OBJECT DETECTION USING SSD AND EFFICIENT NET B7 AS BASE NETWORK

A MAJOR PROJECT-II REPORT

SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF

MASTER OF TECHNOLOGY
IN
INFORMATION SYSTEMS

Submitted by:
ASHISH KUMAR
2K22/ISY/05

Under the supervision of

DR. PRIYANKA MEEL

DEPARTMENT OF INFORMATION AND TECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

May, 2024

DEPARTMENT OF INFORMATION TECHNOLOGY
DELHI TECHNOLOGICAL UNIVERSITY
(Formerly Delhi College of Engineering)
Bawana Road, Delhi-110042

ACKNOWLEDGEMENT

I would like to express my sincere gratitude towards my supervisor Dr. Priyanka Meel for providing her invaluable guidance, comments, and suggestions throughout the course of the project.

The results of this thesis would not have been possible without support from all who directly or indirectly, have lent their hand throughout the course of the project. I would like to thank my parents and faculties of the Department of Information Technology, Delhi Technological University, for their kind cooperation and encouragement which helped me complete this thesis. I hope that this project will serve its purpose to the fullest extent possible.

ASHISH KUMAR
2K22/ISY/05

DEPARTMENT OF INFORMATION TECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CANDIDATE'S DECLARATION

I, Ashish Kumar, 2K22/ISY/05 student of M.Tech in Information Systems, hereby declare that the Major Project-II dissertation titled "**OBJECT DETECTION USING SSD AND EFFICIENT NET B7 AS BASE NETWORK**" which is submitted by me to the Department of Information Technology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is original and not copied from any source without proper citation. This work has not previously formed the basis for the award of any degree, Diploma Associateship, Fellowship, or other similar title or recognition.

Place: Delhi **ASHISH KUMAR**

Date: **2K22/ISY/05**

DEPARTMENT OF INFORMATION TECHNOLOGY

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Bawana Road, Delhi-110042

CERTIFICATE

I hereby certify that the Major Project-II dissertation titled "**OBJECT DETECTION USING SSD AND EFFICIENT NET B7 AS BASE NETWORK**" which is submitted by Ashish Kumar, 2K22/ISY/05, Department of Information Technology, Delhi Technological University, Delhi in partial fulfillment of the requirement for the award of the degree of Master of Technology, is a record of the project work carried out by the student under my supervision. To the best of my knowledge, this work has not been submitted inpart or full for any degree or diploma to this University or elsewhere.

Place: Delhi
Date:

DR. PRIYANKA MEEL
SUPERVISOR
Asst. Professor
Department of Information Technology
DELHI TECHNOLOGICAL UNIVERSITY

ABSTRACT

This thesis addresses the complex challenge of image classification and object detection using advanced deep learning techniques. The research focuses on utilizing EfficientNet-B7 for image classification and implementing the Single Shot MultiBox Detector (SSD) for object detection. The study initially evaluates the performance of four neural networks—VGG-16, ResNet-50, AlexNet, and EfficientNet-B7—on an animal dataset containing 31 distinct classes. For the image classification task, each model was trained and tested to determine its accuracy in recognizing various animal species within the dataset. Among the evaluated models, EfficientNet-B7 emerged as the superior performer, achieving the highest accuracy in both training and testing phases. This outstanding performance underscores the model's capability to effectively handle complex classification tasks and its potential for broader applications in the field of computer vision. Building on the success of EfficientNet-B7 in image classification, the research proceeded to integrate this network as the foundational layer for the SSD framework in object detection tasks. The SSD leverages the feature extraction capabilities of EfficientNet-B7 to detect and localize objects within images. The combination of EfficientNet-B7's robust feature extraction and SSD's efficient detection mechanism resulted in an object detection accuracy of approximately 78%. The findings of this thesis highlight the efficacy of EfficientNet-B7 in both image classification and object detection domains. By demonstrating superior performance in classification tasks and achieving notable accuracy in object detection, this research contributes valuable insights into the application of deep learning models for complex computer vision challenges. The study provides a comprehensive evaluation of neural network models and offers a compelling case for the adoption of EfficientNet-B7 and SSD in practical image analysis and object detection scenarios. This work lays the groundwork for future research and development in the field, emphasizing the importance of selecting and optimizing deep learning models for specific tasks to achieve the best possible outcomes.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER -1
INTRODUCTION

Object detection is a basic task in computer vision and has seen a lot of progress recently because of new deep learning methods and creative algorithms. This delves into several research papers that introduce pioneering methodologies for object detection across specialized domains. Each paper presents novel techniques designed to overcome unique challenges inherent in their respective application areas. The methodologies employed in these papers encompass a wide range of approaches, including deep learning architectures, algorithmic innovations, and dataset curation tailored to specific tasks.

In the realm of precision agriculture, researchers have developed sophisticated object detection systems tailored to identify crop weeds amidst cultivated fields. These systems utilize advanced deep learning models, such as YOLOv5 and RetinaNet, and employ curated datasets with bounding box annotations to train and evaluate model performance accurately. Similarly, in underwater exploration, object detection methodologies leverage deep learning architectures such as YOLOv8, alongside innovative techniques like the phase-shifting coder (PSC), to detect and classify underwater objects effectively. These approaches are validated through extensive experimentation on specialized datasets collected under diverse environmental conditions.

One most important use comes in the traffic management and monitoring. It has a wide range of application in enhancing safety and overall functionality of traffic management. Object detection helps in analysing the flow of traffic by detecting and counting vehicles, which is crucial for traffic management and planning. This data can help improve traffic light schedules and decrease traffic Also, by identifying incidents in real-time, object detection systems can alert authorities to accidents or unusual traffic patterns, enabling quick response and potentially preventing further accidents. Above all these, , there are many other applications like traffic light Enforcement, Pedestrian Safety, Smart Traffic Lights, Public Transport Monitoring, Parking Management

Furthermore, in the domain of flying object detection, researchers introduce real-time detection system specifically designed to identify aerial entities using algorithms such as YOLOv8. These systems rely on curated datasets comprising various classes of flying objects and employ techniques like differentiable angle coding to accurately predict object orientations.

Additionally, in oriented object detection tasks, novel approaches such as the Phase-Shifting Coder (PSC) are introduced, leveraging differentiable angle coding and addressing challenges related to boundary discontinuity and object shape variability.

In the field like Autonomous Vehicles for Obstacle Detection and Traffic Sign Recognition, Surveillance and Security for Intruder Detection and Crowd Monitoring, Smart Cities traffic and parking management, Retail and Inventory Management, Healthcare Medical Imaging and Patient Monitoring, Robotics Object Manipulation and Navigation Object detection systems combining Single Shot Multibox Detector (SSD) with EfficientNet are increasingly popular due to their balance of speed and accuracy.

1.1 Background

Object detection typically involves both classification (determining the class of an object) and localization (finding the spatial location of an object in an image). Object classification, on the other hand, speaks about the task of allocating a class label to an entire image or a specific region within an image.

Two Stage Classification: Feature extraction from the input data is the first step of a conventional two-stage object classification method. Convolutional neural networks and other techniques are frequently used for this (CNNs). A classifier that uses the retrieved characteristics to assign class labels makes up the second stage. Two-stage detectors focus on accuracy by first generating region proposals and then refining object detection through classification and localization.

Single Stage Classification: One network design is usually used in a one-stage object classification strategy, which receives input data and produces class predictions immediately. Deep neural networks, like CNNs, are frequently used for this without a separate feature extraction and classification step. One-stage detectors are known for their efficiency, as they process images in a single forward pass. In real time applications, when there speed is dominated in object detection instead of accuracy, this one stage detection is beneficial.

Two Stage Object Detection:

A Region based Proposal Network is usually used in the initial stage of a two-stage object identification strategy to generate region suggestions. On the basis of these suggested regions, the second step subsequently carries out object categorization and bounding box regression.

Single Stage Object Detection:

A one-stage object detection technique generates region proposals explicitly and completes the item identification process in a single step, encompassing both localization and classification.

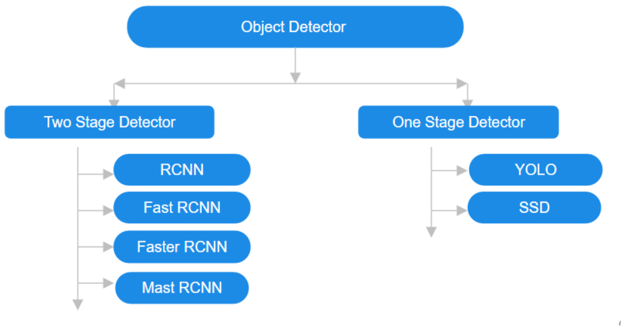

Fig-1.1 Object Detection Methods

1.2 PROBLEM DEFINITION

Object detection is found as one of the valuable task in the field of Computer Vision, as it includes identifying and localizing of objects within an image or video, called classification and identification. Irrespective of many advancement in the domain, a complete solution is still unavailable due to diverse range of application scenarios and their unique requirements. This is primarily due to the complexity and variability in object sizes, shapes, lighting conditions and occlusion. The real time processing requires necessity in the application development of various object detection methods.

Despite of significant advancements in models, there are several challenges that affects the performance and application of detection algorithms and models. One important role in detection is image classification, i.e. the image belongs to which class. And other is localization, i.e. what are the scale variation, aspect ratio variation, occlusion and background of image.

This thesis aims to cover the challenge of image classification and object detection using deep learning methods. Efficient Net B7 is used for the image classification with the single shot multi-box detector for object detection. There are various model already proposed in modern technologies like YOLO and its many versions available. When it comes for the speed SSD and YOLO are almost equivalent, but in terms of Accuracy SSD performs better.

The specific challenge and research questions can be addressed when we use Efficient Net B7 along with SSD for object detection and why efficient net b7.

- What are the modifications to the SSD architecture necessary to fully leverage the feature extraction capabilities of efficient net – B7 and how the network is utilized to improve detection accuracy for object of varying sizes?
- What are the most effective methods to prune the network, so that we can reduce computational requirements while maintaining performance?
- What are the trade-offs between model complexity and detection speed are applicable for real world applications?
- What evaluation metrics and techniques can be used to assess the performance of the model, considering the complex nature of SSD?

By addressing the above mentioned challenges and research questions, the combination of efficient net – B7 with SSD has the potential to advances the state object detection, providing a powerful tool for various practical applications that require high accuracy and efficiency.

1.3 Motivation

The modern technology of computer vision have a significant impact in numerous fields in which object detection is one of the most widely adopted. Many real time applications like autonomous driving, satellite image analysis, surveillance are totally dependent on this. So to get the accurate, fast and effective results in applications presents a compelling challenge.

I am more motivated towards the Balancing accuracy and Efficiency in the detection where the efficient network which is known for its image classification benchmark, offers an exciting opportunity to trigger this balance. The scaling method of this dense network provides an effective scaling of model dimension.

