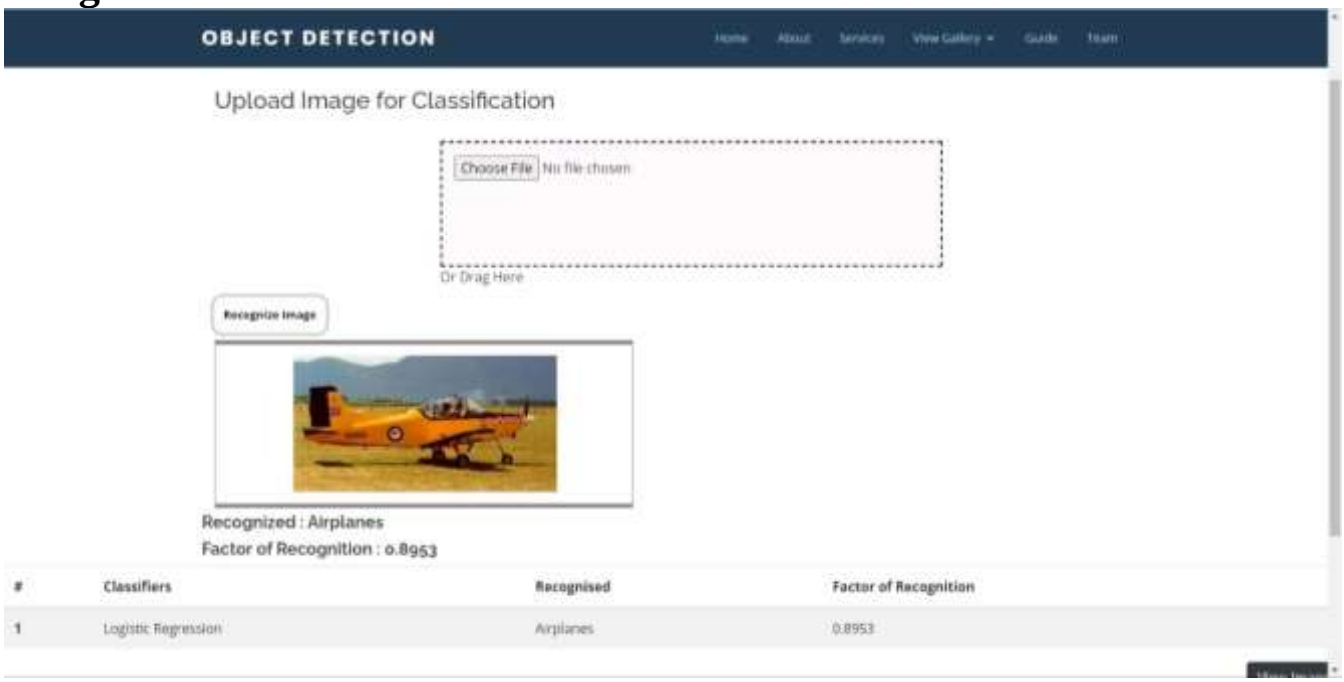
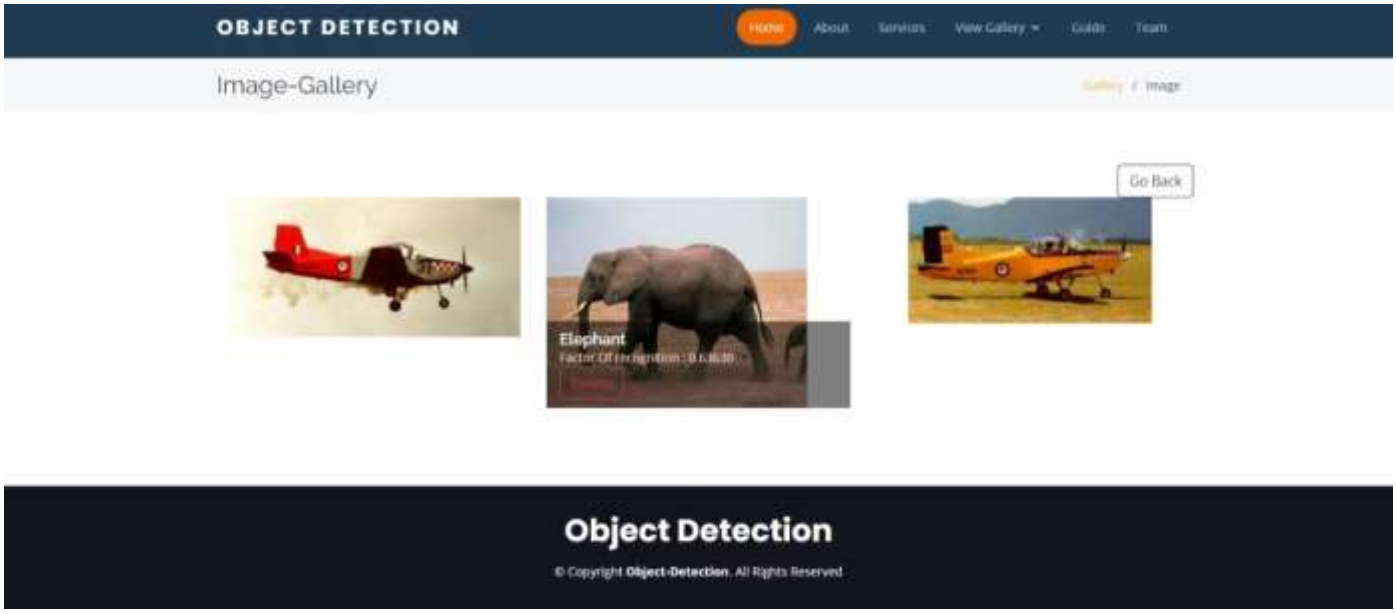

OBJECT DETECTION
WHAT WE DO OFFER
Image Classification
COVID-19 Detection using Radiography

Image classification and covid detection UI's



Gallery


OBJECT DETECTION
Radiography-Gallery
Covid : Positive
Go Back


OBJECT DETECTION
Home
About
Services
View Gallery
Guide
Team
Image-Gallery
Elephant
Go Back
Object Detection
© Copyright Object-Detection. All Rights Reserved

CONCLUSION AND FUTURE WORK

In recent years, most of the literature have highlighted on combining trans-formation techniques for image representation but failed to preserve discriminative features. An ensemble model is proposed which combines DCT, Wavelet and HOG to preserve low frequency co-efficient, point singularities and edge orientations respectively from an image. Proposed ensemble method combined with different distance measure and logistic regression techniques are demonstrated on very large benchmark datasets for classifying images and has obtained competitive results by outperforming recent techniques such as coslet, bag of features, bag of words, spatial pyramid matching techniques. In future, the performance of our model can be examined for different distance measures, support vector machine, neural network architectures and deep learning techniques. The scope of this project lies on real time object detection in cameras.

The case study is designed and developed an intelligent system for the COVID-19 identification with high accuracy and minimum complexity by combining the features extracted by histogram-oriented gradient (HOG) features and convolutional neural network (CNN). Suitable feature selection and classification are absolutely vital in the COVID-19 detection using chest X-ray images. Chest X-ray images were entered into the system in order to produce the output of the marked lung significant region, which was used to identify COVID-19. The proposed feature fusion system showed a higher classification accuracy (99.49%) than the accuracies obtained by using features obtained by individual feature extraction techniques, such as HOG and CNN.

Our model is simple to construct and can be trained directly on full images. Unlike classifier-based approaches, YOLO is trained on a loss function that directly corresponds to detection performance and the entire model is trained jointly. Fast YOLO is the fastest general-purpose object detector in the literature and YOLO pushes the state-of-the-art in real-time object detection. YOLO also

generalizes well to new domains making it ideal for applications that rely on fast, robust object detection.

This idea is deployed in a web application with user friendly user-interface. It is extended to digest multiple feed from video input devices as per user requirement. Where user can monitor the feed from any part.

