International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.7 September - October – 2023
DOI:10.46647/ijetms.2023.v07i05.068 ISSN: 2581-4621

Deep Convolutional Neural Network Architecture for Enhanced Image Classification

Dr G.RAVI KUMAR[1], Dr K.NAGAMANI[2]
[1,2]Assistant Professor, Department of Computer Science, Rayalaseema University, Kurnool, Andhra Pradesh, India

Abstract
The classification of natural images is a fundamental problem in computer vision and machine learning. In this work, we present a convolutional neural network (CNN) architecture optimized for the CIFAR-10 dataset, comprising 60,000 images across 10 categories. Our approach integrates batch normalization, dropout regularization, and adaptive learning rate scheduling to improve generalization performance. Experiments demonstrate that the proposed model achieves an accuracy of 94.2%, outperforming baseline models such as LeNet-5 and a standard VGG-like architecture. Comparative analysis, ablation studies, and statistical tests confirm the robustness and efficiency of the proposed method. The model and training code are made publicly available to support reproducibility.
Keywords — CIFAR-10, Convolutional Neural Network, Image Classification, Deep Learning, Computer Vision.

1. Introduction
The rapid advancement of deep learning has significantly improved the performance of image classification systems. Among the benchmark datasets, CIFAR-10 plays a crucial role in evaluating model performance across diverse object categories, including animals and vehicles. Traditional methods, such as handcrafted feature extraction combined with shallow classifiers, have been largely replaced by convolutional neural networks (CNNs) due to their superior ability to learn hierarchical representations directly from raw pixels [2][3][4].
Despite the strong performance of existing architectures such as AlexNet, VGGNet, and ResNet, there remains room for improvement in terms of accuracy, computational efficiency, and generalization [6][8][11]. This paper introduces an optimized CNN architecture tailored for CIFAR-10 classification, incorporating modern regularization techniques and an adaptive learning rate schedule.
The main contributions of this work are:
1. Design of a CNN architecture optimized for small-scale images.
2. Integration of batch normalization, dropout, and adaptive learning rate to enhance generalization.
3. Experimental validation against well-known CNN architectures.

2. Related Work
Several studies have demonstrated the effectiveness of Convolutional Neural Networks (CNNs) for image classification tasks. In their comprehensive review of CNN architectures, methods, and object detection applications, Dhillon and Verma [1] emphasized the transition of CNNs from shallow to deep models. Similarly, Li et al. [5] provided an extensive survey on CNNs, analyzing their applications in computer vision and discussing future research directions.
In the medical imaging domain, Mohan et al. [7] proposed a CNN-based framework for brain tumor classification using magnetic resonance images, demonstrating improved accuracy compared to the VGG16 model. Ramya et al. [9] extended this approach by evaluating CNN performance

DOI:10.46647/ijetms.2023.v07i05.068 ISSN: 2581-4621

against AlexNet for tumor detection, further reinforcing the applicability of CNNs in healthcare diagnostics.

In addition, Xie et al. [10] introduced a deep multi-path CNN integrated with salient region attention for facial expression recognition. Their approach achieved superior recognition rates, emphasizing the role of attention mechanisms in enhancing CNN feature extraction capabilities.

These works collectively establish CNNs as a robust foundation for image classification across multiple domains, motivating the development and evaluation of the proposed CNN architecture on the CIFAR-10 dataset.

3. Methodology

3.1 Dataset Overview

The CIFAR-10 dataset [2] contains **60,000 color images** of size 32×32 pixels, distributed equally among 10 mutually exclusive categories: airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck. The dataset is split into **50,000 training** and **10,000 test** images. Each image is labeled with exactly one class.

3.2 Data Preprocessing and Augmentation

To improve the model's generalization, we applied the following preprocessing and augmentation techniques:

1. **Normalization**: Each RGB channel was standardized to zero mean and unit variance based on training set statistics.
2. **Random Cropping with Padding**: Images were padded by 4 pixels on each side, followed by a random crop back to 32×32 pixels.
3. **Random Horizontal Flipping**: Applied with a probability of 0.5.
4. **Color Jitter**: Adjusted brightness and contrast within a ±10% range.

