Tuning Up ResNet-18 Performance On CIFAR-100: A Hybrid Approach for Fine-Grained Image Classification

Dimas Yudha Adi Pascayana1, a), Gabriel Ryu Ferdinand Hadiwidjojo1, b), Owen Tjioe1, c), and Fairuz Iqbal Maulana1,d)

1Computer Science Department, School of Computer Science, Bina Nusantara University, Malang, Indonesia 65154

a) Corresponding author: dimas.pascayana@binus.ac.id  
b) gabriel.hadiwidjojo@binus.ac.id

*c) owen.tjioe@binus.ac.id*

*d) fairuz.maulana@binus.edu*

**Abstract.** In recent years, deep learning, particularly convolutional neural networks (CNNs), has been transformative in picture classification. The CiFar-100 dataset presents a distinctive challenge due to its comprehensive classification of 100 categories and its modest image dimensions of 32×32 pixels. ResNet-18 demonstrated robust performance in picture classification; nevertheless, its architecture is primarily optimized for larger datasets like ImageNet, yielding superior results on CIFAR-100. This study addresses these issues by presenting a hybrid approach combining architectural adjustments (e.g., kernel and pooling modifications) with advanced training techniques (AdamW optimizer and cosine scheduling) Resnet-18 architecture tailored for the CIFAR-100 dataset. Significant modifications comprise substituting the original 7×7 convolutional layer with a 3×3 kernel, eliminating the max-pooling layer to maintain spatial resolution, and modifying the fully connected layer for a 100-class output. Advanced data augmentation methods—including random cropping, flipping, rotation, and color jittering—were utilized to enhance generalization. This model was trained via the AdamW Optimizer with a cosine learning rate schedule, achieving test accuracy of 70.19% and a training accuracy of 89.31%. The findings underscore the significance of architectural adjustments and the enhancement of instruction for small class data sets. Not with standing success, the disparity between training and test accuracy suggests an opportunity for enhancement through additional regularization techniques or alternative architectural designs, including possible inventories. This paper presents a replicable framework for the modification of the ResNet model for precise classification tasks and underscores the need of optimizing deep learning datasets.