REFERENCES

[1] K. Mahantesh and Manjunath Aradhya V N,An Impact of Gaussian Mixtures in Image Retrieval System

[2] Jiefa Wu, Sheng Yang, Lingling Zhang, "Detection based on improved HOG feature and Robust adaptive boosting algorithm", Proceedings of 4th International Congress on Image and signal processing.

[3] Shukui Bo, Yongju Jing, "Image Clustering using Mean shift algorithm" Proceedings of Fourth international conference on Computational Intelligence and Communication Networks,(2012).

[4] Meenaakshi N. Munjal, Shaveta Bhatia," A Novel technique for effective image gallery search using content-based image retrieval system", Proceedings of International conference on Machine learning, Big data, Cloud and parallel computing, India, (2019).

[5] Y. Jian, D. Zhang, A. F. Frangi, and Y. Jing-yu, "Two-dimensional PCA: a new approach to appearance-based face representation and recognition," Pattern Analysis and Machine Intelligence, IEEE Transactions on, vol. 26, (2004)

[6] DCT QUANTIZATION NOISE IN COMPRESSED IMAGES, Mark A. Robertson and Robert L. Stevenson, Laboratory for Image and Signal Analysis, University of Notre Dame, Notre Dame, IN 46556.

[7]. N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in Proceedings of the 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05) - Volume 1 - Volume 01, ser. CVPR '05, (2005).

[8]. Q. Zhu, M.-C. Yeh, K.-T. Cheng, and S. Avidan, "Fast human detection using a cascade of histograms of oriented gradients," in Proceedings of the 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition - Volume 2, ser. CVPR '06, (2006)

[9]. K. Mahantesh, V.N. Manjunath Aradhya,"An Exploration of Ridgelet Transform to handle Higher Dimensional Intermittency for Object Categorization in Large Image Datasets", Proceedings of "The 2nd International Conference on Applied Information and Communications Technology" - ICAICT (2014).

[10]. O. Deniz, G. Bueno, J. Salido, and F. De la Torre, "Face recognition using histograms of oriented gradients," Pattern Recogn. Lett., vol. 32, no. 12, pp. 1598–1603, Sep. (2011).

[11]. R. Hu and J. Collomosse, "A performance evaluation of gradient field hog descriptor for sketch based image retrieval," Comput. Vis. Image Underst., vol. 117, no. 7, pp. 790–806, Jul (2013).

[12]. D. G. Lowe, "Object recognition from local scale-invariant features," in Proceedings of the International Conference on Computer Vision-Volume 2 - Volume 2, ser. ICCV '99, (1999), pp. 1150–1157.
[13]. S. Thumfart, R. H. A. H. Jacobs, E. Lughofer, C. Eitzinger, F. W. Cornelissen, W. Groissboeck, and R. Richter, "Modeling human aesthetic perception of visual textures," ACM Trans. Appl. Percept., vol. 8, no. 4, pp. 27:1–27:29, Dec. (2008).
[14].Mahantesh K, Shubha Rao A, "Content based Image Retrieval – Inspired by Computer Vision & Deep Learning Techniques",4th International Conference on Electrical, Electronics, Communication, Computer Technologies and Optimization Techniques (2019)
[15]. J. Puzicha, T. Hofmann, and J. M. Buhmann, "Non-parametric similarity measures for unsupervised texture segmentation and image retrieval," in Proceedings of the 1997 Conference on Computer Vision and Pattern Recognition (CVPR '97), ser. CVPR '97, (1997), pp. 267–272.
[16]. P. Viola and M. Jones, "Rapid object detection using a boosted cascade of simple features," (2001), pp. 511–518.
[17]. Z. Tu, "Probabilistic boosting-tree: Learning discriminative models for classification, recognition, and clustering," in Proceedings of the Tenth IEEE International Conference on Computer Vision - Volume 2, ser. ICCV '05, 2005, pp. 1589–1596.
[18]. B. Yao, A. Khosla, and F.-F. Li, "Combining randomization and discrimination for fine-grained image categorization." IEEE, (2011), pp. 1577–1584.
[19]. H. Zhang and L. Zhao, "Integral channel features for particle filter based object tracking," in Proceedings of the 2013 5th International Conference on Intelligent Human-Machine Systems and Cybernetics - Volume 02, ser. IHMSC '13, 2013, pp. 190– 193.
[20] P. Viola, M. Jones, D. Snow, "Detecting pedestrians using patterns of motion and appearance," Proceedings of IEEE International conference on Computer Vision, Nice, France, (2003).
[21] Takuya Kobayashi, Akinori Hidaka, Takio Kurita. Selection of Histograms of Oriented Gradients Features for Pedestrian Detection.Part II, LNCS 4985, pp. 598-607, (2008).
[22] Jon Shlens. A tutorial on PCA derivation, discussion and SVD. (2003)
[23] Li Fie Fie, RobFergus, and Pietro Perona Learning generative visual models from few training examples: an incremental Bayesian approach tested on 101 object categories. In IEEE CVPR Workshop of Generative Model-Based Vision, (2004).
intro part

[24]. Anusha, T.R., Hemavathi, N., Mahantesh, K., Chetana, R.: An investigation of combining gradient descriptor and diverse classifiers to improve object taxonomy in very large image dataset. In: International Conference on Contemporary Computing and Informatics, IEEE, pp. 581–585 (2014)

[25] Turk, M., Pentland, A.: Eigenfaces for recognition. J. Cogn. Neurosci. 3, 71–86 (1991)

[26] Pankaj, D.S., Wilscy, M.: Comparison of PCA, LDA & Gabor Features for Face Recognition Using Neural Networks, vol. 177, pp. 413–422 (2013)

[27] Preeti Rai, Prithee Khanna.: An Illumination, Expression & Noise Invariant Gender Classifier Using 2D-PCA on Real Gabor Space, pp. 15–28. Elseiver (2015)

[28] Shlens, J.: A Tutorial on PCA Derivation, Discussion and SVD (2003)

Exp result analysis

[29] A. Goldberg, X. Zhu, A. Singh, Z. Xu, R. Nowak. Multi-manifold semi-supervised learning. Proceedings of the International Conference on Artificial Intelligence and Statistics,(2009).

[30] Mahantesh.K,Anusha.T.R, Hemavathi.N, Chetana.R. An Investigation of Combining Gradient Descriptor and Diverse Classifiers to Improve Object Taxonomy in Very Large Image Dataset,(2014)International Conference on Contemporary Computing and Informatics, (IC3I) IEEE.pp.978-1-4799-6629-5/14,(2014)

[31] Bao-DiLiu , Yu-XiongWang, Yu-JinZhang, BinShen. Learning dictio-nary on manifolds for image classification. Pattern Recognition, Elsevier. vol.46, pp.1879-1890, (2013).

[32] B. Liu, Y. Wang, Y. Zhang, Y. Zheng. Discriminant sparse coding for image classification. Proceedings of the International Conference on Acoustics, Speech and Signal Processing,IEEE. pp. 2193-2196, 2012.

