A Proposed Exploratory Study of Object Detectors to Learn the Influence of Datasets on Model Performance

Vidya Kamath, Renuka A.
Department of Computer Science and Engineering, Manipal Institute of Technology, Manipal Academy of Higher Education, Manipal, India
E-mail: vidyakuldeep@gmail.com, renuka.prabhu@manipal.edu

Abstract. The quality of the images used to train the models in the field of object detection using deep learning models is critical in determining the model’s quality. However, there are very few methods for exploring these images in datasets to see what aspects in these images have a significant impact on the model’s performance. This could be one of the reasons why the models don’t match human perceptions. There is a need for more study that can suggest unique methodologies to address the topic at hand because the existing literature overlooks this line of thought. As a result, this paper provides a methodology based on exploratory sequential design, which may be used to identify several aspects of images in the dataset that influence model performance.

1. Introduction
Deep learning has been used for the task of object detection in the past decade. Many researchers have given more importance to the process of training a model to improve its accuracy. Whereas there is not as much focus on the dataset used to train the model [1]. The major reason being that object detection models are treated as black boxes. Another possible reason for this is that image datasets have less tools for data exploration than tabular datasets. Apart from that, the availability of pretrained models is too appealing, therefore one is tempted to choose a pretrained model over worrying about the impact of a dataset on the model’s performance. It has been a trend to use pretrained architectures, hence developing new models from scratch has been not so appealing choice in the current era, let alone fear about the influence of the dataset used.

Nevertheless, image datasets are no excuse when it comes to data exploration and not understanding the level of influence the dataset poses on our model’s learning capability is a serious flaw that must be kept in mind and avoided for achieving true success. Exploring the dataset to get the best of our models is the key in any machine learning process. The same holds to deep learning models as well and hence more research studies are required in finding the impact and influence of the image datasets on the performance of models [2]. There are multiple things to be considered when dealing with an image dataset compared to a normal textual dataset. First is the level of influence an image dataset has on the model’s performance in terms of the quality of the dataset [1]. Second is the fact that the quality of a dataset when it comes to image dataset is not measurable. If at all we do find that the dataset has an impact on the model, finding out which are the major factors of the dataset that are responsible for this impact is a question that is not clarified yet.
This major problem in any deep learning-based object detection might be a possible reason due to which the object detectors present today are not able to perform at human level perceptions. Hence, this work aims to propose a methodology that uses an exploratory study to find the influence of the dataset on the performance of object detection models.

2. Related Work
Image quality assessment has been done in the past for a variety of applications. Assessment of images in terms of metrics like PSNR (Peak Signal to Noise Ratio) or SSIM (Structural Similarity Index) [3], or using algorithms such as full reference image quality assessment (FR-IQA) algorithms, reduced-reference image quality assessment (RR-IQA) algorithms, and no-reference/blind image quality assessment (NR-IQA) algorithms, for judicial image quality assessment [4], or using a deep learning model itself to assess the quality of an image [2, 5, 6, 7] are few among the limited set of works that try to find the measure to quality of an image. Nonetheless, we need methods which find major aspects related to a dataset in terms of its influence on the performance of the model as identified by Tianxing et al. in [1]. He tries to study four aspects of a dataset namely Dataset Equilibrium, Dataset Size, Quality of Label, and Dataset Contamination. However, more aspects of the images in the dataset must be explored, and new factors must be discovered.

The lack of work in this area may possibly be from the fact that there does not exist a theory that would prove that the quality of the dataset influences the quality of the model. With the insights obtained in the recent [8, 9, 10], this theory seems to have a certainty which largely requires an exploratory study that can prove the influence as well as quantitatively be able to measure the factors that influence which can lead to the future research to invent new data exploration tools when it comes to image datasets.

3. Research Gaps Identified
(i) There is a need for a specific study that potentially learns the impact and influence posed by a dataset on the model built using that dataset.
(ii) Limited attempts have been done to study the influence of factors like dataset size, label quality, dataset equilibrium, noise and contamination on the performance of object detection models, but it is not yet clarified, due to which it is difficult to know which factors predominantly affect the model’s performance.