Also the versatility and Robustness to handle wide range of challenge including objects of various aspect ratios, object scales and handling them under same model is another one.

1.4 Objective

The main objective is to develop a model based on deep learning approach that can classify the objects accurately and then pass this classified image through Single Shot Multi-Box detector to get the localization of objects within the image or video frames, aiming to get a balance between high accuracy, speed and computational efficiency. This thesis aims to address generalization and adaptability of the proposed model across different application domain and dataset, ensuring its adaptability to real world applications. The main Objective, if quoted in bullet points are:

- To provide a comprehensive review of current research work by utilizing deep learning approach for detection using one of the largest neural network.
- To design and implement a deep learning based model for classification of object using Efficient Net-B7 and SSD for Localisation of object.
- To evaluate the model performance with other neural network that SSD uses for classification, and compare them.

1.5 Challenges

Due to inherent complexities of both components and their integration, there comes a lot of challenge. Model integration is one of the biggest challenge. Adjusting the SSD architecture to fully leverage the features extracted by efficient net – b7 involves complex architectural changes and hyper parameter tuning. Hardware constraints are another challenge, because efficient net - b7 training requires a high computational demands and time consuming. Dataset is also a challenge in our case, because, training requires a properly labelled dataset and it should be sufficiently diverse.

Chapter - 2

Related Work

Object detection is a task comes under the domain of computer vision, which aims to identify and locate the objects within an image. Over the past many years, various approaches have been developed, evolving from traditional techniques to modern deep learning methods.

Traditional methods relied mostly on Sliding window approach where a classifier scans the image at various scales and positions. One another approach is the use of Histogram of Oriented Gradients combined with SVM. Region Based method uses the concept of selective search and Edge boxes for detection, where those regions which has high edge density and edge energy with bounding boxes present are proposed.

R-CNN method developed by “Ross Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik” in 2014 were the first to create the ground breaking object identification system known as R-CNN, or Region based Convolutional Neural Network. R-CNN uses a few steps to search items in an image [24].

Region Proposal Network (RPN): Initially, a collection of potential bounding boxes, or region proposals, that may include objects, are produced by the algorithm. RPN uses an algorithm called selective search algorithm for the prediction of bounding boxes. Selective search takes into account several indicators like as colors, texture, and size to quickly suggest a large number of bounding boxes that are likely to include.

Feature Extraction: R-CNN extracts fixed-size feature vectors for every suggested area. It extracts deep features from each area separately using a pre-trained convolutional neural network (CNN), like Alex Net, VGG, or Residual Net.

Object Classification: Each region's collected characteristics are sent into a different classifier (SVM) that has been trained to identify objects and differentiate between several classes. These classifiers give each suggested location a class label and forecast the likelihood that items will be present.

Bounding Box Refinement: Following categorization, each suggested region's bounding boxes are adjusted to more closely match the object's true placement inside the region. This refining procedure aids in raising object localization accuracy.

Non-Max Suppression (NMS): R-CNN uses a method called non max suppression get rid of repeated findings and choose the most reliable predictions. Bounding boxes having a maximum overlap (IoU), or intersection over union) along with other bounding boxes that have higher confidence ratings are suppressed by NMS.

Fast R-CNN model is introduced by Ross Girshick in 2015. It's an improvement over the previous R-CNN model, addressing some of its limitations such as slow inference speed. It introduces the RPN, which generates "art boxes" directly from feature maps produced by a CNN. As a result, external region proposal techniques like R-CNN's Selective Search are no longer necessary. Fast R-CNN can be combined with a Feature Pyramid Network (FPN) which helps in handling objects at different scales effectively. FPN generates feature pyramids with semantic information at multiple scales, aiding in accurate object detection. It uses an area designated (RoI) pooling layer, effectively gathers important information from the CNN and each proposed area has its own set of feature maps. This layer helps in aligning spatial locations inside the suggested areas. While training Fast R-CNN, it fine-tunes a loss function that handles multiple tasks at once, and includes terms for classification loss (e.g., soft max or sigmoid cross-entropy) and localization loss (e.g., Smooth L1 loss) for bounding box regression. Fast R-CNN achieves significant improvements in inference speed compared to R-CNN, making it more practical for real-time object detection applications. It maintains competitive accuracy while being computationally efficient.

Faster R-CNN is an improvement of the Fast R-CNN architecture, which addresses some of its limitations while achieving faster computation speeds [2]. Faster R-CNN was suggested by Shaoqing Ren et al. in 2015. It presents a Region Proposal Network (RPN) that is comparable to the detection network similarly like convolutional layers. Object detection comes after a distinct region proposal generating stage in Fast R-CNN, while Faster R-CNN integrates the RPN into the detection network [35]. The model's ability to simultaneously produce region suggestions and execute object recognition, results in its notable speed increases over Fast R-CNN [1]. Advantages over Fast R-CNN includes Improved Speed, End-to-end training, simplicity.

In the paper "Towards Real-Time Object Detection with Region Proposal Networks" by "Ren, S., He, K., Girshick, R., & Sun, J.", proposed the application of detection of infrastructure damage in aerial images captured by drones. Faster R-CNN was applied to aerial surveying for identifying damaged infrastructure post-disaster. It achieved an mAP of 76.8% on a custom dataset, processing images in near real-time, which is crucial for timely disaster response.

In the paper by “Girshick, R.” in 2015, the application of Monitoring customer behaviour and product placement in retail stores. Fast R-CNN was utilized for detecting and tracking customer movements and interactions with products in retail environments. It achieved an mAP of 82.1% on the Retail Interaction dataset, providing valuable insights for optimizing store layouts and product placements [35].

Mask R-CNN is an advancement of the Faster R-CNN architecture, which itself is an improvement over the previous Fast R-CNN model. Mask R-CNN extends Faster R-CNN by including a branch for segmentation in mask prediction with the current branches for object recognition and classification [24]. The architecture has three major components: a base network (like ResNet), a region proposal network (RPN). It has two separate parts that work together to find objects and outline them precisely called finding objects (object detection) and outlining them clearly (instance segmentation). Faster R-CNN gives us anchor boxes and the likelihood of an object's class for everything it finds, Mask R-CNN extends this by additionally producing a dual mask for each detected object, marking the pixels that belong to it. Due to the additional segmentation task, Mask R-CNN is slightly more complex than Faster R-CNN [24].

In the paper by“He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017)”, the application of Detecting and tracking players and equipment in sports footage, have been described. Mask R-CNN used for analyzing basketball game footage achieved an mAP of 88.3% for player detection and equipment tracking, enabling detailed performance analytics and strategy development.

Single Shot Multi box Detector (SSD) is a popular and an efficient deep learning model for object detection in images [7]. It was introduced by Liu et al. in 2016 and has since become widely used due to its balance of accuracy and speed. SSD is used for detecting and localizing objects within images. SSD operates by using a single neural network to generate predictions across multiple scales of feature maps. SSD model is able to find objects with varied sizes and aspect ratios because these feature maps record semantic information at multiple levels. SSD uses anchor boxes, at different positions and scales in the feature maps. SSD gives the result a bit slow but more accurate even in real time application than any other model [8].

In the paper by “Tan, M., Pang, R., & Le, Q. V”, explains the real time detection of vehicles, pedestrians and traffic signs. The paper demonstrates and efficient net – B3 model which was deployed on an autonomous platform achieved an mAP of 78.5% on the KITTI dataset with an inference time of 25ms per frame. It ensures, robust performance in dynamic driving environment [49].

“Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q.” in 2021 has proposed the detection anomalies in medical image, such as tumours in radiographs in his paper “Densely connected convolution networks”. The integration of EfficientNet-B4 as the backbone in SSD for detecting lung nodules in chest X-rays resulted in an mAP of 85.1%, significantly outperforming traditional methods. The model's efficient architecture enabled rapid analysis, crucial for timely medical diagnosis [51].

“Tan, M., & Le, Q. V” in his paper “Efficient Net: Rethinking Model Scaling for Convolutional Neural Networks” has applications in Identifying and classifying crops and detecting diseases in plants. An EfficientNet-B3 SSD model applied to agricultural monitoring achieved an mAP of 80.7% on the Plant Village dataset. The model efficiently processed high-resolution images, aiding in precise crop management and disease detection [52].

In Surveillance system, Zhang, X., Zhou, X., Lin, M., & Sun, J. in his paper An Extremely Efficient Convolutional Neural Network for Mobile Devices, demonstrated the application of Monitoring and detecting suspicious activities in crowded areas. In a smart city surveillance project, the EfficientNet-B5 SSD model demonstrated superior performance with an mAP of 82.3% on the Surveillance Objects dataset. The model maintained real-time detection capabilities, processing 30 frames per second (FPS) on a standard GPU [50].

You Only Look Once (YOLO) is a famous deep learning model for real time detection of objects in images, introduced by “Joseph Redmon et al.” in 2016 and has since go through various improvements and versions, including YOLOv2, YOLOv3, and YOLOv4. YOLO uses one neural network to predict where objects are in an image and how likely they are to be certain classes, all by looking at the whole picture at once [12]. YOLO, divides the entire input image into a grid and makes predictions for each cell in the grid. where objects are and how likely they are to belong to specific classes. Due to this grid approach, it allows YOLO to detect multiple object across the image in a single pass [12][13]. The accuracy of YOLO gets improved because of the use of bounding boxes. These boxes with varying dimensions and aspect ratios are known as anchor boxes, and YOLO predicts offsets and confidence scores for these anchor boxes. YOLO includes multiple prediction layers at different scales in the network architecture [14]. Different layers in YOLO are in charge of finding objects of different sizes and scales, enabling YOLO to handle objects of various sizes effectively [15].

The paper [25,26] meticulously examines the application of YOLOv7, an advanced object detection algorithm, within the realm of precision agriculture, with a specific focus on identifying crop weeds [25] from images captured by Unmanned Aerial Vehicles (UAVs). Through a comprehensive evaluation of YOLOv7's performance on a meticulously curated dataset tailored for this purpose, the study elucidates its efficacy in accurately discerning weeds amidst cultivated crops. This thorough analysis not only underscores the algorithm's robustness but also highlights its potential as a cornerstone technology in modern agricultural practices.

The significance of such technological advancements in precision agriculture is profoundly emphasized throughout the paper, particularly in relation to their implications for sustainable farming practices. By enabling precise weed detection [25], YOLOv7 emerges as a promising tool for facilitating more efficient crop management strategies and reducing reliance on herbicides, thereby aligning with the growing imperative for environmentally conscious agricultural approaches. Moreover, the paper's comparative analysis, juxtaposing YOLOv7 against alternative object detection techniques, further accentuates its superior performance and practical viability in real-world agricultural contexts, solidifying its position as a transformative tool for farmers and agronomists seeking to enhance crop yield and sustainability.

