IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)
An International Open Access, Peer-reviewed, Refereed Journal

Class Academia Entity Classification using CNN

[1]Mahadevan.K, [2]Charulatha B.S, [3]Swaminathan.B

[1]Student, [2]Professor, [3]Professor
[1]Computer Science and Engineering,
[1]Rajalakshmi Engineering College, Chennai, India

Abstract: *Object recognition is the process of identifying a distinguished object in a picture or video sequence. Many different approaches of visual perception using a normal classifier or deep neural network were proposed. The objective is to implement a Convolution Neural Network for object classification. It described the innovative solution that provides efficient object detection through deep learning with convolutional neural networks (CNNs) which has achieved great success in the classification of various types of object. A variety of layer-wise visualization methods were applied using CNN, trained with a publicly available any class object given academia image dataset. So, it observed that neural networks can capture the colors and textures of objects specific to respective type which resembles human decision-making.*

Index Terms - ***Object Detection, Deep Learning, TensorFlow, CNN, Keras***

I. INTRODUCTION

Object recognition is the mechanism of searching and pinpointing a selected object in a digital image or video sequence. Humans can easily recognize an object in an image even though occlusion comes into place which disrupts images in different sizes or scales. However, visual perception from a picture or video remains a challenge for computer vision systems. Even with the assistance of smart algorithms and human assistants, a classifier within the computer remains unable to catch everything in a picture. Many approaches to the task are implemented over multiple decades. Object recognition task is successful if the network system is in a position to label the things which are supported models of known objects. For example, given an image containing one or more different objects in the background, the network system is capable of assigning the labels to a set of regions in the image correctly as they were. The classification accuracy of the network system can be calculated by comparing the result with a set of labels corresponding to a set of objects known to the system. The object recognition has a very close relationship with segmentation as it incurs outline of the object within an image.

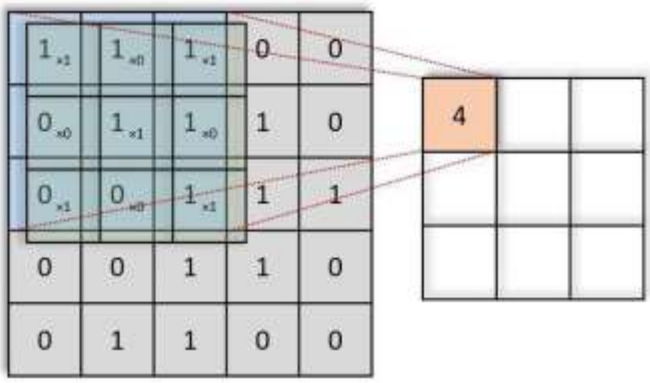
1. DEEP LEARNING:

Deep learning is a subset of machine learning that impersonates the neurons of the neural networks present in the human brain. Computer Vision Deep learning models are trained on a set of images also known as training data, to solve a task. The deep learning models are mainly used in the field of Computer Vision which allows a computer to perceive and visualize the same way as a human does on regular basis. Deep Learning models can be interpreted as a set of points each of which makes a resolution based on the inputs to the node. This network resembles the biological nervous system with each node acting as a neuron within a larger network. So, deep learning models are a class of artificial neural networks. Deep learning algorithms learns in an escalating manner about the image as it goes through each neural network layer. Layers at preliminary levels learn how to detect and extract features like edges, and subsequent layers combine features from earlier layers into a more holistic and complete representation. In neural networks, Convolutional neural network (ConvNets or CNNs) is one of the important algorithms do images recognition, images classification, image tracking and localization. CNN image classifications take an input image, process it and classify it under respective classes as given under subject.