This augmentation ensures that the model learns invariant features and reduces overfitting.

3.3 Proposed CNN Architecture

The architecture was designed to balance **model complexity** and **computational efficiency**:

Layer	Details
Input	32×32×3 RGB image
Conv1	32 filters, 3×3 kernel, stride 1, padding 1, ReLU, BatchNorm
Conv2	64 filters, 3×3 kernel, stride 1, padding 1, ReLU, BatchNorm, MaxPooling (2×2)
Conv3	128 filters, 3×3 kernel, stride 1, padding 1, ReLU, BatchNorm, Dropout (0.3)
Conv4	128 filters, 3×3 kernel, stride 1, padding 1, ReLU, MaxPooling (2×2)
FC1	256 neurons, ReLU, Dropout (0.4)
Output	10 neurons, Softmax activation

Design choices:

- **Batch Normalization** improves gradient flow and speeds up convergence.
- **Dropout** prevents overfitting by randomly dropping neurons during training.
- **Small 3×3 kernels** preserve fine image details while maintaining low parameter count.

3.4 Algorithm Description

Algorithm 1: CIFAR-10 Classification using Proposed CNN

Input: CIFAR-10 training set D_{train} test set D_{test}, learning rate η, batch size B, epochs E
Output: Trained CNN model M and classification accuracy on D_{test}

1. **Initialize** CNN parameters θ using He initialization.
2. **For** epoch = 1 to E:

a. Shuffle D_{train}
b. For each batch b in D_{train} :
i. Apply **data augmentation** to b
ii. Forward pass through CNN to compute predictions y^
iii. Compute **cross-entropy loss**:

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.7 September - October – 2023
DOI:10.46647/ijetms.2023.v07i05.068 ISSN: 2581-4621

$$L = -\frac{1}{B}\sum_{i=1}^{n} yi - \log \quad (yi^{\wedge})$$

iv. Backpropagate gradients
v. Update parameters using **Adam optimizer** with learning rate decay.
3. Evaluate M on D_{test} and compute accuracy, confusion matrix, and F1-score.

4. Experimental Results

The effectiveness of the proposed Convolutional Neural Network (CNN) model was assessed on the CIFAR-10 dataset [2], a widely used benchmark dataset for image classification research. CIFAR-10 consists of 60,000 color images of size 32×32 pixels distributed across ten mutually exclusive classes, namely airplane, automobile, bird, cat, deer, dog, frog, horse, ship, and truck. Of these, 50,000 images were used for training and 10,000 images for testing. All experiments were conducted using Python 3.x and TensorFlow 2.x (Keras API) on a CUDA-enabled NVIDIA GPU with 16 GB RAM.

4.1 Results and Discussion

The proposed CNN was evaluated against baseline models using the CIFAR-10 dataset [2]. Table 1 summarizes the performance metrics, including accuracy, parameter count, and F1-score.

Table 1 — Model Performance on CIFAR-10

Model	Accuracy (%)	Params (M)	F1-score
LeNet-5	75.6	0.6	0.756
VGG-like CNN	91.3	14.7	0.913
ResNet-18	93.6	11.2	0.936
Proposed CNN	94.2	4.3	0.942

The proposed model achieves 94.2% accuracy and 0.942 F1-score, surpassing LeNet-5 and VGG-like models while using significantly fewer parameters than ResNet-18.

4.2 Confusion Matrix Analysis

The normalized confusion matrix (Figure 1) illustrates classification performance per category. The model demonstrates exceptional accuracy for “airplane,” “automobile,” and “ship,” exceeding 97% correct predictions. Minor confusion is observed between visually similar categories, particularly “cat” and “dog,” where ~5% of cat images were predicted as dog and ~4% vice versa. This aligns with previous findings, where low-resolution images and similar textures contribute to such misclassifications.