[33] Griffin, G., Holub, A., Perona, P.: Caltech 256 object category dataset. Technical Report UCB/CSD-04-1366, California Institute of Technology (2007)

[34] Serre, T., Wolf, L., Poggio, T.: Object recognition with features inspired by visual cortex. In: CVPR, San Diego (2005)

[35] Holub, A., Welling, M., Perona, P.: Exploiting unlabelled data for hybrid object classification. In: NIPS Workshop on Inter-Class Transfer, Whistler, BC (2005)

[36] Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: CVPR, vol. 2, pp. 2169–2178 (2006)

[37] Zhang, H., Berg, A.C., Maire, M., Malik, J.: SVM-KNN: Discriminative Nearest Neigh-bor Classification for Visual Category Recognition. IEEE-CVPR 2, 2126–2136 (2006)

[38] Ahammed, K.; Satu, M.S.; Abedin, M.Z.; Rahaman, M.A.; Islam, S.M.S. Early Detection of Coronavirus Cases Using Chest X-ray Images Employing Machine Learning and Deep Learning Approaches. medRxiv 2020. medRxiv 2020.06.07.20124594.
[39] Chowdhury, N.K.; Rahman, M.M.; Kabir, M.A. PDCOVIDNet: A parallel-dilated convolutional neural network architecture for
detecting COVID-19 from chest X-ray images. Health Inf. Sci. Syst. 2020, 8, 1–14. [CrossRef] [PubMed]
[40] Abbas, A.; Abdelsamea, M.M.; Gaber, M.M. Classification of COVID-19 in chest X-ray images using DeTraC deep convolutional neural network. Appl. Intell. 2021, 51, 854–864.
[41] Che Azemin, M.Z.; Hassan, R.; Mohd Tamrin, M.I.; Md Ali, M.A. COVID-19 Deep Learning Prediction Model Using Publicly Available Radiologist-Adjudicated Chest X-Ray Images as Training Data: Preliminary Findings. Int. J. Biomed. Imaging 2020,
[42] El-Rashidy, N.; El-Sappagh, S.; Islam, S.M.R.; El-Bakry, H.M.; Abdelrazek, S. End-To-End Deep Learning Framework for Coronavirus (COVID-19) Detection and Monitoring. Electronics 2020, 9, 1439.
[43] Khan, I.U.; Aslam, N. A Deep-Learning-Based Framework for Automated Diagnosis of COVID-19 Using X-ray Images. Information 2020, 11, 419.
[44] Loey, M.; Smarandache, F.M. Khalifa, N.E. Within the Lack of Chest COVID-19 X-ray Dataset: A Novel Detection Model Based on GAN and Deep Transfer Learning. Symmetry 2020, 12, 651.
[45]Minaee, S.; Kafieh, R.; Sonka, M.; Yazdani, S.; Soufi, G. Deep-COVID: Predicting COVID-19 from chest X-ray images using deep transfer learning. Med. Image Anal. 2020, 65, 101794.
[45] Sekeroglu, B.; Ozsahin, I. Detection of COVID-19 from Chest X-Ray Images Using Convolutional Neural Networks. SLAS Technol. Transl. Life Sci. Innov. 2020, 25, 553–565.
[46] Wang, N.; Liu, H.; Xu, C. Deep Learning for The Detection of COVID-19 Using Transfer Learning and Model Integration. In Proceedings of the 2020 IEEE 10th International Conference on Electronics Information and Emergency Communication (ICEIEC), Beijing, China, 17–19 July 2020; pp. 281–284
[47] Panwar, H.; Gupta, P.K.; Siddiqui, M.K.; Morales-Menendez, R.; Singh, V. Application of deep learning for fast detection of COVID-19 in X-rays using nCOVnet. Chaos Solitons Fractals 2020, 138, 109944.

[48] Ozturk, T.; Talo, M.; Yildirim, E.A.; Baloglu, U.B.; Yildirim, O.; Acharya, U.R. Automated detection of COVID-19 cases using deep neural networks with X-ray images. Comput. Biol. Med. 2020, 121, 103792.
[49] Khan, A.I.; Shah, J.L.; Bhat, M.M. Coronet: A deep neural network for detection and diagnosis of COVID-19 from chest X-ray images. Comput. Methods Prog. Biomed. 2020, 196, 105581.
[50] Apostolopoulos, I.D.; Mpesiana, T.A. Covid-19: Automatic detection from x-ray images utilizing transfer learning with convolutional neural networks. Phys. Eng. Sci. Med. 2020, 43, 635–640.
[51] Mahmud, T.; Rahman, M.A.; Fattah, S.A. CovXNet: A multidilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. Comput. Biol. Med. 2020, 122, 103869.
[52] Benbrahim, H.; Hachimi, H.; Amine, A. Deep transfer learning with apache spark to detect COVID-19 in chest X-ray images. Rom. J. Inform. Sci. Technol. 2020, 23, S117–S129.
[53] Martinez, F.; Martínez, F.; Jacinto, E. Performance evaluation of the NASNet convolutional network in the automatic identification of COVID-19. Int. J. Adv. Sci. Eng. Inform. Technol. 2020, 10, 662. [CrossRef]
[54] Toraman, S.; Alakus, T.B.; Turkoglu, I. Convolutional capsnet: A novel artificial neural network approach to detect COVID-19 disease from X-ray images using capsule networks. Chaos Solitons Fractals 2020, 140, 110122. [CrossRef]
[55] Duran-Lopez, L.; Dominguez-Morales, J.P.; Corral-Jaime, J.; Vicente- Diaz, S.; Linares-Barranco, A. COVID-XNet: A custom deep learning system to diagnose and locate COVID-19 in chest X-ray images. Appl. Sci. 2020, 10, 5683
learning system to diagnose and locate COVID-19 in chest X-ray images. Appl. Sci. 2020, 10, 5683
[56] M. B. Blaschko and C. H. Lampert. Learning to localize objects with structured output regression. In Computer Vision– ECCV 2008, pages 2–15. Springer, 2008.

[57] L. Bourdev and J. Malik. Poselets: Body part detectors trained using 3d human pose annotations. In International Conference on Computer Vision (ICCV), 2009.

[58] H. Cai, Q. Wu, T. Corradi, and P. Hall. The crossdepiction problem: Computer vision algorithms for recognising objects in artwork and in photographs. arXiv preprint arXiv:1505.00110, 2015.

[59] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 1, pages 886–893. IEEE, 2005. 4,

[60] T. Dean, M. Ruzon, M. Segal, J. Shlens, S. Vijayanarasimhan, J. Yagnik, et al. Fast, accurate detection of 100,000 object classes on a single machine. In Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, pages 1814–1821. IEEE, 2013.

[61] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell. Decaf: A deep convolutional activation feature for generic visual recognition. arXiv preprint arXiv:1310.1531, 2013.

[62] J. Dong, Q. Chen, S. Yan, and A. Yuille. Towards unified object detection and semantic segmentation. In Computer Vision–ECCV 2014, pages 299–314. Springer, 2014.

[63] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov. Scalable object detection using deep neural networks. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, pages 2155–2162. IEEE, 2014.

[64] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The pascal visual object classes challenge: A retrospective. International Journal of Computer Vision, 111(1):98–136, Jan. 2015.

[65] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(9):1627–1645, 2010

[66] M. B. Blaschko and C. H. Lampert. Learning to localize objects with structured output regression. In Computer Vision– ECCV 2008, pages 2–15. Springer, 2008.

[67] L. Bourdev and J. Malik. Poselets: Body part detectors trained using 3d human pose annotations. In International Conference on Computer Vision (ICCV), 2009.

[68] H. Cai, Q. Wu, T. Corradi, and P. Hall. The crossdepiction problem: Computer vision algorithms for recognising objects in artwork and in photographs. arXiv preprint arXiv:1505.00110, 2015.

[69] N. Dalal and B. Triggs. Histograms of oriented gradients for human detection. In Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on, volume 1, pages 886–893. IEEE, 2005. 4,

[70] T. Dean, M. Ruzon, M. Segal, J. Shlens, S. Vijayanarasimhan, J. Yagnik, et al. Fast, accurate detection of 100,000 object classes on a single machine. In Computer Vision and Pattern Recognition (CVPR), 2013 IEEE Conference on, pages 1814–1821. IEEE, 2013.

[71] J. Donahue, Y. Jia, O. Vinyals, J. Hoffman, N. Zhang, E. Tzeng, and T. Darrell. Decaf: A deep convolutional activation feature for generic visual recognition. arXiv preprint arXiv:1310.1531, 2013.

[72] J. Dong, Q. Chen, S. Yan, and A. Yuille. Towards unified object detection and semantic segmentation. In Computer Vision–ECCV 2014, pages 299–314. Springer, 2014.

[73] D. Erhan, C. Szegedy, A. Toshev, and D. Anguelov. Scalable object detection using deep neural networks. In Computer Vision and Pattern Recognition (CVPR), 2014 IEEE Conference on, pages 2155–2162. IEEE, 2014.

[74] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The pascal visual object classes challenge: A retrospective. International Journal of Computer Vision, 111(1):98–136, Jan. 2015.

