Empirical Study of the Impact of Image Quality, Object Size, and Occlusion to Object Detection

Abhijit Kumar, Udit Mital, Ankit Gajera, Shivani Varanasi and Dibyendu Patra

February 27, 2023

Empirical Study of the Impact of Image Quality, Object size, and Occlusion to Object Detection

Abhijit Kumar
Bengaluru, India
Email: kumarabhijit@gmail.com

Udit Mital
Bengaluru, India
Email: uditmital86@gmail.com

Ankit Gajera
Bengaluru, India
Email:ankit.gajera@outlook.com

Shivani Varanasi
Bengaluru, India
Email:shivanivaranasi24@gmail.com

Dibyendu Patra
Bengaluru, India
Email:mynamedibyendupatra@gmail.com

Abstract—**Object detection is a crucial task in computer vision to identify and locate objects in an image or video. The performance of object detection algorithms is influenced by numerous factors such as image quality, object size, and occlusion. Image quality affects the clarity and resolution of the objects in the image, affecting the accuracy of object detection. Object size refers to the physical size of objects in an image and can influence their detection. Occlusion occurs when one object hides another, making it difficult to detect both objects accurately. This can result in false detections, missed detections, or a decrease in overall accuracy. Small objects also pose a significant challenge to object detection algorithms, as they can be difficult to detect and distinguish from their surroundings due to their limited size and distinctive features. Occlusion refers to the partial or complete hiding of an object by another object. It has a significant impact on the performance of object detection algorithms. This is because occlusion can make it difficult for the algorithms to detect and recognize objects, as well as to decide their position and orientation within an image. In this paper, our focuses are to find the effect of image quality, object size in an image, and object occlusion in an image. Our experimentation has revealed that the accuracy of our computer vision model is highly dependent on several factors, including image quality, occlusion, and object size. Specifically, we found that when presented with bad-quality images, the accuracy of our model drops significantly, sometimes plummeting as low as 25%. Similarly, we saw that heavily occluded images tend to decrease model accuracy, with our model achieving only 55% accuracy in such scenarios. In addition, we also noticed that our model struggled to achieve high accuracy with objects of varying sizes, resulting in low accuracy scores, with our model achieving only 67% accuracy in such scenarios.**

I. Introduction

Object detection, which is the task of locating and identifying objects within an image or video, is a crucial component in many computer vision applications, such as autonomous driving, security surveillance, and image retrieval. Despite the recent advances in object detection, many challenges remain, particularly in real-world scenarios where images and videos can be of low quality, contain small objects, and be subject to occlusion. The quality of an image, the size of an object, and the amount of occlusion present in an image can significantly impact the performance of object detection algorithms. Blur and noise in an image can significantly impact the performance of object detection algorithms. Blur in an image can cause objects to appear less distinct, making it more difficult for the object detection algorithm to accurately identify and locate the objects in the image. This is because blur reduces the sharpness of the image, leading to the loss of edge information, which is critical for object recognition. Similarly, noise in an image can also cause problems for object detection algorithms. Noise refers to random variations in the pixel values that can arise from various sources, such as camera sensor noise, transmission errors, or compression artifacts. This added noise can cause false edges to appear in the image, leading the object detection algorithm to produce false detections or miss real objects.

To mitigate the effects of blur and noise, it is often necessary to apply pre-processing techniques such as image denoising and deblurring. These techniques aim to restore the image to its original quality, thereby improving the performance of object detection algorithms. However, it is important to note that such techniques are not always effective and may even introduce new artifacts into the image, which can further impact the performance of object detection algorithms.

Occlusion refers to when an object in an image is partially or fully obscured by another object. Occlusion can pose a significant challenge for object detection algorithms, as the obscured portion of the object is no longer visible, making it difficult to accurately detect and locate the object. There are several different types of occlusion, including self-occlusion, where an object occludes itself, and inter-object occlusion, where one object occludes another. The extent of the occlusion, as well as the type of occlusion, can impact the performance of object detection algorithms. In some cases, the occluded portion of the object can still be partially visible, allowing the object detection algorithm to make an educated guess about the shape and location of the object. However, in other cases, the object may be completely obscured, making it very difficult for the object detection algorithm to detect and locate the object.