In the field of intelligent underwater vehicles [27], object detection stands as a cornerstone technology pivotal for a myriad of applications ranging from ocean exploration to salvage operations and military endeavours. In the paper author introduce a pioneering object detection approach named TCYOLO[27,28], representing a notable enhancement over the original YOLOv5 model. TC-YOLO integrates Transformer self-attention and coordinate attention mechanisms within its architecture, thereby enhancing feature extraction for underwater objects. Additionally, the authors incorporate an adaptive histogram equalization algorithm for image enhancement and leverage an optimal transport scheme for label assignment during training. The overarching goal of TC-YOLO is to address the multifaceted challenges posed by blurry underwater images, small and dense targets, and the constrained computational resources typical of underwater platforms. Through validation on the RUIE2020 dataset, TC-YOLO demonstrates superior performance compared to YOLOv5s and analogous networks in detecting underwater objects, all while maintaining a compact model size and computational efficiency conducive to mobile applications.

The significance of this research is profound, as it holds the potential to substantially augment the capabilities of intelligent underwater vehicles by furnishing them with more accurate and robust object detection tools. The integration of attention mechanisms and image enhancement techniques marks a substantial leap forward in the field, enabling enhanced performance even in the challenging underwater environment characterized by limited visibility and complex target characteristics. The findings presented in the paper have far-reaching implications for the future of underwater exploration and autonomous vehicle technology, promising to bolster efficiency, reliability, and safety in various underwater applications [28], ranging from scientific research endeavours to industrial operations and defence missions.

The paper presents a meticulous evaluation of various deep learning object detectors for the specific task of weed detection in cotton fields [29,30], a critical component in the development of non-chemical weed control strategies. The researchers curated a comprehensive three-class weed dataset with bounding box annotations, comprising 848 images collected under diverse field conditions. Across 13 different deep learning models tested, including both one-stage and two-stage object detectors such as YOLOv5, Retina Net, Efficient Net, Fast RCNN, and Faster RCNN, Retina Net (R101-FPN) [28] emerged as the top performer, achieving the highest detection accuracy with a mean average precision (mAP@0.50) of 79.98%. Notably, YOLOv5n exhibited promise for real-time deployment on resource-constrained devices due to its compact model size and fast inference time, while still maintaining a competitive detection accuracy of 76.58% mAP@0.50. Additionally, the study revealed that data augmentation techniques, including geometric and colour transformations, yielded significant improvements in model accuracy, enhancing performance by up to 4.2%.

Chapter – 3

Proposed Method

This paper will analyze the accuracy on classification of image on some most popular modern pre trained networks like VGG, Efficient net B7, Alex net and ResNet. Then we will use the best classification accuracy neural network in SSD for object detection. The first detailed survey on the application of deep learning for object classification was conducted around 2012-2013, shortly after the deep learning revolution gained momentum with the success of CNN in image classification tasks. The survey paper "Deep Learning in Object Detection and Recognition" by Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., & Fu, C. Y. (2016) [23] provided a comprehensive overview of deep learning techniques applied to object detection and recognition tasks. It provided an extensive overview of the various deep learning techniques and methodologies used for object classification tasks, covering topics such as image pre-processing, feature extraction, model architectures, and datasets used in the field. The survey aimed to consolidate the existing research and offer valuable perspectives on cutting-edge methodologies for utilizing deep learning in object classification tasks.

3.1 Neural Network Architecture

VGG :- VGG is famous for being simple and includes several layers with small areas they pay attention to (3x3 CNN filters) and followed by highest-value pooling layers [11]. The VGG16 and VGG19 architectures were first appear in the research paper of "Very Deep Convolutional Networks for Large-Scale Image Recognition" by “Simonyan and Zisserman” in 2014. With proper training and optimization, VGG have achieve an accuracy ranging from 79% to 82% on the dataset with 30 classes and 15,000 images. To prevent overfitting, we employed a dropout rate of 0.5 and included a final Soft Max layer for classification. To avoid overfitting and enable quicker weight updates during training, we further employed golorot uniform as the kernel regularizer and stochastic gradient descent (SGD) as the optimizer [10].

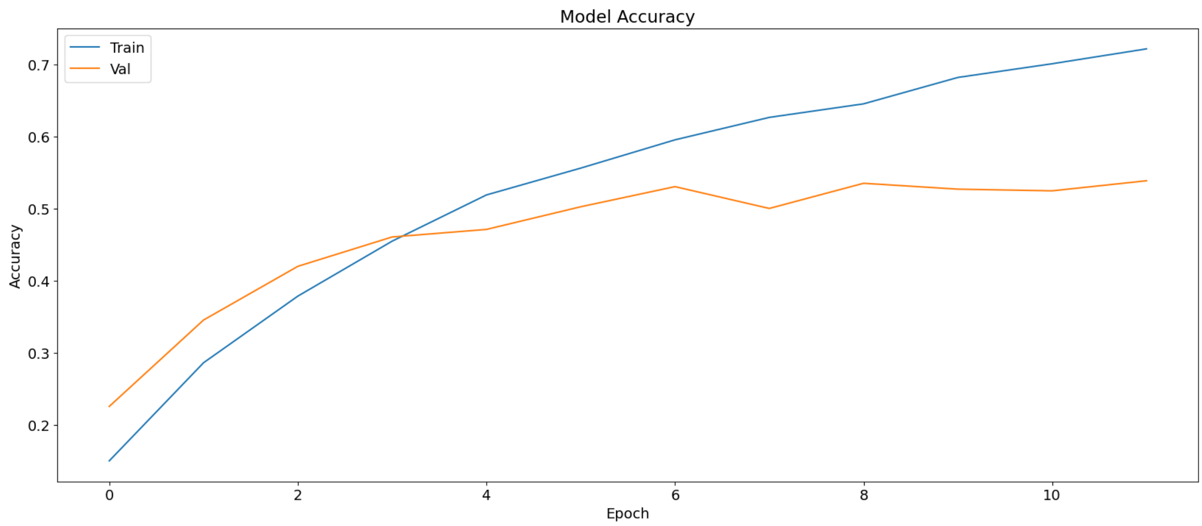

Fig 3.1 Model Accuracy in VGG 16

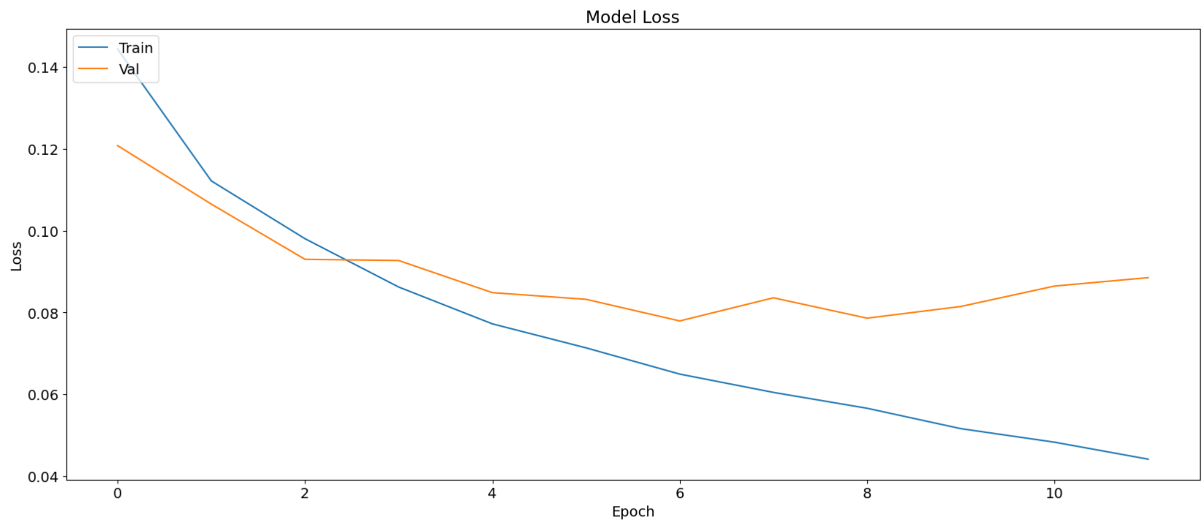

Fig 3.2 Model Loss in VGG 16

ResNet :- ResNet introduced the concept of residual learning, utilizing skip connections to tackle the issue of disappearing gradients encountered in deep neural networks. It has deeper architectures compared to VGG. ResNet excels in classifying objects in images with high accuracy, especially in challenging datasets like ImageNet. ResNet was introduced in the paper "Deep Residual Learning for Image Recognition" by He et al. in 2015. The ResNet architecture addressed the problem of training very deep networks by adding residual blocks. ResNet architectures, such as ResNet50 or ResNet101 is renowned for its capability to effectively train extremely deep networks. On a dataset of 30 classes and 15,000 images, ResNet have achieve accuracy levels similar to VGG or slightly higher, often surpassing 81% to 82% accuracy with proper training and tuning.