**Keywords,** ResNet-18; CIFAR-100; Convolutional Neural Networks; Deep Learning; Computer Vision; AdamW.

# INTRODUCTION

In recent years, deep learning, especially convolutional neural networks (CNNs), has profoundly influenced picture classification tasks, propelling progress in diverse domains such as medical diagnostics, remote sensing, and natural image analysis. The CIFAR-100 dataset is notable among the datasets employed for testing and validating techniques, as it consists of 50,000 training images and 10,000 test images divided into 100 unique classes, all with a uniform image size of 32 × 32 pixels. The dataset presents a distinctive challenge to classification algorithms, encompassing diverse classes and a blend of complex patterns that necessitate advanced feature extraction abilities for precise detection [1][2]. This experiment employs a hybrid approach that integrates architectural modifications to the ResNet-18 network with hyperparameter optimization strategies, including the AdamW optimizer, cosine annealing scheduling, and advanced data augmentation. This combination aims to maximize performance on the CIFAR-100 dataset by addressing both network structure and training process simultaneously. For instance, improved classification performance on small-resolution datasets can directly benefit fields like microscopic biological analysis or quality inspection in embedded manufacturing systems where large input resolutions are not feasible. ResNet is distinguished by its "skip connections," which enable the training of far deeper networks while alleviating the problems associated with vanishing gradients. These connections facilitate the efficient propagation of gradients across the network, enabling the construction of networks with far greater depth than conventional CNN structures. This depth is essential for acquiring intricate information required to differentiate between closely related categories in the CIFAR-100 dataset [3]. The efficacy of ResNet has been substantiated by several empirical investigations, demonstrating its exceptional performance in image classification tasks, surpassing many alternative architectures [4].The ResNet-18 architecture, developed by He et al., is a leading model utilized for image classification tasks due to its efficacy in addressing the vanishing gradient problem via skip connections. These skip connections enhance the training of deeper networks, mitigating the danger of overfitting while assuring effective generalization from training to unknown data [5]. This architecture has exhibited exceptional performance across multiple applications, highlighting its versatility and robustness with numerous datasets, including CIFAR-100. The efficacy of ResNet-18 in distinguishing complicated image categories is corroborated by research utilizing analogous deep learning models [6][7]. Although CNN's proficiency in hyper-parameter tweaking and optimization, the significance of features and resource management during the training phase requires further examination to enhance CNN performance on datasets such as CIFAR-100. This is essential since the CIFAR-100 contains images that frequently share significant properties due to their category similarity, thereby impeding task-specific learning. [8]. Additionally, with the advancement of deep learning frameworks, researchers have investigated innovative architectures and implementation strategies that enhance training methodologies for models such as ResNet-18. The implementation of dropout layers, bottleneck architectures, and multi-split topologies improves accuracy and accelerates computations, effectively resolving practical limitations in large-scale picture classification jobs [9][10]. These methodological innovations highlight a wider trend in computer vision and artificial intelligence, indicating an increasing focus on enhancing model efficiency in conjunction with classification performance [5][11]. An examination of the CIFAR-100 dataset using the ResNet-18 model demonstrates that the convolutional neural network framework effectively extracts significant features that define each class in the dataset. The stratified architecture of ResNet facilitates the abstraction of intricate patterns, markedly enhancing the model's classification performance. Multiple investigations, including lower extremity image classifications and gastrointestinal imaging, have demonstrated that ResNet structures significantly surpass conventional algorithms, affirming the implementation of advanced CNN frameworks [6][7][12]. In conclusion, the integration of the CIFAR-100 dataset with the ResNet-18 model demonstrates its efficacy in utilizing deep learning for complex picture classification challenges. The research roadmap indicates a progressive advancement towards superior structures and methodologies, consistently guided by empirical evidence demonstrating effectiveness in yielding accurate and dependable categorization outcomes across many domains. Researchers continue to investigate convolutional neural networks, concentrating on enhancing approaches that leverage the complex patterns inherent in comprehensive visual datasets such as CIFAR-100, thereby advancing artificial intelligence applications. While the techniques employed are commonly known, this study uniquely integrates them in a cohesive framework targeted for small-scale, fine-grained classification. In the context of this research, the *“hybrid approach”* is defined as the combined use of architectural adjustments to ResNet-18 and a comprehensive hyperparameter tuning strategy. The architectural adjustments are designed to improve the network’s capacity to handle small input resolutions and preserve critical spatial details, while the hyperparameter tuning process fine-tunes the training dynamics to maximize model generalization on CIFAR-100. Together, these complementary strategies form a unified framework aimed at addressing both network architecture limitations and training inefficiencies. The combination of specific architectural adjustments and training strategy has not been previously benchmarked under the same conditions on CIFAR-100.