[75] P. F. Felzenszwalb, R. B. Girshick, D. McAllester, and D. Ramanan. Object detection with discriminatively trained part based models. IEEE Transactions on Pattern Analysis and Machine Intelligence, 32(9):1627–1645, 2010

An Ensemble Model to Extract Discriminative Features for Semantic Image Classification in Large Datasets

Pranesh B, Nitin T, ShreeCharan, D P Tejash, Dr. Mahantesh K
{praneshbadarinath, nitint1999, shreecharan0507, creativetejash.d.p, mahantesh.sjbit}@gmail.com
Department of Electronics and Communication Engineering, SJB Institute of Technology, India

Abstract. An efficient image representation and extracting discriminative features in compressed domain is attracting researchers in computer vision and pattern recognition to develop efficient algorithms to classify and annotate images in large datasets. In this paper, an ensemble model combining DCT, Wavelet and HOG is developed to represent an image in compresses domain. DCT is useful to find low frequency coefficients of an image which express visual features, further subjected to multi-resolution analysis using wavelets with an advantage of developing robust and geometrically invariant structured object visual features through spectral analysis and finally PCA is used to map lower dimensional feature space with the transformational matrix in which a set of observed data is infused highlighting the edge orientation along with histogram of oriented gradients as its feature vectors. For classification purpose, different distance measure techniques and machine learning algorithm is used to obtain average classification rate. Proposed ensemble model is demonstrated on Caltech-101 and Caltech-256 datasets and compared the results with several benchmarking techniques in literature.

Keywords: Discrete cosine transforms (DCT), Distance Metrics (DM), Histogram of Oriented Gradients (HOG), Image Classification, Logistic regression (LR), Wavelets.

1. Introduction

The innovation and research and development sectors are growing rapidly due to the digital collection of data. The development of the internet and digital media techniques increased the demand for data. Hence, classification is an essential and required process for searching and efficient indexing of data. This leads to the existence of data classification and detection techniques and significant interest in the research community. The advancement of the internet and image acquisition techniques has resulted in a huge increase in digital image collections produced by research, educational, medical, industrial, and other applications. The standard approach for retrieving images from large datasets is to describe the image using terms and then use the keywords to retrieve the images from databases, but manually annotating the images takes a long time. Furthermore, due to the natural inability of keywords to bridge the semantic divide. For massive image collections that are often specific to human experience, context-sensitive, and incomplete, manually annotating images for a wide

variety of images is a time-consuming and costly process. In the early 1980s, content-based image retrieval (CBIR) was implemented as an important solution to text-based approaches because they struggled to accommodate a range of task-dependent queries.

CBIR indexes images based on their visual quality [1], such as texture, color and shape. Pattern recognition methods are used in image retrieval systems to explain visual information with partial semantics, while the process of extracting images has remained predominantly statistical in nature. While dealing with vast content image databases, some sophisticated algorithms are designed to interpret color, shape, and texture features which fails to exhibit image semantics and depict limitations [2]. In this paper, the distance measures and machine learning algorithms are used to build computer programs that produce solutions and evolve over time to solve problems that cannot be solved using enumerative or calculus-based approaches.

2. Related works

To accomplish the tasks for the classification of images [3], approaches associated with the depiction of low image functionality are used, including shape, size, color, etc. Many papers have been written on image processing using various ensemble approaches. It has many advantages and uses in a wide range of areas and recommended HOG strategies for image classification [4]. In the histogram of oriented gradients (HOG), the effects of normalizing are investigated [5]. The normalizing function has been incorporated into the gradient creation stage. PCA was first used to remove substantial features from a function space by Turk and Pentland. Pankaj and Wilscy [6] suggested a system for comparing PCA [7] and features for face recognition. PCA is a computational technique for converting a huge amount of linked data into a smaller amount of congruent data [8]. The feature vectors are fed into PCA, which uses feature vectors derived from HOG to rewrite HOG features in terms of new variables. The image will be first preprocessed to improve its quality, such as by standardizing and smoothing it [9]. The gradient histogram (HoG) and wavelet properties are utilized to make predictions using machine learning. Extraction techniques examine images and remove the most significant attributes representing different kinds of objects [10]. Features are used to attribute items to the class where they use as inputs for classifiers. The extraction process allows to minimize the amount of data having through analyzing certain characteristics of images which contain pertinent information or functionality which distinguish one pattern from the other. The several variations of features available including shape, color and texture-based features. Texture is amongst the most crucial attributes for distinguishing objects or areas of image enhancement is the process. These coefficients are decomposed into multi-resolution sub-bands by wavelets [11], allowing spectral analysis to create geometrically invariant structural object visual features. DCT eliminates low-frequency components that convey image visual features, and wavelets decompose these coefficients through multi-resolution sub-bands, causing spectral analysis to generate geometrically symmetric structural object visual features [12]. Using PCA, a dynamic model is

used to map a collection of observed data into a lower-dimensional spatial domain [6]. In general, very less amount of significant work has been noticed in literature, where Caltech 101/256 datasets have been used extensively used for demonstration and benchmarking algorithms with standard train-test procedures as mentioned in [15, 16, 17]. Finally, various distance measuring techniques are used in order to derive an overall consistency score for object categorization. We demonstrated the proposed methodology in Section 3.4 on two very popular and challenging datasets and achieved the highest classification thresholds in addition to other benchmarking approaches discussed in the literature. Factors to consider when assessing the accuracy of models and the conclusions drawn from them objectively are discussed.

3. Methods

This section includes the frequency domain (DCT) technique and the methods for mixed time-frequency domains by isolating the borders and gathering the full energy of the image. This section goes into the details of combining DCT and wavelets and later HOG to obtain semantic features, followed by feature extraction and classification techniques.

3.1 Discrete Cosine Transforms

Since PCA is found to have weaknesses in terms of lacking class separability and localization [13]. In the majority of images, the signal intensity at low frequency is high, and based on visual properties; DCT effectively divides image spectral bandwidth. The frequency domain of a signal is changed. The length of the f(x) sequence for x = 0,1, 2,...,N is indicated where N is the length of the sequence.

The DCT of an image f(x,y) is computed using equation (1):

$$C(u,v) = \frac{2}{\sqrt{mn}}\,\alpha(u)\alpha(v)\sum_{x=1}^{m}\sum_{y=1}^{n} f(x,y)\,cos\left[\frac{(2x+1)u\pi}{2m}\right]\,cos\left[\frac{(2y+1)v\pi}{2n}\right] \quad (1)$$

where u = 1, 2,...,m and v=1,2,...,n are variables of scaling, as opposed to a 1D signal.

Fig.1(a) depicts a color aero plane image with a scale of 256 X 256 pixels, Fig.1(b) shows the color map of the DCT's log magnitude, showing high energy compaction at the origin, and Fig.1(c) shows the standard zigzag scanning scrutinize for DCT factors for an image.

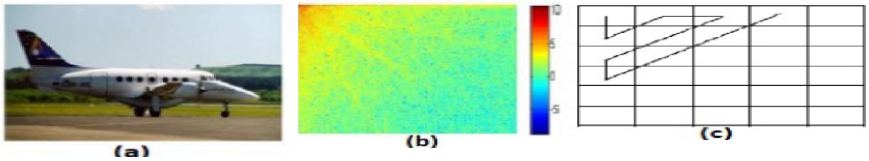


Fig. 1. (a) original image; (b) color map of cipher degree of DCT;(c) scanning strategy of DCT coefficients

3.2 Multi resolution analysis using wavelets

dbN wavelets are the Daubechies' extreme phase wavelets. The Daubechies function extraction is discussed using scalar products with scaling signals and wavelets, similar to the Haar wavelet transform.

Daubechies combine the principles of multi-resolution analysis and pyramid coding techniques, then outlined their similarities and differences, and initially incorporated wavelet transform with the conventional definition of filter banks. when a source image is decomposed, four equal-sized output images are produced as shown in Fig.2.