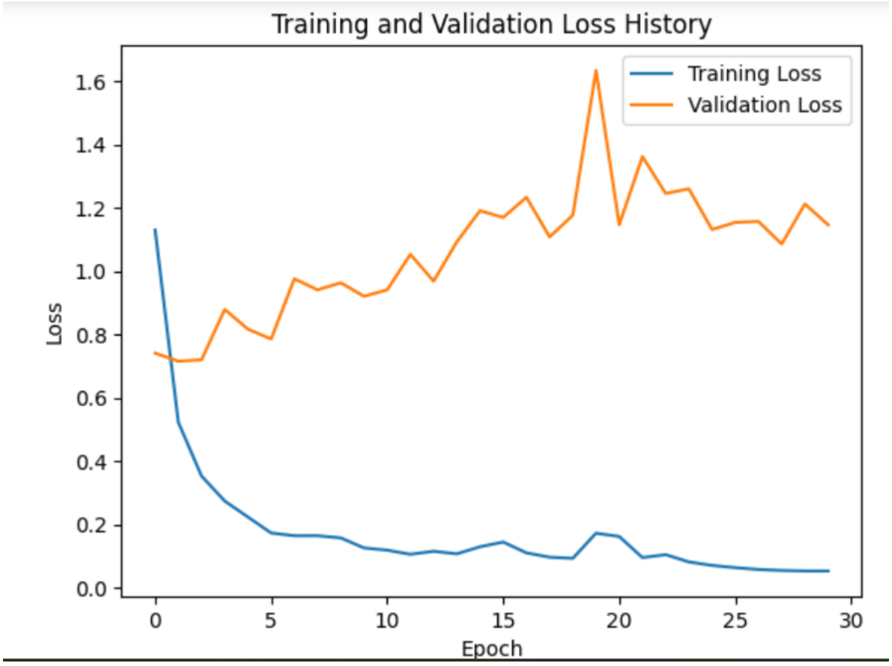

Fig 3.3 Training and Validation Loss History

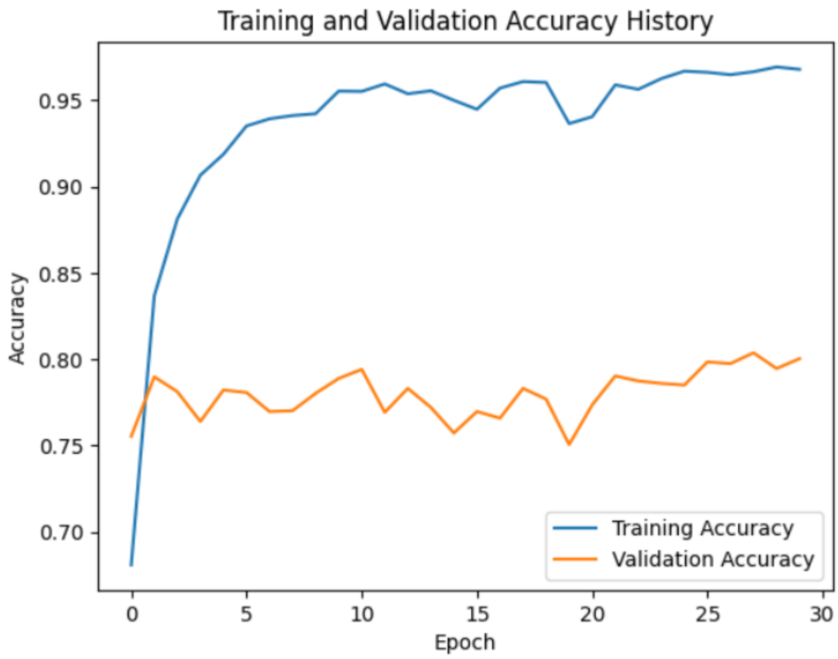

Fig 3.4 Training and Validation Accuracy History

AlexNet :-Alex Net is one of the pioneering deep convolutional neural networks (CNNs) gained recognition by following its victory in ILSVRC in 2012. Alex Net was developed by Alex Krizhevsky, Ilya Sutskever, and Geoffrey Hinton. It gained significant attention after winning the ILSVRC in 2012, marking a milestone in deep learning. Alex Net, while an older architecture compared to the others mentioned, can still achieve respectable accuracy on image classification tasks. On a dataset with 30 classes and 15,000 images, Alex Net can typically achieve accuracy levels ranging from 77% to 78% with appropriate training and optimization. Alex Net is relatively lightweight compared to later architectures like Residual Net or Efficient Net. Training Alex Net requires less computational resources, making it working better in applications where there is a constraints on resources.

Efficient Net B7:- The Efficient Net networks utilizes a compound scaling technique, harmonizing model depth, width, and resolution [16]. Efficient Net B7 is among the larger variants of the Efficient Net family. Efficient Net was proposed in the paper "Efficient Net: Rethinking Model Scaling for Convolutional Neural Networks" by Tan and Le in 2019. Efficient Net introduced a novel scaling method that balanced model depth, width, and resolution to improve efficiency [17]. Efficient Net B7, being one of the larger variants, can achieve excellent accuracy on image classification tasks. On a dataset with 30 classes and 15,000 images, Efficient Net B7 can often surpass 85% accuracy and even approach or exceed 85% to 87% accuracy with careful hyper parameter tuning and data augmentation.

Efficient Net B7 – Object Recognition with Efficient net b7, we have used SGD optimizer with a learning rate of 0.1 and momentum of 0.9. Below graphs visualize the Training and validation loss and accuracy of efficient net b7 on object classification before and after fine tuning are shown in Fig – 3.5 and Fig – 3.6 respectively.

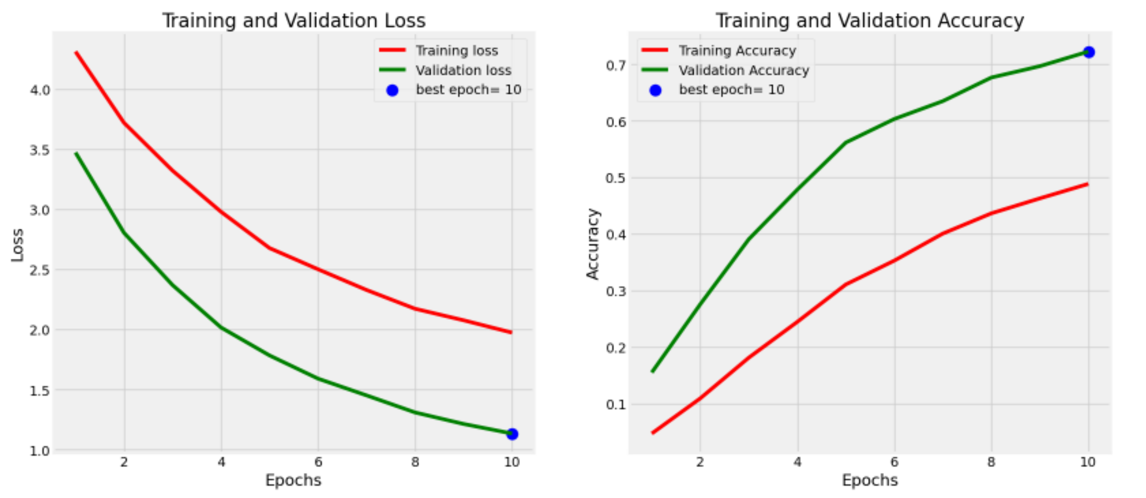

Fig 3.5

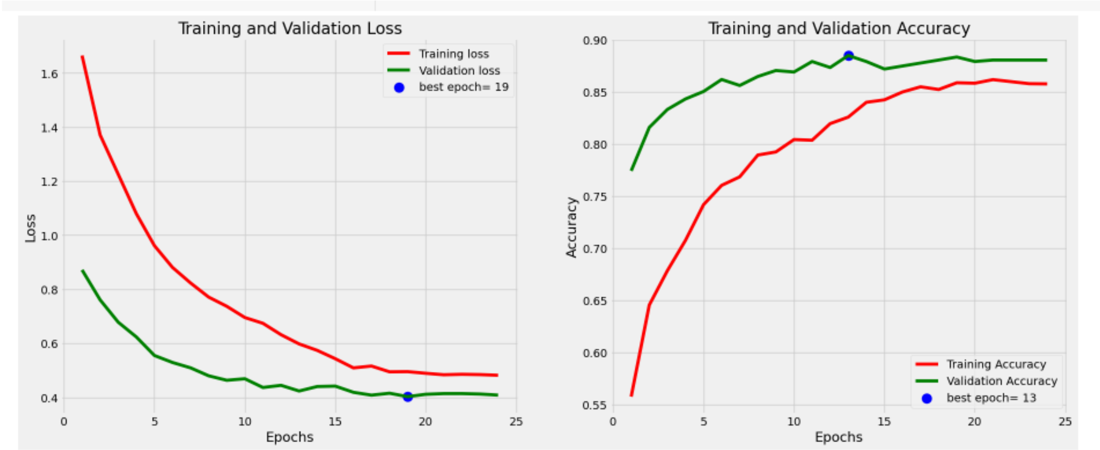

Fig 3.6

	VGG-16	RESNET	Efficient Net	Alex Net
ImageNet	88%	92%	94%	80%
CIFAR-10	92%	94%	97%	80%
CIFAR - 100	67%	72%	77%	55%

Table – 3.1: Results on pre trained Model on famous Dataset

3.2 Data Augmentation Approach

Data augmentation is a crucial approach in deep learning for object recognition, especially in those circumstances when the training model has limited data. It helps in generating new training samples through various transformations, helping to improve the model's robustness.

- The data will be split into three different categories: Training, Validation and Testing. The training data will be used to train the deep learning CNN model and its parameters will be fine-tuned with the validation data.
- The model images will be subjected to a pre-trained CNN model **Batch size**: 32 **Epochs:** 30, **Input Shape**: (224, 224, 3), **Output layer**: 10.
- Geometric Transformation like Rotation of image by 90. By doing this we can rotate image by 0 to 360 degree. By rotation, simply the image pixels gets rotated.

- Scaling helps model better handle real world applications, prevents overfitting, and improves detection. We have performed scaling ranged from 0.8 to 1.2 times of its original size.
- Brightness helps in introducing variability in lightning condition, brightness adjustment helps model better handle real world scenarios, prevents overfitting and improves detection, and recognition in diverse environments. We have adjusted the brightness factor from range of 0.8 to 1.2.

3.3 Dataset description

The dataset used in the paper is an Object dataset. This dataset is a combination of different Objects which are subdivided into various groups. Each group have a separate class and each class has a specific type of Objects. There are total of 31 classes. Each class contains about 300 images and each image has a separate label associated with it. Our dataset is categorized into two parts, Train and Test Dataset. Train Dataset contains about 9000 images. Test class contains 150 images in each class. We have also performed some data augmentation to enhance brightness, rotation and scaling purpose.

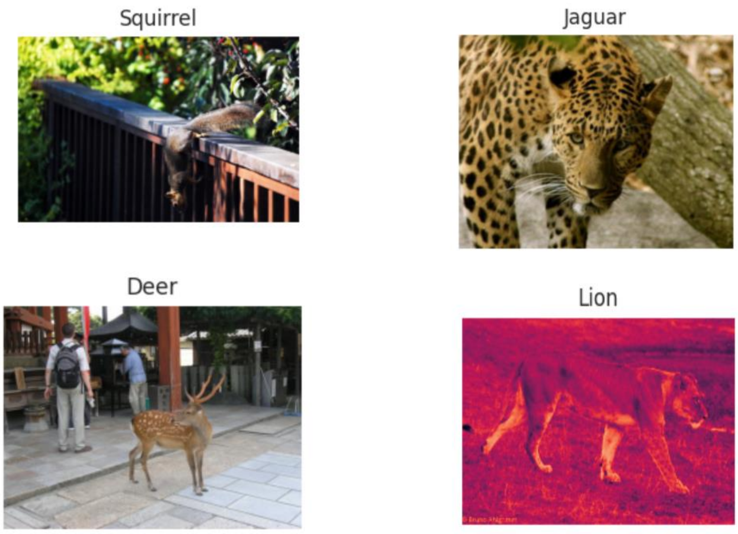

Fig-3.7 – Some sample images from Object-31 dataset

3.4 Training results on Neural Networks for Object Recognition

Model	Parameterization	Epochs	Train Accuracy
VGG	Learning Rate – 0.001 Batch Size – 32 Output Layer – 10 Optimiser – SGD input size = (32, 3, 224, 224) L2 Regularization and Dropout of around 0.5	**30**	**~ 79.00%**
Efficient – Net B7	Learning Rate – 0.01 Batch Size – 32 Output Layer – 10 Optimiser – Adam input size = (224,224,3) L2 Regularization and Dropout of around 0.1	**30**	**~ 85.00%**
ResNet - 51	Learning Rate – 0.01 Batch Size – 32 Optimiser – Adam input Size - (224*224*3)	**30**	**~ 81.00%**
Alex Net	Learning Rate – 0.01 Batch Size – 32 Output Layer – 10 Optimiser – Adam input size = (224,224,3) L2 Regularization and Dropout of around 0.1.	**30**	**~ 77.00%**

Table – 3.2: Description of the models trained

3.5 Experimental Setup

We are working on the classification of animal using Efficient Net B7 neural Network, which results best accuracy out of previous pre trained neural network, when tested with proper parameterization. For our experiment, we first imported the necessary libraries. We subsequently organize everything to replicate results for future reference cases with seed value of 42. Then we Load and Transform the data with Batch Size of 32 and Target Size of (224,224), and placing the data in correct data frame. Here we apply a check mark also to check the corrupted image, and computed the error rate. Next steps includes data pre-processing, data Augmentation which includes rescaling, resizing and random flip. After training, next step includes model evaluation. Accuracy measures the probability of the applied model of being right. Some of the other standards to measure accuracy are as follows:

Precision (P): Ratio of correct predictions, or true positives (TP), to the entire number of results; that is, the total of TP and false positives (FP). In classification issues, when there is involvement of more than one classes, P is the average of all the classes.

