

# Effectiveness of Data Resampling in Mitigating Class Imbalance for Object Detection

Michał Tomaszewski<sup>1</sup>, Jakub Osuchowski<sup>1</sup>

<sup>1</sup> Opole University of Technology, Prószkowska 76 St., Opole, 45-758, Poland

## Abstract

Mitigating class imbalance for object detection on digital images is a critical challenge in the field of computer vision. This problem stems from the uneven distribution of object classes within image datasets, where some classes are significantly more prevalent than others. In object detection tasks, the primary goal is to identify and locate various objects within images accurately. However, when dealing with imbalanced datasets, several significant issues arise. Firstly, the training of machine learning models on imbalanced data can result in bias, where models tend to perform well on majority classes but struggle to recognize minority classes effectively. This bias is due to the disproportionate number of samples available for each class during training, leading to inadequate learning for underrepresented classes. Secondly, imbalanced datasets can lead to reduced detection accuracy, particularly for the minority classes. Models trained on such data may exhibit high overall accuracy but perform poorly when it comes to identifying rare objects or those from underrepresented classes. Moreover, the problem of class imbalance can lead to the loss of valuable information. Minority classes may include objects or instances that are crucial for the specific application, yet they are often overlooked due to their scarcity in the dataset.

Furthermore, models trained on imbalanced data may struggle to generalize effectively to real-world scenarios where class distributions are more balanced. This limitation can hinder the practical applicability of object detection systems.

The article aims to investigate the impact of data resampling methods on improving the object detection model YOLOv8m when dealing with an imbalanced image dataset.

The described initial research involves using an object detection dataset with skewed class distributions and applying various resampling techniques like oversampling and undersampling to balance the data. The research used The Insulator Defect Image Dataset (IDID) representing power line insulators, which contains a large class depicting undamaged insulators and two other, relatively small classes depicting two types of damaged insulators.

The implications of this research are practical, as it guides practitioners and researchers in selecting the most suitable resampling approach to address class imbalance in object detection tasks. Ultimately, this knowledge contributes to the development of more robust and reliable computer vision systems for real-world applications.

## Keywords

object detection, class imbalance, resampling, YOLO

## 1. Introduction

The problem of imbalanced classes in image datasets refers to a situation where the distribution of different classes or categories within the dataset is highly skewed. In other words, some classes have significantly more instances or samples compared to others, leading to an imbalance in class representation. This issue is particularly prevalent in image datasets used for machine learning and computer vision tasks.

Imbalanced classes can pose several challenges in machine learning and image analysis:

---

Proceedings ITTAP'2023: 3rd International Workshop on Information Technologies: Theoretical and Applied Problems, November 22–24, 2023, Ternopil, Ukraine, Opole, Poland

EMAIL: m.tomaszewski@po.edu.pl (A. 1); j.osuchowski@po.edu.pl (A. 2)

ORCID: 0000-0001-6672-3971 (A. 1); 0000-0002-9404-966X (A. 2)



© 2020 Copyright for this paper by its authors.

Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).

CEUR Workshop Proceedings (CEUR-WS.org)

- Model bias: Machine learning models trained on imbalanced datasets may develop biases towards the majority class, as they have more examples to learn from. This can result in poor performance for minority classes.
- Reduced accuracy: Imbalanced datasets can lead to misleading accuracy metrics. A model may achieve a high accuracy by simply predicting the majority class most of the time, while failing to correctly classify minority class instances.
- Limited generalization: Models trained on imbalanced data may struggle to generalize well to real-world scenarios where class distributions are more balanced. They may perform poorly on underrepresented classes.
- Data collection cost: In some cases, collecting more data for underrepresented classes can be challenging, time-consuming, or expensive.

To address this problem, researchers and practitioners employ various techniques, including:

- Resampling: Oversampling the minority classes (adding more instances) or undersampling the majority classes (removing instances) to balance class distribution.
- Cost-Sensitive Learning: Assigning different misclassification costs to different classes to make the model more sensitive to minority classes.
- Transfer Learning: Leveraging pre-trained models or features from large datasets to improve the performance on imbalanced data.
- Synthetic Data Generation: Creating synthetic samples for underrepresented classes to increase their representation in the dataset.
- Evaluation Metrics: Using appropriate evaluation metrics, such as precision, recall, F1-score, or Average Precision (AP), that consider the imbalanced nature of the dataset.

These strategies aim to improve the model's ability to recognize and classify minority classes effectively, leading to more balanced and accurate image analysis results. Mitigating class imbalance for object detection is essential for building robust and effective computer vision systems that can accurately detect and locate objects across a wide range of classes, regardless of their representation in the dataset.