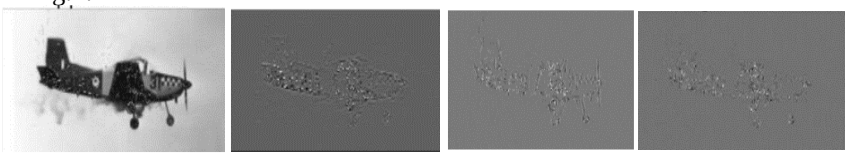


Fig.2. ApproximationElement, Horizontal Element, Vertical Element, Diagonal Element

3.3 HOG

HOG is a feature descriptor that is often used to extract features from image data [5]. The approximation element from the previous section is used as base image for feature extraction. Fig.3. gives HOG representation along with its original image data.

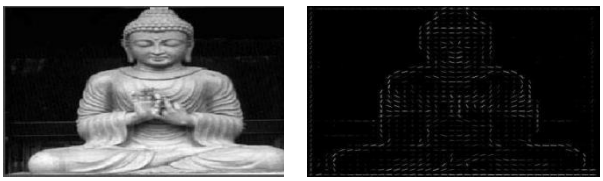


Fig .3. (a) Original image (b) HOG representation

The configuration of this work is somewhat different from that of Dalal and Triggs [14]. Linear-gradient voting into nine orientation bins ranging from 0 to 180 degrees, 16x16 pixel blocks with four 8x8 pixel cells, and 16x16 pixel blocks with four 8x8 pixel cells are used in the selected configuration. The size of the descriptor can be computed using equation (2):

$$Descriptor\ size = b_s^2\, n_b\, (v_{cells} - b_s + \sigma))(h_{cells} - b_s + \sigma) \qquad (2)$$

The detector window has a resolution of 64x128 pixels. Converging the gradient masks [1 0 1] and $[1\ 0\ 1]^T$ with the raw image I yields the HOG descriptor by computing the 1-D p1oint derivatives Gx and Gy in the x and y directions. The equations (3) describes the gradient computation (being I the image and (i, j) the pixel coordinates):

$$G_x(i,j) = \frac{\partial I}{\partial x}(i,j) \qquad G_y(i,j) = \frac{\partial I}{\partial y}(i,j) \qquad (3)$$

The derivatives G_x and G_y are then applied to each pixel to calculate the gradient degree |G (x, y)| and path angle (x,y). The equation (4) indicates the magnitude of the gradient at a certain pixel:

$$M(i,j) = \sqrt{G_x^2(i,j) + G_y^2(i,j)} \tag{4}$$

The gradient degree of the direction histogram is simply used as a weighting factor. The gradient direction angle can be calculated using equation (5):

$$\theta(i,j) = arctan\left(\frac{G_y(i,j)}{G_x(i,j)}\right) \tag{5}$$

Depending on whether the angle is signed or unsigned, this histogram represents angles that are equally spaced between 0 and 180 or 0 and 360. As a result, the equation (6) is used for determining the kth bin of a histogram:

$$h_k = \sum_{i,j} M(i,j) 1[\Phi(i,j) = k] \tag{6}$$

In our case, k is a positive integer between 1 and 9. The vote will most likely take place between two bin centers at any given angle, if the previous angle division is contained in bins, so the voting is decided by bilinear interpolation between two bin centers mentioned in equation (7):

$$f(x|x_1, x_2) = f(x_1) + \frac{f\big(f(x_2) - f(x_1)\big)}{x_2 - x_1}\,(x - x_1) \tag{7}$$

As the variations in gradient intensity show, some kind of illumination normalization is needed for a robust descriptor. Let the k-norm of v, with k ∈ 1, 2, and a small constant, as the k-norm of v, as a vector containing all the histogram for the block $||v||_k$. L1-norm, L1 squared-norm and L2-norm of a block v descriptor vectors are mentioned in equations (8),(9) and (10) respectively:

$$L1 - norm: v \rightarrow \frac{v}{\big|\big|v\big|\big|_2 + \epsilon} \tag{8}$$

$$L1 - squared\ norm: v \rightarrow \sqrt{\frac{v}{\big|\big|v\big|\big|_2 + \epsilon}} \tag{9}$$

$$L2 - norm: v \rightarrow \frac{v}{\sqrt{\big|\big|v\big|\big|_2^{\,2} + \epsilon^2}} \tag{10}$$

The values of v are constrained to 0.2 after an L2-norm, clipping, and re-normalization. This can be achieved by reducing, cutting and normalizing the product of the L2-norm. The tests revealed that L1-norm-squared, L2-norm, and L2-norm-Hys

all behave equally and produce positive outcomes, but L1-norm degrades efficiency by 5%. If you do not normalize, the score is penalized by approximately 27 percent.

3.4 Classification

The HOG features obtained from the previous section are used for classification purpose. The visual similarities are calculated using different distance measure techniques like Euclidean, Manhattan, KL Divergence and JS Divergence which can be computed using equations from Table 1. Further, logistic regression classifier is also used to compliment the results.

Table 1. Mathematical representation of different distance measures

Distance measures	Equations
Euclidean distance	$d(A, \mathrm{B}) = \lvert \mathrm{a} - \mathrm{b} \rvert = \sqrt{(a_1 - b_1)^2 + (a_2 - b_2)^2 + \cdots + (a_n - b_n)^2}$
Manhattan distance	$M_{dist}(A, \mathrm{B}) = \lvert a_1 - b_1 \rvert + \lvert a_2 - b_2 \rvert + \cdots \ldots + \lvert a_n - b_n \rvert$
Kullback-Leibler Divergence	$\mathrm{KL(P} \,\|\| \mathrm{Q)} = \sum_x XP(x) * \log\left(\frac{P(x)}{Q(x)}\right)$
Jensen-Shannon Divergence	$JS(P \,\|\| Q) = {}^{\mathrm{KL(P} \,\|\| \mathrm{M)} + \mathrm{KL(Q} \,\|\| \mathrm{M)}}/_2$ $M = {}^{(P+Q)}/_2$

In the above Table 1, A and B are the two-dimensional points with coordinates (a1, a2,..),(b1, b2,..) respectively. Where A is considered as train data features and B is query data feature. Divergence methods calculate the difference between two probability distributions. The log of the event in P over the event likelihood in Q.

KL divergence score is based on the idea that the odds of an occurrence from P vary greatly from those of the same event from Q. It may be used to differentiate between discrete and continuous probability distributions, which determine the integrality of outcomes rather than the number of discrete events that are expected to occur.

The JS divergence, a symmetrical normalized rating, is calculated using the KL divergence. It provides a smoothed and normalized variant of KL divergence, with scores ranging from 0 (identical) to 1 (extremely different), making it more functional as a metric by using the base-2 algorithm.

3.4.1 Logistic Regression

It is broadly utilized in ML and statistics. In both binary and multiclass grouping, it performs admirably. A LR model predicts P(Y=1) as an X-function mathematically. LR employs the log odds ratio rather than probabilities to predict category, and an incremental multivariate technique rather than a least squares technique to build the final model. Assumption of logistic regression includes (1) For each instance, each

independent variable should have a single value; (2) collinearity is expected to be minimal, although not necessarily totally independent; and (3) Irrelevant Independent Alternatives (IIA). The probabilities of preferring one class over another, despite the apparent presence of numerous unrelated and irrelevant choices are defined by IIA. In order to prevent multicollinearity in the model, the variables must be independent of one another. We need associated variables and a broad sample size for logistic regression. Using this approach it is possible to predict the average classification rate.

