iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 Issue: VI Month of publication: June 2024

DOI: https://doi.org/10.22214/ijraset.2024.61627

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue VI June 2024- Available at www.ijraset.com

Image Classification Using CNN with CIFAR-10 Dataset

Y. Lokesh[1], M. Rithvik[2], M. Madhavi[3], S. Madhu[4], Prof. Dr. K. Rajeshwar Rao[5]
[1, 2, 3, 4]Undergraduate students, [5]Prof. Department of Computer Science, Mallar Reddy University

Abstract: Image classification is a very important issue in digital image analysis. CNN is type of Deep Neural Networks (DNN) that contains multiple layerssuch as Conv layers, integration layer and fully integrated layer. Convolutional neural networks (CNNs) are widely used in pattern-and picture- recognition problems as they have a number of advantages compared to others strategies. CIFAR-10is a very popular computer vision database. Thisdatabase is well read in it many types of in-depth study of object recognition. This database contains 60,000 images separated by 10 target classes, each a section containing 6000 images of 32 * 32 shapes. This database contains images of low-resolution (32 * 32), which allows researchers to experiment with new algorithms. Image classification is a fundamentaltask in computer vision that involves assigning a labelor category to an image. Convolutional Neural Networks (CNNs) have become the state-of-the-art method for image classification due to their ability toautomatically learn features from raw pixel values. Inthis project, we aim to build a CNN model that can accurately classify images in the CIFAR-10 dataset.
Keywords: Convolutional Neural Network (CNN), CIFAR-10 dataset, Image classification, Deep Learning, Computer Vision, Feature extraction,Accuracy, Data augmentation.

I. INTRODUCTION

Our project focuses on the development of an interactive image classification system using Convolutional Neural Networks (CNNs). Image classification is a fundamental task in computer vision with numerous applications across various domains, including healthcare, security, and transportation. Leveraging the CIFAR-10 dataset, which comprises 60,000 32x32 color images distributed across 10 classes, our objective is to create a user-friendly application that allows individuals to upload images and obtain real-time predictions regarding the class ofthe uploaded image. In today's digital era, the ability to understand and interpret images is crucial for various applications, ranging from autonomous driving to medical diagnosis. Image classification, a fundamental task in computer vision, involves categorizing images into predefined classes based on their visual content. Convolutional Neural Networks (CNNs) have emerged as powerful tools for image classification, capable of learning intricate patterns and features directly from pixel values. Our project focuses on leveraging CNNs to classify images from the CIFAR-10 dataset accurately. The primary objective is to develop a robust CNN architecturecapable of effectively learning and extracting relevantfeatures from the dataset while generalizing well to unseen images. By doing so, we aim to showcase the potential of deep learning techniques in image classification tasks and contribute to advancements incomputer vision research

II. LITERATURE SURVEY

The development of deep convolution neural networks [1] and large-scale hand-labeled datasets like ImageNet [2] are contributing factors to the advancement of picture categorization. Good resultsare obtained from recent work that applies deepconvolutional neural networks to multilabel classification. Optimizing a deep convolutional ranking [3]. Top-k ranking goal, which gives the loss of a positive label a lower weight. Max pooling is used by Hypotheses-CNN-Pooling [4] to combine the predictions from several hypothesis region suggestions. These approaches mostly disregard the correlations between labels and handle each label independently. Another method for achieving multilabel classification is to learn a joint image/label embedding. A three-way canonical analysis that maps the image, label, and semantics into the same latent space is called Multiview canonical correlation analysis [5]. WASABI [6] andDEVISE [7] employ the learning to rank frameworkwith WARP loss to acquire the joint embedding. Toquantify the similarity between an image and a label,metric learning [8] trains a discriminative metric. Aslabel encodings, bloom filters [9] and matrix completion [10] can also be used. While these approaches successfully take advantage of the semantic redundancy of the label, they are not able to represent the label co-occurrence dependency.Various approaches have been proposed to exploit the label co-occurrence dependency for multi-label image classification. [11] learns a chain of binary classifiers, where each classifier predicts whether the current label exists given the input feature and the already predicted labels.

The label co-occurrence dependency can also be modeled by graphical models, such as Conditional Random Field [12], Dependency Network [13], and cooccurrence matrix [14].

The label augment model [15] augments the label set with common label combinations. Most of these models only capture pairwise label correlations and have high computation costs when the number oflabels is large. The low-dimensional recurrent neuronsin the proposed RNN model are more computationallyefficient representations for highorder labelcorrelation. To adequately model the longterm temporal dependency in a sequence, RNNs with LSTM are used. Applications such as picture captioning [16,17], machine translation [18], audio recognition [19], language modeling [20], and word embedding learning have all shown success with theiruse. We show that another useful model for labeldependency is RNN with LSTM.