Research aimed at mitigating the class imbalance problem in image datasets holds substantial significance in the realm of computer vision and machine learning. This importance can be better understood through several key aspects.

Firstly, addressing the challenge of imbalanced datasets directly translates into improved model accuracy. Imbalanced datasets often lead to models excelling at recognizing majority classes while struggling with minority or underrepresented ones. Research in this area strives to create more balanced models that achieve higher overall accuracy by rectifying this bias toward majority classes.

Secondly, these studies enhance the capability to detect rare or uncommon damage in critical infrastructure, which holds paramount importance in various real-world applications. For instance, in the realm of power infrastructure, the identification of rare instances of damage or anomalies is crucial for ensuring the integrity and safety of power systems. Research aimed at mitigating data imbalance ensures that machine learning models can effectively recognize and address these infrequent cases, thereby enhancing the overall resilience and maintenance of critical infrastructure. Similarly, for instance, in medical imaging [1], identifying rare diseases or anomalies is vital for patient care. The research on mitigating dataset imbalance ensures that machine learning models can effectively recognize and handle these infrequent instances, thereby improving the quality of healthcare and diagnostics.

Furthermore, the generalization of models to real-world scenarios is a crucial consideration. Models trained on imbalanced datasets might falter when applied to scenarios with more balanced class distributions. This limitation can hinder the practical applicability of computer vision systems. Research in this field aims to enhance model generalization, making them more robust and reliable across a spectrum of real-world settings.

Another vital aspect is the prevention of data loss. Imbalanced datasets often lead to the underrepresentation of minority classes, resulting in the loss of valuable information. This research endeavours to ensure that all classes are adequately considered in the learning process, thereby harnessing the full potential of available data.

Moreover, in applications where fairness and ethical considerations are paramount, such as facial recognition technology, addressing class imbalance becomes crucial. It helps prevent the unfair favouring or disadvantaging of particular groups or classes, mitigating biases and ethical concerns associated with imbalanced datasets.

Cost efficiency is another notable benefit. Collecting and annotating datasets can be resource-intensive. Research on balancing imbalanced datasets can lead to more efficient data collection by prioritizing efforts in areas that require representation the most, optimizing resource allocation. Ongoing research in this area drives advancements in state-of-the-art techniques and methodologies for addressing class imbalance. This benefits the broader machine learning community by continually improving the toolkit available to researchers and practitioners, ultimately advancing the field as a whole.

The objective of the presented initial study is to assess the impact of data resampling methods on object detection tasks performed by deep learning neural network. Specifically, the study aims to determine whether resampling techniques, such as different variants of oversampling or undersampling, can mitigate the challenges posed by class imbalance in object detection. As mentioned earlier, this problem applies in particular to technical systems and methods of their visual inspection, during which a lot of image material is collected illustrating correctly operating components and attempts to identify relatively rare failures, which, due to their nature, are usually represented by few instances in data sets teaching.

Therefore, the described study concerns the detection of damage in power line insulators. Overhead power lines are used to transmit electricity, which means that they are also one of the key elements of each country's security. Because the transmission of electricity plays a critical role in supporting the industry, economy, and defence, power disruptions resulting from failures in overhead power lines can lead to severe consequences for the entire country. Additionally, the safety and reliability of high-voltage lines can significantly impact the quality of life for citizens. To minimize interruptions in electricity supply and reduce their duration, electricity distribution companies should conduct frequent and comprehensive inspections of power network[2]. Despite the ongoing research efforts in the field of automated inspection of high-voltage lines, for example [3],[4],[5],[6],[7],[8],[9] there are still existing challenges and limitations that require addressing. To establish a dependable power grid, it is imperative to develop robust methods for the automated detection of various defects in its components.

An overhead power line consists of three essential parts: conductive wires, transmission towers, and power insulators. The primary functions of overhead line insulators are twofold: they provide insulation between the conductor and the ground and tower structure while also offering mechanical support to the conductive wires. Particularly in the case of high-voltage lines, insulators used in overhead lines must undergo regular inspection due to their susceptibility to degradation. The analysis was carried out using The Insulator Defect Image Dataset (IDID) [10], which depicts power line insulators. The dataset contains images with undamaged insulators and different examples of damaged insulators.

## 2. Related works

The problem of data imbalance is widely described in the literature on machine learning methods. Publication [11] showed that there are many challenges in the vast field of imbalanced data problems that require attention from the research community and intensive study. There are still many unaudited directions to be investigated in this branch of machine learning. The work is based on earlier review articles such as [12],[13],[14] on different approaches to the problem of data imbalance.