4. Results and Performance Analysis

The results and analysis of the Caltech-101 and Caltech-256 datasets are covered in this section. Caltech - 101 is a collection of 9,144 images from 101 distinct sorts of nature scenes (animals, butterflies, chandeliers, Garfield, cars, flowers, human faces, etc.) [15]. Most of the pictures are centered, occluded, influenced by corner artifacts, and display high-intensity variability, making this dataset very problematic [16]. The Caltech-256 dataset, which includes image data from 256 item categories and a total of 30,607 pictures, is an augmentation of the Caltech-101 dataset [17]. It has higher intensity, clutter, item size, location, and posture variations, as well as more categories with at least 40 images per category and higher inter-class and intra-class variability. Table.2 gives classification rates of different distance metrics for 15 & 30 train per class/category respectively. Sample images of datasets are as shown in Fig.4

Fig.4. Sample images of Caltech 101 and 256 datasets

Table 2. Results of different distance metrics using HOG

Distance metrics	Caltech 101 15 Train	Caltech 101 30 Train	Caltech 256 15 Train	Caltech 256 30 Train
	Classification Rate in %			
Euclidean Distance	35.2246	38.1712	12.3379	14.5542
Manhattan Distance	40.5264	44.2474	17.5962	20.5483
KL Method	23.8639	25.5652	-	-
JS Method	28.1980	28.1980	-	-

To evaluate the performance of the proposed methodology, the conventional experimental procedure outlined in [19] is considered, labeling the first 15 & 30 images in each category as training to generate feature vectors and the remaining as testing.

Table 3. Analysis of different transformation techniques with logistic regression

Methodology	Caltech 101 15 Train	Caltech 101 30 Train	Caltech 256 15 Train	Caltech 256 30 Train
	Classification Rate in %			
Coslets[12]	37	43.1	16.9	21
Holub et al. [17]	37	43	-	-
Serre et al. [18]	35	42	-	-
HOG	35.1316	43.3317	11.1489	15.1768
DCT+HOG	40.5596	46.9413	13.6233	15.2541
Wavelet+HOG	39.6711	46.1189	12.8623	15.9632
DCT+Wavelet+HOG	48.3920	53.0854	16.6244	19.0058
PCA+Wavelet+HOG	44.0699	48.0862	14.0952	17.2614

Using the image, a transformation strategy is performed to extract lower frequency coefficient information in the compressed DCT domain using 2D db wavelets. The approximated element is used as a base image for HOG feature extraction. Vectorization procedure (DCT Coefficients in zig-zag pattern) is initiated to generate the feature vectors. The trained feature dataset and the query feature vector use different distance measures and machine learning technique to obtain similarity. As a consequence, the performance of HOG descriptors in combination with transformation techniques are found to be efficient in comparison other descriptors, and the overall performance is increased.

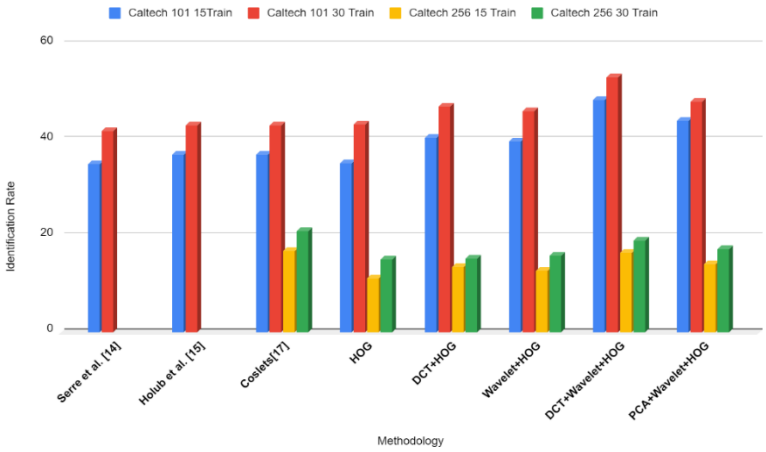


Fig. 5. Comparative analysis of Ensemble Methodologies

Table 3 signifies that the proposed model's performance to a few state-of-the-art approaches mentioned in the relevant research. The histogram descriptor technique with transformation techniques has turned out to be very interesting method for improving classification rate when compared to other works. On manifolds, Fig.5 compares the performance of the proposed method to current dictionary learning tech-

niques. In contrast to [18, 20, 12, 21], where grids of HOG features are proven to be very primitive and significant in outperforming the other strategies, the suggested method achieves a remarkable identification rate merely by employing a single visual descriptor. In comparison to conventional classifiers mentioned in [7, 19, 22], the proposed methodology with HOG has obtained leading classification rates of 48.4% & 53.08% for 15 & 30 images per category respectively, and was found consistent when compared with spatial pyramid feature technique [19] & method based on sparse localized features [23]. The proposed methodology using HOG outperformed [17] by 3.84% and produced very competitive results in comparison to the methodologies indicated in [6, 24, 13], which employed just 12,800 (i.e., no more than 50 images per category) images for testing.

5. Discussion and Conclusion

In recent years, most of the papers in literature have highlighted on combining transformation techniques for image representation but failed to preserve discriminative features. An ensemble model is proposed which combines DCT, Wavelet and HOG to preserve low frequency co-efficient, point singularities and edge orientations respectively from an image. Proposed ensemble method combined with different distance measure and logistic regression techniques are demonstrated on very large benchmark datasets for classifying images and has obtained competitive results by outperforming recent techniques such as coslet, bag of features, bag of words, spatial pyramid matching techniques. In future, the performance of our model can be examined for different distance measures, support vector machine, neural network architectures and deep learning techniques. The scope of this paper lies on real time object detection in cameras.

References

1. Munjal, Meenaakshi & Bhatia, Shaveta. A Novel Technique for Effective Image Gallery Search using Content Based Image Retrieval System. (2019), pp.25-29.
2. K. Mahantesh and S. Rao A., Content Based Image Retrieval - Inspired by Computer Vision & Deep Learning Techniques, IEEE (2019), pp. 371-377.
3. Liu, Bao-Di & Wang, Yu-Xiong & Zhang, Yu-Jin & Zheng, Yin. (2012). Discriminant sparse coding for image classification. Acoustics, Speech, and Signal Processing, 1988. ICASSP-88.,1988InternationalConferenceon. 2193-2196. 10.1109/ICASSP.2012.6288348.
4. O. Deniz, G. Bueno, J. Salido, and F. De la Torre, Face recognition using histograms of oriented gradients, Pattern Recogn. Lett., vol. 32, no. 12,Sep. (2011) pp. 1598–1603.
5. L. Zhang, W. Zhou, J. Li, J. Li and X. Lou, "Histogram of Oriented Gradients Feature Extraction Without Normalization," 2020 IEEE Asia Pacific Conference on Circuits and Systems (APCCAS), 2020, doi: 10.1109/APCCAS50809.2020.9301715, pp. 252-255.
6. Pankaj, D.S., Wilscy, M.: Comparison of PCA, LDA & Gabor Features for Face Recognition. Using Neural Networks, vol. 177, pp. 413–422 (2013).
7. Jon Shlens. A tutorial on PCA derivation, discussion and SVD. (2003)