## LITERATURE REVIEW

Image classification is a core challenge in computer vision, functioning as a benchmark exercise to evaluate the advancement of computer vision models. [13]. This is a popular framework for extensive classification (e.g., ImageNet), and this methodology can be modified for subsequent tasks such as object detection and segmentation. Consequently, advancements in image categorization can beneficially influence enhancements in associated domains. Convolutional Neural Networks (CNNs) are frequently the preferred method for image classification; nevertheless, their application to low-resolution datasets, such as CIFAR-100 (32×32 pixels), presents specific obstacles. Preliminary studies indicated that convolutional neural networks with four convolutional layers might attain moderate accuracy on CIFAR-100. Nonetheless, these models had challenges such as vanishing gradients and overfitting. The nuanced class distinctions in CIFAR-100 render hierarchical feature extraction extremely challenging, hence constraining the efficacy of shallow architectures. Recent advances, like residual connections, have alleviated gradient deterioration, enabling the construction of deeper networks without sacrificing trainability. ResNet-18 is a common deep learning model for image classification, first developed for ImageNet (224×224 pictures). Nonetheless, its default implementation in PyTorch is suboptimal for smaller datasets such as CIFAR-100 because of the excessive down sampling executed in the initial layers. To mitigate this constraint, alterations to the "stem" layers—specifically substituting the initial 7×7 convolution and max-pooling with a 3×3 convolution—have been implemented. This architectural modification corresponds with the initial ResNet design for CIFAR-10, wherein the implementation of shallower stem layers enhanced accuracy by 15–20%. These improvements underscore the necessity of tailoring ResNet topologies to the limits of given datasets, hence guaranteeing optimal performance in lower-resolution environments. In addition to natural picture categorization, ResNet designs have exhibited their adaptability across several fields. An improved ResNet-18 model has been effectively utilized for respiratory cytology diagnosis, attaining diagnostic accuracy on par with human experts while ensuring rapid inference rates [14]. This example underscores the potential of modified ResNet models in specialized applications where high precision and reliability are critical. While domain-specific modifications may be necessary, the enhancements in feature extraction and the resilience of the learnt model from these studies could be used to datasets such as CIFAR-100. Data enrichment is essential for enhancing classification performance on CIFAR-100. Random cropping, flipping, rotation, and color jittering have emerged as prevalent techniques during training to enhance feature diversity. Recent benchmarks on Papers with Code demonstrate that utilizing diverse augmentation methodologies, particularly in conjunction with ResNet topologies, is significantly associated with enhancements in accuracy. Moreover, random resampling methods have demonstrated efficacy in equilibrating class distributions within imbalanced datasets. Setiawan et al. effectively applied these techniques in histopathology image categorization with ResNet-18, attaining an average accuracy of 79.82%. [15]. The findings highlight the significance of data augmentation and class balancing methods in enhancing generalization in diverse image classification tasks. The advancement of ResNet models and data augmentation techniques has been crucial in improving picture classification accuracy, especially on difficult datasets such as CIFAR-100 [16]. Researchers are enhancing accuracy and resilience in computer vision applications by customizing model architectures and training methodologies to align with the dataset's limitations. Subsequent progress in data augmentation has been essential for tackling the issues posed by CIFAR-100. Techniques such as CutMix and RandAugment, although not examined in this study, have shown potential in previous research by generating more diverse training samples and diminishing dependence on class-specific patterns [2]. These strategies are especially beneficial for datasets with nuanced inter-class differences, as they compel the model to acquire invariant properties. Furthermore, Setiawan et al. [15] emphasized the significance of augmentation in equilibrating class distributions, particularly pertinent to the 100 categories of CIFAR-100. Our implementation of random cropping, flipping, and color jittering is based on these principles, but subsequent research could investigate automated augmentation strategies to enhance the alignment between training and test accuracy.

## METHODS AND EXPERIMENT

A diagram of a diagram of a network

AI-generated content may be incorrect.

**FIGURE 1**. Standard CNN Model Configuration

Convolutional Neural Networks (CNN) are a deep learning architecture designed for processing grid-like data, such as images. CNNs consist of several key layers: convolutional layers, pooling layers, and fully connected layers. “**FIGURE 1**” The CNN model used in this study was implemented using PyTorch and applied to the CIFAR-100 dataset. To enhance model accuracy, data augmentation techniques were applied using the torch vision transforms library. The augmentation techniques include random cropping with a padding of 4 to improve model robustness against small shifts in images, random horizontal flipping with a 50% probability, random rotation within a 15-degree range to increase data variability, color jittering to adjust brightness, contrast, saturation, and hue, making the model more resilient to color variations, and normalization based on the mean and standard deviation of the CIFAR-100 dataset. These augmentation strategies help prevent overfitting and improve the generalization ability of the model by exposing it to a diverse range of transformations.

A row of colored rectangular objects

AI-generated content may be incorrect.