In a survey [15] a lot of studies are summarized and discussed, exploring a number of advanced techniques for learning from imbalanced data with deep neural networks. It has been shown that traditional machine learning techniques for handling class imbalance can be extended to deep learning models with success. The survey also finds that nearly all research in this area has been focused on computer vision tasks with convolutional neural networks.

Study [1] investigates the existing deep learning techniques to address class imbalanced data, a critical challenge in real-world applications like fraud detection and cancer diagnosis. The authors state that despite the growing popularity of deep learning, limited empirical research exists in this area. The survey examines existing studies, highlighting the effectiveness and limitations of deep learning

models, particularly in computer vision tasks, and identifies areas for future research to bridge the gap in this important field of study. Several areas for future work are outlined. Utilizing the various approaches across a broader range of datasets and varying degrees of class imbalance, assessing their performance using multiple complementary metrics, and presenting statistical evidence will aid in selecting the most suitable deep learning method for upcoming applications dealing with class imbalance. Experimenting with deep learning methods in handling class imbalance within the realms of big data and rare class scenarios holds significant benefits for the advancement of big data analytics. Additionally, the authors showed that further investigations involving non-convolutional deep neural networks are necessary to establish the generalizability of the presented methods to alternative architectural frameworks.

In [16] presents a review of various learning issues due to the imbalanced distribution of data and different approaches to handle the problem of imbalanced data in classification. It was noted that the impact of imbalance on classification is baleful and the mentioned effect increases with the extent of a task, concluding that the classification of imbalanced data is an extensive research subject in the field of machine learning.

In [17] Authors investigate complexity measures effectiveness on real imbalanced datasets and how they are affected by applying different data imbalance treatments. The issue of classification with imbalanced datasets is also extensively presented in [18]. The authors additionally provide a website [19] containing many materials on the discussed subject.

Despite a large amount of research on data imbalance, the fact of creating new machine learning techniques, in particular object detection algorithms, makes it necessary to conduct empirical research showing the effectiveness of using different techniques for balancing different data sets for new algorithms.

Relatively few works concern the problem of data imbalance in fault detection in particular technical systems. For example, the paper [20] delves into the impact of data sampling techniques on improving cross-project defect prediction (CPDP) models. Employing eight data resampling methods, they resampled datasets and integrated them into CPDP model training after applying the Nearest Neighbour filter. The results demonstrated that data resampling methods effectively improved recall and performance measures but had limited success in terms of AUC performance. While these methods helped mitigate class imbalance issues, further research is needed to enhance prediction performance.

In [21] proposed a framework that addresses the problem of data imbalance in supervised classification techniques for non-technical losses (NTL) in electrical power grids detection through resampling techniques. The Authors stated that an issue that other studies received not enough attention in the research is the imbalance between fraudulent and non-fraudulent data, which can have a significant negative impact on the performance of supervised learning methods. The same NTL issue was described in the paper [22], where deep reinforcement learning (DRL) was used to solve the data imbalanced problem. The advantage of the proposed method is that the classification method is adopted to use the partial input features without a pre-processing method for input feature selection.

Work [23] describes the potential of hierarchical Federated Learning (FL) in the Internet of Things (IoT) heterogeneous systems. In particular, the Authors proposed an optimized solution for user assignment and resource allocation over hierarchical FL architecture for IoT heterogeneous systems. This work focuses on a generic class of machine learning models that are trained using gradient-descent-based schemes while considering the practical constraints of non-uniformly distributed data across different users.

In summary, the existing literature on addressing class imbalance in image datasets primarily focuses on general applications of proposed methods, overlooking the specific challenges posed by technical systems such as industrial automation, robotics, and critical infrastructure. This gap in the literature hinders the development of tailored solutions for real-world technical systems where misclassification can have significant consequences. Research in this area needs to prioritize the adaptation of algorithms to meet the unique data characteristics and operational constraints of technical systems. Additionally, there is a need for investigations into the scalability, efficiency, interpretability, and adaptability of class imbalance solutions in specific technical environments. Bridging this gap will be essential to ensure the robust and reliable deployment of image-based systems in critical technical domains.

### 3. Methodology

#### 3.1. Oversampling

The study concerned the use of various variants of oversampling and undersampling in order to check the impact of such action on the efficiency of object detection. The mathematical model for oversampling (upsampling) individual classes in a dataset to the size of the largest class using random sampling with replacement can be defined as follow.

