Journal of Engineering Faculty

| cumfad.cumhuriyet.edu.tr | Founded: 2023 Available online, ISSN:3023-7203 Publisher: Sivas Cumhuriyet Üniversitesi

Analyzing the Side Effects of Blur in Image Classification with Convolutional Neural Networks

Ahmet Fırat Yelkuvan[1,a,*]

[1] *Sivas Cumhuriyet University, Faculty of Engineering, Department of Computer Engineering, 58140, Türkiye*

**Corresponding author*

Research Article

History

Received: 07/04/2025
Accepted: 02/05/2025

ABSTRACT

Blur is one of the common factors that deteriorate image quality and can be caused by various factors such as motion, defocus, or environmental conditions. The presence of partially or globally blurred images in a dataset can make object recognition challenging, thereby reducing the effectiveness of image classification models. To mitigate this issue, blurred images must either be removed from the dataset or processed using deblurring techniques. In this project, the impact of blurred images on the performance of deep learning-based image classification models investigated. Specifically, the goal was to analyze how different levels of image blur affect classification accuracy. To achieve this, a convolutional neural network (CNN) model was trained using the CIFAR-10 dataset, with varying proportions of blurred images: 0%, 25%, 50%, and 100%. The experiment results demonstrated that increasing the proportion of blurred images in the training dataset led to a decline in validation accuracy. The model achieved validation accuracies of 67.53%, 65.50%, 63.90%, and 55.74% when trained with datasets containing 0%, 25%, 50%, and 100% blurred images, respectively. These findings highlight the adverse effects of image blur on classification performance, emphasizing the importance of high-quality image data in deep learning applications.

Keywords: convolutional neural network, blurred images, object recognition, image classification

[a] *aftyelkuvan@cumhuriyet.edu.tr* *ORCID* | [b] *mail* *ORCID*

How to Cite: Yelkuvan F, (2025) Analyzing the Side Effects of Blur in Image Classification with Convolutional Neural Networks, Journal of Engineering Faculty, 3(1): 32-37

Introduction

Many digital images contain blurred regions caused by motion or defocus. The automatic identification and classification of these blurred regions are crucial for various multimedia processing tasks [1]. To address these issues, most existing studies focus on detecting blurred images, classifying them, and removing them if users wish to curate high-quality photo albums. However, these approaches primarily aim to identify and filter blurred images rather than classifying the objects within them.

Image blur continues to be a prominent challenge in computer vision, particularly in tasks such as object recognition and image classification. Initial efforts in this area concentrated on detecting and characterizing blur within images. Early studies introduced statistical and gradient-based methods for identifying blurred regions and classifying them accordingly [1,2].

With the growing influence of deep learning, a shift occurred from hand-crafted feature-based methods to data-driven approaches. Neural networks began to be employed for blind image blur correction, utilizing latent semantic features to restore image quality [3]. Deep learning-based restoration methods also demonstrated strong performance in recovering fine details lost due to blur [4].

More sophisticated frameworks soon emerged, such as variational Bayesian models that account for uncertainty in the deblurring process, resulting in improved blind deblurring performance [5]. Surveys by Zhang et al. [6] highlighted the rise of architectures that integrate attention mechanisms and residual learning for robust deblurring. Similarly, neural blind deconvolution methods using deep priors have shown strong performance across varied scenarios [7].

Parallel to these developments, researchers explored the assessment of blur without reference images. One method utilized wavelet transform features for classifying blur severity in an unsupervised manner [8].

In large-scale datasets such as CIFAR-10 and ImageNet, convolutional neural networks have been shown to be vulnerable to quality distortions, including blur, which significantly impacts classification accuracy [9,10]. These findings underscore the importance of building blur-resilient models for real-world applications.

Recent studies have further extended the field with generative and transformer-based approaches. A generative diffusion model, BD-Diff, was introduced to address deblurring in unknown domains by decoupling structural and blur-specific components through a blur-aware learning process [11]. Transformer-based networks have also been optimized for high-resolution motion deblurring, offering reduced computational complexity without sacrificing performance [12]. Additionally, a comprehensive review by Xiang et al. [13] categorized the deep learning-based deblurring literature and emphasized the need for interpretability and standardized benchmarks.

Table 1 presents a comparative summary of prior works addressing image blur in classification tasks. In contrast to the broader corruption robustness studies or region-based blur detection, our study provides a fine-grained evaluation of performance degradation at multiple blur levels using a consistent dataset and model.