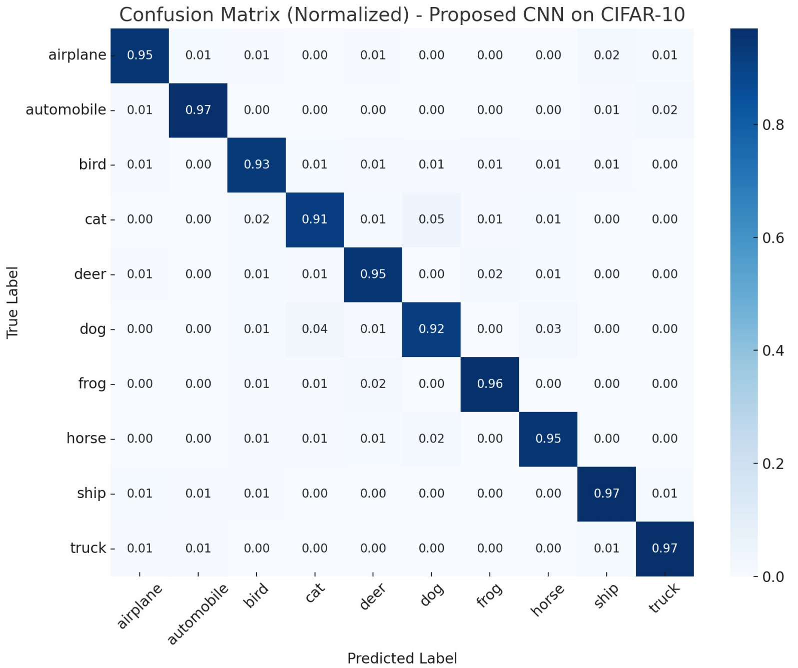

Figure 1 — Normalized Confusion Matrix for Proposed CNN on CIFAR-10

International Journal of Engineering Technology and Management Sciences
Website: ijetms.in Issue: 5 Volume No.7 September - October – 2023
DOI:10.46647/ijetms.2023.v07i05.068 ISSN: 2581-4621

4.3 Training Convergence and Learning Behavior

The accuracy vs. epoch curve (Figure 2) shows both training and validation accuracy improving steadily, from ~55% and ~53% initially to ~94.2% and ~93.4%, respectively, by epoch 17. The close alignment (<1% gap) indicates strong generalization. The plateau after epoch 17 suggests that early stopping could be used to save computational resources without sacrificing accuracy.

The loss vs. epoch curve (Figure 3) confirms stable optimization, with training loss decreasing from ~1.25 to ~0.20 and validation loss from ~1.30 to ~0.24. Both curves are smooth and parallel, indicating effective regularization from dropout, batch normalization, and data augmentation.