Let's assume we have a dataset  $D$  consisting of  $N$  samples, and each sample is associated with a class label  $y_i$ , where  $i$  denotes the index of the sample. We aim to oversample each class to the size of the largest class.

We define  $K$  as the target size of samples in each class, which is equal to the size of the largest class in the dataset. Our goal is to generate a new dataset  $D'$  where each class  $i$  will have exactly  $K$  samples.

The mathematical model for oversampling classes using random sampling with replacement can be described as follows:

1. Find the size of the largest class:  $max\_class\_size = \max(N_i)$  for all classes  $i$
2. For each class  $i$ :
  - a. If  $N_i < K$ , randomly select  $K - N_i$  samples from class  $i$  with replacement until the target size is reached.
  - b. If  $N_i \geq K$ , include all samples from class  $i$  in  $D'$ .

#### 3.2. Undersampling

The mathematical model for undersampling (downsampling) individual classes in a dataset to the size of the smallest class using random sampling without replacement shown below:

Let's assume we have a dataset  $D$  consisting of  $N$  samples, and each sample is associated with a class label  $y_i$ , where  $i$  denotes the index of the sample. We aim to undersample each class to the size of the smallest class.

We define  $K$  as the target size of samples in each class, which is equal to the size of the smallest class in the dataset. Our goal is to generate a new dataset  $D'$  where each class  $i$  will have exactly  $K$  samples.

The mathematical model for undersampling classes using random sampling without replacement can be described as follows:

1. Find the size of the smallest class:  $min\_class\_size = \min(N_i)$  for all classes  $i$
2. For each class  $i$ :
  - a. If  $N_i > K$ , randomly select  $K$  samples from class  $i$  without replacement.
  - b. If  $N_i \leq K$ , include all samples from class  $i$  in  $D'$ .

### 4. The deep learning architecture used for object detection

YOLOv8m [24] was used to perform this experiment. It is the newest state-of-the-art You Only Look Once (YOLO) model that can be used for object detection, image classification, and instance segmentation tasks. YOLOv8 was developed by Ultralytics, who also created the influential and industry-defining YOLOv5 model. YOLOv8 includes numerous architectural and developer experience changes and improvements over YOLOv5. YOLO has been nurtured by the computer vision community since its first launch in 2015 by Joseph Redmond. In the early versions, YOLO was maintained in C code in a custom deep learning framework Darknet. Subsequent versions were developed in PyTorch - a deep learning Python framework. In addition to its robust model foundation, the YOLO maintainers have shown a dedicated effort to foster a thriving software ecosystem for the model. They proactively address concerns and enhance the repository's functionalities in response to the community's needs.

There are five different models of YOLOv8 models made for object detection: YOLOv8n (nano), YOLOv8s (small), YOLOv8m (medium), YOLOv8l (large), YOLOv8x (extra-large). YOLOv8 Nano (YOLOv8n) is the fastest and smallest, while YOLOv8 Extra Large (YOLOv8x) is the most accurate yet the slowest.

More information about the current version of YOLO is given in [25],[26],[27].

## 5. Evaluation metrics

The following metrics were used to evaluate the study results: Precision, Recall, F1 score and mAP (with different variations). Precision measures the accuracy of the positive predictions made by the model. In the context of image detection, precision represents the proportion of correctly detected objects (true positives) out of all objects that the model predicted as positive (true positives + false positives). Recall measures the model's ability to correctly detect all instances of the object in the dataset. In the context of image detection, recall represents the proportion of correctly detected objects (true positives) out of all actual instances of the object in the dataset (true positives + false negatives). The F1 score is the harmonic mean of precision and recall. It provides a balance between precision and recall and is especially useful when you want to consider both false positives and false negatives in your evaluation. The F1 score ranges from 0 to 1, where a higher score indicates better performance. A perfect F1 score of 1 means that the model has achieved both high precision and high recall, implying that it can detect all instances of the object with no false positives.

Mean average precision (mAP) is a crucial evaluation metric used in object detection tasks. It provides a comprehensive and quantitative measure of the accuracy and reliability of models employed in these tasks. Average precision (AP) is calculated for each class by computing the precision-recall curve. These measures are computed at different confidence thresholds to generate multiple data points forming the curve. mAP is then obtained by averaging AP values across all object classes. It signifies the overall accuracy and performance of the model across various classes, thereby providing a single, consolidated evaluation score.