2. METHODOLOGY:

2.1 FEATURE EXTRACTION:

In this part, the system will discharge successive convolutions and pooling operations during which the features are detected. The features are extracted using Convolution layer.



$$(1\times1+0\times1+1\times1) + (0\times0+1\times1+1\times0) + (1\times0+0\times0+1\times1) = 4$$

Fig 1.1 Extracting features

2.2 DIMENSIONALITY REDUCTION:

The Pooling layer is responsible for reducing the spatial size of the Convolved Feature for the purpose of decreasing the computational power required to process the data through dimensionality reduction and henceforth also useful for deracinating superior features which are rotational and positional invariant, thus maintaining the process of effectively training of the model. Pooling curtails the training time and controls over-fitting.

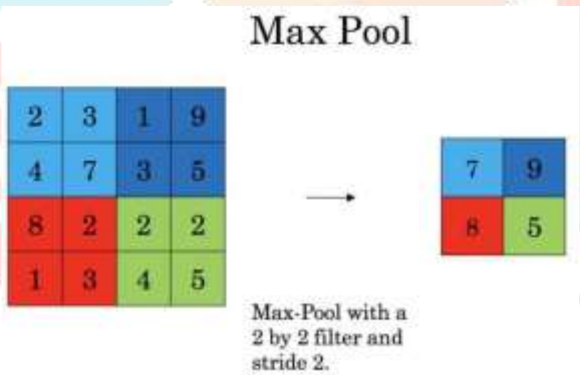


Fig 1.2 Dimensionality reduction using Max Pooling

2.3 CLASSIFICATION:

Fully Connected Layer is a Multi-Layered feed forward neural network. The input to this layer is the actual output from the ultimate Pooling or Convolutional Layer, which is flattened and then fed into the fully connected layer.

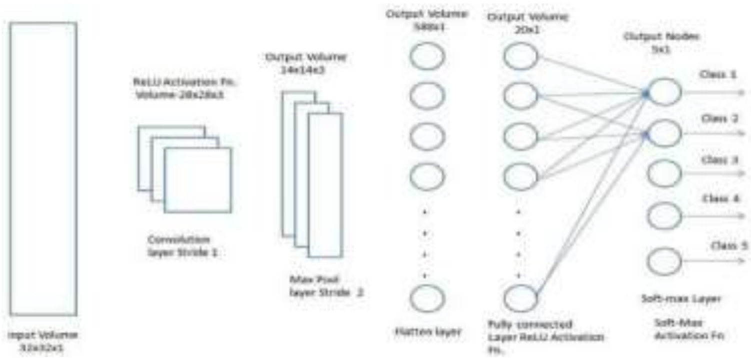


Fig 1.3 Fully Connected model

3. PRE-PROCESSING AND TRAINING THE MODEL:

The dataset is pre-processed such as Image reshaping, resizing and conversion to an array form on the training data. Similar processing is also done on the testing data. A dataset consisting of about 6 different objects (Book, Bottle, Chair, Duster, Pen, Tiffin box) is obtained, out of which any image can be used as a test image for the software.

The training dataset is used to train the model (CNN) so that it can identify the test image and the object it has. CNN has different layers that are Dense, Dropout, Activation, Flatten, Convolution2D, and MaxPooling2D. After the model is trained successfully, the software can identify the object contained in the dataset. After successful training and pre-processing, comparison of the test image and trained model takes place to classify the object.

```
Model: "sequential"
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
conv2d (Conv2D)              (None, 21, 21, 32)        896
_________________________________________________________________
max_pooling2d (MaxPooling2D) (None, 10, 10, 32)        0
_________________________________________________________________
flatten (Flatten)            (None, 3200)              0
_________________________________________________________________
dense (Dense)                (None, 128)               409728
_________________________________________________________________
dense_1 (Dense)              (None, 6)                 774
=================================================================
Total params: 411,398
Trainable params: 411,398
Non-trainable params: 0
```

Fig 1.4 CNN Model Summary details

II. LITERATURE SURVEY

Yunhang Shen, Rongrong Ji, Xi Li, Changhu Wang, Xuelong Li, "Weakly Supervised Object Detection via Object-Specific Pixel Gradient" 2018 [1] - It uses Weakly Supervised Object Detection (WSD) algorithm, which utilizes a novel Object-Specific Pixel Gradient (OPG) map to model the relation between image pixels and classification scores of individual object categories. High values in the OPG map indicate that the corresponding pixels are discriminative. An ICM scheme is further proposed to extract bounding boxes based on the OPG map, which iteratively mines all components or instances of the target object. It first extracts an OPG map to reveal the contributions of individual pixels to a given object category, upon which an iterative mining scheme is further introduced to extract instances or components of this object.