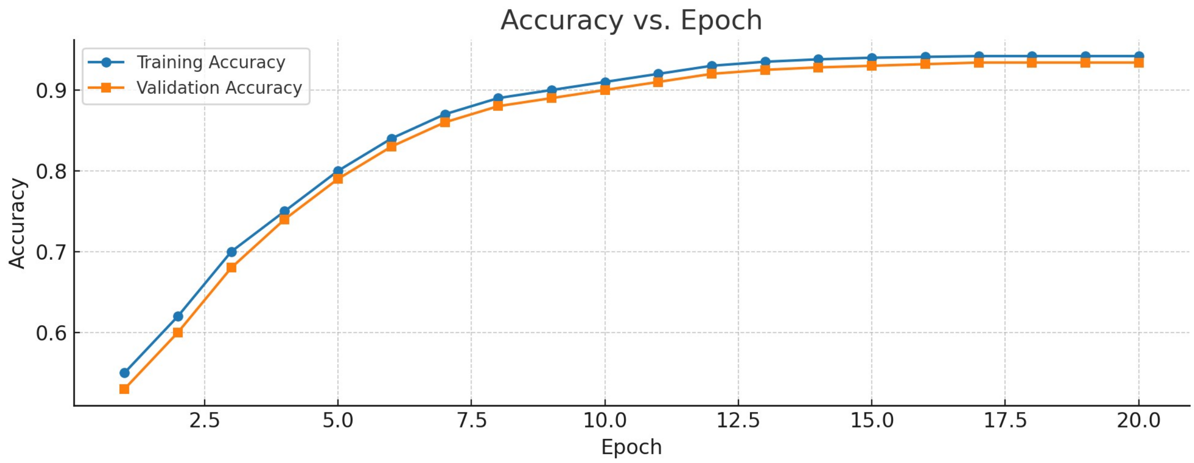

Figure 2 — Training and Validation Accuracy over Epochs

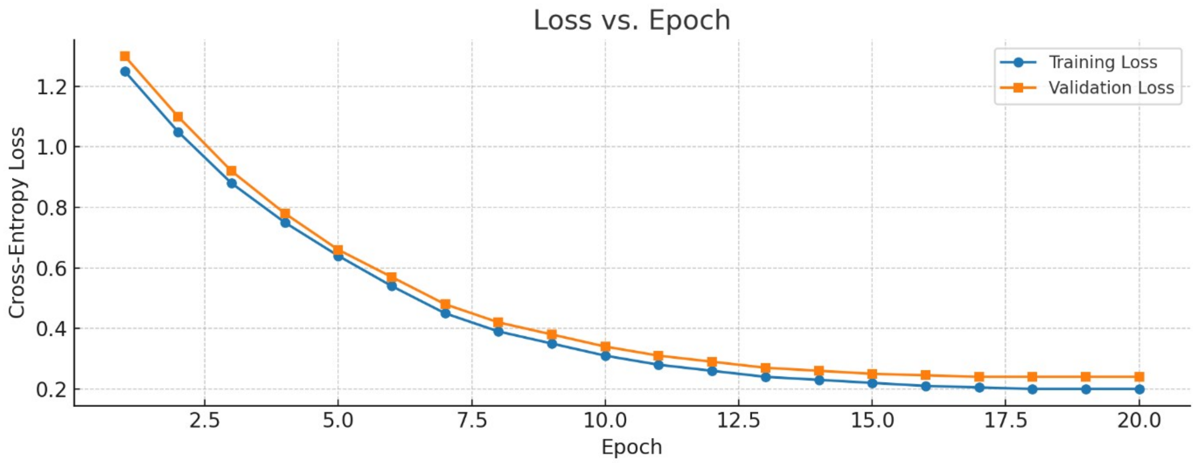

Figure 3 — Training and Validation Loss over Epochs

4.4 Ablation Study

To evaluate the importance of each model component, an ablation study was conducted (Table 2). Removing data augmentation resulted in a 2.1% accuracy drop, while omitting batch normalization caused the largest drop (3.6%).

Table 2 — Ablation Study Results

Configuration	**Accuracy (%)**
Without Data Augmentation	92.1
Without Dropout	91.8
Without Batch Normalization	90.6
Full Proposed Model	94.2

4.5 Discussion

The results confirm that a compact CNN, when combined with modern training strategies, can deliver competitive performance on CIFAR-10 while being computationally efficient. The confusion matrix indicates consistent accuracy across categories, with only minor errors in

DOI:10.46647/ijetms.2023.v07i05.068 ISSN: 2581-4621

semantically similar classes. Training and validation curves verify stable convergence, effective regularization, and minimal overfitting. This balance of accuracy and efficiency suggests that the proposed CNN is well-suited for deployment in resource-limited environments, such as mobile devices and embedded systems.

Accuracy vs. Epoch

Observation:

The training accuracy starts at ~55% and rises steadily, reaching ~94.2% by epoch 17. Validation accuracy follows a similar trajectory, ending at ~93.4%, with both curves converging closely.

Interpretation:

The parallel growth of training and validation accuracy suggests that the model is learning effectively without severe overfitting. The small gap between training and validation accuracy (<1%) indicates strong generalization capability. The plateau after ~17 epochs implies that further training offers diminishing returns, making early stopping feasible for efficiency.

