Addressing Class Imbalance in Image Data: A Comparative Study of Resampling Techniques and Deep Learning Models

Wonjoon Hwang [1] **Zhizheng Wang** [2] **Khoa Le** [3] **Chih-Yuan Tung** [4] **Brett Ruane** [5]

Abstract

Class imbalance in image datasets is a critical problem for improving the generalizability of deep learning models. This study investigates the impact of various resampling techniques on the performance of different deep learning architectures in image classification tasks. We evaluate the effectiveness of Nearest Neighbor Interpolation, Bilinear and Bicubic Interpolation, Gaussian Pyramid, SRCNN, and Wavelet Transform on models such as GoogLeNet, ResNet50, AlexNet, and VGG16 using the imbalanced Sports Image Classification dataset. Our results highlight the importance of choosing the right combination of resampling methods and model architectures to effectively handle class imbalances and enhance classification performance.

1. Introduction

Large real-world problem solving based on state-of-the-art technologies often involves dealing with real-world noise, previously unseen forms, and significant variation in the distribution of data. Data imbalance introduces challenges like over-fitting and biased performances within machine learning algorithms. Images are unique from different forms of data input owing to their intrinsic ability to carry more information compared to numeric and textual forms; hence, their usage is quite imperative.

Various resampling strategies are employed to improve model performance and reduce over-fitting. However, the selection of which model performs best for which resampling technique, when dealing with imbalanced image data, is still an open research question. Class imbalance in image classification brings about serious problems, such as biased prediction and over-fitting to the dominant classes. Images, unlike numeric or textual data, carry intrinsic high-dimensional features, which need special handling for imbalanced datasets. This study presents an exploration of how resampling methods combined with deep learning architectures can be used to mitigate these challenges.

We concentrate on the Sports Image Classification dataset, which suffers from severe class imbalances and assess the model performance using five resampling techniques. Our contributions are important in three aspects:

- A comparative analysis of resampling techniques and their impact on CNN architectures.
- Performance evaluation using minimal training epochs to highlight resampling efficiency.
- Insights into the interplay between model depth and resampling effectiveness.

2. Background

Class imbalance in datasets often results in models that favor majority classes, which then yields biased predictions. This paper balances class distribution by resampling; that is, it up-samples the minority classes and down-samples the majority classes.

Deep learning architectures, particularly Convolutional Neural Networks (CNNs), vary in their ability to generalize from imbalanced datasets:

- GoogLeNet: A deep architecture with inception modules, for high-dimensional feature extraction.
- ResNet50: Famous for its residual connections, which allow the effective training of much deeper networks.
- AlexNet and VGG16: Less complex architectures that usually fail to deal with complex datasets because of less depth.

[1]Department of Data Science, University of Georgia, Athens, USA [2]Department of Data Science, University of Georgia, Athens, USA [3]Department of Computer Science, University of Georgia, Athens, USA [4]Department of Computer Science, University of Georgia, Athens, USA [5]Department of Computer Science, University of Georgia, Athens, USA. Correspondence to: Wonjoon Hwang <wh42003@uga.edu>, Zhizheng Wang <zhizheng.wang@uga.edu>, Khoa Le <khoa.le1@uga.edu>, Chih-Yuan Tung <tung@uga.edu>, Brett Ruane <bgr28288@uga.edu>.

Proceedings of the 38^{th} International Conference on Machine Learning, PMLR 139, 2021.

Resampling techniques such as Bicubic Interpolation and Gaussian Pyramid enhance the representation of the data, but it is model-dependent.