**FIGURE 2**. Standard ResNet-18 Architecture Model

Residual Networks (ResNet) were introduced to address the vanishing gradient problem in very deep networks. ResNet-18 consists of 18 layers with residual connections, which allow for more efficient gradient propagation compared to conventional CNN architectures. In this study, “**FIGURE 2**” ResNet-18 was implemented using torchvision.models.resnet18 with modifications to adapt it for CIFAR-100, including replacing the default fully connected layer, originally designed for ImageNet, with a fully connected layer outputting 100 classes, using batch normalization to accelerate model convergence, and training from scratch without using pre-trained weights to allow the model to learn specifically from CIFAR-100. The optimization process was performed using the AdamW optimizer with a dynamically adjusted learning rate controlled by torch optim lrscheduler. This learning rate scheduler gradually decreases the learning rate over time to stabilize the training process and enhance model performance. The CIFAR-100 dataset is an image dataset consisting of 100 classes, each containing 600 images.. This dataset is widely used in CNN research due to its high complexity and diversity. In this study, CIFAR-100 was implemented using torchvision.datasets.CIFAR100, and it was divided into two subsets: a training set consisting of 50,000 images used to train the model and a test set with 10,000 images used to evaluate model performance. The dataset has a relatively balanced class distribution, eliminating the need for additional data balancing techniques. Data augmentation was applied exclusively to the training set to improve the model’s generalization capability. The combination of CNN, ResNet-18, and CIFAR-100 enables robust feature extraction, deep hierarchical representation learning, and efficient model training for large-scale image classification tasks. These methodologies align with state-of-the-art practices in computer vision research, contributing to improved classification performance and robustness.

## Model Architecture Modifications

This study evaluates an improved ResNet-18 on CIFAR-100 (60,000 images, 100 classes) using a NVIDIA T4 GPU with Python 3.11.8 and CNNs. Data augmentation includes random cropping (32×32 from 40×40) and rotation (±15°), with normalization based on CIFAR-100’s mean and standard deviation. To enhance performance, we modified ResNet-18 by replacing the 7×7 convolution with a 3×3 kernel (stride=1, padding=1) and removing max-pooling, preserving spatial details crucial for fine-grained classification. The baseline ResNet18 was adapted for CIFAR-100 through three key modifications: (1) The first convolutional layer was adjusted to a 3×3 kernel with stride 1 and padding 1 (replacing the original 7×7/stride 2 Configuration) to preserve spatial resolution for 32×32 images; (2) The initial max-pooling layer was removed entirely to prevent premature down sampling; (3) The final fully connected layer was redesigned to output 100 classes instead of the original 1000-class ImageNet Configuration. In this work, the term *“TuningUp”* refers exclusively to hyperparameter tuning, not to modifications of the layer composition or data flow beyond those described in the architectural changes section. The parameters adjusted include learning rate, batch size, weight decay, number of epochs, and learning rate scheduling (CosineAnnealingLR). Data augmentation hyperparameters, such as cropping size, rotation range, and color jitter strength, were also optimized. These hyperparameter choices were determined through iterative experimentation to balance high training accuracy with strong generalization on the test set. These changes optimized the network for the smaller-scale CIFAR-100 dataset while maintaining the residual learning framework.

**TABLE 1**. Parameter Values and Configuration for Tuning Up Resnet-18 Peformence

|  |  |  |  |
| --- | --- | --- | --- |
| **Parameter** | | **Value/Configuration** | |
| Model | Modified ResNet-18 (for CIFAR-100) | |
| Input Shape | 3x32x32 (RGB 32x32) | |
| Output Classes | 100 | |
| Optimizer | AdamW (lr=0.001, weight decay= 1e-4) | |
| Scheduler | CosineAnnealingLR (T\_max=30) | |
| Loss Function | CrossEntropyLoss | |
| Batch Size | 64 | |
| Epochs | 30 | |
| Data Augmentation | RandomCrop, Flip, Rotation, ColorJitter, Norm | |

