International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 14, Issue 02, NCRTCS - 2026

Performance Analysis of CNN Models in Noisy and Distorted Visual Environments

Shreyash Madhukar Palse
Master's Student Data Science
Dr. D. Y. Patil Arts, Commerce, Science College, Pimpri
Pune, India

Nilesh Kundlik Kumbhar
Master's Student Data Science
Dr. D. Y. Patil Arts, Commerce, Science College, Pimpri
Pune, India

***Abstract* - Deep Convolutional Neural Networks (CNNs) achieve high accuracy in image classification under clean conditions; however, their reliability in distorted real-world environments remains a critical concern. This study evaluates the robustness of two widely used architectures, ResNet-18 and EfficientNet-B0, using the CIFAR-10 dataset comprising 60,000 images across 10 classes. Both models are trained on clean images and tested under controlled distortions including Gaussian noise, blur, rotation, brightness variation, and JPEG compression at multiple intensity levels. Performance is measured using classification accuracy and accuracy degradation rate to quantify robustness. Experimental results show that although both models achieve above 90% accuracy on clean data, their performance degrades differently under varying distortions. The findings highlight the importance of robustness evaluation for reliable real-world deployment of CNN-based vision systems.**

The findings provide practical insights into model selection for real-world computer vision systems where image quality cannot be guaranteed. This work highlights the importance of robustness evaluation beyond standard benchmark testing and contributes to the development of more reliable deep learning models for deployment in dynamic visual environments.

Keywords - Convolutional Neural Networks, Robustness Analysis, Image Distortion, CIFAR-10, ResNet-18, EfficientNet-B0,Deep Learning, Image Classification.

I. INTRODUCTION

Image classification is a core task in computer vision with applications in healthcare diagnostics, intelligent transportation, smart surveillance, and industrial automation. The emergence of deep Convolutional Neural Networks (CNNs) has significantly improved classification accuracy by enabling automated feature extraction through hierarchical learning.

In our earlier research, we performed a comparative evaluation of ResNet-50 and EfficientNet-B0 using the VGG-Flowers dataset, focusing primarily on performance efficiency, model complexity, and classification accuracy under ideal conditions. Although that study provided insight into architectural efficiency, it did not address performance behavior under degraded image conditions.

In practical environments, images captured by cameras may suffer from environmental noise, low lighting, motion blur, or compression artifacts. Therefore, evaluating model robustness becomes equally important as evaluating accuracy. This study extends the previous work by examining how CNN architectures respond to controlled image distortions.

II. OBJECTIVES

The main objectives of the enhanced research are:

1. To train ResNet-18 and EfficientNet-B0 on clean CIFAR-10 images.
2. To evaluate performance under multiple image distortions.
3. To measure accuracy degradation across increasing distortion levels.
4. To compare robustness behavior between architectures.
5. To provide practical recommendations for real-world deployment.

III. PROBLEM STATEMENT

While CNN models demonstrate high accuracy on clean benchmark datasets, their behavior under distorted visual conditions is less studied. Real-world deployment environments introduce uncertainties that may significantly degrade performance.

The central question addressed in this study is:

How do ResNet-18 and EfficientNet-B0 perform under varying levels of controlled image distortions, and which architecture demonstrates greater robustness?

IV. METHODLOGY

This study adopts a structured experimental framework to evaluate the robustness of deep convolutional neural network architectures under controlled image distortions. The CIFAR-10 dataset was selected for experimentation due to its balanced distribution and moderate visual complexity. The dataset consists of 60,000 color images of size 32 × 32 pixels categorized into 10 object classes, including airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck. A total of 50,000 images were used for training, while 10,000 images were reserved for testing. Standard preprocessing techniques such as normalization and random horizontal flipping were applied during training to improve generalization performance. Importantly, no distortions were introduced during training, ensuring that the evaluation phase accurately reflects the models' inherent robustness rather than learned adaptation to corrupted data.

Two widely used CNN architectures were selected for evaluation:

1. ResNet-18

ResNet-18 utilizes residual (skip) connections that allow gradient information to flow directly across layers. This architectural design mitigates vanishing gradient issues and stabilizes deep network training. The 18-layer configuration was selected due to its computational efficiency and suitability for conference-scale experimentation.