These developments collectively demonstrate progression from conventional blur detection to advanced, learning-based techniques capable of enhancing image clarity and maintaining classification performance under challenging conditions. The present study builds upon this foundation by systematically analyzing the impact of different levels of blur in training data on model performance, offering insights into how blur-resilient strategies can be integrated into deep learning pipelines.

Table 1: Comparison of related studies analyzing the impact of blur on image classification tasks:

Study	Dataset	Blur Type	Key Contribution	Limitation
Su et al. (2011)	Real-world images	Local blur	Detection and classification of blurred regions	Did not study classification impact
Vasiljevic et al. (2016)	ImageNet	Synthetic Gaussian blur	Evaluated CNN robustness to blur	Focused on trained vs. test image blur mismatch
Zhang et al. (2018)	CIFAR-10, ImageNet-C	Multiple corruptions including blur	Proposed benchmark for corruption robustness	Did not isolate blur's individual effects
This study	CIFAR-10	Global Gaussian blur at 4 levels (0%, 25%, 50%, 100%)	Quantifies classification accuracy degradation with blur ratio	Limited to one dataset and basic CNN architecture

Materials and Methods

Dataset

CIFAR-10 Dataset was used in this study as the main dataset. The CIFAR-10 dataset consists of 60000 32x32 color images in 10 classes, with 6000 images per class. There are 50000 training images and 10000 test images. The dataset is divided into five training batches and one test batch, each with 10000 images. The test batch contains exactly 1000 randomly selected images from each class [10]. The training batches contain the remaining images in random order, but some training batches may contain more images from one class than another. Between them, the training batches contain exactly 5000 images from each class. To conduct a CNN model based on train, test and validation have splitted train test into two parts with 38000 and 12000 images then transfer the part which contains 12000 into validation. As a result, a training data containing 38000 images, test data containing 10000 images, and finally validation data containing 12000 images.

Besides the CIFAR-10 dataset, blurred images were generated via algorithm which use gaussian kernel with 1 radius. The blurred images were used as training and test dataset, validation part was only consisting of non-blurred images.

A Gaussian kernel with radius 1 was selected to simulate mild to moderate blur while preserving object structure, ensuring the classification task remained challenging but not impossible. This value is also consistent with prior work on controlled degradation in image classification.

Algorithms and Models

A CNN model was conducted which contains two 2D convolutional (Conv2D) layers, a max pooling with the dimension of 2 by 2 used between them, and a flatten layer followed the second Conv2D layer. There was also a Dense layer at the end of the architecture. The model is shown in

Fig 1. The input of the model 32 x 32 x 3 was used since the CIFAR-10 dataset has 32 x 32 RGB images. Some deeper architecture was performed, and the accuracy result did not change significantly. That is why I used the mentioned model with lower trainable parameters.

In this project, the model was trained with a dataset which consists of blurred and original images from CIFAR-10 dataset. While both blurred and original images were used for training and testing, only non-blurred images were used for validation. The model was trained using CIFAR-10 dataset of which different percentages (%0, %25, %50 and %100) were blurred.

The prediction results for each experimental condition—corresponding to 0%, 25%, 50%, and 100% blurred training data—were further analysed by generating confusion matrices. These matrices which are presented in Results section provided a detailed view of class-wise prediction accuracy and misclassification patterns, allowing us to evaluate how varying degrees of blur influenced the model's ability to correctly distinguish between specific object categories. This class-level analysis offered deeper insights into which categories were most susceptible to performance degradation under blur and revealed trends in model behaviour across different blur scenarios.

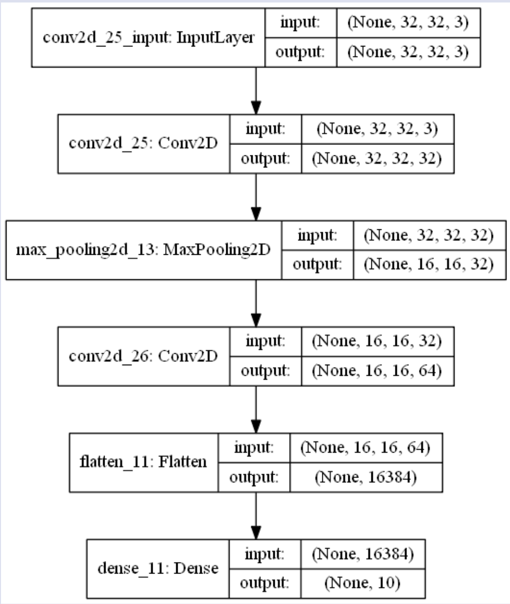

Figure 1: CNN model of the study.