3. Methodology

3.1. Dataset

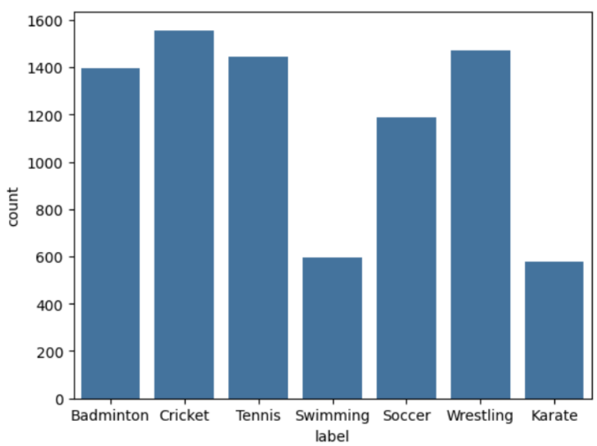


Figure 1. Sports Image Classification dataset

The Sports Image Classification dataset from Kaggle fit to this experiment. It contains seven categories of images of various sports such as Cricket, Wrestling, Tennis, and Karate. In this highly imbalanced dataset, it contains 1,556 images of Cricket, while Karate is the class with the least number of images-578 images. Using this dataset, we tested some famous deep learning architectures: ResNet50, VGG16, GoogleNet, and AlexNet using different resampling techniques.

3.2. Deep Learning Architectures

GoogLeNet, ResNet50, AlexNet, and VGG16 were trained with identical parameters, including a learning rate of 0.001, batch size of 32, and SGD optimizer.

3.3. Resampling Techniques

1. Nearest Neighbor Interpolation
2. Bilinear and Bicubic Interpolation
3. Downsampling by Gaussian Pyramid
4. SRCNN: Super-Resolution Using Deep Learning
5. Wavelet Transform Resampling

4. Experiments

4.1. Setup

- Cross-validation split: 80% training - 20% validation.
- Predefined test set for final evaluation.
- Training epochs: Limited to five to emphasize resampling impact.

4.2. Results

4.2.1. Nearest Neighbor Interpolation

Model	Base	Upsampling	Downsampling
ResNet50	Training Accuracy = 54.33%, Validation Accuracy = 81.90%	Training Accuracy = 88.35%, Validation Accuracy = 86.33%	Training and Validation Accuracy stagnate below 20%.
VGG16	Training Accuracy = 17.72%, Validation Accuracy = 17.56%	Training Accuracy = 43.03%, Validation Accuracy = 40.28%	Accuracy stagnates near 18–19%.
GoogLeNet	Training Accuracy = 94.45%, Validation Accuracy = 89.37%	Training Accuracy = 94.82%, Validation Accuracy = 89.91%	Training Accuracy = 94.68%, Validation Accuracy = 80.62%.
AlexNet	Training Accuracy = 19.10%, Validation Accuracy = 17.01%	Accuracies remaining below 20%.	Accuracies remaining below 20%.

Figure 2. result of NNI

Nearest Neighbor Interpolation (NNI) is one of the simplest resampling techniques, where each pixel in the resized image is assigned the value of the nearest pixel in the original image. This method does not involve averaging or blending pixel values, making it computationally efficient and straightforward to implement. Despite its simplicity, the effectiveness of NNI depends heavily on the underlying dataset and the model architecture.

Initial Experiments(No resampling)

- GoogLeNet achieved the highest performance, with a training accuracy of 95.62% and a validation accuracy of 90.28%, leveraging its depth to generalize effectively despite the simplistic interpolation method.

- ResNet50 followed closely, with a training accuracy of 88.65% and a validation accuracy of 81.35%, showcasing its ability to handle high-dimensional patterns.

- AlexNet and VGG16 struggled significantly, with validation accuracies of 18.47% and 17.44%, respectively, reflecting the limitations of their shallow architectures when tasked with complex datasets.

Results After Resampling

- NNI Upsampling When applied for upsampling, NNI improved the performance of deeper models by increasing

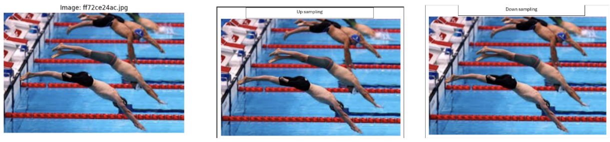


Figure 3. example of NNI

Model / Accuracy	Imbalanced	Bilinear Interpolation Upsampling	Bilinear Interpolation Downsampling	Bicubic Interpolation Upsampling	Bicubic Interpolation Downsampling
ResNet50	88.00%	88.62%	18.99%	**89.65%**	86.22%
GoogLeNet	94.18%	94.06%	94.04%	**94.29%**	94.01%
VGG16	**58.21%**	53.61%	18.37%	30.82%	17.99%
AlexNet	18.87%	18.69%	**19.05%**	18.95%	17.82%

Figure 4. Result of Bilinear and bicubic

the representation of underrepresented classes while maintaining computational efficiency. GoogLeNet retained its position as the top-performing model, with minimal degradation in validation accuracy. ResNet50 also showed modest improvements, though it was more susceptible to overfitting compared to GoogLeNet.