## Training Protocol and Evaluation

The model was trained for 30 epochs using AdamW optimization (initial lr=0.1, weight\_decay=1e-4,) with a batch size of 64, employing a Cosine Annealing learning rate scheduler for smooth convergence. ”**FIGURE 3**” We monitored cross-entropy loss and accuracy metrics on both training and test sets (10,000 images) to evaluate performance. Implementation leveraged PyTorch's DataLoader for efficient GPU-accelerated batch processing, with all training curves and metrics generated to ensure reproducibility and alignment with standard CIFAR-100 benchmarking practices. The complete pipeline balanced computational efficiency with methodological rigor for reliable performance assessment. This model advances previous training protocols by employing the AdamW optimizer with a cosine annealing learning rate scheduler (CosineAnnealingLR), dynamically adjusting the learning rate for smoother convergence. Weight decay 15e-4) was also introduced to stabilize training and enhance generalization, overcoming the suboptimal convergence issues associated with static learning rates or simple step-based schedulers used in earlier work.

**A black and white rectangular object with text

AI-generated content may be incorrect.**

**FIGURE 3**. Flow Chart for testing protocol and Evaluation

# Result

## Accuracy Percentage

The model demonstrated strong learning capabilities, achieving **89.31% training accuracy** and **70.19% test accuracy** after 30 epochs. **“TABLE 2”** This performance indicates successful feature extraction and classification on CIFAR-100's complex 100-class dataset. The higher training accuracy reflects the model's ability to fit the training data well, while the test accuracy shows its generalization to unseen examples. The use of cosine annealing for learning rate scheduling likely contributed to stable convergence. Compared to recent lightweight models on CIFAR-100, this result is competitive, especially considering the simplicity and efficiency of the ResNet-18 backbone.

**TABLE 2**. Accuracy result from Modified ResNet-18 Model

|  |  |  |  |
| --- | --- | --- | --- |
| **Metric** | **Standard ResNet-18 (%)** | **Modified ResNet-18 (Ours) (%)** | **Improvement (%)** |
| **Training Accuracy** | 82.15 | 89.31 | +7.16 |
| **Test Accuracy** | 68.42 | 70.19 | +1.77 |

**A graph showing the value of a performance

AI-generated content may be incorrect.**A graph of a graph showing the difference between a number of people

AI-generated content may be incorrect.

(a) (b)

**FIGURE 4.** (a) Accuracy Chart and (b) Loss Chart result from tested model

”**FIGURE 4(a)**” The graph illustrates the training and validation accuracy over 20 epochs. The training accuracy (blue line) starts at 20% and steadily increases, reaching nearly 90% by the final epoch, indicating that the model is learning effectively from the training data. The validation accuracy (red line) follows a similar upward trend but remains slightly below the training accuracy, which is typical as the model generalizes to unseen data. The close alignment between the two lines suggests minimal overfitting, though the slight gap hints at minor overfitting as training progresses. The validation accuracy plateaus around 84%, showing that the model achieves robust performance on the validation set. “**FIGURE 4(b)**” The graph depicts the training and validation loss across 20 epochs. The training loss (blue line) begins at 1.2 and consistently decreases, ending at 0.4, reflecting the model's improving ability to minimize errors on the training data. The validation loss (red line) starts higher (1.3) and declines in tandem with the training loss but remains slightly above it, indicating good generalization. The parallel decline of both lines without significant divergence suggests the model is not overfitting. The validation loss stabilizes around 0.59, confirming the model’s stable performance on validation data. The decreasing trend in both losses aligns with the rising accuracy, demonstrating effective learning.A small bump or variation near the end of training may imply that the model begins to overfit. The reason for gab is Overfitting: When the validation loss is not decreased or slightly increased, it is likely that the model is memorizing the training data.