Results

To determine the optimal number of epochs for the model, experiments were conducted using different epoch values. It was observed that around the 85th epoch, the accuracy began to reach a saturation point, indicating diminishing improvements with additional training.

Figure 2 presents the validation accuracy graph for the model trained with 100 epochs using only non-blurred images. As shown in the graph, the accuracy levels off between 80 and 90 epochs, confirming that the model reaches its saturation accuracy within this range. Based on this observation, 85 epochs were chosen as the optimal value for subsequent experiments in this study.

The training and validation accuracy values are presented in Figure 3. All accuracy values were averaging over five iterations to ensure reliability. As shown in the figure, the fluctuations in validation accuracy increased as the percentage of blurred images in the dataset rose. Notably, the stability of validation accuracy significantly decreased when the model was trained with 100% blurred images, indicating a more erratic learning process.

The final validation accuracies of the model, trained with different proportions of blurred images, were as follows:

- 0% blurred images: 67.53%
- 25% blurred images: 65.50%
- 50% blurred images: 63.90%
- 100% blurred images: 55.74%

These results, illustrated in Figure 4, demonstrate a near-linear decline in classification accuracy as the proportion of blurred images in the dataset increases, clearly indicating that the presence of blur—even at moderate levels—can substantially degrade the performance of deep learning-based image classification models. This trend validates the hypothesis that image blur poses a significant challenge to object recognition tasks and underscores the importance of developing robust, blur-resilient strategies to maintain model reliability and accuracy in real-world scenarios where image quality may be compromised.

Figure 5 presents the confusion matrix, illustrating the class-based performance of the model trained for 100 epochs using only original (non-blurred) images.

According to the results, the highest levels of misclassification occurred between the dog and cat classes, as well as between the airplane and ship classes. These misclassifications can be attributed to the visual similarities between certain categories. The confusion between dogs and cats arises due to their similar appearance, as both are four-legged animals with comparable body structures. Similarly, airplanes and ships were frequently misclassified because their images share common visual elements, such as the general shape of the main body and backgrounds featuring open skies or water, which may introduce ambiguity into the classification process. These findings highlight the challenges associated with distinguishing visually similar objects in image classification tasks.

Across all experiments, some class-based prediction performances exhibited instability. The number of correctly predicted instances for the deer class varied between iterations, with 726 correct predictions in the first iteration and 906 in the second iteration. Similarly, for the automobile class, the model achieved 947 correct predictions in the first iteration, but this number decreased to 855 in the second iteration.

The average class-based prediction values from the confusion matrices, aggregated over five iterations, are presented in Figure 6. The results indicate that the airplane, automobile, and frog classes consistently achieved higher prediction performance across all experiments. Additionally, while the ship class exhibited high prediction accuracy in three of the experiments, its prediction performance declined significantly in the experiment where 100% of the images were blurred. Notably, the greatest performance declines due to increasing blur was observed in the cat, deer, and dog classes. This suggests that these classes are particularly susceptible to misclassification when image quality is degraded, highlighting the significant impact of blur on object recognition accuracy.