Recall (R): The percentage of false negatives (FN) and TP from the overall amount of TP. when there is involvement of more than one classes, then R is calculated as the average of all the classes.

F1 Score (F1): The precision and recall harmonic mean. When there is involvement of more than one classes, F1 is the average of all the classes.

Below is the formulae for the Precision (P), Recall (R) and F1 Score (F1) are:

$$\text{Precision} = \text{True Positive} / (\text{True Positive} + \text{False Positive})$$

$$\text{Recall} = \text{True Positive} / (\text{True Positive} + \text{False Negative})$$

$$\text{F1 Score} = 2 * (\text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall})$$

After all setup, we got the accuracy of around 80.89% which was highest among all the pre-trained neural network that we are using in this paper.

Now it's the time to include this neural network in SSD. We have added Efficient net B7 as base network for SSD with 5 convolution layer. Input image size is of 224 *224

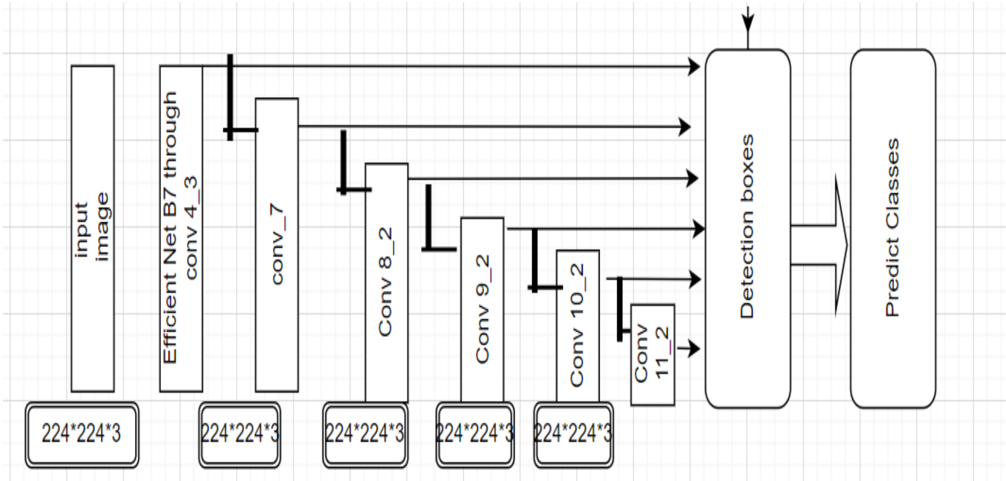

Fig 3.8 Integration of Efficient Net B7 with SSD

Working of Single shot multi box detector starts with base layer (Efficient net b7). It takes the input image and extracts the features at different spatial resolutions through various convolution layers [9]. There are few more layers were added on the top of network to create a feature pyramid. These extra layers captures features at different scales and allows model to search objects of various size in the image [10]. Each layer is responsible for detecting objects at a specific scales and generating multi scale feature maps by applying additional convolutional layers. These feature maps retain spatial information and semantic details at different resolutions. After that, SSD uses anchor boxes to propose region of interest (ROIs) for object detection. Anchor boxes are predefined boxes of various aspect ratios and scales that are placed at different aspect ratios and scales, positioned at various locations on feature maps and serve as reference boxes for predicting bounding boxes and object classes [11].

Now for each anchor boxes, SSD will predict two types of information:

Localization - redirects the offsets (deltas) for adjusting the anchor box.

Classification- Predicting the probability scores for each class (object category) within the anchor box.

During training, SSD with Efficient Net b7 optimizes the network parameters with the help of techniques such as backpropagation and SGD [19]. The loss function includes localization loss (smooth L1 loss) and classification loss (soft max cross-entropy loss) to guide the network towards accurate object localization and classification. In the inference phase, SSD with Efficient Net generates bounding box predictions added with confidence scores for each class

[20]. Next, non-maximum suppression (NMS) is used to eliminate unnecessary bounding boxes while keeping the most reliable detections.

Epoch	Learning Rate
1-5	0.001
6-10	0.0002
11-15	0.00002
16-30	0.0000005

Table 3.3 Learning Rate

The model implemented have a dropout of 0.5 preventing overfitting and promoting the independence of neurons. Activation function we have used is softmax to introduce Non Linearity and affecting the gradient to learn from data.

Loss function is categorical crossEntropy loss, weight used is image net, and shape of image is 224*224*3. Total Parameters were 4,406,722, Trainable Parameters 4,364,699, and Non trainable parameters are 42,023.

To make our neural network model more robust and capable of generalizing to new data, we use a set of image augmentation techniques. These techniques, implemented using the ImageDataGenerator class from the Keras library, help increase the size and variability of our training dataset. This way, our model can learn more effectively from the data.

- Rotation Range: We randomly rotate the images up to 40 degrees. This helps the model recognize objects no matter how they are oriented.
- Width Shift Range: We randomly move the images left or right by up to 20% of their width. This teaches the model to recognize objects even if they are shifted horizontally.
- Height Shift Range: We randomly move the images up or down by up to 20% of their height. This helps the model recognize objects even if they are shifted vertically.
- Rescale: We scale the pixel values of the images by 1/255. This changes the pixel values to a range between 0 and 1, which is better for training neural networks.
- Shear Range: We apply a shear transformation of up to 20%. This means we slant the image, helping the model learn to recognize objects from different angles.
- Zoom Range: We randomly zoom in or out of the images by up to 20%. This helps the model learn to recognize objects at different scales.

- Horizontal Flip: We randomly flip the images horizontally. This helps the model learn to recognize objects even if they appear as mirror images.
- Fill Mode: When these transformations create empty areas in the image, we fill them with the nearest pixel values from the image.

CHAPTER 4

RESULTS AND DISCUSSION

In this section we will be discussing on the results that we obtained when we use efficient net b7 as the base network with SSD. We will plot the image of testing and training accuracy that we got, and also validation loss and validation accuracy.

The dataset we have used has 31 different classes or categories of animals and each categories contains around 200 images.

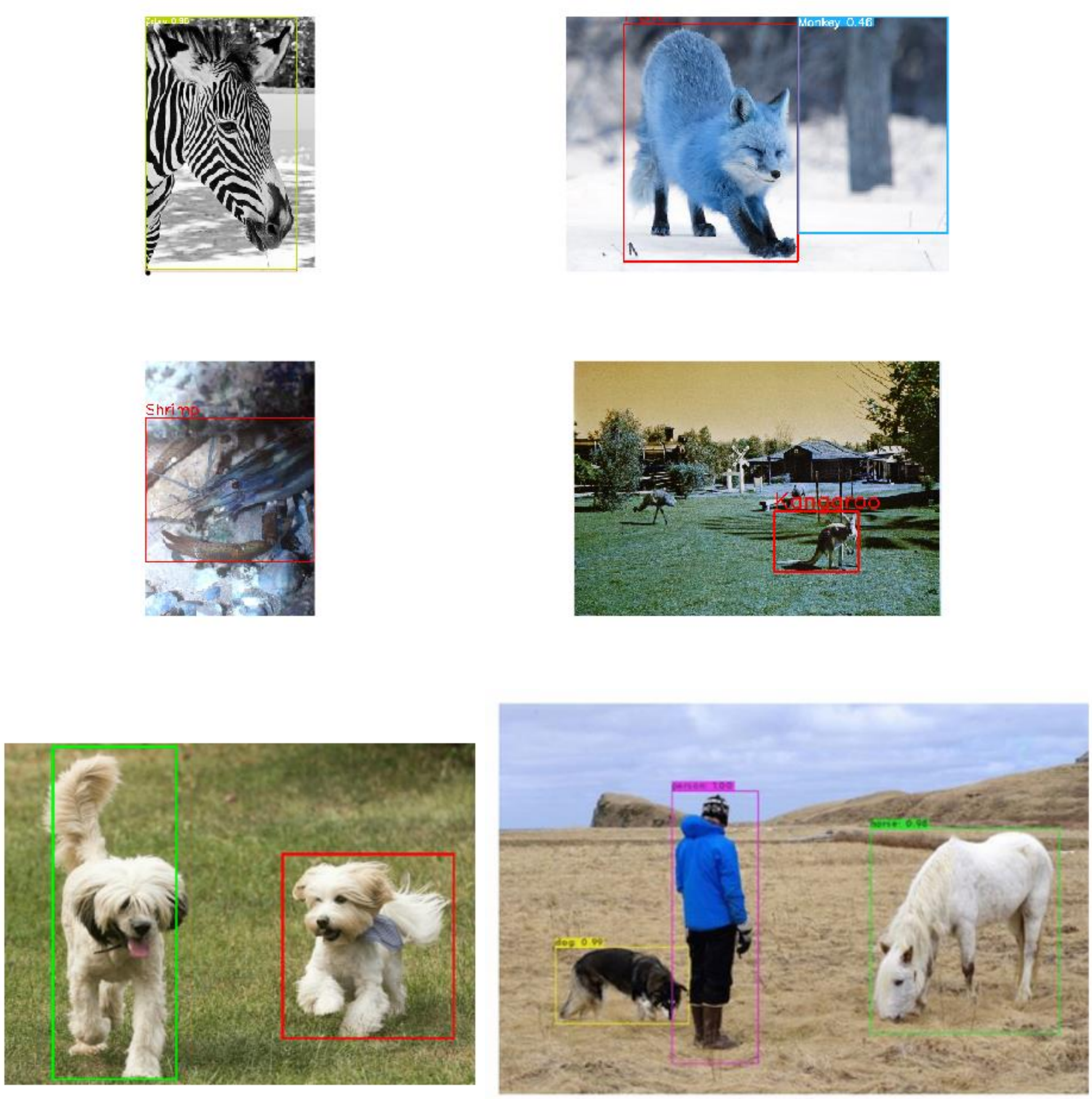

Fig 4.1 – Animal detected under bounding boxes

We can see from output that, the image which contains animal, are marked under bounding boxes and their corresponding classes also shown. We have run the model for 50 epochs, and noted the training accuracy, training loss, and validation accuracy, validation loss graph.

```
Epoch 5/50
91/91 [==============================] - ETA: 0s - loss: 1.1288 - accuracy: 0.6775
Epoch 5: val_loss did not improve from 3.44860
91/91 [==============================] - 405s 4s/step - loss: 1.1288 - accuracy: 0.6775 - val_loss: 3.6241 - val_accuracy: 0.0451 - lr: 0.0010
Epoch 6/50
91/91 [==============================] - ETA: 0s - loss: 1.0163 - accuracy: 0.7038
Epoch 6: val_loss did not improve from 3.44860
91/91 [==============================] - 404s 4s/step - loss: 1.0163 - accuracy: 0.7038 - val_loss: 3.6080 - val_accuracy: 0.0389 - lr: 2.0000e-04
Epoch 7/50
...
Epoch 50/50
91/91 [==============================] - ETA: 0s - loss: 0.8180 - accuracy: 0.7518
Epoch 50: val_loss did not improve from 0.71475
91/91 [==============================] - 462s 5s/step - loss: 0.8180 - accuracy: 0.7518 - val_loss: 0.7259 - val_accuracy: 0.7961 - lr: 5.0000e-07
```