2. EfficientNet-B0

EfficientNet-B0 applies compound scaling, which proportionally scales network depth, width, and input resolution. This strategy achieves competitive accuracy with significantly fewer parameters. EfficientNet-B0 serves as a parameter-efficient baseline for robustness comparison.

Both models were initialized using pre-trained ImageNet weights and fine-tuned on the CIFAR-10 training dataset under identical training conditions to ensure fair comparison. Training was conducted using the Adam optimizer with a learning rate of 0.001, a batch size of 64, and a cross-entropy loss function for 20 epochs. Validation accuracy and loss were monitored throughout the training process.

After training on clean data, robustness evaluation was performed by introducing controlled distortions to the test dataset. Distortions were applied independently to simulate realistic environmental degradations encountered in practical applications. Gaussian noise with increasing variance levels was added to simulate sensor noise and low-light interference. Gaussian blur with varying kernel sizes was applied to replicate motion blur and focus inconsistencies. Image rotation was introduced to assess geometric robustness, while brightness reduction simulated illumination variations. Additionally, JPEG compression at reduced quality levels was applied to evaluate resilience against transmission artifacts. Each distortion scenario generated a modified test set, enabling systematic performance comparison across different corruption intensities.

Model performance was assessed using classification accuracy as the primary metric, computed as the proportion of correctly predicted samples relative to the total number of test images. To quantify robustness, an accuracy degradation rate was calculated by measuring the percentage drop in performance relative to clean test accuracy. Furthermore, a robustness index was defined as the ratio of distorted accuracy to clean accuracy, providing a normalized measure of resilience across models. Confusion matrices were generated for both clean and distorted datasets to analyze class-level prediction behavior and identify patterns of misclassification under increasing distortion severity. In addition, robustness curves were plotted to visualize accuracy trends as distortion intensity increased, enabling a comparative assessment of stability between the two architectures.

The overall experimental pipeline consisted of dataset preprocessing, model training on clean images, evaluation on clean test data, application of controlled distortions to the test set, computation of performance metrics, and comparative robustness analysis. This methodology ensures a fair, systematic, and reproducible evaluation of CNN reliability under simulated real-world visual degradation conditions.

V. MODEL EVALUATION AND PERFORMANCE

The proposed robustness evaluation framework was assessed using standard performance metrics, including classification accuracy, precision, recall, F1-score, confusion matrix analysis, and robustness index. Both ResNet-18 and EfficientNet-B0 were trained on clean CIFAR-10 images and subsequently evaluated on both clean and systematically distorted test datasets.

Table 1: Clean Image Performance

Model	Accuracy (%)
ResNet-18	91.4
EfficientNet-B0	92.1

Both models achieved above 90% accuracy under clean conditions. Under clean test conditions, EfficientNet-B0 achieved an accuracy of 92.1%, slightly outperforming ResNet-18, which achieved 91.4%. Precision, recall, and F1-score values for both models remained consistently above 90%, indicating balanced classification performance across the 10 object categories.

When evaluated under distorted conditions, performance degradation was observed in both models; however, the rate of degradation varied depending on distortion type and severity. Under high Gaussian noise ($\sigma = 0.3$), ResNet-18 accuracy decreased to 70.2%, while EfficientNet-B0 maintained a relatively higher accuracy of 74.8%. Similarly, under strong blur (5×5 kernel), EfficientNet-B0 achieved 82.3% compared to ResNet-18's 78.5%. Brightness reduction and compression artifacts also affected performance, though EfficientNet-B0 demonstrated comparatively smoother degradation trends.

Table 2: Distortion Performance

Gaussian Noise ($\sigma = 0.3$)

Model	Accuracy (%)	Degradation (%)
ResNet-18	70.2	23.2
EfficientNet-B0	74.8	18.8

Blur (5×5)

Model	Accuracy (%)	Degradation (%)
ResNet-18	78.5	14.1
EfficientNet-B0	82.3	10.6

Brightness Reduction (40%)

Model	Accuracy (%)	Degradation (%)
ResNet-18	75.9	17.0
EfficientNet-B0	79.6	13.6

Overall, the experimental findings indicate that while both architectures achieve strong classification performance under ideal conditions, EfficientNet-B0 demonstrates slightly superior robustness across most distortion types. The results emphasize that evaluation under controlled corruptions provides deeper insight into model reliability beyond conventional clean-dataset benchmarking.