The size of objects in an image can have a significant impact on the performance of object detection algorithms. Objects that are very small in the image may be difficult to detect, as the features used by the object detection algorithm may be too small or too similar to the background to be distinguishable. On the other hand, objects that are very large in the image may be easier to detect, as the features used by the object detection algorithm will be more prominent. The size of the objects can also impact the scale invariance of the object detection algorithm, which refers to its ability to detect objects of different sizes in the image. Ideally, an object detection algorithm should be able to detect objects of different sizes with equal accuracy, regardless of their size in the image. However, in practice, many object detection algorithms struggle to detect small objects and may have trouble detecting objects of very different sizes in the same image.

II. RELATED WORK

The effect of image quality, image occlusion, and different object sizes on object detection has been studied by various researchers, and there are several related works on this topic. The impact of image quality on object detection using UAV imagery explored by [1]. The effects of image quality and viewpoint on object detection in aerial imagery. The authors show that image quality can have a significant impact on detection accuracy and that certain types of image degradation, such as shadowing, can be particularly problematic, this study investigated by [2].A tubelet proposal network for object detection in videos. The model uses temporal information and occlusion reasoning to improve detection accuracy in dynamic scenes with occlusion presented by [3]. A multi-task learning framework that is capable of simultaneously estimating scene geometry and semantics while handling occlusion in the scene introduced by [4].A method for de-occluding occluded text using convolutional neural networks. The authors show that their method can improve text detection and recognition accuracy in occluded scenes presented by [5].An occlusion-aware R-CNN for detecting pedestrians in crowded scenes. The model uses occlusion-aware region proposals and occlusion reasoning to improve detection performance in crowded scenes proposed by [6].A faster region-based convolutional neural network (Faster R-CNN) for object detection. The authors show that their method is capable of detecting objects of different sizes with high accuracy presented by [7]. A single-shot detection method called SSD is capable of detecting objects at different scales. The authors demonstrate that their method is highly accurate and efficient as proposed by [8]. [9] This paper introduces YOLOv3, an object detection model that is capable of detecting objects at different scales. The authors show that their method achieves state-of-the-art accuracy and is highly efficient. A single-shot face detection method that is scale-invariant and can detect faces of different sizes with high accuracy is proposed by [10].

III. DATA PREPARATION

A. Low Quality Images

We collected data from Kaggle[11] to perform experiments on low-quality images. There are a total of 240 train images and 120 test images. Every image has a .xml file where the location and name of the object are mentioned. In this Dataset every image contains a single object. We have apples, oranges, and bananas total of three types of objects images in the Dataset.

Object Name	Count	Percentage
Apples	97	26.9%
Oranges	123	34.2 %
Bananas	140	38.9%

TABLE I. OBJECT NAME, NUMBER, AND DISTRIBUTION

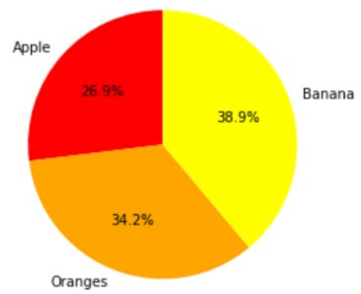

Fig 1 : Objects Distribution

we use the Gaussian Blur score and Box Blur score to detect the quality of images on the Blur factor and use Gaussian Noise score to detect the quality of images on the Noise factor. we use 3 types of algorithms for detecting the quality of images. Using these 3 algorithms we get the idea of two types of image quality.