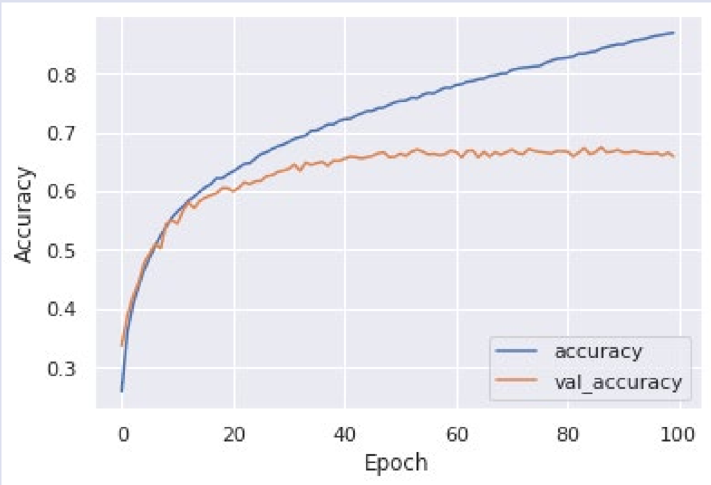

Figure 2: Validation Accuracy with epoch 100 non-blurred images

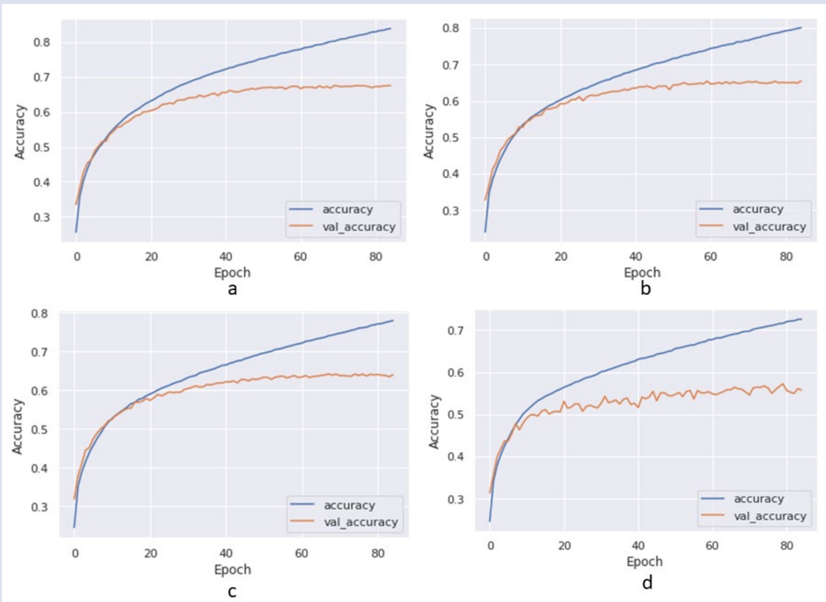

Figure 3: Validation accuracy results for a) %0, b) %25, c) %50 and d) %100

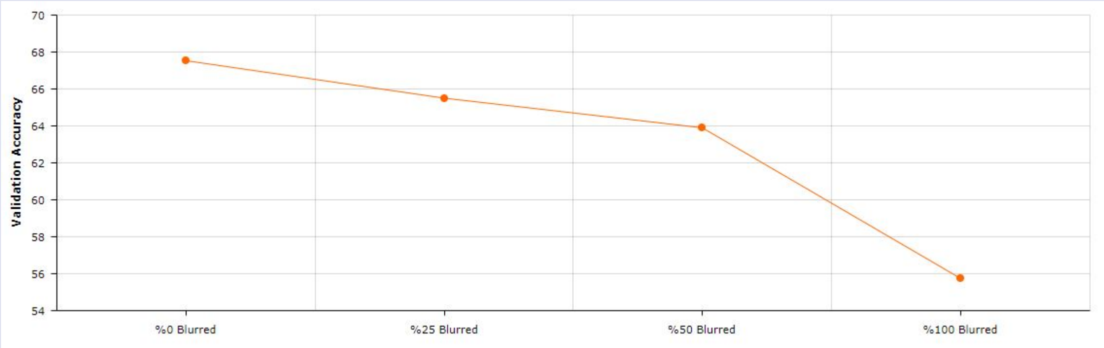

Figure 4: Validation accuracies for all percentages

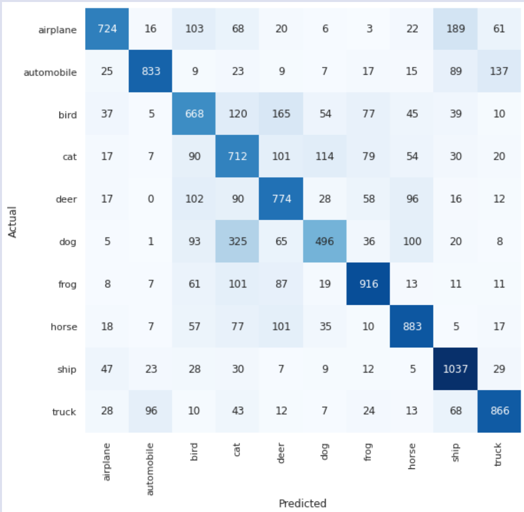

Figure 5: Heatmap of the confusion matrix with 100 epoch and %0 blur.