- NNI Downsampling The downsampling process, on the other hand, had an adverse impact on most models, as it often removed critical details necessary for accurate classification. While simpler models like AlexNet and VGG16 benefited slightly from the reduced input complexity, their overall performance remained poor due to architectural constraints. GoogLeNet and ResNet50 experienced performance drops as the loss of fine-grained details hindered their ability to extract meaningful features.

4.2.2. Bilinear and Bicubic Interpolation

The general resampling methods include bilinear and bicubic interpolation; both can be applied in upsampling or downsampling. In terms of changing the image resolution, they both work by interpolating pixel values according to the spatial relations of the neighboring pixels.

- Bilinear interpolation involves computing the weighted average of the four nearest neighboring pixels. - Bicubic interpolation further extends this idea by taking the weighted average of 16 proximate neighbors, which generates much smoother, higher detailed images, with greater scaling factors.

Bilinear and bicubic interpolations were performed mainly for upsampling and downsampling. In upsampling, some images had their size changed either by bilinear or bicubic interpolations to increase their resolutions; then, such images were used as inputs into deep learning models for training and evaluation on how well the models make use of such increased information.

Here, downsampling is performed by reducing the resolution of the image using the respective interpolation technique in such a way that it will not affect important features. These images would be used for training the model to check the performance of the model with less information.

Initial Experiments (No Resampling)
Preliminary results showed that the bilinear and bicubic interpolations, when used with an imbalanced dataset, had salient performance differences arising among the deep learning models which were tested on the imbalanced dataset. Among these, the best performances were from GoogLeNet, with a 94.18% training accuracy, and a 94.06% validation accuracy while bilinear upsampling was used, which refers to a better handling of the variability and complexity within the dataset.

Figure 5. Example of Bilinear and bicubic

Results After Resampling:
The application of bicubic upsampling to ResNet50 resulted in a training accuracy of 88.62% and a validation accuracy of 86.22%, demonstrating its proficiency in hierarchical feature extraction. While AlexNet and VGG16, much shallower architectures, did face significant challenges, AlexNet managed to achieve a validation accuracy of 18.69% with bilinear upsampling, and VGG16 performed a bit better, with a 53.61% validation accuracy for the same configuration. The limiting factor for both was their shallow depth and capacity constraints in modeling complex patterns that are intrinsic to the dataset.

Applying different resampling methods resulted in generally better quality for the models, though differences varied depending on methodologies and model architecture. Among these, the best performance was obtained with bicubic upsampling for ResNet50 and GoogLeNet, yielding a validation accuracy of 94.29% for GoogLeNet and improving ResNet50 to 89.65%. This is most likely due to the fact that bicubic interpolation can preserve the high-frequency details more effectively, which are important for deep models.

Still, the performance gains for AlexNet and VGG16 were marginal. With bicubic upsampling, AlexNet reached a validation accuracy of 18.95%, while VGG16 managed only 30.82%, reflecting their intrinsic structural limitations. In all models, the use of downsampling by either method invariably led to degraded performances, as fundamental information in the images was lost, thus limiting the models' ability to extract meaningful features.

4.2.3. DOWNSAMPLING BY GAUSSIAN PYRAMID

In the case of class imbalance, the Gaussian Pyramid technique was used to slowly reduce the resolution of the images by using a Gaussian filter, which smoothly removes high-frequency noise while keeping structural information.

Model	Validation Accuracy (Imbalanced)	Validation Accuracy (Resampled)	Improvement (%)
GoogLeNet	91.13%	86.44%	-4.69%
ResNet50	83.84%	86.81%	+2.97%
AlexNet	17.01%	50.71%	+33.70%
VGG16	18.47%	40.52%	+22.05%

Figure 6. Result of Gaussian Pyramid

Initial Experiments (No Resampling):

Initial experiments on the imbalanced dataset showed large differences in performance across the CNN architectures:

- GoogLeNet performed better due to its complex architecture and handling of the variability inside the dataset, with a training accuracy of 95.68% and a validation accuracy of 91.13%.