Fig 4.2 – Epochs run for model

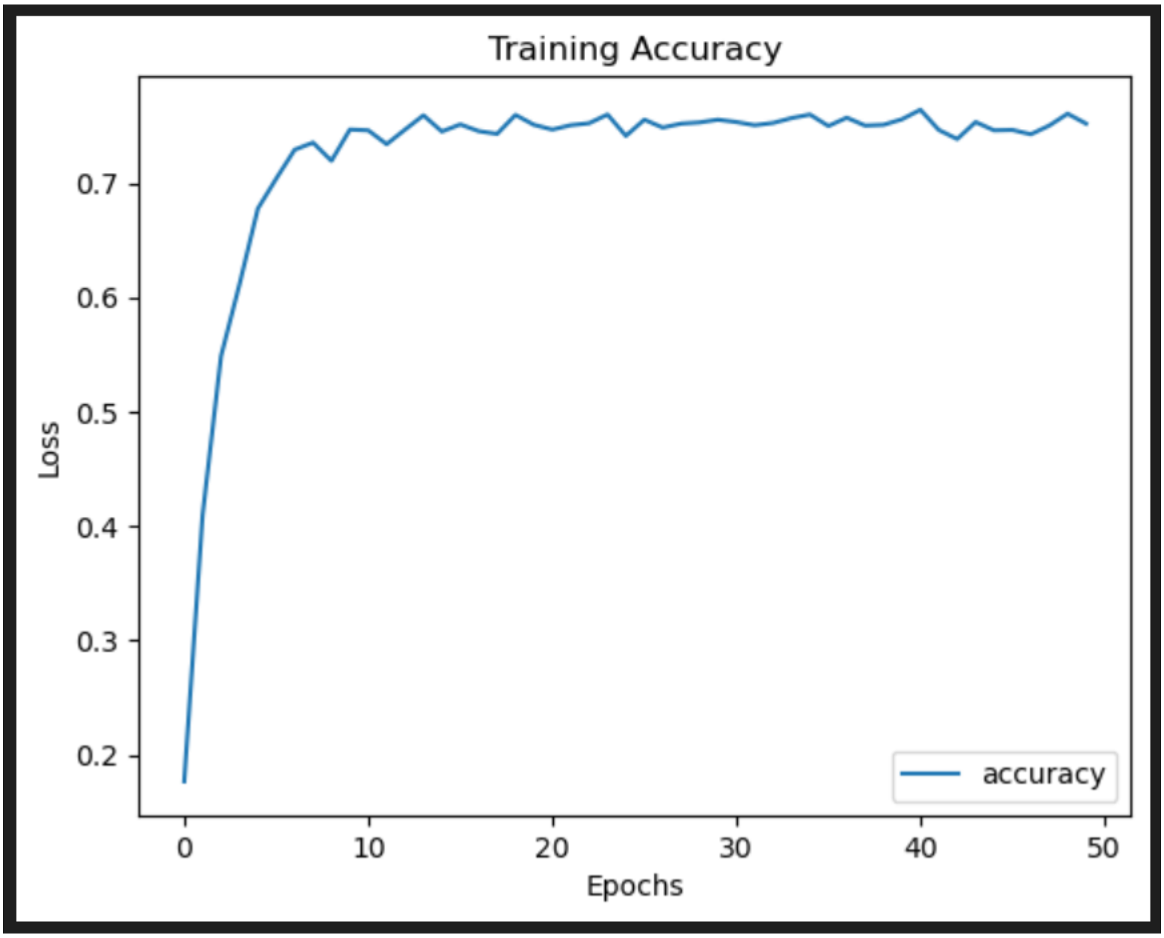

Fig 4.3 Training Accuracy of SSD using Efficient net – B7

From the model, we have achieved a training accuracy of 86.77% after end of 50 epochs. As we have checked in the neural network section that efficient net b7 have achieved highest accuracy among all.

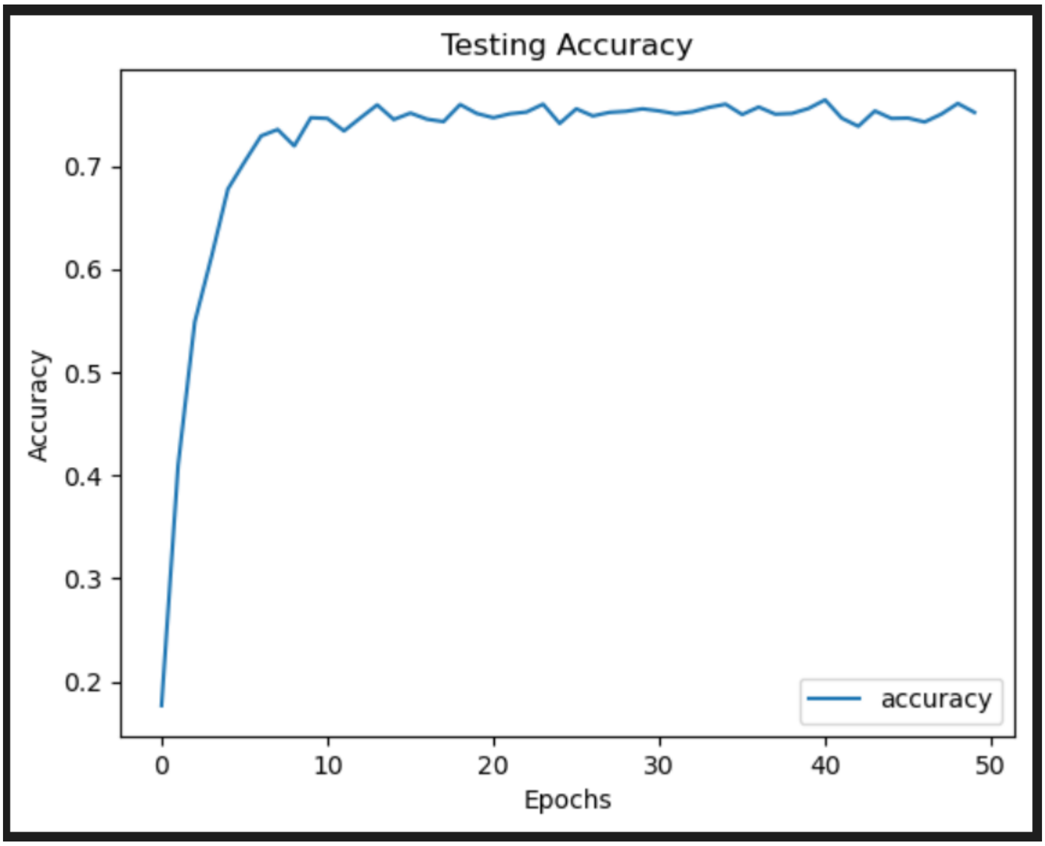

Fig 4.4 Testing Accuracy of SSD using Efficient net – B7

We have achieved testing accuracy of around 84.77% after 30 epochs. Below two graphs will show Validation Accuracy and Training Loss respectively.

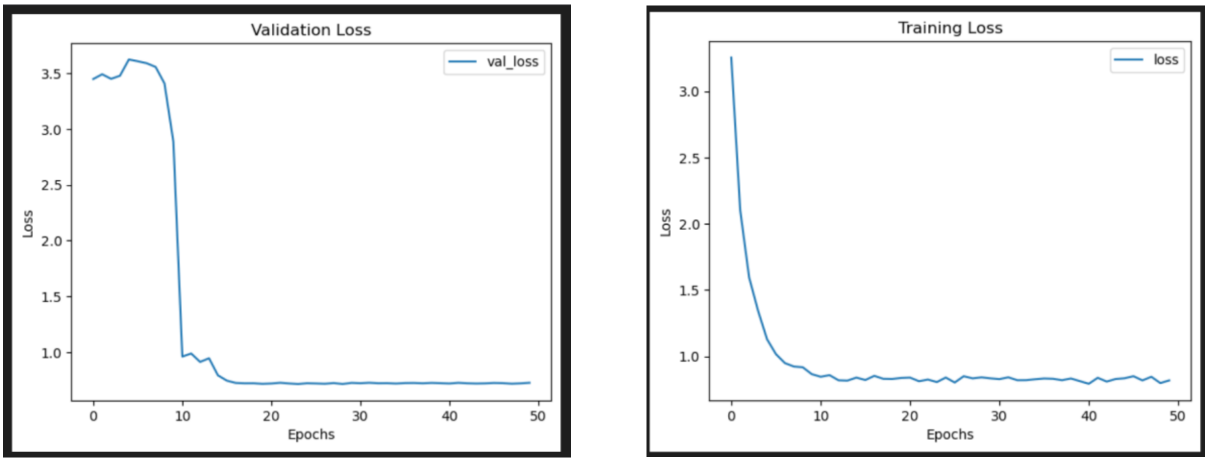

Fig 4.5 Validation Loss and Training Loss

Chapter 5

CONCLUSION

In this paper, we have investigated four deep neural network for object classification, and found that Efficient net b7 works best among all. So, integrating Efficient Net b7 into SSD brings the advantages of efficient feature extraction, improved feature representation across scales, and potentially higher accuracy due to the superior capabilities of Efficient Net b7 in capturing complex visual patterns. This combination is particularly beneficial for object detection tasks where accuracy, speed, and model efficiency are critical factors.