8. Preeti Rai, Pritee Khanna,An illumination, expression, and noise invariant gender classifier using two-directional 2DPCA on real Gabor space,Journal of Visual Languages & Computing,Volume 26,(2015),SSN 1045-926X,pp. 15-28.
9. H. Zhang and L. Zhao, Integral channel features for particle filter based object tracking, in Proceedings of the 2013, Volume 02, ser. IHMSC ’13, 2013, pp. 190– 193.
10. N. S. J. Shyla and W. R. S. Emmanuel, "Automated Classification of Glaucoma Using DWT and HOG Features with Extreme Learning Machine," 2021 Third International Conference on Intelligent Communication Technologies and Virtual Mobile Networks (ICICV), 2021, pp. 725-730, doi: 10.1109/ICICV50876.2021.9388376.
11. K. Mahantesh, V.N. Manjunath Aradhya, An Exploration of Ridgelet Transform to handle Higher Dimensional Intermittency for Object Categorization in Large Image Datasets, ICAICT (2014),pp. 515-521.
12. Mahantesh K., Aradhya V.N.M., Niranjan S.K. (2015) Coslets: A Novel Approach to Explore Object Taxonomy in Compressed DCT Domain for Large Image Datasets. Advances in Intelligent Systems and Computing, vol 320. Springer.
13. Hemavathi N., Anusha T.R., Mahantesh K., Manjunath Aradhya V.N. An Investigation of Gabor PCA and Different Similarity Measure Techniques for Image Classification. Advances in Intelligent Systems and Computing, vol 381. Springer, New Delhi.(2016).
14. N. Dalal and B. Triggs,Histograms of oriented gradients for human detection, 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, doi: 10.1109/CVPR.2005.177. pp. 886-893 vol. 1.
15. B. Yao, A. Khosla and L. Fei-Fei, Combining randomization and discrimination for fine-grained image categorization, CVPR 2011, 2011, doi: 10.1109/CVPR.2011.5995368, pp. 1577-1584.
16. Bao-DiLiu , Yu-XiongWang, Yu-JinZhang, BinShen. Learning dictio-nary on manifolds for image classification. Pattern Recognition, Elsevier. vol.46, (2013),pp.1879-1890.
17. Holub, A., Welling, M., Perona, P.: Exploiting unlabelled data for hybrid object classification. In: NIPS Workshop on Inter-Class Transfer, Whistler, BC (2005)
18. Serre, T., Wolf, L., Poggio, T.: Object recognition with features inspired by visual cortex.In: CVPR, San Diego (2005)
19. Li Fei-Fei, R. Fergus and P. Perona, Learning Generative Visual Models from Few Training Examples: An Incremental Bayesian Approach Tested on 101 Object Categories,2004 Conference on Computer Vision and Pattern Recognition Workshop, 2004, pp. 178-178.
20. Kihyuk Sohn, Dae Yon Jung, Honglak Lee, Alfred O.Hero III. Efficient Learning of Sparse,Distributed,Convolutional Feature Representation for Object Recognition(2011). pp.215-223.
21. T. R. Anusha, N. Hemavathi, K. Mahantesh and R. Chetana, An investigation of combining gradient descriptor and diverse classifiers to improve object taxonomy in very large image dataset, IC3I, 2014, doi: 10.1109/IC3I.2014.7019774, pp. 581-585.
22. Griffin, Gregory and Holub, Alex and Perona, Pietro Caltech-256 Object Category Dataset. California Institute of Technology .(2007)
23. T. Serre, L. Wolf and T. Poggio, "Object recognition with features inspired by visual cortex," 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), 2005, vol. 2, doi: 10.1109/CVPR.2005.254, pp. 994-1000.
24. K. Mahantesh and Manjunath Aradhya V, An Exploration of Neural Networks & Transformation Techniques for Image Classification, International Journal of Advanced Research in Computer Science and Software Engineering,Vol. 5, Issue-11, (2015).

Certificates achieved in International Conference on Distributed Computing and Optimization Techniques (ICDCOT 2021)

(Awarded as a best paper)

|| JAI SRI GURUDEV ||

Sri Adichunchunagiri Shikshana Trust®

International Conference on Distributed Computing and Optimization Techniques

ICDCOT-2021

Organized by

Department of Electronics & Communication Engineering
and
Department of Electrical & Electronics Engineering
SJB Institute of Technology, Bengaluru, India

In Association with

Technical Institute for Engineers

CERTIFICATE OF APPRICIATION

Dr. Mahantesh. K

For the paper titled An Ensemble Model to Extract Discriminative Features for Semantic Image Classification in Large Datasets presented by Dr. Mahantesh. K which has been selected as the Best Paper amongst presented papers in the International Conference on Distributed Computing and Optimization Techniques (ICDCOT-2021), organized by the Department of Electronics & Communication Engineering and Department of Electrical & Electronics Engineering, SJB Institute of Technology, Bengaluru, India in association with Technical Institute for Engineers (T.I.E.), Bengaluru on 25th-26th June, 2021.

Dr. Hemalatha K.L.	Dr. K. V. Mahendra Prashanth	Dr. Babu N. V.	Dr. Ajai Chandran.C. K
President, Technical Institute for Engineers	Professor & HoD, Electronics & Communication Engineering, SJB Institute of Technology	Professor & HoD, Electrical & Electronics Engineering, SJB Institute of Technology	Principal, SJB Institute of Technology

|| JAI SRI GURUDEV ||

Sri Adichunchunagiri Shikshana Trust®

International Conference on Distributed Computing and Optimization Techniques

ICDCOT-2021

Organized by

Department of Electronics & Communication Engineering
and
Department of Electrical & Electronics Engineering
SJB Institute of Technology, Bengaluru, India

In Association with

Technical Institute for Engineers

CERTIFICATE OF APPRICIATION

Nitin. T

For the paper titled An Ensemble Model to Extract Discriminative Features for Semantic Image Classification in Large Datasets presented by Nitin. T which has been selected as the Best Paper amongst presented papers in the International Conference on Distributed Computing and Optimization Techniques (ICDCOT-2021), organized by the Department of Electronics & Communication Engineering and Department of Electrical & Electronics Engineering, SJB Institute of Technology, Bengaluru, India in association with Technical Institute for Engineers (T.I.E.), Bengaluru on 25th-26th June, 2021.

Dr. Hemalatha K.L.	Dr. K. V. Mahendra Prashanth	Dr. Babu N. V.	Dr. Ajai Chandran.C. K
President, Technical Institute for Engineers	Professor & HoD, Electronics & Communication Engineering, SJB Institute of Technology	Professor & HoD, Electrical & Electronics Engineering, SJB Institute of Technology	Principal, SJB Institute of Technology

|| JAI SRI GURUDEV ||

Sri Adichunchunagiri Shikshana Trust®

International Conference on Distributed Computing and Optimization Techniques

ICDCOT-2021

Organized by

Department of Electronics & Communication Engineering
and
Department of Electrical & Electronics Engineering
SJB Institute of Technology, Bengaluru, India

In Association with

Technical Institute for Engineers

CERTIFICATE OF APPRICIATION

ShreeCharan

For the paper titled An Ensemble Model to Extract Discriminative Features for Semantic Image Classification in Large Datasets presented by ShreeCharan which has been selected as the Best Paper amongst presented papers in the International Conference on Distributed Computing and Optimization Techniques (ICDCOT-2021), organized by the Department of Electronics & Communication Engineering and Department of Electrical & Electronics Engineering, SJB Institute of Technology, Bengaluru, India in association with Technical Institute for Engineers (T.I.E.), Bengaluru on 25th-26th June, 2021.

Dr. Hemalatha K.L.	Dr. K. V. Mahendra Prashanth	Dr. Babu N. V.	Dr. Ajai Chandran.C. K
President, Technical Institute for Engineers	Professor & HoD, Electronics & Communication Engineering, SJB Institute of Technology	Professor & HoD, Electrical & Electronics Engineering, SJB Institute of Technology	Principal, SJB Institute of Technology

II JAI SRI GURUDEV II

Sri Adichunchunagiri Shikshana Trust®

International Conference on Distributed Computing and Optimization Techniques

ICDCOT-2021

Organized by

Department of Electronics & Communication Engineering
and
Department of Electrical & Electronics Engineering
SJB Institute of Technology, Bengaluru, India

In Association with

Technical Institute for Engineers

CERTIFICATE OF APPRICIATION

D. P. Tejash

For the paper titled An Ensemble Model to Extract Discriminative Features for Semantic Image Classification in Large Datasets
presented by D. P. Tejash which has been selected as the Best Paper amongst presented papers in the International Conference on Distributed Computing and Optimization Techniques (ICDCOT-2021), organized by the Department of Electronics & Communication Engineering and Department of Electrical & Electronics Engineering, SJB Institute of Technology, Bengaluru, India in association with Technical Institute for Engineers (T.I.E.), Bengaluru on 25th-26th June, 2021.