4. Problem Statement, Research Questions and Objectives
4.1. Problem Statement
The models for object detection using deep learning architectures have been improving its performance over time but it is still too low compared to a human level perception. The quality of models has been assessed mainly with the accuracy, speed and size of the model and lot of research has happened to improve the model performance by increasing the training time and dataset quantity but very less importance has been given to the quality of the dataset used for training as well as the major aspects of a dataset which can influence the model’s performance. There is a need for research that identifies the reasons for the low performance of deep learning models for object detection in relation to the dataset used by the model and its quality.

4.2. Research Questions
RQ1. Does the quality of the dataset used have any impact on the quality of the model trained on it?
RQ2. What are the major aspects of a dataset which can be used to measure the quality of that dataset?
RQ3. How do the factors of a dataset influence the performance of the model trained on it?
RQ4. Can a deep learning model be trained to assess the quality of data when the model itself depends on the data used to train it?
4.3. Objectives of Study
(i) To propose a method for qualitative study that can identify the major factors related to a dataset which influences the object detection model in terms of human level perception.
(ii) To perform an experimental analysis of the effect of major factors related to a dataset on the performance of the object detection model in terms of accuracy and speed.

5. Rationale of the Research
There are several reasons that can prompt and motivate the inception of any research work. This section tries to justify the choices made in the proposed research plan to solve the current problem at hand. Firstly, the modern applications such as smart homes, autonomous vehicle navigation, medical image processing, robotic applications, smart devices and IoT edges, style transfer, face recognition, person re-identification and video surveillance monitoring [11, 12, 13] to name a few, have extensively been relying on the deep learning models for the primary task of detecting an object. Being able to make these models better in terms of accuracy has been the primary concern of the researchers. It is known that these models have been developed on the principles of artificial neural networks that try to replicate the model of the human nervous system. Even though it is true that the computer neuron is much faster than the human neuron [14], the models have not yet been able to match human level perception. The factors that mainly affect the model have not yet been clearly listed and hence this work tries to propose a methodology that can first explore the independent variables in the model and then find the dependency relation among them [1]. This type of research will not only add to the existing body of knowledge in the field, but it will also have a significant impact on the progress of all associated subdomains and applications.

Since the research questions going to be addressed are one of the first of its types that attempts to explore and find the relation between quality of the dataset used with the model, and due to the lack of literature available in its support, the work first tries to find what the independent and dependent variables in the system are on an ontological stance. Once the variables are identified, the experimental analysis on an epistemological stance tries to find how the variables in the system relate to each other [15]. Hence, a mixed method approach is chosen to perform an exploratory study [16] that is going to identify possible directions for the future research works. Finally, on a higher note, this work tries to fill the research gaps identified in this field by proposing a methodology for conducting studies related to the problem statement, which is one of its kind and attempts to answer the unanswered questions in a systematic way and opens new insights for the future researchers.

6. Research Methodology

![Figure 1](image.png)

**Figure 1.** The overall methodology of the exploratory study.
The major steps of the exploratory study are briefly depicted in Figure 1. It mainly tries to explore the variables of the system. This study majorly assumes that the quality of the dataset affects the quality of the model and tries to reason it inductively through the outcome of the qualitative step. This way the study explores new variables and even though it is subjective in nature, it tries to test it on the system in an experimental fashion to discover the truth deductively in the following quantitative step. So, the first objective focuses on building the hypothesis inductively whereas the second objective tries to test the hypothesis deductively. Even though the study involves both qualitative and quantitative approaches, the decision making, and conclusion drawing is highly based on the quantitative results obtained and hence it is more quantitatively driven with some aid of qualitative inputs. Therefore, from the philosophical view, this work addresses the problem in hand in a positivist paradigm by giving maximum focus to discover and examine the variables in a scientific way [17].