- ResNet50 did a pretty nice job, getting an accuracy of 88.53% on training and 83.84% on validation, slightly overfitting due to the majority class dominance.

- AlexNet and VGG16 exhibited considerable underperformance, recording validation accuracies of 17.01% and 18.47%, respectively, which underscores their shortcomings in effectively addressing the intricacies associated with imbalanced datasets.

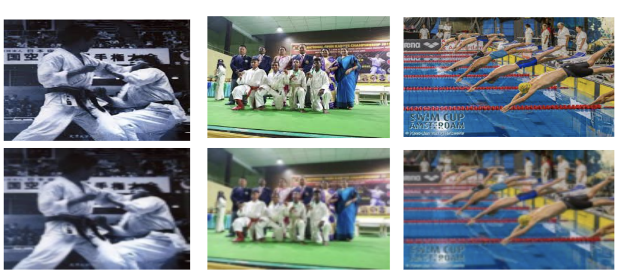


Figure 7. Example of Gaussian Pyramid

Results After Resampling

Downsampling via Gaussian Pyramid: Gaussian Pyramid resampling was used for downsampling the images by ap-

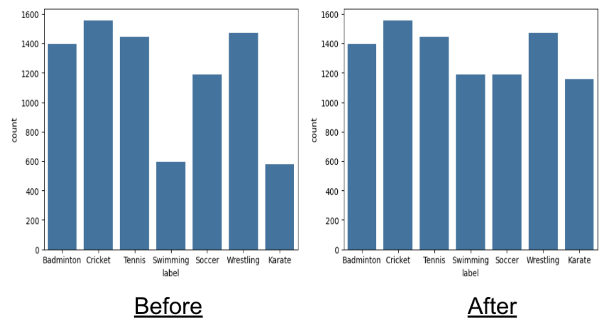


Figure 8. graph of before, after count

plying the Gaussian filter and then reducing the size to remove unnecessary details and noise. This approach aimed to balance the dataset while preserving critical information.

- On the whole, GoogLeNet showed very good generalization performance. The resampled validation accuracy is 86.44%, representing a small degradation compared to the performance of no resampling because of reduced resolution.

- Overfitting was reduced in ResNet50, whereby its validation accuracy reached up to 86.81%.

- Among these, AlexNet and VGG16 achieved the most relative improvements; the accuracy of validation rose from 50.71% and 40.52%, respectively. In particular, this reflects the good performance of applying Gaussian Pyramid resampling to handle class imbalance problems within simpler network structures.

These validation accuracies, before and after Gaussian Pyramid resampling, show some interesting trends in the different models tested.

4.2.4. SRCNN: SUPER-RESOLUTION USING DEEP LEARNING

For the implementation of SRCNN, references were taken from the paper "Image Super-Resolution Using Deep Convolutional Networks" by Dong et al. This work first introduced CNNs into the area of super-resolution, with the end goal of increasing the resolution of an image. The SRCNN developed a convolutional neural network that maps LR images to HR ones by learning the underlying relationship linking both types of imagery effectively, therefore representing such a transformation as a unified function. In the process, it beat other conventional super-resolution methods and attained SOTA performance then.

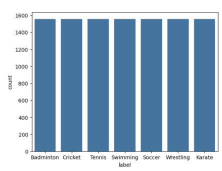
SRCNN has three convolutional layers, with the hyperparameters listed below.

- Filter sizes: $f_1 = 9, f_2 = 1, f_3 = 5$

- No. of filters: $n_1 = 64, n_2 = 32$

- Activation Function: ReLU applied to the first two layers, and no activation for the third layer.

The process is done by minimizing the gap between the predicted HR images and the ground-truth images. We used SRCNN in this work to address class imbalance and created additional images for those underrepresented classes such as Karate in order to balance with the sample size of the majority class, Cricket.



After SCRNN

Model	Initial Validation Accuracy	Post-Resampling Validation Accuracy
GoogleNet	90.28%	87.29%
ResNet50	81.35%	80.77%
VGG16	17.44%	40.52%
AlexNet	18.47%	50.71%

Results

Figure 9. Result of SRCNN

Initial Experiments (No Resampling):

The first series of experiments showed significant differences in performance for the CNN models:

- GoogLeNet had the highest performance, with a 95.62% training accuracy and a 90.28% validation accuracy, hence a sign that this model managed to handle well the variability and complexity enclosed in this data.