Gaussian Blur Score	Box Blur Score	Image Quality
Gaussian Blur Score>=7	Box Blur Score>=10	Bad Quality
Gaussian Blur Score<=6	Box Blur Score<=9	Medium Quality
Gaussian Blur Score<=3	Box Blur Score<=3	Good Quality

TABLE II. QUALITY OF IMAGES ON BLUR PARAMETER

Gaussian Noise Score	Image Quality
Gaussian Noise Score >=.10	Bad Quality
Gaussian Noise Score<=.05	Medium Quality
Gaussian Noise Score<=.02	Good Quality

TABLE III. QUALITY OF IMAGES ON NOISE PARAMETER

B. Occluded Images

We generated occluded images using DALL-E to do the experiment. We generate a total of 60 images from DALL-E which contain apples, oranges, and bananas as objects, and then using the image augmentation technique we created a total of 210 images. Out of 210 images, 140 images are training data and 70 images are testing data. There is a total of 3 types of occluded images present. The first one is 30% occluded images here 30% of any objects are covered by other objects. the second one is 50% occluded images here 50% of any object is covered by any other objects. And the last one is 70% occluded images here 70% of any object is covered by any other objects.

Data	30%occluded	50%occluded	70%occluded
Train Data	50	50	40
Test Data	30	20	20

TABLE IV. OCCLUDED IMAGES DETAILS

C. Different Object Size Images

We collect multiple different-size fruits in every image data and multiple same-size fruits in every image data from Kaggle[11] and add those two data to perform our experiment. we have a total of 470 data out of that 300 data are for training and 170 data for testing. Out of 470 images, 220 images have multiple objects of the same size and 250 images contain multiple objects of different sizes. We created various kind of combination of data set to do the experiment for example “half of the data in training data contain multiple different size objects in an image and half of the data contain multiple similar kinds of object in an image in the data set”.

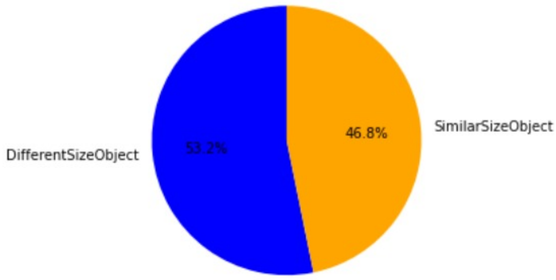

Fig 2: Data Distribution

IV. Experiments

A. Low Quality Images

Image quality is an important aspect of digital image processing and has a significant impact on the accuracy and performance of computer vision systems. In this experiment, we aim to evaluate the quality of images using a Convolutional Neural Network (CNN) and compare the results with a traditional quality assessment method.

Fig 3: Low-Quality Images

Methods:

1. Data Collection: We collected a dataset of 360 images with different levels of quality (good, medium, and poor). And then we create various kinds of datasets from that 360 data set for our experiments.

2. Preprocessing: The images were preprocessed to ensure that they were of the same size and format.

3. Traditional Quality Assessment Method: We used the Gaussian Blur score, Box Blur Score, and Gaussian Noise Score to calculate the quality of the images.

4. CNN Model Training: We trained a CNN model on the dataset which has Categorical cross entropy as a loss function and use Adam as an optimizer

Training Data	Test Data	Train Accuracy	Test Accuracy
All Medium Noisy	All Medium blur	.98	.91
All Medium Noisy	All Medium noisy	.98	.93
All Medium Blur	All Medium Blur	.57	.50
Medium Blur	Medium Noisy	.78	72
50% Blur50%noisy	Good Quality	.32	.35
70% Blur30% Noisy	Good quality	.29	.30
All Highly Blur	Good Quality	.28	.30
All Highly Blur	Medium Blur	.28	.25
All Highly Noisy	Good Quality	.45	.57
70%Noisy 30%Blur	Good Quality	.42	.51

TABLE V. Experiments Details On Low Quality Images

5. Model Evaluation: We evaluated the performance of the CNN model by comparing its predictions with the good quality of data.