VI. RESULT ANALYSIS

The comparative results demonstrate that clean accuracy alone does not fully characterize model performance. Although both networks perform competitively under ideal conditions, robustness evaluation reveals meaningful architectural differences. EfficientNet-B0 exhibits improved stability under multiple distortion scenarios, making it more suitable for deployment in dynamic and visually uncertain environments.

Results indicate:

- EfficientNet-B0 generally shows slower degradation under blur and brightness variations.
- ResNet-18 shows competitive resilience under moderate noise.
- Severe distortions significantly reduce performance in both models.
- Clean accuracy alone is insufficient for model selection.
- Distortion type influences performance differently.
- EfficientNet-B0 shows better stability under multiple corruption types.

Table 3: Comparative robustness index across distortion types

Distortion Type	ResNet-18 Avg. Degradation (%)	EfficientNet-B0 Avg. Degradation (%)
Gaussian Noise	19.4	15.7
Blur	13.2	9.8
Brightness	15.5	12.3
Compression	11.8	9.1

While both models achieve comparable results on clean data, EfficientNet-B0 consistently exhibits lower degradation percentages and higher robustness index values across most distortion types. In contrast, ResNet-18, although competitive, shows relatively sharper performance decline under severe noise and blur conditions. The aggregated robustness metrics therefore confirm that EfficientNet-B0 provides a more balanced trade-off between accuracy and stability.

Overall, the robustness evaluation establishes EfficientNet-B0 as the more reliable architecture for practical vision systems operating in uncertain and dynamically changing environments.

VII. CONCLUSION

This study investigated the robustness of ResNet-18 and EfficientNet-B0 under controlled image distortions using the CIFAR-10 dataset. While both models achieved strong performance on clean images with accuracy above 90%, their behavior under degraded conditions revealed meaningful differences. As distortion severity increased, classification accuracy declined for both architectures; however, EfficientNet-B0 consistently maintained lower degradation rates and higher robustness index values across most distortion types.

The results demonstrate that evaluating models solely on clean benchmark datasets does not fully reflect their reliability in real-world environments. Robustness analysis provides deeper insight into model stability when exposed to noise, blur, illumination changes, and compression artifacts. Among the evaluated architectures, EfficientNet-B0 showed comparatively better resilience, making it a more suitable choice for deployment in practical computer vision systems where image quality cannot always be controlled.

Overall, this work emphasizes that robustness should be considered an essential evaluation criterion alongside accuracy when selecting deep learning models for real-world applications.

REFERENCES

[1] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 770–778.

[2] M. Tan and Q. Le, “EfficientNet: Rethinking model scaling for convolutional neural networks,” in *Proceedings of the 36th International Conference on Machine Learning (ICML)*, 2019, pp. 6105–6114.

[3] A. Krizhevsky and G. Hinton, “Learning multiple layers of features from tiny images,” University of Toronto, Technical Report, 2009. (CIFAR-10 dataset)

[4] S. Bianco, R. Cadene, L. Celona, and P. Napoletano, “Benchmark analysis of representative deep neural network architectures,” *IEEE Access*, vol. 6, pp. 64270–64277, 2018.

[5] R. Wightman, H. Touvron, and H. Jégou, “ResNet strikes back: An improved training procedure in timm,” *arXiv preprint arXiv:2110.00476*, 2021.

[6] D. Hendrycks and T. Dietterich, “Benchmarking neural network robustness to common corruptions and perturbations,” in *International Conference on Learning Representations (ICLR)*, 2019.

[7] I. J. Goodfellow, J. Shlens, and C. Szegedy, “Explaining and harnessing adversarial examples,” in *International Conference on Learning Representations (ICLR)*, 2015.

[8] S. Palse and N. Kumbhar, “Image classification using CNN architectures: ResNet vs EfficientNet,” *International Journal of Engineering Development and Research(IJEDR)*, vol. 13, issue no. 4, October-2025.