6.1. Qualitative Study

In the qualitative step, the raw image data must first be collected and then modified to change the major aspects of an image that can possibly be a candidate for independent variables. Data batches must be created based on the acquired data, which may then be physically analyzed for quality. Now, if a questionnaire can be created using these images to record ranks from a focus group of choice, it gives a human level perception to each of the images. The questionnaire can be used to conduct a survey using Google Forms or any other mode of convenience, with applicants ranking the images on a scale of 1 to 5 according to their perception. The survey’s responses can then be used to examine and discover the primary factors that influence human perception of images. In this work, we have synthetically generated values for a questionnaire to measure the ranks for four major factors of the images, namely resolution, blur, contrast and noise. Note that in case of actual implementation of the questionnaire survey, it will require a preliminary study to optimally choose the focus group for the survey depending upon the factors chosen for testing. We chose obvious parameters like resolution, noise etc., for the sake of explaining how the process works. The metrics that assess image quality, such as PSNR or SSIM, are among the few other factors that may be examined in this way.

6.2. Quantitative Analysis
The major factors identified in the previous step will be tested on a few sample models in deep learning-based object detection to see if they have influenced the performance of these models on a quantitative scale, as assessed by dependent variables like speed and accuracy of detection. We tested two popular models of object detection in this stage using the outcome of our synthetic data analysis done in step two, as input. The objective is to find a relation among the dependent and independent variables by quantitative measures through an ablation study. The major models considered for this study are the tiny YOLOv4 [18] and the Mask-RCNN [19] where tiny YOLOv4 is one of the recent one among one stage object detectors and Mask-RCNN among the two stage object detectors.
The ablation study involves changing the independent variables of the system one by one in order to find how it affects the performance of the model. The outcome of each trial is supposed to be recorded to finally conduct an in-depth analysis. The relation between the variables when found, will not only help to test our hypothesis but will also help the future researchers in choosing the right datasets for their models.

7. Data Collection
Since this work follows an exploratory study, the data collection happens in several stages of the study. From the initial step of collecting raw image data in the beginning up until the collection of the outcomes of the final experiments, the data collection must be done in several stages. The Figure 4 gives a broader conception of the various forms of the data being handled in the proposed study.

In the first stage the images captured must have a variety of characteristics, such as resolution, pixel quality, aspect ratios, noise, and so on, so that these aspects of the data may be investigated and potentially used as an independent variable in the system. The data used as input for this study is collected both from primary as well as secondary sources. The primary source mainly contains images captured from mobile phones. These are particularly used as candidates for images that have low resolution and high noise. The secondary data used in this work are the image datasets popularly available on the internet for training the deep learning models. These images usually come with good quality.

Stage two is the qualitative data collection step where selected images from the previous stage need to be used to prepare a questionnaire that aims to address the quality aspects of the data by considering the human insights. This mainly involves choosing the focus group and collecting and storing their responses. Apart from this, the data collection process continues throughout the stage while conducting, gathering results, storing the answers and analyzing the outcomes by taking several forms.

In stage three, the train and test sets for experimentally analyzing the models for object detection needs to be formed using the data collected from stage one. The batches must be formed to test the factors identified in the qualitative step. Apart from this, the process of collecting data continues and takes the form of observation to collect quantitative data, where the results of the experiments need to be tabulated for further analysis.
8. Data Analysis, Results and Inference
In this section, we will look at how to evaluate the data that was obtained in the three stages and how to make conclusions from this exploratory study’s analysis.

8.1. Analysis of the Image datasets and raw images.
The first step here is to identify some factors that can be tested in next stage. Since the major factors that degrade the quality of any given image are the noise and the blur, we select these two for our analysis. Figure 5 and 6 shows how noise and blur can influence the quality of the image.

| Dataset        | No. of Images | Contrast | Avg Resolution | Blur | Noise |
|----------------|---------------|----------|----------------|------|-------|
| ImageNet       | ≈1164(tiny 200 class) | High     | 482X418        | Low  | Low   |
| Corel 1000     | 1000          | Low      | 300X200        | Medium | Medium |
| MS COCO        | ≈1600 (tiny 80 class) | High     | 600X400        | Low  | Low   |
| Caltech 101    | ≈9146         | Low      | 300X200        | Medium | Medium |
| Caltech 256    | ≈5160 (tiny 256 class) | Medium   | 351X351        | Medium | Medium |
| Custom data    | ≈1570         | Low      | 120X260        | High  | High  |

Figure 4. The various stages of the Data Collection Process.