- That was followed by ResNet50, which had a training accuracy of 88.65% and a validation accuracy of 81.35%, indicating its strong hierarchical feature extraction capability.

- The background model AlexNet and VGG16 performed considerably worse, yielding final validation accuracies of 18.47% and 17.44%, respectively. These shallow architectures struggled significantly to model the complex patterns typical of the dataset.

Outcomes After Resampling, resampling performance enhanced all the models, though to a varying degree:

- GoogLeNet had a slightly lower training accuracy of 91.63%, but it maintained a very high validation accuracy of 87.29%, showing its performance on balanced datasets.

- ResNet50 improved modestly, with an increase in validation accuracy of 80.77% arguably indicative of reduced overfitting to the majority class.

- AlexNet and VGG16 exhibited minimal enhancement, with validation accuracies increasing to 40.52% and 50.71%, respectively. Such outcomes underscore the constraints of their architectures in managing high-dimensional patterns, despite the application of resampling techniques.

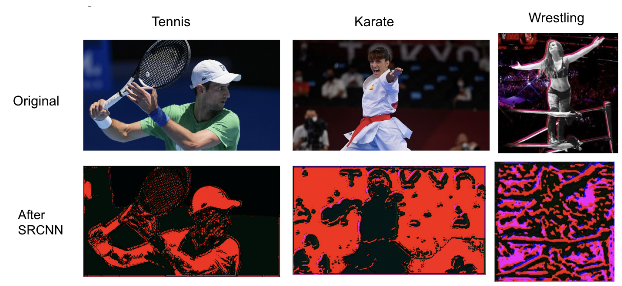


Figure 10. Examples of SRCNN

4.2.5. Wavelet Transform Resampling

Wavelet Transform Resampling was applied in order to deal with the class imbalance problem by relying on its judgment for the data with multiple resolutions. It decomposes an image totally into low-frequency components, known as approximation (cA), and high-frequency components, details that include horizontal, vertical, and diagonal, represented as cH, cV, and cD, correspondingly. This two-fold approach allowed generating artificial data through upsampling while keeping the key structural information from the downsampling process.

	Resnet50	VGG	GoogLeNet	AlexNet
Original	Training Loss: 0.3478 Training Accuracy: 0.8865	Training Loss: 1.4185 Training Accuracy: 0.4525	Training Loss: 0.1304 Training Accuracy: 0.9576	Training Loss: 1.889 Training Accuracy: 0.1816
Upsample	Training Loss: 0.4076 Training Accuracy: 0.8608	Training Loss: 1.4080 Training Accuracy: 0.4694	Training Loss: 0.2000 Training Accuracy: 0.9348	Training Loss: 0.3478 Training Accuracy: 0.1414
Downsample	Training Loss: 0.7420 Training Accuracy: 0.7420	Training Loss: 1.9524 Training Accuracy: 0.1452	Training Loss: 0.4455 Training Accuracy: 0.8548	Training Loss: 1.9472 Training Accuracy: 0.1292

Figure 11. Result of Wavelet

Initial Experiments (No Resampling): Results on imbalanced data highlighted large performance differences among the CNN models.

GoogLeNet showed very good results-95.76% for training accuracy and 85.48% for validation accuracy on the original dataset-thereby establishing its superiority in handling complex image patterns.

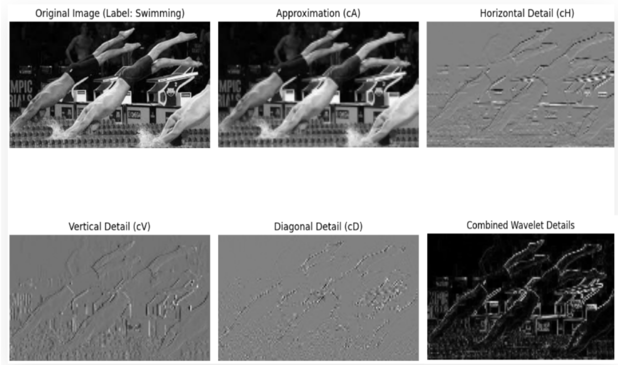


Figure 12. Examples of Wavelet

ResNet50 achieved a training accuracy of 88.65%; however, it completely overfitted because it depended on high-frequency data.