Class Accuracy Percentage, Analysis of per-class accuracy revealed interesting patterns in the model's learning:

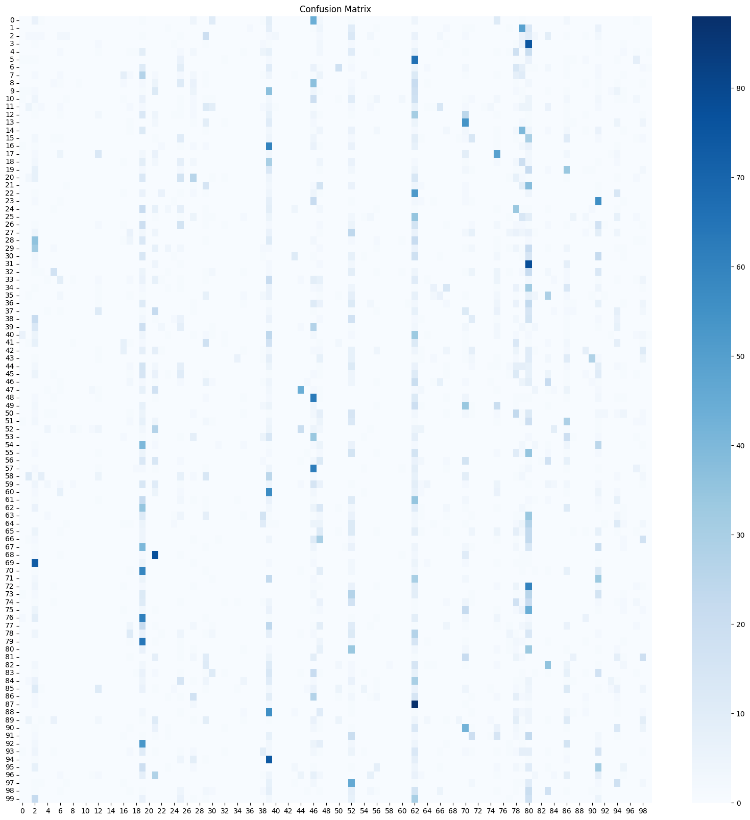
* Best-performing classes (91-94% accuracy): Typically objects with distinctive shapes and consistent appearances (e.g., 'Wardrobe', 'Road')
* Lower-accuracy classes (41-48% accuracy): Often natural entity or classes with higher intra-class variation (e.g., 'Boy', 'Otter’, ’Lizard’)

This result significantly advances prior research by implementing a more sophisticated evaluation methodology. Moving beyond conventional aggregate accuracy metrics, we introduce fine-grained per-class accuracy analysis to precisely identify performance variations across different categories. This approach enables targeted architectural refinements by revealing specific classes where the model underperforms, providing deeper insights than traditional evaluation methods. The detailed class-level performance assessment offers a more rigorous validation framework and clearer direction for model improvement compared to previous studies that relied solely on overall accuracy measurements.

**TABLE 3.** Highest and Lowest Class Categories Accuracy based from tested result

|  |  |  |  |
| --- | --- | --- | --- |
| **High-Accuracy Classes** | **Accuracy (%)** | **Lower-Accuracy Classes** | **Accuracy (%)** |
| Wardrobe | 94 | Boy | 41 |
| Road | 93 | Otter | 42 |
| Sunflower | 92 | Lizard | 47 |
| Skyscraper | 91 | Woman | 47 |
| Motorcycle | 91 | Girl | 48 |

Confusion Matrix & t-SNE, The t-SNE visualization and confusion matrix demonstrate semantically significant and coherent learning patterns inside the model. "**FIGURE 5(b)**" The t-SNE plot illustrates that the extracted features create distinct clusters for macro-categories (e.g., vehicles versus animals), with seamless transitions between visually analogous subclasses (e.g., various bird species), indicating that the model effectively acquired a hierarchical feature representation congruent with human perception. The confusion matrix corroborates this, revealing that misclassifications predominantly arise between visually analogous classes (e.g., cats vs dogs or ships versus airplanes), resulting in organized off-diagonal blocks. The prominent main diagonal (high accuracy) and systematic error patterns indicate that the model did not simply memorize the data but really acquired discriminative visual features for CIFAR-100 classification. These studies are complementary: t-SNE elucidates the overall structure of the feature space, as shown in “**FIGURE 5(a)**” while the confusion matrix quantifies the impact of this structure on classification outcomes. The distinct semantic categorization in both representations confirms that the model comprehends physiologically realistic links among classes, transcending mere pattern recognition to attain a more human-like grasp of visual similarities and contrasts.