Figure 5. How noise degrades the performance of the model.

Figure 6. Blur as a mediating variable that reduces the image quality.
Resolution and contrast are two more criteria that can influence the quality of an image, thus they were considered for our study. These factors give a positive influence on the quality of the image. Higher resolutions usually go with high contrasts.

The details about the image datasets collected through primary and secondary sources is tabulated in Table 1. The size of the original data set is too huge in some cases where a class-based sampling was used to reduce the number of images per class. The custom data is the frames from the recorded low-resolution videos with high noise and part of it is labeled for objects present using the labelImg tool.

8.2. Prepare, analyze and evaluate the qualitative questionnaire

The analysis of qualitative data collection basically must consider two aspects. First, the formation of the questionnaire, target audience and the number of questions used, how well the questionnaire generalizes to help find our independent variables; should be analyzed to make sure the questionnaire meets the standards and does not fail to deliver its task. Secondly, it is required to analyze the results obtained through the questionnaire to eliminate any bias due to the subjective qualitative data collection. This can be done by ensuring that the findings follow a pattern of normal distribution.

To form the questionnaire, selection of the images from the available dataset can be done using a non-probability sampling technique on each dataset that requires the knowledge of the researcher. This does not guarantee that every image has an equal chance of being selected, but due to its simplicity and cost effectiveness, this method can be helpful in speeding up the questionnaire formation step.

The synthetically generated data in our work assumed a questionnaire consisting of ten questions on the selected images. The ranks were generated using our perceptions for the images and then compared with the original quality of the data sample to see if there exists any correlation between the choices made from the metrics compared to the one made by a human. The human brain can easily perform classification on an ordinal scale and hence the test metrics are converted to ordinal values of high, medium and low for easy comparisons. If a questionnaire is created in the real world, the results will be more trustworthy than our synthetically generated ranks; however, the generated data is thought to be sufficient for evaluating the process.

| Variable       | No. of Samples | No. of Responses |
|----------------|----------------|------------------|
| blur low       | 9              | 558              |
| blur medium    | 8              | 496              |
| blur high      | 3              | 186              |
| noise low      | 9              | 558              |
| noise medium   | 6              | 372              |
| noise high     | 5              | 310              |
| contrast low   | 10             | 620              |
| contrast medium| 3              | 186              |
| contrast high  | 7              | 434              |
| resolution low | 8              | 496              |
| resolution medium | 5       | 310              |
| resolution high| 7              | 434              |

A sample size of 20 images was chosen across the datasets and the corresponding values of contrast, blur, resolution and noise were noted from the sample image. The number of samples used for the test against each variable is shown in Table 2. Through a random sampling, 10 images were selected for the preparation of the questionnaire. To reduce bias, the procedure was repeated four times resulting in four different questionnaires with different images. The questionnaires were labeled as set A, set B, set C and set D by ensuring that each sample gets selected twice, and the responses for
the same were generated and ranks were recorded. After tabulating the questionnaire results with the independent values of the variables, the variable that has the greatest impact on the questionnaire results needs to be identified, which is done using a statistical test to check for correlation between each independent variable and the questionnaire result.

The total responses obtained through the questionnaire was 124 which leads to a total of 1240 rank values. Next, the frequencies of the ranks generated across the samples from all the questionnaire responses were calculated for the tabulated responses. The total responses for each of the samples was 62 which led to a total of 1240 responses. This matches the expected ranks from the questionnaire and hence verifies that no value is missed from the response. Figure 7 below shows the histogram for frequencies obtained for the rank against each of the variables considered by plotting the ranks from 1-5 along the x axis and their corresponding frequencies along the y axis.

**Table 3.** Formulated hypotheses for the correlation tests conducted.

- **H10:** There is no linear relationship between rank given in the questionnaire for an image and the blur level in that image.
- **H20:** There is no linear relationship between rank given in the questionnaire for an image and the contrast level of that image.
- **H30:** There is no linear relationship between rank given in the questionnaire for an image and the noise present in that image.
- **H40:** There is no linear relationship between rank given in the questionnaire for an image and the resolution of that image.