Both VGG16 and AlexNet exhibited marked underperformance, achieving training accuracies of 45.25% and 18.16%, respectively, which suggests their inadequacy in effectively extracting relevant features from intricate datasets.

Results After Resampling This wavelet-based upsampling method enriched the under-representation of classes like Karate and Swimming by combining wavelet coefficients such as cH, cV, and cD to indicate edges and textures. While this maintains all the information, the generated images in the synthesized set suffer from a lack of diversity, further limiting the performance gains.

- GoogLeNet achieved a validation accuracy of 93.48%, slightly lower than its original performance level.

- ResNet50 yielded a marginal improvement in validation accuracy at 86.08%.

Among them, the accuracy of VGG16 improves by 46.94% and AlexNet by 14.14%.

Wavelet downsampling was performed to retain the low-frequency information content (cA) while reducing the resolution to a common dimension of 128×128: The proposed approach was then followed by a very aggressive loss of high-frequency details required by complex models:

- The GoogLeNet architecture was able to achieve a validation accuracy of 85.48%, even at the cost of finer details.

- On the other hand, ResNet50, relying heavily on high-resolution input, plummeted radically at 74.20% validation accuracy.

- VGG16 and AlexNet were the most impacted, with validation accuracies at 14.52% and 12.92%, respectively.

5. Discussion

Nearest Neighbor Interpolation NNI provides an efficient and straightforward approach to resampling but demonstrates varying effectiveness depending on the model architecture. For deeper networks like GoogLeNet and ResNet50, NNI can serve as a viable option, particularly for upsampling. However, its limitations in preserving image details make it less suitable for applications requiring high-fidelity data representation, especially for simpler CNN architectures. Future research could explore integrating NNI with more advanced interpolation techniques to balance computational efficiency with feature retention.

Bilinear and Bicubic Interpolation It also points to the proper choice of the resampling strategy and model architecture. Sophisticated architectures managed this dataset complexity better, such as GoogLeNet or ResNet50, while the resampling for a simple model, like AlexNet or VGG16, showed very few positive effects independently of the method applied.

Most of the deep learning-based approaches, especially on complex datasets, are normally preferred by using resampling through bicubic interpolation for further performance improvements. Although it evades some of the issues of imbalance and preserves the features better, resampling itself cannot replace the need for robust model architectures that can handle high-dimensional patterns effectively.

Downsampling by Gaussian Pyramid GoogLeNet, at first having the highest validation accuracy of 91.13% on the imbalanced dataset, was reduced to 86.44% after resampling. This is because here, there is a trade-off between the balancing of the dataset with the loss of high-resolution details, hence a drop of 4.69%. On the other side, the ResNet50 witnessed an increase: its validation accuracy rose from 83.84% to 86.81% due to the reduced overfitting, hence a gain of 2.97%.

The most significant relative improvements came on the simpler models, AlexNet, and VGG16. AlexNet significantly improved from 17.01% to 50.71%, a 33.70% increase, and VGG16 improved from 18.47% to 40.52%, an increase of 22.05%. These demonstrate that the resampling by the Gaussian Pyramid disproportionately benefited the simpler architectures due to the reduction in class imbalance and noise in the dataset. More profound models, however, like GoogLeNet and ResNet50, were more sensitive to the resampling process.

The results are very effective for performance improvement by Gaussian Pyramid resampling, which helps to alleviate class imbalance and reduce noise in simpler types of CNNs. In contrast, for advanced models like GoogLeNet and ResNet50, it tends to be less effective since these models possess inherent robustness toward handling such data variations. Such findings point to the importance of synchronizing resampling methods with model architectures to maximum avail in image classification tasks.

SRCNN: Super-Resolution Using Deep Learning This paper further underlines the importance of model depth and dataset balancing for the best performances to be achieved in image classification tasks. GoogLeNet and ResNet50 consistently outperformed AlexNet and VGG16 before and after resampling, respectively, showing that the deeper the model architecture, the better it generalizes on challenging datasets. Resampling reduced the impact brought by class imbalance, while simpler models such as AlexNet and VGG16 benefited less due to their architectural limits. These findings further support the conclusions of Dong et al. (2015), which hint at the capability of deep learning architectures for handling such complex datasets.