Dr. Hemalatha K.L.	Dr. K. V. Mahendra Prashanth	Dr. Babu N. V.	Dr. Ajai Chandran.C. K
President, Technical Institute for Engineers	Professor & HoD, Electronics & Communication Engineering, SJB Institute of Technology	Professor & HoD, Electrical & Electronics Engineering, SJB Institute of Technology	Principal, SJB Institute of Technology

II JAI SRI GURUDEV II

Sri Adichunchunagiri Shikshana Trust®

International Conference on Distributed Computing and Optimization Techniques

ICDCOT-2021

Organized by

Department of Electronics & Communication Engineering
and
Department of Electrical & Electronics Engineering
SJB Institute of Technology, Bengaluru, India

In Association with

Technical Institute for Engineers

CERTIFICATE OF APPRICIATION

Pranesh. B

For the paper titled An Ensemble Model to Extract Discriminative Features for Semantic Image Classification in Large Datasets
presented by Pranesh. B which has been selected as the Best Paper amongst presented papers in the International Conference on Distributed Computing and Optimization Techniques (ICDCOT-2021), organized by the Department of Electronics & Communication Engineering and Department of Electrical & Electronics Engineering, SJB Institute of Technology, Bengaluru, India in association with Technical Institute for Engineers (T.I.E.), Bengaluru on 25th-26th June, 2021.

Dr. Hemalatha K.L.	Dr. K. V. Mahendra Prashanth	Dr. Babu N. V.	Dr. Ajai Chandran.C. K
President, Technical Institute for Engineers	Professor & HoD, Electronics & Communication Engineering, SJB Institute of Technology	Professor & HoD, Electrical & Electronics Engineering, SJB Institute of Technology	Principal, SJB Institute of Technology

Certificates achieved in 5th National Level IEEE Project Competition - 2021

Geetha Shishu Shikshana Sangha
(R)

GSSS INSTITUTE OF ENGINEERING & TECHNOLOGY FOR WOMEN

(Affiliated to VTU, Belagavi, Approved by AICTE, New Delhi & Govt. of Karnataka)
Accredited with Grade 'A' by NAAC
Accredited Branches by NBA, New Delhi; UG – ECE, CSE, ISE & TE.
(Validity: 01.07.2017 – 30.06.2020 & 01.07.2020 – 30.06.2023)
K R S ROAD | METAGALLI | MYSURU - 570016 | KARNATAKA | INDIA

IEEE STUDENT BRANCH

5th National Level IEEE Project Competition-2021

Co-Sponsors

Participation

Certificate of Participation

This is to certify that NITIN T

of SJB Institute of Technology has

presented the project entitled Impact of Transformation Techniques for Classifying Images on Large Datasets

in the **"5th National Level IEEE Project Competition-2021"** Organized by IEEE Student Branch (STB17861), GSSSIETW in association with IEEE Bangalore Section and IEEE Mysore Subsection on **26th June 2021.**

Dr. Parameshachari B D
IEEE Student Branch Counselor,
Professor and Head, Dept. of TCE,
GSSSIETW, Mysuru

Dr. Shivakumar M
Principal,
GSSSIETW, Mysuru

Made for free with Certify'em

Geetha Shishu Shikshana Sangha (R)

GSSS INSTITUTE OF ENGINEERING & TECHNOLOGY FOR WOMEN

(Affiliated to VTU, Belagavi, Approved by AICTE, New Delhi & Govt. of Karnataka)
Accredited with Grade 'A' by NAAC
Accredited Branches by NBA, New Delhi; UG – ECE, CSE, ISE & TE.
(Validity: 01.07.2017 – 30.06.2020 & 01.07.2020 – 30.06.2023)
K R S ROAD | METAGALLI | MYSURU - 570016 | KARNATAKA | INDIA

IEEE STUDENT BRANCH

5th National Level IEEE Project Competition-2021

Co-Sponsors

Participation

Certificate of Participation

This is to certify that SHREE CHARN

of SJB Institute of Technology has

presented the project entitled Impact of Tranformation Techniques for classifying images on large data sets

in the **"5th National Level IEEE Project Competition-2021"** Organized by IEEE Student Branch (STB17861), GSSSIETW in association with IEEE Bangalore Section and IEEE Mysore Subsection on **26th June 2021.**

Dr. Parameshachari B D
IEEE Student Branch Counselor,
Professor and Head, Dept. of TCE,
GSSSIETW, Mysuru

Dr. Shivakumar M
Principal,
GSSSIETW, Mysuru

Geetha Shishu Shikshana Sangha
(R)

GSSS INSTITUTE OF ENGINEERING & TECHNOLOGY FOR WOMEN

(Affiliated to VTU, Belagavi, Approved by AICTE, New Delhi & Govt. of Karnataka)
Accredited with Grade 'A' by NAAC
Accredited Branches by NBA, New Delhi; UG – ECE, CSE, ISE & TE.
(Validity: 01.07.2017 – 30.06.2020 & 01.07.2020 – 30.06.2023)
K R S ROAD | METAGALLI | MYSURU - 570016 | KARNATAKA | INDIA

IEEE STUDENT BRANCH

5th National Level IEEE Project Competition-2021

Co-Sponsors

Participation

Certificate of Participation

This is to certify that D P TEJASH

of SJB INSTITUTE OF TECHNOLOGY has

presented the project entitled Impact of Transformation Techniques for Classifying Images on Large Datasets

in the **"5th National Level IEEE Project Competition-2021"** Organized by IEEE Student Branch (STB17861), GSSSIETW in association with IEEE Bangalore Section and IEEE Mysore Subsection on **26th June 2021.**

Dr. Parameshachari B D
IEEE Student Branch Counselor,
Professor and Head, Dept. of TCE,
GSSSIETW, Mysuru

Dr. Shivakumar M
Principal,
GSSSIETW, Mysuru

Geetha Shishu Shikshana Sangha
(R)

GSSS INSTITUTE OF ENGINEERING & TECHNOLOGY FOR WOMEN

(Affiliated to VTU, Belagavi, Approved by AICTE, New Delhi & Govt. of Karnataka)
Accredited with Grade 'A' by NAAC
Accredited Branches by NBA, New Delhi; UG – ECE, CSE, ISE & TE.
(Validity: 01.07.2017 – 30.06.2020 & 01.07.2020 – 30.06.2023)
K R S ROAD | METAGALLI | MYSURU - 570016 | KARNATAKA | INDIA

IEEE STUDENT BRANCH

5th National Level IEEE Project Competition-2021

Co-Sponsors

Participation

Certificate of Participation

This is to certify that Pranesh B

of SJB Institute Of Technology has

presented the project entitled Impact of Transformation Techniques for Classifying Images on Large Datasets

in the **"5th National Level IEEE Project Competition-2021"** Organized by IEEE Student Branch (STB17861), GSSSIETW in association with IEEE Bangalore Section and IEEE Mysore Subsection on **26th June 2021.**

Dr. Parameshachari B D
IEEE Student Branch Counselor,
Professor and Head, Dept. of TCE,
GSSSIETW, Mysuru

Dr. Shivakumar M
Principal,
GSSSIETW, Mysuru

Certificate Achieved in InterCollegiate Project Competition (IPC) 2021

DAYANANDA SAGAR ACADEMY OF TECHNOLOGY AND MANAGEMENT
(Affiliated to VTU ,Belagavi& Approved by AICTE,New Delhi)
(All B.E Branches are accredited by NBA, New Delhi.)
Bangalore-560082

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
(Accredited for 3 years by NBA, New Delhi (Validity: 26-07-2018 to 30-06-2021))

INTERCOLLEGIATE PROJECT COMPETITION (IPC) CERTIFICATE

July 17th, 2021

This is to certify that the following students

NAMES

NITIN T

PRANESH B

SHREECHARAN

D P TEJASH

have participated from ***SJBIT, Bangalore*** with project titled **An Ensemble Model to Extract Discriminative Features for Semantic Image Classification in Large Datasets** in **"INTERCOLLEGIATE PROJECT COMPETITION"** Conducted by the Department of Electronics and Communication Engineering, Dayananda Sagar Academy of Technology and Management, Bangalore-560082.

Dr. Karibasappa K, Professor &
Dr. Sumaiya MN, Assoc. Professor
IPC Coordinators
DSATM

Dr.ManjunathaPrasad R
HOD/ECE
DSATM