An mAP of 1.0 indicates perfect accuracy, while lower values (like 0.95, 0.75 or 0.5) imply potential inaccuracies in IOU calculation. IOU, or Intersection over Union, is a commonly used metric in object detection tasks to measure the accuracy of bounding box predictions made by a model. It assesses the overlap between the predicted bounding box and the ground truth (actual) bounding box for an object in an image.

Refer to the YOLOv8 documentation for more information on each metric [28].

## 6. Data Collection and Preprocessing

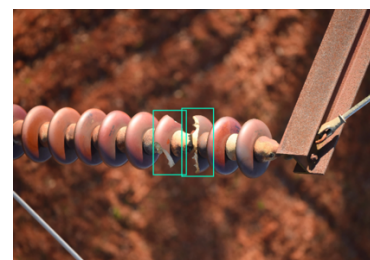
The Insulator Defect Image Dataset (IDID) [14] representing power line insulators, which contains a large class depicting undamaged insulators and two other, relatively small classes depicting two types of damaged insulators was used in the research. The dataset IDID contains 1596 images in total. Figure 1: Examples representing each of the listed classes (GIS, FDIS and BIS). shows examples representing each of the listed classes.



GIS



FDIS



BIS

**Figure 1:** Examples representing each of the listed classes (GIS, FDIS and BIS).

Annotations describing damaged and undamaged individual disks - components of power insulators - have been added to individual images of the dataset. The purpose of the detection was precisely these individual components, along with assigning them to the appropriate class.

Publicly available dataset IDID is divided into two parts: “train dataset” and “test dataset”. For the purposes of this study, 1,000 examples representing undamaged isolator disks (“good insulator shell” – GIS class) and 200 examples of damaged isolator disks (100 examples for “flashover damage insulator shell” – FDIS class and 100 examples for “broken insulator shell” – BIS class) were selected from the part “train dataset”. These images in different resampling variants were used to train the detector.

All images from the part “test dataset” were used to test the effectiveness of the detector. It contains 930 examples (instances) of class GIS, 70 examples of class FDIS and 66 examples of class BIS.

## 7. Experimental Results and Discussion

Based on the images selected from the IDID dataset, described in the "Data Collection and Preprocessing" section, and using various resampling methods, 6 variants of the training sets listed below were prepared:

- Variant I: 1000 GIS, 100 FDIS, 100 BIS, starting case, imbalanced dataset,
- Variant II: 200 GIS, 100 FDIS, 100 BIS, undersampling so that the train dataset contains the same number of undamaged objects (GIS class) and damaged objects (FDIS class + BIS class),
- Variant III: 100 GIS, 100 FDIS, 100 BIS, undersampling to the smallest classes FDIS, BIS,
- Variant IV: 1000 GIS, 500 FDIS, 500 BIS, oversampling so that the train dataset contains the same number of undamaged objects (GIS class) and damaged objects (FDIS class + BIS class),
- Variant V: 1000 GIS, 600 FDIS, 600 BIS, oversampling the smallest number of classes to triple their size and, as a result, reduce disproportions between classes,
- Variant VI: 1000 GIS, 1000 FDIS, 1000 BIS, oversampling to the largest class GIS.

The listed numbers indicate the number of instances in each class. For each of the presented variants, the learning process was carried out. The mean results obtained for the test dataset are presented in the Table 1.

**Table 1**  
Summary of mean results for individual resampling variants.

Variant	Variant I	Variant II	Variant III	Variant IV	Variant V	Variant VI
<b>Class size</b>	1000 GIS 100 FDIS 100 BIS	200 GIS 100 FDIS 100 BIS	100 GIS 100 FDIS 100 BIS	1000 GIS 500 FDIS 500 BIS	1000 GIS 600 FDIS 600 BIS	1000 GIS 1000 FDIS 1000 BIS
<b>Mean AP at IoU threshold of 0.5 for all classes</b>	0.498	0.535	0.452	0.580	0.501	0.455
<b>Mean AP at IoU threshold of 0.75 for all classes</b>	0.414	0.462	0.382	0.478	0.400	0.362
<b>Mean AP at IoU threshold of 0.5 to 0.95 for all classes</b>	0.336	0.373	0.318	0.409	0.341	0.307

The best efficiency results were obtained for Variant IV - oversampling so that the train dataset contains the same number of undamaged objects (GIS class) and damaged objects (FDIS class + BIS class) for each of the analyzed mean metrics. This variant did not reduce the number of examples for

the most numerous GIS class, at the same time performing five-fold oversampling of the least numerous classes allowed for better representation of these classes (FDIS and BIS).