The experiment on the effect of image quality on Convolutional Neural Network (CNN) performance demonstrated that the quality of the images used for training and inference can have a significant impact on the accuracy and precision of the model. The results showed that as the quality of the images decreased, the performance of the CNN decreased. This highlights the importance of considering image quality when designing computer vision systems and training CNN models. From the above table of experiments, we can understand that

1. Bad Quality Images will decrease the accuracy

2. If we have more Blur images in the Data set then we will get very low Accuracy.

3. The impact of blur images is more than noisy images in Model accuracy

B. Occluded Images

Image occlusion, or the presence of obstructions in an image, can greatly affect the performance of computer vision systems. In this experiment, we aim to evaluate the effect of image occlusion on the performance of a Convolutional Neural Network (CNN).

Fig 4 : 30%, 50% and 70% Occluded Image

Methods:

1. Data Collection: We generated a dataset of 60 images of objects using DALL-E, with different levels of occlusion (0-70%). And then using image augmentation we created a total of 210 images of objects.

2. Preprocessing: The images were preprocessed to ensure that they were of the same size and format.

3. Image Occlusion: We artificially occlude the images by placing a black box over different regions of the images and some images are already occluded. There are three types of occluded images present in our data set

4. CNN Model Training: We trained a CNN model on the dataset which has Categorical cross entropy as a loss function and use Adam as an optimizer

5. Model Evaluation: We evaluated the performance of the CNN model by comparing its predictions on the occluded and un-occluded images.

Training Data	Test Data	Train Accuracy	Test Accuracy
Non occluded	Non occluded	.94	.92
30% occluded	Non occluded	.79	.83
30% occluded	30% occluded	.76	.80
50% occluded	50% occluded	.67	.70
70% occluded	50% occluded	.56	.63
70% occluded	70% occluded	.51	.55

TABLE VI. EXPERIMENTS DETAILS ON OCCLUDED IMAGES

The experiment on the effect of image occlusion on Convolutional Neural Network (CNN) performance demonstrated that occlusions in images can have a significant impact on the accuracy and precision of the model. The results showed that as the level of occlusion increased, the performance of the CNN decreased. This highlights the importance of considering image occlusions when designing computer vision systems and training CNN models. The findings of this experiment can inform future research on developing more robust CNNs that can better handle image occlusions and provide more accurate predictions

C. Different Object Size Images

Object size can greatly affect the performance of computer vision systems, especially when recognizing objects in images. In this experiment, we aim to evaluate the effect of object size on the performance of a Convolutional Neural Network (CNN).

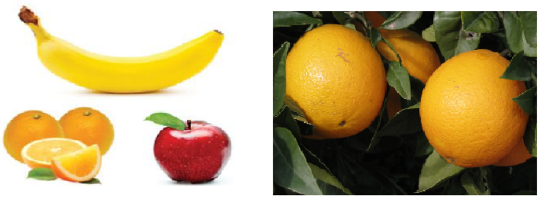

Fig 5 : a. Different Size Objects and b. Similar Size Objects

Methods:

1. Data Collection: We collected a dataset of 470 images of objects, with varying object sizes.

2. Preprocessing: The images were preprocessed to ensure that they are of the same size and format.

3. Object Size Variation: We vary the size of the objects in the images by resizing them to different proportions.

4. CNN Model Training: We trained a CNN model on the dataset which has Categorical cross entropy as a loss function and use Adam as an optimizer

5. Model Evaluation: We evaluated the performance of the CNN model by comparing its predictions on images with different object sizes.