![Figure 7. The frequency distribution of the values from the image recording.](image-url)
The variables that need to be possibly tested against a deep learning object detection model must be the ones that majorly influence human perception. Hence the correlation of the variable with the human perception (ranks from the questionnaire) needs to be tested as the next step. In order to execute it, the SPSS software was used to perform Karl Pearson’s correlation test of each of the four variables with the ranks obtained. This test was chosen due to the requirement of finding the correlation that answers our research question, as well as the fact that these variables are both non-metric where a KP test is the best possible test usually done. Hypotheses formulated are listed in Table 3. The values for low, medium and high for each of the variables were assigned ordinal values 1, 2, 3 respectively while entering the values in the SPSS.

According to the results of the test, all the null hypotheses were rejected at 0.05 significance level. This implies that these variables have a high influence on human perception. Also, the results show that noise and blur have a positive correlation with the rank while contrast and resolution have negative correlation. This implies that more the resolution of an image less is the value of the rank. As rank 1 is considered the best during the questionnaire, it is obvious that resolution and contrast of an image play a dominant role in the decision of ranking while blur and noise are equally dominant while giving a rank of 5. Hence the dataset with more resolution must improve the object detection model’s performance while the other way must decrease it.

Hence, the qualitative study draws light to the fact that the variables like resolution and contrast must be considered as independent variables while training any deep learning model to make the model match human perception. The results can be tested practically as shown in the next step where quantitative analysis of the model is done based on these chosen factors.

8.3. Experiment and Results

Finally, we need to analyze the model’s performance metrics measured by training and testing them for those variables identified from the previous step. The dataset level exploration decides what factors of the dataset are most relevant whereas the model exploration in this phase does the model’s performance assessment. By taking models for object detection like YOLO-v4 and Mask-RCNN which are generalizable, the models must perform similarly for all the datasets considered for the experiment. The failure of this will prove that the dataset quality has an impact on the model’s performance. Also, the factors considered in each dataset in the experiment can be compared in the ablation study to find out which factor has the most effect on the performance of the models.

| Table 4. Hypotheses formulated for the paired sample t-tests to test whether there is increase in speed and average precision in the two deep learning models. (* -contrast and resolution) |
|---------------------------------------------------------------|
| H50: There is no increase in the speed of MaskRCNN by changing the independent variables*. |
| H60: There is no increase in the Average Precision of MaskRCNN by changing the independent variables*. |
| H70: There is no increase in the speed of YOLOv4 by changing the independent variables*. |
| H80: There is no increase in the Average Precision of YOLOv4 by changing the independent variables*. |

The measures that can be used to evaluate the overall performance of any object detection model in terms of the two major dependent variables, are namely the model’s accuracy in terms of its average precision (AP) and the speed of the model given in frames per second (fps).

The experiments were carried out using Google Colab GPU. Each model was trained with the sample dataset consisting 98 images for an epoch of 10 and was tested for the dependent variables which are the metrics like accuracy and precision. The same model was retrained with the same dataset sample a second time by changing the independent variables of the images in the dataset.
(resolution and contrast) after which the steps were repeated, and the metrics were again noted. The comparison of these dependent metrics before and after the change of the independent variable promises to answer the research questions in an accurate way.

Since it is required to test whether the change in the contrast and resolution of the dataset affected the model or not, a paired sample t-test was done on the values of the dependent variables before and after the change in the independent variables. The hypotheses formulated are shown in Table 4.

The results for the test reject the null hypotheses for the AP in both the models at 0.05 significance level, whereas while considering speed, the null hypothesis is rejected only in case of YOLOv4, but it fails to reject the null hypothesis (H50) in case of MaskRCNN.

8.4. Inference drawn
The inference drawn from the first objective is majorly statistical and hence it requires more knowledge on the subject matter to make the decisions. Even though the inferences drawn through this qualitative method does not give evidence to be said entirely true, it does help us in moving further with the testing phase where the inference drawn can be more significant and can also explain the reasons for how the identified variables affect the model’s performance.