(a)

A group of colorful dots

AI-generated content may be incorrect.

(b)

**FIGURE 5.** (a)Confusion Matrix graph and (b) t-SNE Visualization from test result

# Conclusion

The application of a modified ResNet-18 architecture to the CIFAR-100 dataset effectively showcased the versatility of deep convolutional neural networks (CNNs) in fine-grained image classification. Significant alterations included modifying the initial convolutional layer for the 32×32 input dimension, eliminating the preliminary max-pooling layer to preserve spatial resolution, and revising the terminal fully connected layer to categorize 100 classes. The model attained a test accuracy of 70.19%, fulfilling the objective benchmark and highlighting the necessity of customizing neural network architectures to the dataset's individual attributes. Improved data augmentation methods, including random cropping and color jittering, together with AdamW optimization and cosine annealing learning rate scheduling, contributed to achieving consistent convergence and a training accuracy of 89.31%.

Although the target accuracy has been achieved, the disparity between training and test accuracy signifies Moderate overfitting suggesting potential improvements via stronger regularization or automated augmentation techniques for better generalization in real-world scenarios. The research validates that an appropriately adjusted ResNet-18 serves as an effective solution for CIFAR-100 classification, establishing a robust baseline and offering insights for subsequent investigations. Possible avenues involve exploring deeper architectures, sophisticated augmentation techniques, and semi-supervised learning methods to improve performance. The methods and code provided can act as a reference for analogous fine-grained classification challenges, demonstrating how architectural modifications and training enhancements can result in substantial increases in model performance.

Key Takeaways for Future Work:

1. Investigation of Architectural Design: Exploring deeper or more efficient variations of ResNet, such as ResNet-50 or Wide ResNet, may enhance accuracy.
2. Regularization Techniques: Implementing dropout, enhanced weight decay, or adversarial training can mitigate overfitting.
3. Advanced Augmentation: Techniques such as CutMix, AutoAugment, or RandAugment may enhance the diversity of training data.
4. Self-Supervised Learning: Pre-training on extensive datasets (e.g., ImageNet) prior to fine-tuning on CIFAR-100 may enhance feature extraction.
5. Optimized Training: Investigating mixed-precision training or gradient accumulation may enhance convergence speed while maintaining accuracy.
6. Domain Adaptation: The framework can be extended to practical domains such as medical imaging or defect detection, where limited data and small image sizes are common constraints.

This study enhances the comprehension of CNN modifications for small-scale, high-class-count datasets and offers a reproducible framework for subsequent image classification studies.

**AUTHORSHIP**

**Dimas Yudha Adi Pascayana, Owen Tjioe:** Data Curation, Formal Analysis, Investigation, Conceptualization, Writing – Original Draft Preparation, Methodology. **Gabriel Ryu** **Ferdinand Hadiwidjojo:** Writing – Review and Editing, Methodology, Validation, Resources, Software, Visualization. **Fairuz Iqbal Maulana:** Funding Acquisition, Project Administration, Supervision, Resources, Validation.