Shantanu Deshmukh, Teng-Sheng Moh, "Fine Object Detection in Automated Solar Panel Layout Generation" 2018 [2] - A solar panel layout is a diagram of the roof, with the roof edges and obstacles marked. Using the layout diagram, one can determine the exact location to install solar panels and can also predict approximate installation cost and power generation. Using a fine object detector, it can highlight the exact edges of an obstacle, and then use it to create a layout which will be presented to the user as a default auto generated layout. Object detection API like TensorFlow, YOLO, SSD are used to detect the objects after that edge detection algorithms like Prewitt operator, Laplacian operator canny edge detector are used with the detection API.

Nashwan Adnan Othman, Mehmet Umut Salur, Mehmet Karakose, Ilhan Aydin, "An Embedded Real-Time Object Detection and Measurement of its Size" 2018 [3] - A powerful real time object measurement method was proposed for industrial systems. Computer Vision used to detect and measure objects. The system consists of two parts which are object detection and object measurement. In the first part, raspberry pi camera used to achieve the frames. In the second part, computer vision module will be applied to the captured frames to determine the objects, then, we will measure each object. The detected object of the current frame immediately will be processed to extract dimensions of objects. Objects are detected via canny edge detector algorithm. we need to preprocess our image. The camera will capture a frame and the frame will convert to grayscale to increase quickness and accuracy. The system can detect and measure objects in a real time video. After the object has been detected by using canny edge detector, the size is obtained for each object by using OpenCV functions.

Stefania Matteoli, Laura Zotta, Marco Diani, Giovanni Corsini, "POSEIDON: An Analytical End-to-End Performance Prediction Model for Submerged Object Detection and Recognition by Lidar Fluorosensors in the Marine Environment" 2017 [4] –
Performance prediction model for Submerged object detection and recognition by lidar fluorosensors in the marine environment (POSEIDON) is to model the end-to-end underwater object recognition chain in specific operational scenarios and to abstract the user from implementation details of complex models for signal propagation, lidar sensor, and water optical behavior, with the final goal of providing estimated performance measures. The purpose was to assist the user during tasks such as lidar system design and mission planning by

assessing the parameter tradeoffs as concerns the ability of the system to recognize objects submerged in the water column by their fluorescence spectral signature. This is accomplished by making use of three modules: 1) the simulator module; 2) the signal processing module; and 3) the performance prediction module. Test cases examples encompassing simulated operational scenarios of interest obtained by varying lidar system deployment, sensor noise, object kind and depth, water, and illumination conditions have been described and examined, showing how POSEIDON can be used in the analysis of underwater object recognition performance prediction.

Hao Chen, Yali Wang, Guoyou Wang, Xiang Bai, and Yu Qiao, "Progressive Object Transfer Detection" 2019 [5] - It focused on detecting salient objects from a video and proposed a framework using STD features together with SpatioTemporal Conditional Random Field (STCRF). This method takes into account temporal information in a video as much as possible in different ways, namely, feature extraction and saliency computation. The proposed SpatioTemporal Deep (STD) feature utilizes local and global contexts in both spatial and temporal domains. STCRF is capable to capture temporal consistency of regions over frames and spatial relationship between regions. This method significantly outperforms state-of-the-art methods on publicly available datasets. The method was also applied to the video object segmentation task, showing that the method outperforms existing unsupervised VOS methods on the DAVIS dataset. An input video was segmented at multiple scale levels and compute a saliency map at each scale level at each frame, and then aggregate all saliency maps at different scale levels at each frame into a final saliency map.