The results from the quantitative analysis showed that resolution and contrast increase the performance of the model in terms of average precision. This shows that deep learning models for object detection must give maximum importance to the resolution and contrast of their train/test set when compared to the validation set. Nevertheless, it could partially answer the RQ1 on an ontological stance because it was found that the quality of the image could not affect the performance of one of the two models in terms of speed.

While answering RQ2, even though only resolution and contrast were chosen and tested due to their nature of negative correlation with ranks (positive correlation with human perception), blur and noise can also be considered to effect the change in the model’s performance by trying to reduce these variables inferring from their nature of positive correlation with the ranks. Limited to this work, this again partially answers the RQ3. The possibility again needs to be further tested using quantitative analysis in order to consider these variables on a stronger ground. Hence this work only lists a few of the factors that could possibly affect the model’s performance but not all. The process, on the other hand, can be used to evaluate any other feature of an image to see how much it influences a model’s performance.

While considering the pace of this study, it is found that such an analysis using qualitative preliminary work involving questionnaires seems to be a lot of time consuming while not to mention the other limitations that a qualitative study can bring in. But using a deep learning model to access the quality of an image, while the model itself is not matching the human level perception does not seem to be a good choice. It is hence advisable to carry out qualitative studies whenever human level perceptions are considered. This answers the RQ4, but the choice is always left open to the researcher.

The resolution and contrast proved to affect the performance of the object detection models chosen for the experiment, positively. Even though the result was obvious from our qualitative study, the experimental quantitative analysis strengthened the results. Similarly, the experiment alone would not be able to prove the correlation between these independent variables like resolution and contrast with the human level perception, for which the qualitative study done prior made loads of ground. This justifies the requirement of an exploratory study done in this work without which the results of both the qualitative study as well as quantitative analysis would individually fail to answer the research questions.

9. Conclusion and Future Scope
Even though tremendous work has been done in the field of object detection using deep learning, this work has proposed a new methodology for exploring image datasets and discovered new research questions that need to be addressed in this field, which had a huge lack of literature. Hence, the need
for an exploratory study was met by this work. New gaps were identified, and the objectives were set which attempted to fill the gaps to their best.

The quality of an image chosen in the dataset sample while training and testing a model has an impact on the performance of the model. This was evidently shown throughout, along with justifying results and inferences. This addition to the body of knowledge can be a driving force to the future models of object detection in making them more realistic and perspective.

The work tested only a few of the factors that can decide the image quality with only two prominent models in object detection. Moreover, the ranks of the questionnaire were synthetically generated. Many other factors like blur and noise can also be experimentally evaluated to check for its influence on the model’s performance, which is left as future work. Also, the work can be repeated with different object detection models to see if the obtained results differ in any aspects due to the limitations of this study.

Other than this, the major factors in the quantitative analysis of the model were chosen based on results obtained from the qualitative study. This can possibly make the work more subjective and dependent on the focus group to which the questionnaire gets shared as well as the errors in the non-probability sampling done to choose the samples for the test. Such an error however would be proved itself in the following quantitative analysis. Based on this assumption, this exploratory study stands good.

Altogether, this work could bring light to those aspects of the deep learning model while dealing with its datasets, which were never touched before and hence the work would possibly give a new direction for research with new research questions to be answered.