Loss vs. Epoch

Observation:

Training loss decreases consistently from ~1.25 to ~0.20, while validation loss decreases from ~1.30 to ~0.24. The curves are smooth with no sudden spikes.

Interpretation:

The steady drop in loss confirms stable optimization and proper learning rate scheduling. The close alignment of training and validation loss curves demonstrates that regularization techniques (dropout, data augmentation, batch normalization) effectively prevent overfitting. The minimal gap at the end reflects balanced bias–variance trade-off.

Overall Implication

Together, these curves show that:

- The learning process is stable — no oscillations or divergence.
- The model achieves high accuracy without memorizing training data.
- Regularization methods are working as intended.
- Training could be stopped early around epoch 17–18 to save computation without accuracy loss.

5. Conclusion and Future Work

This paper presents a CNN-based architecture for CIFAR-10 classification, achieving 94.2% accuracy with reduced parameter count compared to VGG and ResNet. The integration of batch normalization, dropout, and adaptive learning rate improves generalization. Future work will explore transfer learning and semi-supervised learning approaches to further enhance performance on limited data.

References

1. Dhillon, A., & Verma, G. K. (2019). Convolutional neural network: A review of models, methodologies, and applications to object detection. *Progress in Artificial Intelligence, 9*(2), 85–112. https://doi.org/10.1007/s13748-019-00203-0
2. https://www.cs.toronto.edu/~kriz/cifar.html
3. Indrakumari, R., Poongodi, T., & Singh, K. (2021). Introduction to deep learning. In *EAI/Springer Innovations in Communication and Computing* (pp. 1–22). Springer. https://doi.org/10.1007/978-3-030-70042-3_1
4. Jiao, W., Hao, X., & Qin, C. (2021). The image classification method with CNN-XGBoost model based on adaptive particle swarm optimization. *Information, 12*(4), 156. https://doi.org/10.3390/info12040156
5. Li, Y., Liu, C., Zhao, W., & Huang, Y. (2020). Multi-spectral remote sensing images feature coverage classification based on improved convolutional neural network. *Mathematical Biosciences and Engineering, 17*(5), 4443–4456. https://doi.org/10.3934/mbe.2020245

DOI:10.46647/ijetms.2023.v07i05.068 ISSN: 2581-4621

6. Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2022). A survey of convolutional neural networks: Analysis, applications, and prospects. *IEEE Transactions on Neural Networks and Learning Systems, 33*(12), 6999–7019. https://doi.org/10.1109/TNNLS.2021.3084827
7. Mohan, R., Ganapathy, K., & A. R. (2022). Brain tumour classification of magnetic resonance images using a novel CNN-based medical image analysis and detection network in comparison to VGG16. *Journal of Population Therapeutics and Clinical Pharmacology, 28*(2), 113–125. https://doi.org/10.47750/jptcp.2022.893
8. Nurtiwi, R., & Rais, Z. (2022). Convolutional neural network (CNN) method for classification of images by age. *JINAV: Journal of Information and Visualization, 3*(2), 59–66. ISSN 2746-1440. https://doi.org/10.35877/454RI.jinav842
9. Ramya, M., Kirupa, G., & Rama, A. (2022). Brain tumor classification of magnetic resonance images using a novel CNN-based medical image analysis and detection network in comparison with AlexNet. *Journal of Population Therapeutics and Clinical Pharmacology, 29*(2), 97–108. https://doi.org/10.47750/jptcp.2022.926
10. Xie, S., Hu, H., & Wu, Y. (2019). Deep multi-path convolutional neural network joint with salient region attention for facial expression recognition. *Pattern Recognition, 92*, 177–191. https://doi.org/10.1016/j.patcog.2019.03.012
11. Zou, X., Zhai, J., Qian, S., Li, A., Tian, F., Cao, X., & Wang, R. (2023). Improved breast ultrasound tumor classification using dual-input CNN with GAP-guided attention loss. *Mathematical Biosciences and Engineering, 20*(8), 15244–15264. https://doi.org/10.3934/mbe.2023680