Using EfficientNet B7 as the main network for object detection greatly improves how well it performs and how efficiently it uses resources compared to older networks like VGG-16, ResNet-50, and AlexNet. It's really good at getting accurate results without needing too much computing power, which is great for today's object detection tasks. This study shows that EfficientNet B7 could become the new standard in object detection because it's so good at balancing accuracy, efficiency, and adaptability.

References

[1] Cui-jin Li, Zhong Qu, Sheng-ye Wang, Ling Liu, A method of cross-layer fusion multi-object detection and recognition based on improved faster R-CNN model in complex traffic environment.

[2] Jun Deng1, Xiaojing Xuan2, Weifeng Wang1, Zhao Li3, Hanwen Yao4 and Zhiqiang Wang2, A review of research on object detection based on deep learning

[3] Christian Eggert, Stephan Brehm, Anton Winschel, Dan Zecha, Rainer Lienhart,A Closer Look: Small Object Detection In Faster R-CNN

[4] Y. Liu, "An Improved Faster R-CNN for Object Detection," 2018 11th International Symposium on Computational Intelligence and Design (ISCID), Hangzhou, China, 2018, pp. 119-123, doi: 10.1109/ISCID.2018.10128.

[5] Hui Zhao1,2 · Zhiwei Li1,2 · Lufa Fang1,2 · Tianqi Zhang1, A Balanced Feature Fusion SSD for Object Detection

[6] Hossain, Faysal & Masud, Raihan & Mondal, M. Rubaiyat. (2020). Effectiveness of Deep Learning in Real Time Object Detection. International Journal of Computer Applications. 176. 55-60. 10.5120/ijca2020920551.

[7] Kanishk Wadhwa1, Jay Kumar Behera2, Accurate Real-Time Object Detection using SSD. International Research Journal of Engineering and Technology (IRJET).

[8] Kumar, A., Zhang, Z.J. & Lyu, H. Object detection in real time based on improved single shot multi-box detector algorithm. J Wireless Com Network 2020, 204 (2020).

[9] Qianjun Shuai, Xingwen Wu . Object detection system based on SSD algorithm .

[10] Ayesha Younis,Li Shixin, Shelembi JN, Zhang Hai. Real-Time Object Detection Using Pre-Trained Deep Learning Models MobileNet- SSD

[11] JINLING LI1, QINGSHAN HOU, JINSHENG XING2, AND JIANGUO JU 3. SSD Object Detection Model Based on Multi-Frequency Feature Theory

[12] Viswanatha V, Chandana R K, Ramachandra A.C Real Time Object Detection System with YOLO and CNN Models: A Review

[13] Y. Lu, L. Zhang and W. Xie, "YOLO-compact: An Efficient YOLO Network for Single Category Real-time Object Detection," 2020 Chinese Control And Decision Conference (CCDC), Hefei, China, 2020, pp. 1931-1936, doi: 10.1109/CCDC49329.2020.9164580.

[14] Lee, J., Hwang, Ki. YOLO with adaptive frame control for real-time object detection applications. Multimed Tools Appl 81, 36375–36396 (2022).

[15] Joseph Redmon, Santosh Divvala, Ross Girshick, Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection.

[16] Hou, Q., Cheng, M. M., Hu, X., Borji, A., Tu, Z., & Torr, P. H. (2019). Cross-Domain Correspondence Learning for Exemplar-Based Object Detection. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).

[17] Zhang, X., Zhou, X., Lin, M., & Sun, J. (2018). Shufflenet: An extremely efficient convolutional neural network for mobile devices. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[18] Zoph, B., & Le, Q. V. (2017). Neural Architecture Search with Reinforcement Learning. In International Conference on Learning Representations (ICLR).

[19] Li, Z., & Hoi, S. C. (2019). Gsnet: Jointly learning to guide depthwise feature maps for efficient arbitrary-shape text spotting. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[20] Wu, J., Leng, C., Wang, Y., Hu, Q., & Cheng, J. (2020). Winograd non-uniform quantization for efficient cnns. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

[21] Pham, H., Guan, M. Y., Zoph, B., Le, Q. V., & Dean, J. (2018). Efficient neural architecture search via parameter sharing. In International Conference on Machine Learning (ICML).

[22] Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.

[23] Liu, W., Anguelov, D., Erhan, D., Szegedy, C., Reed, S., & Fu, C. Y. (2016). SSD: Single Shot MultiBox Detector.

[24] He, K., Gkioxari, G., Dollár, P., & Girshick, R. (2017). Mask R-CNN. In Proceedings of the IEEE International Conference on Computer Vision (ICCV).

[25] Gallo, Ignazio, Anwar Ur Rehman, Ramin Heidarian Dehkordi, Nicola Landro, Riccardo La Grassa, and Mirco Boschetti. ”Deep object detection of crop weeds: Performance of YOLOv7 on a real case dataset from UAV images.” Remote Sensing 15, no. 2 (2023): 539.

[26] Lin, Zhihao, Yongtao Wang, Jinhe Zhang, and Xiaojie Chu. ”Dynamicdet: A unified dynamic architecture for object detection.” In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 6282-6291. 2023.

[27] Liu, Kun, Lei Peng, and Shanran Tang. ”Underwater object detection using TC-YOLO with attention mechanisms.” Sensors 23, no. 5 (2023): 2567.

[28] Guo, Qianren, Yuehang Wang, Yongji Zhang, Hongde Qin, Hong Qi, and Yu Jiang. ”AWF-YOLO: enhanced underwater object detection with adaptive weighted feature pyramid network.” OAE Publishing Inc., 2023.

[29] Rahman, Abdur, Yuzhen Lu, and Haifeng Wang. ”Performance evaluation of deep learning object detectors for weed detection for cotton.” Smart Agricultural Technology 3 (2023): 100126.

[30] Alchanatis, Victor, Leonid Ridel, Amots Hetzroni, and Leonid Yaroslavsky. ”Weed detection in multi-spectral images of cotton fields.” Computers and electronics in agriculture 47, no. 3 (2005): 243-260. [7] Reis, Dillon, Jordan Kupec, Jacqueline Hong, and Ahmad Daoudi. ”Real-time flying object detection with YOLOv8.” arXiv preprint arXiv:2305.09972 (2023).

[31] Yu, Yi, and Feipeng Da. ”Phase-shifting coder: Predicting accurate orientation in oriented object detection.” In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 13354- 13363. 2023.

[32] Shi, Furong, Tong Zhang, and Tao Zhang. ”Orientation-aware vehicle detection in aerial images via an anchor-free object detection approach.” IEEE Transactions on Geoscience and Remote Sensing 59, no. 6 (2020): 5221-5233.

[33] Source: Ren, S., He, K., Girshick, R., & Sun, J. (2015). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks.

[34] Ruiz-Ponce, Pablo, David Ortiz-Perez, Jose Garcia-Rodriguez, and Benjamin Kiefer. ”Poseidon: A data augmentation tool for small object detection datasets in maritime environments.” Sensors 23, no. 7 (2023): 3691.

[35] Moosbauer, Sebastian, Daniel Konig, Jens Jakel, and Michael Teutsch. ”A benchmark for deep learning based object detection in maritime environments.” In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition workshops, pp. 0-0. 2019.

[36] Amjoud, Ayoub Benali, and Mustapha Amrouch. ”Object detection using deep learning, CNNs and vision transformers: a review.” IEEE Access (2023).

[37] Li, Yanghao, Hanzi Mao, Ross Girshick, and Kaiming He. ”Exploring plain vision transformer backbones for object detection.” In European Conference on Computer Vision, pp. 280-296. Cham: Springer Nature Switzerland, 2022.

[38] Diwan, Tausif, G. Anirudh, and Jitendra V. Tembhurne. ”Object detection using YOLO: Challenges, architectural successors, datasets and applications.” Multimedia Tools and Applications 82, no. 6 (2023): 9243- 9275.

[39] Krishnachaithanya, N., Gurdit Singh, Smita Sharma, Rangisetti Dinesh, Sumeet Ramsingh Sihag, Kamna Solanki, Abhishek Agarwal, Mrinalini Rana, and Ujjwal Makkar. ”People Counting in Public Spaces using Deep Learning-based Object Detection and Tracking Techniques.” In 2023 International Conference on Computational Intelligence and Sustainable Engineering Solutions (CISES), pp. 784-788. IEEE, 2023.

[40] Hong, Chee Jia, and Muhammad Hazli Mazlan. ”Development of Automated People Counting System using Object Detection and Tracking.” International Journal of Online & Biomedical Engineering 19, no. 6 (2023).

[41] Alahmadi, Tahani Jaser, Atta Ur Rahman, Hend Khalid Alkahtani, and Hisham Kholidy. ”Enhancing Object Detection for VIPs Using YOLOv4 Resnet101 and Text-to-Speech Conversion Model.” Multimodal Technologies and Interaction 7, no. 8 (2023): 77.

[42] Ma, Pingli, Chen Li, Md Mamunur Rahaman, Yudong Yao, Jiawei Zhang, Shuojia Zou, Xin Zhao, and Marcin Grzegorzek. ”A state-ofthe-art survey of object detection techniques in microorganism image analysis: from classical methods to deep learning approaches.” Artificial Intelligence Review 56, no. 2 (2023): 1627-1698.

[43] Tan, M., Pang, R., & Le, Q. V. (2020). EfficientDet: Scalable and Efficient Object Detection.

[44] Zhang, X., Zhou, X., Lin, M., & Sun, J. (2022). ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices.

[45] Huang, G., Liu, Z., Van Der Maaten, L., & Weinberger, K. Q. (2021). Densely Connected Convolutional Networks.

[46] Tan, M., & Le, Q. V. (2019). EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks.

Accepted Papers

Gmail **Ashish Keshri <keshriashish07@gmail.com>**

I. Acceptance Notification 1st IEEE ICAC2N-2024 & Registration: Paper ID 687 @ ITS Engineering College, Greater Noida

1 message

Microsoft CMT <email@msr-cmt.org> Mon, May 13, 2024 at 12:15 AM
Reply-To: "Dr. Vishnu Sharma" <vishnu.sharma@its.edu.in>
To: Ashish Kumar <keshriashish07@gmail.com>

Dear Ashish Kumar ,

Delhi Technological

University Greetings

from ICAC2N-2024 ...!!!

Congratulations....!!!!!

On behalf of the ICAC2N-2024 organising Committee, we are delighted to inform you that the submission of "Paper ID- 687 " titled " Object Detection using SSD and Efficient net B7 as base Network " has been accepted for presentation and further publication with IEEE at the ICAC2N- 24. All accepted papers will be submitted for inclusion into IEEE Xplore subject to meeting IEEE Xplore's scope and quality requirements.

Registration/Fee Payment related details are available at https://icac2n.in/register.

For early registration benefit please pay your fee and complete your registration by clicking on the following Link: https://forms.gle/E7RuvuQQPxPZQnJU6 by 20 May 2024.