Huaxin Xiao, Jiashi Feng, Yunchao Wei, Maojun Zhang, Shuicheng Yan, "Deep Salient Object Detection with Dense Connections and Distraction Diagnosis" 2018 [6] - A novel salient object detection network and a new self-diagnosis component for distraction localization and removal. The overall architecture of the model consists of two components: saliency detection and distraction detection. The S-Net encodes an input image into a pixel-level saliency map where each element indicates the probability of the corresponding pixel belonging to a salient object. The SNet employs a fully convolutional architecture to make the dense saliency prediction. First, the S-Net passes the input image through a pre-trained image classification network and extracts features at multiple levels. A carefully designed context block within S-Net then encodes these features into saliency-diagnostic context features. Such encoded features are up-sampled and concatenated with lower level features recursively via newly proposed short- and long-range dense connections until the resulted features have the same resolution as the input image.

Xinyi Zhou, Wei Gong, WenLong Fu, Fengtong Du, "Application of Deep Learning in Object Detection" 2017 [7] - This paper deals with the field of computer vision, mainly for the application of deep learning in object detection task. On the one hand, there is a simple summary of the datasets and deep learning algorithms commonly used in computer vision. On the other hand, a new dataset is built according to those commonly used datasets, and choose one of the network called faster r-cnn to work on this new dataset. Through the experiment to strengthen the understanding of these networks, and through the analysis of the results learn the importance of deep learning technology, and the importance of the dataset for deep learning.

Mukesh Tiwari, Dr. Rakesh Singhai, "A Review of Detection and Tracking of Object from Image and Video Sequences" 2017 [8]- Object detection and tracking is one of the critical areas of research due to routine change in motion of object and variation in scene size, occlusions, appearance variations, and ego-motion and illumination changes. Specifically, feature selection is the vital role in object tracking. It is related to many real time applications like vehicle perception, video surveillance and so on. In order to overcome the issue of detection, tracking related to object movement and appearance. Most of the algorithm focuses on the tracking algorithm to smoothen the video sequence. On the other hand, few methods use the prior available information about object shape, color, texture and so on. Tracking algorithm which combines above stated parameters of objects is discussed and analyzed in this research. The goal of this paper is to analyze and review the previous approach towards object tracking and detection using video sequences through different phases. Also, identify the gap and suggest a new approach to improve the tracking of object over video frame.

Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, and Xindong Wu, "Object Detection with Deep Learning: A Review" 2019 [9] - Due to object detection's close relationship with video analysis and image understanding, it has attracted much research attention in recent years. Traditional object detection methods are built on handcrafted features and shallow trainable architectures. Their performance easily stagnates by constructing complex ensembles which combine multiple low-level image features with high-level context from object detectors and scene classifiers. With the rapid development in deep learning, more powerful tools, which are able to learn semantic, high-level, deeper features, are introduced to address the problems existing in traditional architectures. These models behave differently in network architecture, training strategy and optimization function, etc. In this paper, we provide a review on deep learning-based object detection frameworks. Our review begins with a brief introduction on the history of deep learning and its representative tool, namely Convolutional Neural Network (CNN). Then we focus on typical generic object detection architectures along with some modifications and useful tricks to improve detection performance further. As distinct specific detection tasks exhibit different characteristics, we also briefly survey several specific tasks, including salient object detection, face detection and pedestrian detection. Experimental analyses are also provided to compare various methods and draw some meaningful conclusions. Finally, several promising directions and tasks are provided to serve as guidelines for future work in both object detection and relevant neural network -based learning systems.