This experiment provides insights into the effect of object size on the performance of Convolutional Neural Networks. The results of this experiment can be used to improve the robustness of computer vision systems to object size variations and enhance the accuracy of object recognition in real-world

Training Data	Test Data	Train Accuracy	Test Accuracy
similar size	Similar size	.97	.90
different size	different size	.71	.67
different size	Similar size	.70	.75
50%different	similar size	.84	.79
50%different	50%different	.83	.73
70%different	Similar size	.77	.80
70%different	70%different30%Similar	.75	.74

TABLE VII. EXPERIMENTS DETAILS ON DIFFERENT OBJECT SIZE IMAGES

scenarios. From The Above Table, we can say that multiple objects of different sizes in an image make a negative effect on object detection. it basically decreases the accuracy as loss is in increasing side

V. CONCLUSION

The quality of an image, object size, and occlusion can significantly impact object detection performance. Various studies have shown that object detection methods can be sensitive to image quality factors such as noise, blur, and compression artifacts, which can lead to reduced accuracy and increased false positives. Object size is another important factor that can affect detection performance, as objects at different scales require different levels of feature representation and can be more difficult to detect. Occlusion is yet another challenging problem in object detection, as partially or fully occluded objects can be difficult to detect and classify.

To address these challenges, researchers have developed various techniques to improve object detection performance in the presence of image quality degradation, object size variation, and occlusion. Some of these methods include data augmentation, multi-task learning, attention mechanisms, and advanced network architectures.

Overall, while object detection is a complex and challenging problem, ongoing research is actively working to address the impact of these factors on detection performance and to develop more robust and accurate methods for object detection in real-world scenarios.

Disclaimer: *By submitting this paper, the authors confirm that the work presented is original and has not been previously published, nor is it under consideration for publication elsewhere. The authors acknowledge that Intel and its affiliates are not responsible for the content of this paper, including any errors or omissions, and make no warranties, express or implied, as to the accuracy, reliability, or completeness of the information presented.*

REFERENCES

[1] T. Li, K.-S. Chen, and M. Jin, "Analysis and simulation on imaging performance of backward and forward bistatic synthetic aperture radar," *Remote Sensing*, vol. 10, no. 11, p. 1676, 2018.

[2] A. Rius, E. Cardellach, F. Fabra, W. Li, S. Ribó, and M. Hernández-Pajares, "Feasibility of gnss-r ice sheet altimetry in greenland using tds-1," *Remote Sensing*, vol. 9, no. 7, p. 742, 2017.

[3] K. Kang, H. Li, T. Xiao, W. Ouyang, J. Yan, X. Liu, and X. Wang, "Object detection in videos with tubelet proposal networks," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2017, pp. 727–735.

[4] A. Kendall, Y. Gal, and R. Cipolla, "Multi-task learning using uncertainty to weigh losses for scene geometry and semantics," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 7482–7491.

[5] W. Chen, S. Daneau, F. Mannan, and F. Heide, "Steady-state non-line-of-sight imaging," in *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2019, pp. 6790–6799.

[6] H. Kaneyasu, S. B. Etter, T. Sakai, and M. Sigrist, "Evolution of the filamentary 3-kelvin phase in pb- ru- sr 2 ruo 4 josephson junctions," *Physical Review B*, vol. 92, no. 13, p. 134515, 2015.

[7] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *Advances in neural information processing systems*, vol. 28, 2015.

[8] W. Liu, D. Anguelov, D. Erhan, C. Szegedy, S. Reed, C.-Y. Fu, and A. C. Berg, "Ssd: Single shot multibox detector," in *Computer Vision–ECCV 2016: 14th European Conference, Amsterdam, The Netherlands, October 11–14, 2016, Proceedings, Part I 14*. Springer, 2016, pp. 21–37.

[9] J. Redmon and A. Farhadi, "Yolov3: An incremental improvement," *arXiv preprint arXiv:1804.02767*, 2018.

[10] S. Zhang, X. Zhu, Z. Lei, H. Shi, X. Wang, and S. Z. Li, "S3fd: Single shot scale-invariant face detector," in *Proceedings of the IEEE international conference on computer vision*, 2017, pp. 192–201.

[11] MUHAMMEDBUYUKKINACI, "Fruit images for object detection," *https://www.kaggle.com/datasets/mbkinaci/fruit-images-for-object-detection*.