Note:

1. All figures and equations in the paper must be clear.
2. Final camera ready copy must be strictly in IEEE format available on conference website.
3. Transfer of E-copyright to IEEE and Presenting paper in conference is compulsory for publication of paper in IEEE.
4. If plagiarism is found at any stage in your accepted paper, the registration will be cancelled and paper will be rejected and the authors will be responsible for any consequences. Plagiarism must be less then 15% (checked through Turnitin).
5. Change in paper title, name of authors or affiliation of authors will not be allowed after registration of papers.

Conference Application Details

Status: Approved

Conference Information

Conference Record Number:	63387
Conference Title:	2024 1st International Conference on Advances in Computing, Communication and Networking (ICAC2N)
Conference Acronym:	ICAC2N
Start Date:	16-Dec-2024
End Date:	17-Dec-2024
Venue Name:	ITS Engineering College
City:	Greater Noida
State/Province:	
Country/Region:	India

Submitter Information

Salutation:	
Given name/First name:	Vishnu
Middle name:	
Last/Family/Surname:	Sharma
Suffix:	

Sponsor Information	MOU Approval Status
MOU Signatory 1:	Approved
Uttar Pradesh Section (0.00%)	
MOU Signatory 2:	Approved
ITS Engineering College (100.00%)	

Back

onal Conference on Advances in Computing, Communication and Networking- ICAC2N • 1st International Conf

ICAC2N

Submit Paper

HOME

ABOUT

GUIDELINES

SPEAKERS

COMMITTEE

CALL FOR PAPERS

REGISTER

CONTACT

CONFERENCE WILL BE HELD IN BLENDED MODE (ONLINE AND OFFLINE BOTH)

1st IEEE ICAC2N-24 || 16th & 17th December 2024

16th and 17th December 2024

Conference Record Number #63387

IEEE XPLORE COMPLIANT ISBN No. 979-8-3503-5681-6

1st International Conference on Advances in Computing, Communication and Networking- ICAC2N

Computer Science and Engineering Department

ITS Enginerring College

Knowledge Park III, Greater Noida

Submit Paper

Note: All conference papers included in IEEE Xplore will be indexed in the SCOPUS database.

About ICACCN

ICAC2N is a prestigious international conference that brings together top researchers, scientists, engineers, and scholars from around the world to share their latest research findings and experiences in computing, communication and networking. Featuring keynote speeches, technical sessions, and workshops, the conference covers a wide range of topics such as cloud computing, AI, wireless communication systems, IoT, and cybersecurity.

Gmail

Ashish Keshri <keshriashish07@gmail.com>

Acceptance Notification 1st IEEE ICAC2N-2024 & Registration: Paper ID 687 @ ITS Engineering College, Greater Noida

2 messages

Microsoft CMT <email@msr-cmt.org> Mon, May 13, 2024 at 12:15 AM
Reply-To: "Dr. Vishnu Sharma" <vishnu.sharma@its.edu.in>
To: Ashish Kumar <keshriashish07@gmail.com>

Dear Ashish Kumar ,
Delhi Technological University

Greetings from ICAC2N-2024 ...!!!

Congratulations....!!!!!

On behalf of the ICAC2N-2024 organising Committee, we are delighted to inform you that the submission of "Paper ID- 687 " titled " Object Detection using SSD and Efficient net B7 as base Network " has been accepted for presentation and further publication with IEEE at the ICAC2N- 24. All accepted papers will be submitted for inclusion into IEEE Xplore subject to meeting IEEE Xplore's scope and quality requirements.

Registration/Fee Payment related details are available at https://icac2n.in/register.

For early registration benefit please pay your fee and complete your registration by clicking on the following Link: https://forms.gle/E7RuvuQQPxPZQnJU6 by 20 May 2024.

Note:
1. All figures and equations in the paper must be clear.
2. Final camera ready copy must be strictly in IEEE format available on conference website.
3. Transfer of E-copyright to IEEE and Presenting paper in conference is compulsory for publication of paper in IEEE.
4. If plagiarism is found at any stage in your accepted paper, the registration will be cancelled and paper will be rejected and the authors will be responsible for any consequences. Plagiarism must be less then 15% (checked through Turnitin).
5. Change in paper title, name of authors or affiliation of authors will not be allowed after registration of papers.
6. Violation of any of the above point may lead to rejection of your paper at any stage of publication.
7. Registration fee once paid will be non refundable.

If you have any query regarding registration process or face any problem in making online payment, write us at icac2n.ieee@gmail.com.

Regards:
Organizing committee
ICAC2N – 2024

ASHISH KUMAR <keshriashish07@gmail.com> Thu, May 30, 2024 at 3:08 AM
To: "Dr. Vishnu Sharma" <vishnu.sharma@its.edu.in>

Gmail

Ashish Keshri <keshriashish07@gmail.com>

Registration Confirmation 1st IEEE ICAC2N-2024 : Paper ID 687 @ ITS Engineering College, Greater Noida

2 messages

Microsoft CMT <email@msr-cmt.org> Sat, Jun 1, 2024 at 3:32 PM
Reply-To: "Dr. Vishnu Sharma" <vishnu.sharma@its.edu.in>
To: Ashish Kumar <keshriashish07@gmail.com>

Dear Ashish Kumar ,
Delhi Technological University

Greetings from ICAC2N-2024 ...!!! Thanks for Completing your registration...!!

Paper ID- "687 “
Paper Title- “ Object Detection using SSD and Efficient net B7 as base Network “

This email is to confirm that you have successfully completed your registration for your accepted paper at ICAC2N-2024. We have received your registration and payment details. Further, your submitted documents will be checked minutely and if any action will be required at your end you will be informed separately via email.

For further updated regarding conference please keep visiting conference website www.icac2n.in or write us at icac2n.ieee@gmail.com.

Regards:
Organizing committee
ICAC2N – 2024

Note:
1. Transfer of E-copyright to IEEE and Presenting paper in conference is compulsory for publication of paper in IEEE. (For this you will be informed separately via email well before conference)
2. If plagiarism is found at any stage in your accepted paper, the registration will be cancelled and paper will be rejected and the authors will be responsible for any consequences. Plagiarism must be less then 15% (checked through Turnitin).
3. Change in paper title, name of authors or affiliation of authors is not allowed now.
4. Violation of any of the above point may lead to cancellation of registration.
5. Registration fee once paid is non-refundable.

Microsoft CMT <email@msr-cmt.org> Sat, Jun 1, 2024 at 4:00 PM
Reply-To: "Dr. Vishnu Sharma" <vishnu.sharma@its.edu.in>
To: Ashish Kumar <keshriashish07@gmail.com>

[Quoted text hidden]

M

To MANJEET SINGH

₹8,000

Pay again

Split with friends

Completed

31 May 2024, 10:56 am

State Bank of India
4320

UPI transaction ID
415226804240

To: MANJEET SINGH
icac2n@ybl

From: ASHISH KUMAR (State Bank of India)
keshriashish07@okaxis

Google transaction ID
CICAgPDqob-JYg

POWERED BY UPI

G Pay

Gmail **Ashish Keshri <keshriashish07@gmail.com>**

Acceptance Notification 1st IEEE ICAC2N-2024 & Registration: Paper ID 107 @ ITS Engineering College, Greater Noida

1 message

Microsoft CMT <email@msr-cmt.org> Mon, May 13, 2024 at 12:15 AM
Reply-To: "Dr. Vishnu Sharma" <vishnu.sharma@its.edu.in>
To: Ashish Kumar <keshriashish07@gmail.com>

Dear Ashish Kumar ,
Delhi Technological University

Greetings from ICAC2N-2024 ...!!!

Congratulations....!!!!!

On behalf of the ICAC2N-2024 organising Committee, we are delighted to inform you that the submission of “Paper ID- 107 “ titled “ Exploring Novel Strategies for Object Detection: Insights and Innovations “ has been accepted for presentation and further publication with IEEE at the ICAC2N- 24. All accepted papers will be submitted for inclusion into IEEE Xplore subject to meeting IEEE Xplore’s scope and quality requirements.

Registration/Fee Payment related details are available at https://icac2n.in/register.

For early registration benefit please pay your fee and complete your registration by clicking on the following Link: https://forms.gle/E7RuvuQQPxPZQnJU6 by 20 May 2024.

Note:
1. All figures and equations in the paper must be clear.
2. Final camera ready copy must be strictly in IEEE format available on conference website.
3. Transfer of E-copyright to IEEE and Presenting paper in conference is compulsory for publication of paper in IEEE.
4. If plagiarism is found at any stage in your accepted paper, the registration will be cancelled and paper will be rejected and the authors will be responsible for any consequences. Plagiarism must be less then 15% (checked through Turnitin).
5. Change in paper title, name of authors or affiliation of authors will not be allowed after registration of papers.
6. Violation of any of the above point may lead to rejection of your paper at any stage of publication.
7. Registration fee once paid will be non refundable.

If you have any query regarding registration process or face any problem in making online payment, write us at icac2n.ieee@gmail.com.

Regards:
Organizing committee
ICAC2N – 2024

M

To MANJEET SINGH

₹8,000

paper id 107

Pay again

Split with friends

Completed

6 Jul 2024, 12:51 pm

State Bank of India
4320

UPI transaction ID
418880522421

To: MANJEET SINGH
icac2n@ybl

From: ASHISH KUMAR (State Bank of India)
keshriashish07@oksbi

Google transaction ID
CICAgPDR9ZGGdg

POWERED BY UPI

G Pay

ANNEXURE-IV

DELHI TECHNOLOGICAL UNIVERSITY

(Formerly Delhi College of Engineering)

Shahbad Daulatpur, Main Bawana Road, Delhi-42

PLAGIARISM VERIFICATION

Title of the Thesis OBJECT DETECTION USING SSD AND EFFICIENT NET B7 AS BASE NETWORK Total Pages 43 Name of the Scholar Ashish Kumar Supervisor (s)

(1) Dr. Priyanka Meel

(2)____

(3)____

Department INFORMATION TECHNOLOGY, DELHI TECHNOLOGICAL UNIVERSITY

This is to report that the above thesis was scanned for similarity detection. Process and outcome is given below:

Software used: Turnitin Similarity Index: 14 % , Total Word Count: 8232

Date: 30/05/2024

Candidate's Signature **Signature of Supervisor(s)**

PAPER NAME

Ashish Thesis.docx

WORD COUNT	CHARACTER COUNT
8232 Words	**46999 Characters**
PAGE COUNT	FILE SIZE
35 Pages	**2.3MB**
SUBMISSION DATE	REPORT DATE
May 29, 2024 5:23 PM GMT+5:30	**May 29, 2024 5:24 PM GMT+5:30**

14% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

- 8% Internet database
- 7% Publications database
- Crossref database
- Crossref Posted Content database
- 10% Submitted Works database

Excluded from Similarity Report

- Bibliographic material
- Small Matches (Less then 8 words)