In most cases, the worst results were obtained for Variant VI - oversampling to the largest GIS class. Only a slightly worse (but very similar) result was obtained for the "Mean AP at IoU threshold of 0.5 for all classes" metric calculated for Variant III. Thus, it can be seen that the multiple duplication of examples from a small class brings effects only up to a certain threshold (Variant IV), while a significant multiplication of the same instances leads to the deterioration of the detection results. Results for individual resampling variants for each class were shown in Table 2.

**Table 2**

Summary of results for individual resampling variants for each class.

	Variant I			Variant II			Variant III		
	GIS	FDIS	BIS	GIS	FDIS	BIS	GIS	FDIS	BIS
<b>Class size</b>	1000	100	100	200	100	100	100	100	100
<b>F1 score</b>	0.747	0.408	0.310	0.625	0.328	0.247	0.554	0.235	0.199
<b>Precision</b>	0.782	0.345	0.428	0.799	0.235	0.167	0.905	0.141	0.131
<b>Recall</b>	0.719	0.5	0.242	0.513	0.543	0.47	0.399	0.714	0.41
<b>mAP</b>	0.6	0.272	0.138	0.573	0.34	0.205	0.595	0.285	0.073
	Variant IV			Variant V			Variant VI		
	GIS	FDIS	BIS	GIS	FDIS	BIS	GIS	FDIS	BIS
<b>Class size</b>	1000	500	500	1000	600	600	1000	1000	1000
<b>F1 score</b>	0.785	0.413	0.423	0.723	0.29	0.349	0.818	0.280	0.218
<b>Precision</b>	0.879	0.359	0.458	0.833	0.26	0.313	0.742	0.510	0.188
<b>Recall</b>	0.710	0.486	0.393	0.639	0.329	0.394	0.912	0.193	0.258
<b>mAP</b>	0.644	0.354	0.227	0.622	0.232	0.168	0.637	0.206	0.077

During the analysis of the results for individual classes, it was noticed that the F1 score metric obtained the best values for Variant III, as well as for mean metrics. On the other hand, undersampling of the entire dataset to the least numerous class (Variant III) resulted in a significant reduction in the size of the entire training dataset and, as a result, a decrease in the F1 score metric (the worst result).

In contrast, the Precision and Recall metrics yielded mixed results. The multiplication of the number of instances improved the Precision parameter, but at the same time caused a deterioration of the Recall parameter value because the detector was unable to find objects that differed from the given pattern (problems related to the generalization of the neural network).

The mAP metric calculated for each class separately was the best for Variant IV, similar to the collectively calculated mAP and F1 score. In other cases, worse results were obtained, however, for this metric, the negative impact of resampling in individual variants cannot be unequivocally assessed, because ambiguous results of the instance detection efficiency for different classes were obtained - the analysis requires more research in this area.

Addressing the class imbalance in object detection is indeed a critical challenge, and it's essential to acknowledge the potential limitations and challenges encountered during the described experiment. Here are some limitations and challenges that should be considered.

The effectiveness of resampling techniques heavily relies on the quality and representativeness of the dataset. If the dataset does not capture the real-world distribution of objects accurately, the results may not generalize well to practical scenarios. The success of resampling methods on the specific dataset used (IDID) does not guarantee similar outcomes on other datasets with different object classes and distributions. The research should assess the generalizability of the findings. The choice of hyperparameters for resampling methods and object detection models can significantly impact the results. Optimizing these hyperparameters is crucial for achieving the best performance.

For practical applications, it's important to assess the impact of resampling on the real-time inference speed of object detection systems, especially if used in latency-sensitive environments. In practical applications, there may be resource constraints, such as limited labelled data for rare classes. Research should explore strategies for addressing class imbalance in resource-constrained settings.



Resampling techniques may affect the interpretability of object detection models. Ensuring that the models provide meaningful explanations for their detections is essential, particularly in critical applications.

Addressing these limitations and challenges is crucial for advancing the field of object detection in the presence of class imbalance and for ensuring the applicability of research findings to real-world computer vision applications. The article describes the preliminary research carried out for only one detection algorithm, five variants of data resampling and one dataset. In order to face the limitations and challenges, the authors plan to continue the described research on limiting the impact of class imbalance on object detection efficiency.

## 8. Conclusion and Future Work

In summary, the article investigates how data resampling techniques can enhance the performance of object detection models in the presence of class imbalance, offering insights and guidance for improving the effectiveness of these models in practical applications.