Sandeep Kumar, Aman Balyan, Manvi Chawla, "Object Detection and Recognition in Images" 2017 [10] - Object Recognition is a technology in the field of computer vision. It is considered to be one of the difficult and challenging tasks in computer vision. Many approaches have been proposed in the past, and a model with a new approach which is not only fast but also reliable. Easynet model has been compared with various other models as well. Easynet model looks at the whole image at test time so its predictions are informed by global context. At the prediction time, our model generates scores for the presence of the object in a particular category. It makes predictions with a Single network evaluation. Here object detection is a regression problem to spatially separated bounding boxes and associated class probabilities.

III. PROPOSED SYSTEM

A. Problem Statement

We planned to design deep learning technique so that a person with lesser expertise in software should also be able to use it easily. It proposed system to detect the object and classify it. It explains about the experimental analysis of our methodology. Different objects are collected (Book, Bottle, Chair, Duster, pen and tiffin box). Splitting of each object images into three parts (Training, Testing, Validation) is done. The Training Dataset is used to train and build the model. The primary attributes of the image are relied upon the shape and texture-oriented features. Our model which is built on CNN algorithm takes the input test image and it extracts the features of the input image and classifies the objects into its respective type as all objects have different features.

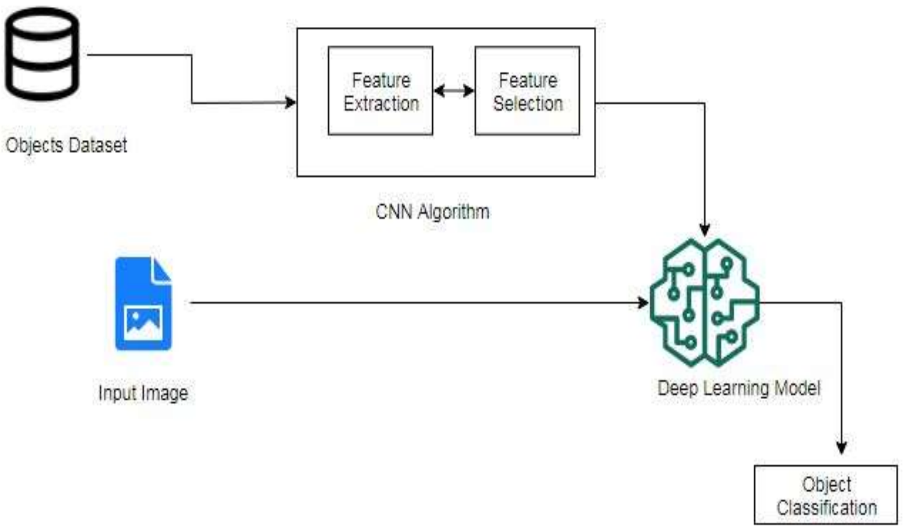


Fig 1.5 System Architecture

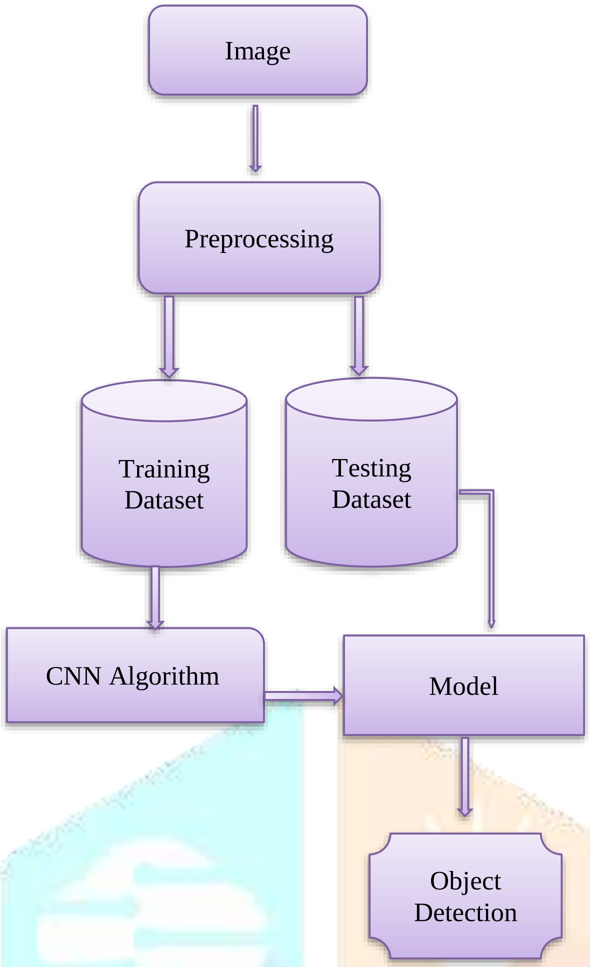


Fig 1.6 Work flow diagram of the Class Object detection

IV. RESULTS AND DISCUSSIONS

We proposed the system to detect objects and classify them based on their respective classes. The implementation has been done with the help of python language and its necessary deep learning modules/packages to curtail computational complexity. The following graphs accounts for the model's accuracy and loss through several experimental results.

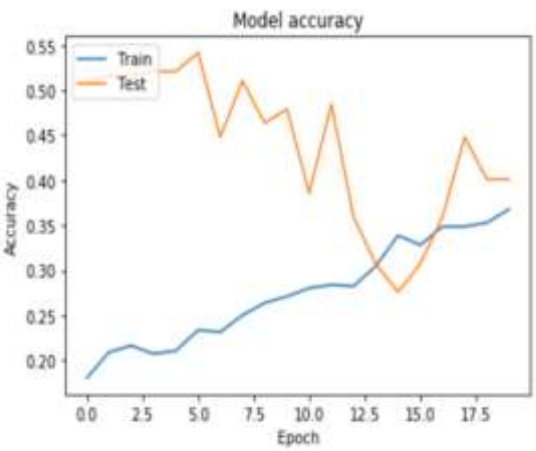


Fig 1.7 depicts the model accuracy as a plot of Accuracy vs Epoch between Training data and Testing data.

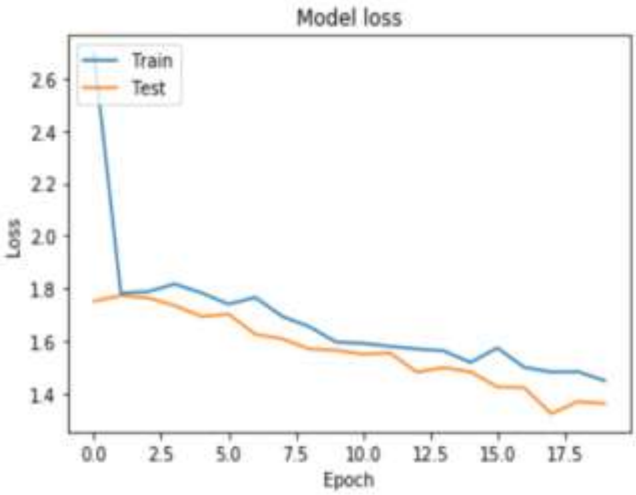


Fig 1.8 depicts model loss as a plot of Loss vs Epoch between Training data and Testing data.