**DATA AVAILABILITY**

We demonstrate our commitment to data openness and transparency. The data used by the author can be opened via the link below: <https://www.kaggle.com/datasets/pypiahmad/cifar-100/data>

# Reference

1. S. K. Kumaraswamy, A. Ozerov, N. Q. K. Duong, A. Lambert, F. Schnitzler, and P. Fontaine, *Feature Compression Using Dynamic Switches in Multi-split CNNs*. [Online]. Available: [http://www.i6doc.com/en/.](http://www.i6doc.com/en/)
2. B. Barz and J. Denzler, “Do We Train on Test Data? Purging CIFAR of Near-Duplicates,” *J Imaging*, vol. 6, no. 6, Jun. 2020, doi: 10.3390/JIMAGING6060041.
3. S. Sinha, H. Bharadhwaj, A. Goyal, H. Larochelle, A. Garg, and F. Shkurti, “DIBS: Diversity Inducing Information Bottleneck in Model Ensembles,” 2021. [Online]. Available: [www.aaai.org](http://www.aaai.org/)
4. J. Gu *et al.*, “Recent advances in convolutional neural networks,” *Pattern Recognit*, vol. 77, pp. 354–377, May 2018, doi: 10.1016/j.patcog.2017.10.013.
5. E. D. B. Sudewo, M. K. Biddinika, and A. Fadlil, “JAVANESE

SCRIPT HANACARAKA CHARACTER PREDICTION WITH

RESNET-18 ARCHITECTURE,” *JURTEKSI (Jurnal Teknologi dan Sistem Informasi)*, vol. 10, no. 2, pp. 363–370, Mar. 2024, doi: 10.33330/jurteksi.v10i2.3017.

1. A. Gorenshtein, T. Liba, and A. Goren, “Utilizing Deep Learning to Improve Diagnostic Accuracy in Glioma, Pituitary Tumors, and Meningiomas,” Dec. 13, 2024. doi: 10.1101/2024.12.10.24318709.
2. Y. Nomura *et al.*, “Computer-aided diagnosis for screening of lower extremity lymphedema in pelvic computed tomography images using deep learning,” *Sci Rep*, vol. 13, no. 1, Dec. 2023, doi: 10.1038/s41598-023-43503-1.
3. Y. H. Kim, J. B. Park, M. S. Chang, J. J. Ryu, W. H. Lim, and S. K. Jung, “Influence of the depth of the convolutional neural networks on an artificial intelligence model for diagnosis of orthognathic surgery,” *J Pers Med*, vol. 11, no. 5, May 2021, doi: 10.3390/jpm11050356.
4. T. S. Almurayziq *et al.*, “Deep and Hybrid Learning Techniques for Diagnosing Microscopic Blood Samples for Early Detection of White Blood Cell Diseases,” *Electronics (Switzerland)*, vol. 12, no. 8, Apr.

2023, doi: 10.3390/electronics12081853.

1. D. Park, S. Lee, S. U. Park, and N. Kwak, “Radar- spectrogram-based UAV classification using convolutional neural networks,” *Sensors (Switzerland)*, vol. 21, no. 1, pp. 1– 18, Jan. 2021, doi: 10.3390/s21010210.
2. S. Sadu, S. R. Dubey, and S. Sreeja, “Moment Centralization based Gradient Descent Optimizers for Convolutional Neural Networks,” Jul. 2022, [Online]. Available: <http://arxiv.org/abs/2207.09066>
3. F. Ramzan *et al.*, “A Deep Learning Approach for Automated Diagnosis and Multi-Class Classification of Alzheimer’s Disease Stages Using Resting-State fMRI and Residual Neural Networks,” *J Med Syst*, vol. 44, no. 2, Feb. 2020, doi: 10.1007/s10916-019-1475-2.
4. Y. Lecun, E. Bottou, Y. Bengio, and P. Haffner, “Gradient- Based Learning Applied to Document Recognition,” 1998.
5. K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, IEEE Computer Society, Dec. 2016, pp. 770– 778. doi: 10.1109/CVPR.2016.90.
6. A. Krizhevsky, “Learning Multiple Layers of Features from Tiny Images,” 2009.
7. J. Chen and C. Zhu, “Research on Image Recognition Based on Improved ResNet,” in *2020 IEEE 6th International Conference on Computer and Communications, ICCC 2020*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 1422–1426. doi:

10.1109/ICCC51575.2020.9345181.