The most efficient outcomes were achieved with variant, where oversampling was applied to ensure an equal number of undamaged (GIS class) and damaged objects (FDIS class + BIS class) in the training dataset for each of the average metrics under investigation. This approach preserved the abundance of GIS class examples while significantly improving the representation of the less common FDIS and BIS classes through a five-fold oversampling. In the remaining scenarios, diverse outcomes were observed, and these are elaborated upon in greater detail within the article.

In this article, it was described initial research on the influence of various methods of resampling on detection efficiency. The potential future research directions are:

- Testing other known methods of resampling training data,
- Development of a new resampling method using the knowledge gained during the described experiments,
- Testing other algorithms for detecting and classifying objects on digital images and their effectiveness depending on the use of different data resampling methods,
- Study of imbalance reduction in test dataset,
- Investigation of the use of semi-supervised learning approaches to leverage both labelled and unlabeled data, potentially reducing the reliance on extensive labelled data for minority classes,
- Applying Generative Adversarial Networks (GANs) to generate synthetic samples for minority classes in object detection datasets. This approach can potentially improve model performance by providing more diverse training data,
- Investigate the use of ensemble models that combine multiple object detection models trained on resampled datasets to improve overall performance and robustness.

The findings of this research have practical implications for improving the accuracy and reliability of object detection models in real-world applications.

Understanding the effectiveness of data resampling can guide practitioners and researchers in selecting the appropriate approach to handle class imbalance in object detection. This knowledge can lead to more efficient computer vision systems, particularly in scenarios where imbalanced classes are common. Mitigating class imbalance for object detection is vital for developing robust computer vision systems capable of accurately identifying and locating objects across a diverse range of classes, regardless of their representation in the dataset. This research area continues to evolve, contributing to advancements in object detection and broader applications of computer vision in fields like autonomous driving, medical imaging, and more.

## 9. References

- [1] D.P. Rana, R.G. Mehta Data Preprocessing, Active Learning, and Cost Perceptive Approaches for Resolving Data Imbalance. IGI Global; 2021. doi:<https://doi.org/10.4018/978-1-7998-7371-6>
- [2] M. Tomaszewski, R. Gasz, J. Osuchowski, Detection of Power Line Insulators in Digital Images Based on the Transformed Colour Intensity Profiles. *Sensors*. 2023;23(6):3343-3343. doi:<https://doi.org/10.3390/s23063343>
- [3] L. Yang, J. Fan, Y. Liu, E. Li, J. Peng, Z. Liang, A Review on State-of-the-Art Power Line Inspection Techniques. *IEEE Transactions on Instrumentation and Measurement*. 2020;69(12):9350-9365. doi:<https://doi.org/10.1109/tim.2020.3031194>
- [4] X. Liu, X. Miao, H. Jiang, J. Chen, Data analysis in visual power line inspection: An in-depth review of deep learning for component detection and fault diagnosis. *Annual Reviews in Control*. Published online October 2020. doi:<https://doi.org/10.1016/j.arcontrol.2020.09.002>
- [5] A. Raza, A. Benrabah, T. Alquthami, M. Akmal, A Review of Fault Diagnosing Methods in Power Transmission Systems. *Applied Sciences*. 2020;10(4):1312. doi:<https://doi.org/10.3390/app10041312>
- [6] W. Liu, Z. Liu, A. Nunez, Z. Han, Unified Deep Learning Architecture for the Detection of All Catenary Support Components. *IEEE Access*. 2020;8:17049-17059. doi:<https://doi.org/10.1109/access.2020.2967831>
- [7] VN. Nguyen, R. Jenssen, D. Roverso, Automatic autonomous vision-based power line inspection: A review of current status and the potential role of deep learning. *International Journal of Electrical Power & Energy Systems*. 2018;99:107-120. doi:<https://doi.org/10.1016/j.ijepes.2017.12.016>
- [8] M. Tomaszewski, P. Michalski P, J. Osuchowski, Evaluation of Power Insulator Detection Efficiency with the Use of Limited Training Dataset. *Applied Sciences*. 2020;10(6):2104. doi:<https://doi.org/10.3390/app10062104>
- [9] M. Tomaszewski, P. Michalski, J. Osuchowski, Object Description Based on Local Features Repeatability. *Advances in intelligent systems and computing*. Published online January 1, 2021:255-267. doi:[https://doi.org/10.1007/978-3-030-72254-8\\_28](https://doi.org/10.1007/978-3-030-72254-8_28)
- [10] P. Kulkarni, D. Lewis, Insulator Defect Detection. Accessed September 6, 2023. <https://iee-dataport.org/competitions/insulator-defect-detection>
- [11] B. Krawczyk, Learning from imbalanced data: open challenges and future directions. *Progress in Artificial Intelligence*. 2016;5(4):221-232. doi:<https://doi.org/10.1007/s13748-016-0094-0>
- [12] H He, EA. Garcia, Learning from Imbalanced Data. *IEEE Transactions on Knowledge and Data Engineering*. 2009;21(9):1263-1284. doi:<https://doi.org/10.1109/tkde.2008.239>
- [13] He. Haibo, Y. Ma, J. Wiley, *Imbalanced Learning : Foundations, Algorithms, and Applications*. John Wiley & Sons, Cop; 2013.
- [14] P. Branco, L. Torgo, R. Ribeiro, A Survey of Predictive Modelling under Imbalanced Distributions. *arXiv.org*. doi:<https://doi.org/10.48550/arXiv.1505.01658>
- [15] J.M. Johnson, T.M. Khoshgoftaar. Survey on deep learning with class imbalance. *Journal of Big Data*. 2019;6(1). doi:<https://doi.org/10.1186/s40537-019-0192-5>
- [16] P. Kumar, R. Bhatnagar, K. Gaur, A. Bhatnagar, Classification of Imbalanced Data: Review of Methods and Applications. *IOP Conference Series: Materials Science and Engineering*. 2021;1099(1):012077. doi:<https://doi.org/10.1088/1757-899x/1099/1/012077>
- [17] VH. Barella, LPF. Garcia, MCP de Souto, AC. Lorena, de Carvalho ACPLF. Assessing the data complexity of imbalanced datasets. *Information Sciences*. 2021;553:83-109. doi:<https://doi.org/10.1016/j.ins.2020.12.006>