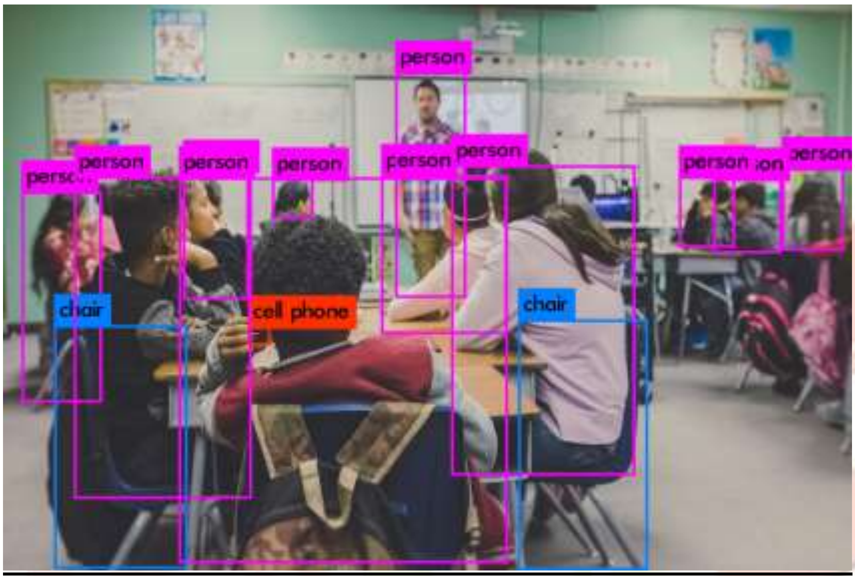


Fig 1.9 The model's prediction as a visual representation (Classroom Scenario).

The above model displayed an accuracy of 99.5% with 0.5% residual as error.

V. CONCLUSION AND FUTURE WORK

It focused how image from the given dataset (trained dataset) and past data set used predict the object features using CNN model. This brings some of the insights about objects classification. The object classification will be accurate according to our trained dataset. Object detection and classifying it will be very useful for security check purposes. The project can be further extended to several real-time objects. Taking into account of occlusion and providing algorithms to get rid off it, will be a considerable work in the future.

REFERENCES

[1] Yunhang Shen, Rongrong Ji, Xi Li, Changhu Wang, Xuelong Li, "Weakly Supervised Object Detection via Object-Specific Pixel Gradient" IEEE Transactions on Neural Networks and Learning Systems 2018

[2] Shantanu Deshmukh, Teng-Sheng Moh," Fine Object Detection in Automated Solar Panel Layout Generation" 17th IEEE International Conference on Machine Learning and Applications 2018

[3] Nashwan Adnan Othman, Mehmet Umut Salur, Mehmet Karakose, Ilhan Aydin, "An Embedded Real-Time Object Detection and Measurement of its Size" International Conference on Artificial Intelligence and Data Processing (IDAP) 2018

[4] Stefania Matteoli, Laura Zotta, Marco Diani, Giovanni Corsini, "POSEIDON: An Analytical End-to-End Performance Prediction Model for Submerged Object Detection and Recognition by Lidar Fluorosensors in the Marine Environment" IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, VOL. 10, NO. 11, NOVEMBER 2017

[5] Hao Chen, Yali Wang, Guoyou Wang, Xiang Bai, and Yu Qiao, "Progressive Object Transfer Detection" IEEE Transactions on Image Processing PP(99):1-1 · September 2019

[6] Huaxin Xiao, Jiashi Feng, Yunchao Wei, Maojun Zhang, Shuicheng Yan, "Deep Salient Object Detection with Dense Connections and Distraction Diagnosis" IEEE Transactions on Multimedia Volume: 20 , Issue: 12 , Dec. 2018

[7] Xinyi Zhou, Wei Gong, WenLong Fu, Fengtong Du, "Application of Deep Learning in Object Detection" IEEE/ACIS 16th International Conference on Computer and Information Science (ICIS).

[8] Mukesh Tiwari, Dr. Rakesh Singhai, "A Review of Detection and Tracking of Object from Image and Video Sequences" International Journal of Computational Intelligence Research ISSN 0973-1873 Volume 13, Number 5 (2017), pp. 745-765

[9] Zhong-Qiu Zhao, Peng Zheng, Shou-tao Xu, and Xindong Wu, "Object Detection with Deep Learning: A Review" IEEE Transactions on Neural Networks and Learning Systems 2018

[10] Sandeep Kumar, Aman Balyan, Manvi Chawla, "Object Detection and Recognition in Images" International Journal of Engineering Development and Research (IJEDR) DECEMBER 2017