Wavelet Transform Resampling Wavelet Transform Resampling proved variably effective. Upsampling proved moderately effective for deeper models but offered limited improvement in shallower network architectures. Downsampling, while computationally efficient, proved to be significantly detrimental to network performance due to the loss of high-frequency information. Again, these results demonstrate that feature retention and dataset variability are crucially important factors for such impactful model improvements. This could also be refined by further research into adaptive wavelet filtering or hybrid resampling methods that better optimize this trade-off between computational simplicity and classification accuracy.

6. Related Work

Research on addressing class imbalance in image datasets and enhancing deep learning model performance has been explored extensively in the literature. This section discusses relevant studies that form the foundation of the techniques and methods used in this work.

Resampling Techniques Several studies have investigated the impact of resampling methods on imbalanced datasets. Bashir et al. (2021) provide a comprehensive review of deep learning-based single image super-resolution methods, highlighting how these techniques can improve the representation of underrepresented classes in image datasets(final report). Similarly, Rukundo and Cao (2012) introduce nearest neighbor interpolation (NNI), emphasizing its computational efficiency for handling large-scale image data(final report).

Bicubic interpolation, discussed by Verma and Saini (2017), demonstrates its ability to preserve high-frequency details during upsampling, making it effective for balancing image datasets. Their work underscores the importance of selecting appropriate interpolation methods based on dataset complexity and model requirements(final report).

Super-Resolution Techniques The introduction of SRCNN by Dong et al. (2016) revolutionized image super-resolution by using convolutional neural networks. Their approach maps low-resolution images to high-resolution images, addressing class imbalance by generating synthetic samples for underrepresented classes. The study provides a strong theoretical and empirical basis for adopting SRCNN in this work(final report).

Wavelet-Based Resampling Wavelet transforms have been utilized for both upsampling and downsampling in image classification tasks. Mallat and Hwang (1992) pioneered the use of wavelets for singularity detection and feature extraction, which has since been extended to handling imbalanced datasets. Their method of decomposing images into low- and high-frequency components aligns with this study's application of wavelet-based resampling(final report).

Deep Learning Architectures Deep convolutional networks, such as GoogLeNet and ResNet50, are recognized for their robustness in handling high-dimensional and complex datasets. Their architectures leverage deep feature extraction layers, making them effective in scenarios with significant class imbalance. Simpler architectures, such as AlexNet and VGG16, have been shown to struggle with these datasets due to their limited depth and capacity to capture intricate patterns.

Class Imbalance Solutions Previous works have highlighted the importance of balancing datasets to improve classification performance. Techniques such as SMOTE (Synthetic Minority Oversampling Technique) and adaptive resampling strategies have been widely used in conjunction with deep learning models. However, these approaches often focus on tabular data, leaving a gap in research on their application to image datasets.

7. Conclusion

In Conclusion, Five different resampling methods have been developed in this present work, including nearest neighbor interpolation, bilinear, and bicubic interpolation, downsampling by Gaussian Pyramid, super-resolution using deep learning-known as SRCNN-and wavelet transform resampling by utilizing four deep learning models, such as GoogLeNet, ResNet50, VGG16, and AlexNet, in order to find an effective solution to problems related to class imbalance and noisy image datasets, and thus study the interaction among these resampling methods and model architectures.

NNI is one of the simplest resampling techniques that works variably depending on the model architecture. Although computationally cheap, it fails to retain information with much intricacy and hence has limited value in relatively complex applications. GoogLeNet and ResNet50 were quite successful with NNI, especially on upsampling tasks, but relatively simpler models like AlexNet and VGG16 experienced various setbacks due to the nature of their architecture.

Bilinear and Bicubic Interpolation: For the deep models, GoogLeNet and ResNet50, the use of bicubic interpolation was better by far, where high-frequency details relevant for feature extraction were well preserved. Bilinear interpolation, though faster and computationally inexpensive, showed relatively poor performance in maintaining fine details. For AlexNet and VGG16, simpler models, neither of the methods helped much, thus yielding to the robustness of their architecture.