- [18] V. López, A. Fernández, S. García, V. Palade, F. Herrera, An insight into classification with imbalanced data: Empirical results and current trends on using data intrinsic characteristics. *Information Sciences*. 2013;250:113-141. doi:<https://doi.org/10.1016/j.ins.2013.07.007>
- [19] Classification with Imbalanced Datasets, *Soft Computing and Intelligent Information Systems*. sci2s.ugr.es. <https://sci2s.ugr.es/imbanced>
- [20] KE. Bennin, A. Tahir, SG. MacDonell, J. Börstler, An empirical study on the effectiveness of data resampling approaches for cross-project software defect prediction. *IET Software*. Published online November 28, 2021. doi:<https://doi.org/10.1049/sfw2.12052>
- [21] G. Figueroa, YS. Chen, NF. Avila, CC. Chu, Improved practices in machine learning algorithms for NTL detection with imbalanced data. Published online July 1, 2017. doi:<https://doi.org/10.1109/pesgm.2017.8273852>
- [22] J. Lee, YG. Sun, I. Sim, SH. Kim, DI. Kim, JY. Kim, Non-Technical Loss Detection Using Deep Reinforcement Learning for Feature Cost Efficiency and Imbalanced Dataset. *IEEE Access*. 2022;10:27084-27095. doi:<https://doi.org/10.1109/access.2022.3156948>
- [23] AA. Abdellatif, N. Mhaisen, A. Mohamed, et al., Communication-efficient hierarchical federated learning for IoT heterogeneous systems with imbalanced data. *Future Generation Computer Systems*. 2022;128:406-419. doi:<https://doi.org/10.1016/j.future.2021.10.016>
- [24] GitHub - ultralytics/ultralytics at blog.roboflow.com. GitHub. Accessed September 7, 2023. <https://github.com/ultralytics/ultralytics?ref=blog.roboflow.com>
- [25] J. Solawetz, Francesco, What is YOLOv8? The Ultimate Guide. Roboflow Blog. Published January 11, 2023. <https://blog.roboflow.com/whats-new-in-yolov8/>
- [26] G. Wang, Y. Chen, P. An, H Hu, J. Hu, T Huang, UAV-YOLOv8: A Small-Object-Detection Model Based on Improved YOLOv8 for UAV Aerial Photography Scenarios. *Sensors*. 2023;23(16):7190-7190. doi:<https://doi.org/10.3390/s23167190>
- [27] J. Terven, D. Cordova-Esparaza, a comprehensive review of yolo: from yolov1 to yolov8 and beyond under review in *acm computing surveys*.; 2023. <https://arxiv.org/pdf/2304.00501.pdf>
- [28] Ultralytics.com, Ultralytics metrics - Accessed September 7, 2023. <https://docs.ultralytics.com/reference/utils/metrics/#ultralytics.utils.metrics.Metric>