References
[1] T. He, S. Yu, Z. Wang, J. Li, and Z. Chen, “From data quality to model quality: An exploratory study on deep learning,” in Proceedings of the 11th Asia-Pacific Symposium on Internetware, ser. Internetware ’19. Association for Computing Machinery, 2019.
[2] L. Kong, A. A. Ikusan, R. Dai, and J. Zhu, “Blind image quality prediction for object detection,” 2019 IEEE Conference on Multimedia Information Processing and Retrieval (MIPR), pp. 216–221, 2019.
[3] Z. Wang, A. Bovik, H. Sheikh, and E. Simoncelli, “Image quality assessment: from error visibility to structural similarity,” IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600–612, 2004.
[4] Q. Gu, W. Cai, S. Yu, and Z. Chen, “An exploratory study on judicial image quality assessment based on deep learning,” in 2019 IEEE 19th International Conference on Software Quality, Reliability and Security (QRS), 2019, pp. 300–305.
[5] W. Hou, X. Gao, D. Tao, and X. Li, “Blind image quality assessment via deep learning,” IEEE Transactions on Neural Networks and Learning Systems, vol. 26, no. 6, pp. 1275–1286, 2015.
[6] S. Bianco, L. Celona, P. Napoletano, and R. Schettini, “On the use of deep learning for blind image quality assessment,” Signal, Image and Video Processing, vol. 12, no. 2, pp. 355–362, 2018.
[7] S. Bosse, D. Maniry, K.-R. Muller, T. Wiegand, and W. Samek, “Deep neural networks for no-reference and full-reference image quality assessment,” IEEE Transactions on Image Processing, vol. 27, no. 1, pp. 206–219, 2018.
[8] H. Talebi and P. Milanfar, “Nima: Neural image assessment,” IEEE Transactions on Image Processing, vol. 27, no. 8, pp. 3998–4011, 2018.
[9] M. Aliaksei, “Deep image quality assessment,” A story in Towards Data Science, 2021. [Online]. Available: https://towardsdatascience.com/deep-image-quality-assessment-30ad71641fac
[10] M. Eric, “Image classification with variable input resolution in keras,” A Story in MindBoard, 2020. [Online]. Available: https://medium.com/mindboard/image-classification-with-
variable-input-resolution-in-keras-cbfbef576126f

[11] C. B. Murthy, M. F. Hashmi, N. D. Bokde, and Z. W. Geem, “Investigations of object detection in images/videos using various deep learning techniques and embedded platforms—a comprehensive review,” Applied Sciences, vol. 10, no. 9, 2020.

[12] P. Li, D. Wang, L. Wang, and H. Lu, “Deep visual tracking: Review and experimental comparison,” Pattern Recognit., vol. 76, pp. 323–338, 2018.

[13] Z. Zou, Z. Shi, Y. Guo, and J. Ye, “Object detection in 20 years: A survey,” Arxiv, 2019.

[14] M. Tom-M, Machine Learning: Mc-Graw-Hill- International Student Edition, 1997

[15] M. M. Al-Ababneh, “Linking ontology, epistemology and research methodology,” Science and Philosophy, vol. 8, no. 1, pp. 75–91, 2020.

[16] J. W. Creswell, Research design: Qualitative, quantitative, and mixed methods approaches. Thousand Oaks, California: SAGE Publications, 2014.

[17] Y. Park, L. Konge, and A. Artino, “The positivism paradigm of research,” Academic medicine : journal of the Association of American Medical Colleges, vol. 95, p. 690–694, 2020.

[18] C.-Y. Wang, A. Bochkovskiy, and H.-y. Liao, “Scaled-yolov4: Scaling cross stage partial network,” 11 2020.

[19] K. He, G. Gkioxari, P. Dollar, and R. Girshick, “Mask r-cnn,” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 42, no. 2, pp. 386–397, 2020.

[20] O. Russakovsky, J. Deng, H. Su et al., “ImageNet Large Scale Visual Recognition Challenge,” International Journal of Computer Vision (IJCV), vol. 115, no. 3, pp. 211–252, 2015.

[21] Corel-1K and Corel-10K databases, [online] Available: http://wang.ist.psu.edu/docs/related.shtml.

[22] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollar, and C. Zitnick, “Microsoft coco: Common objects in context,” Lecture Notes in Computer Science , vol. 8693 LNCS, no. PART 5, pp. 740–755, 2014.

[23] L. Fei-Fei, R. Fergus, and P. Perona, “Learning generative visual models from few training examples: An incremental bayesian approach tested on 101 object categories,” in 2004 Conference on Computer Vision and Pattern Recognition Workshop, 2004, pp. 178–178.

[24] G. Griffin, A. Holub, and P. Perona, “Caltech-256 object category dataset,” CalTech Report, 03 2007.