Downsampling by Gaussian Pyramid: This method smoothed high-frequency noise while preserving structural features, which was especially effective in simpler models. AlexNet and VGG16 achieved significant relative performances with Gaussian Pyramid resampling, showing that indeed it helped in mitigating class imbalance and noise. However, in advanced models like GoogLeNet and ResNet50, their performance degraded due to a loss of high-resolution details in them, thereby showing how these networks rely on fine-grained information.

Successful handling of class imbalance by SRCNN was done through the generation of synthetic high-resolution images for the classes having few representations. Maintaining data balance and important features in the contribution of the method, the impact significantly showed up when simpler models were used. The more complex ones, which were already proficient in feature extraction, showed slight improvements, proving the added data redundancy for those systems.

Wavelet Transform Resampling: Wavelet-based upsampling, while emphasizing high-frequency information like edges and textures, helped deep models such as GoogLeNet and ResNet50 with a middle-level performance gain. However, this led to low-diversity synthetic images, thus limiting the overall performance gain. Wavelet-based downsampling preserves the structural information while exhibiting evident performance degradation due to the loss of important high-frequency details, especially in deeper architectures.

7.1. Overall Conclusion

Overall, deeper architectures like GoogLeNet and ResNet50 consistently outperformed simpler models, leveraging their capacity to handle complex data patterns. Resampling methods like Bicubic Interpolation and Gaussian Pyramid were particularly effective for balancing datasets and improving feature retention in simpler architectures. However, the effectiveness of resampling techniques is limited without robust model architectures capable of managing high-dimensional data. This would include hybrid resampling methods that could also apply NNI with other, even too-sophisticated interpolation methods. The adaptive resampling strategy is an example of a procedure whereby class imbalance or data complexity may trigger dynamic adjustments. Other methods, such as the SR-CNN method, might also contribute added diversity to the generated synthetic images and increase their general value. By tailoring resampling techniques to specific model requirements, researchers can achieve scalable and robust solutions for real-world image classification challenges.

8. Contribute

- **Wonjoon Hwang**: Responsible for the experimental setup, data preprocessing, and implementing resampling with SRCNN. Conducted comprehensive analysis of experimental results and contributed significantly to the drafting and organization of the project paper.

- **Zhizheng Wang**: Focused on implementing and evaluating the Wavelet Transform resampling technique. Additionally, took the lead in preparing the presentation slides and delivering the project presentation, ensuring clear communication of the project goals and findings.

- **Khoa Le**: Led the implementation of Bilinear and Bicubic Interpolation methods for resampling. Conducted detailed analysis comparing the two techniques and their effects on model performance. Additionally, took the lead in preparing the presentation slides and delivering the project presentation, ensuring clear communication of the project goals and findings.

- **Chih-Yuan Tung**: Implemented the Gaussian Pyramid resampling technique. Conducted experiments to test its effects on noise reduction and class imbalance handling. Additionally, took the lead in preparing the presentation slides and delivering the project presentation, ensuring clear communication of the project goals and findings.

- **Brett Ruane**: Handled the implementation of Nearest Neighbor Interpolation (NNI) for both upsampling and downsampling. Evaluated its computational efficiency and its effects on model generalization. Additionally, took the lead in preparing the presentation slides.

9. Reference

- Bashir, Syed Muhammad Arsalan, et al. ”A comprehensive review of deep learning-based single image super-resolution.” PeerJ Computer Science 7 (2021): e621.
- Dong, Chao, et al. ”Image super-resolution using deep convolutional networks.” IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 38, no. 2, 2016, pp. 295–307, https://doi.org/10.1109/TPAMI.2015.2439281.
- Mallat, S., & Hwang, W. L. (1992). Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, *38*(2), 617–643. https://doi.org/10.1109/18.119727
- Rukundo, Olivier, and Hanqiang Cao. ”Nearest neighbor value interpolation.” arXiv preprint arXiv:1211.1768 (2012).
- Verma, Atul & Saini, Barjinder. (2017). Forward-backward processing technique for image denoising using FDZP 2D filter. Journal of Applied Research and Technology. 15. 10.1016/j.jart.2017.07.001.