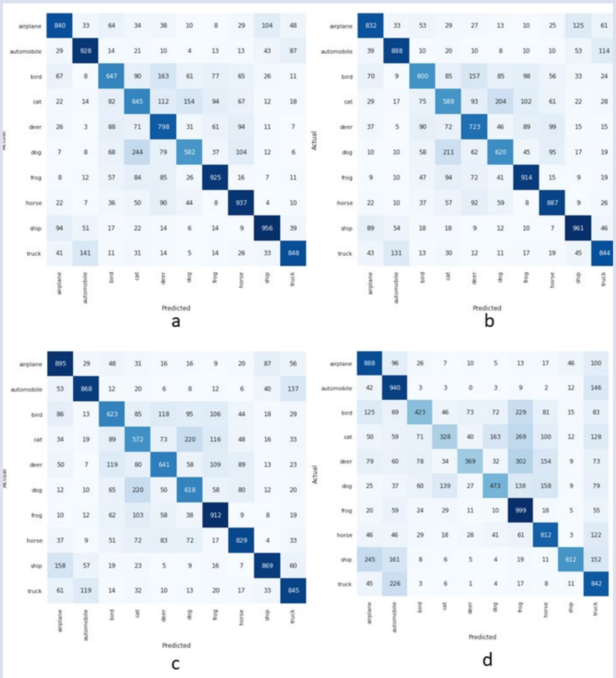

Figure 6: Average confusion matrices for a) %0, b) %25, c) %50 and d) %100

Conclusion

In this study, the impact of training a deep learning model with datasets containing different proportions of blurred images was analyzed. The dataset images were blurred at four levels: 0%, 25%, 50%, and 100%. One key observation was that class-based prediction stability was poor between iterations, particularly for the deer and automobile classes. Additionally, the misclassification between dogs and cats as well as airplanes and ships persisted across successive iterations. This issue was largely attributed to the low resolution of the CIFAR-10 images, which made it difficult for the model to distinguish between visually similar objects.

The most significant accuracy declines due to increasing blur were observed in the cat, deer, and dog classes. Another side effect of low resolution was the loss of object-specific features, which contributed to classification confusion, especially for dog-cat and airplane-ship pairs. Furthermore, in cases where objects were relatively small within the images, their distinguishing features became indistinguishable after blurring, leading to further classification errors. The validation accuracy values exhibited increased fluctuations (ripples) as the percentage of blurred images in the dataset increased. Moreover, as expected, overall classification accuracy decreased proportionally to the amount of blur in the dataset.

From a practical standpoint, we recommend that datasets intended for deep learning classification tasks maintain a blur proportion below 25% to preserve model robustness. For critical applications, incorporating deblurring mechanisms or blur-aware training strategies is advisable.

In future research, the impact of blurred images on class-based confusion in deep learning classification can be further analyzed using statistical methods such as mean, standard deviation, and sum of squared differences (SSD). These analyses could provide deeper insights into how blur affects individual class predictions and overall model performance.

References

[1] Su B, Lu S, Tan CL. Blurred image region detection and classification. In: Proceedings of the 19th ACM International Conference on Multimedia; 2011. New York: Association for Computing Machinery. p. 1397–1400.

[2] Liu R, Li Z, Jia J. Image partial blur detection and classification. In: 2008 IEEE Conference on Computer Vision and Pattern Recognition; 2008 Jun; Anchorage, AK. p. 1–8.

[3] Gowthami S, Harikumar R. Conventional neural network for blind image blur correction using latent semantics. Soft Comput 2020;1–15.

[4] De Vylder J, et al. Image restoration using deep learning. In: Annual Machine Learning Conference of Belgium and The Netherlands (Benelearn 2016); 2016.

[5] Wang H, Yue Z, Zhao Q, Meng D. A deep variational Bayesian framework for blind image deblurring. IEEE Trans Pattern Anal Mach Intell 2021;43(6):1971–1984.

[6] Zhang K, Ren W, Luo W, Lai WS, Stenger B, Yang MH, Li H. Deep image deblurring: A survey. IEEE Trans Pattern Anal Mach Intell 2023;45(1):1–20.

[7] Ren D, Zhang K, Wang Q, Hu Q, Zuo W. Neural blind deconvolution using deep priors. IEEE Trans Image Process 2020;29:4379–4391.

[8] Yoneyama A, Minamoto T. No-reference image blur assessment in the DWT domain and blurred image classification. In: 2015 12th Int Conf on Information Technology – New Generations; 2015 Apr; Las Vegas, NV. p. 329–334.

[9] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. Commun ACM 2017;60(6):84–90.

[10] Krizhevsky A, Hinton GE. Learning multiple layers of features from tiny images. Tech Rep, University of Toronto; 2009.

[11] Cheng J, Chen WT, Lu X, Yang MH. BD-Diff: Generative diffusion model for image deblurring on unknown domains with blur-decoupled learning. arXiv preprint arXiv:2502.01522; 2025.

[12] Amanturdieva A, Zafar MH. Efficient transformer for high-resolution image motion deblurring. arXiv preprint arXiv:2501.18403; 2025.

[13] Xiang Y, Zhou H, Li C, Sun F, Li Z, Xie Y. Deep learning in motion deblurring: current status, benchmarks and future prospects. arXiv preprint arXiv:2401.05055; 2024.