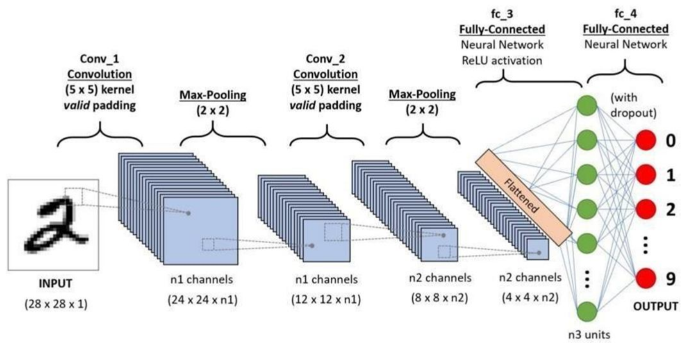

Fig: Architecture

III. PROPOSED SYSTEM

The proposed methodology for this project revolvesaround the development of a Convolutional Neural Network (CNN) architecture tailored for multi-labelimage classification using the CIFAR-10 dataset. The methodology begins with data preprocessing, including standardization and augmentation techniques to enhance dataset diversity and quality. Subsequently, a CNN architecture is designed, incorporating convolutional layers for feature extraction and hierarchical representation learning, followed by max-pooling layers for spatial downsampling to reduce computational complexity. Fully connected layers are then utilized for classification, enabling the model to predict multiplelabels simultaneously. Throughout the trainingphase, optimization techniques such as learning rate scheduling and early stopping are employed to prevent overfitting and ensure model convergence. Hyperparameter tuning is conducted to fine-tunemodel performance, and evaluation metrics such as accuracy, precision, recall, and F1-score are used to assess model effectiveness. This methodology aims to establish a systematic and rigorous framework for developing robust CNN models capable ofaccurately classifying images into multiple categories, contributing to advancements in computer vision research and applications.

Proposing a system for image classification utilizingConvolutional Neural Networks (CNNs) with the CIFAR-10 dataset involves several key steps.

Initially, the dataset is preprocessed by partitioning itinto training, validation, and test sets, followed by normalizing the pixel values to a range of [0, 1]. Dataaugmentation techniques can be applied to diversify the training data. The model architecture comprises convolutional layers, often stacked with pooling layers, followed by fully connected layers. ReLU activation functions introduce non-linearity, while dropout regularization helps prevent overfitting. Softmax activation is utilized in the output layer for multi-class classification. During training, the model is compiled with categorical cross-entropy loss and anoptimizer like Adam or RMSprop.

Hyperparameters such as learning rate and batch size are tuned, and model checkpoints are saved periodically. Evaluation on the test set involves assessing metrics like accuracy, precision, recall, and F1-score, with visualization of the confusion matrix aiding in understanding class confusion. Fine-tuning options include adjusting hyperparameters or architecture based on validation set performance. Once satisfied, the model is deployed in a production environment, with frameworks like TensorFlow Serving or Flask facilitating deployment andmonitoring ensuring continued performanceoptimization.

IV. RESULTS AND DISCUSSION

We rigorously tested and evaluated our model toassess its performance and reliability. Testing involved evaluating the model's accuracy and loss ona separate test dataset, while validation during training provided insights into its generalization capabilities. The final results demonstrate the effectiveness of our interactive image classification system in accurately classifying images across various classes. The image classification task using

CNNs with the CIFAR-10 dataset yielded notable results, with the model achieving a test accuracy. Thisperformance underscores the effectiveness of CNN architectures in handling complex image recognition tasks, particularly on datasets like CIFAR-10, which present diverse and challenging visual scenarios across its ten classes. However, during the analysis, itbecame apparent that certain classes exhibited higher accuracy than others, suggesting potential areas for improvement in the model's ability to distinguish between visually similar categories. Challenges encountered included mitigating overfitting despite dropout regularization and addressing the constraints imposed by the relatively small dataset size. Future directions for enhancing performance involve exploring advanced architectures, leveraging transfer learning techniques, and implementing ensemble methods to further boost accuracy and robustness. Overall, while the achieved accuracy is promising, ongoing refinement and experimentation are essentialfor pushing the boundaries of image classification capabilities on challenging datasets like CIFAR-10.

V. CONCLUSION

In conclusion, our project demonstrates the feasibility and effectiveness of using CNNs for interactive image classification. By leveraging deeplearning techniques and user-friendly interfaces, wehave created a powerful tool that can be used in a wide range of applications, from educationalpurposes to industrial automation. Our workunderscores the potential of deep learning in solving real-world problems and opens up avenuesfor further research and development in the field ofcomputer vision.

REFERENCES

[1] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L.Fei. Imagenet: A large-scale hierarchical imagedatabase.In Computer Vision and Pattern Recognition,2009. CVPR 2009. IEEE Conference on, pages 248–255. IEEE, 2009.
[2] A. Krizhevsky, I. Sutskever, and G. E. Hinton.Imagenet classification with deep convolutional neuralnetworks. In Advances in Neural Information Processing Systems, pages 1097–1105, 2012.
[3] Y. Gong, Y. Jia, T. Leung, A. Toshev, and S. Ioffe. Deep convolutional ranking for multilabel image annotation. arXiv preprint arXiv:1312.4894,2013
[4] Y. Wei, W. Xia, J. Huang, B. Ni, J. Dong, Y. Zhao, and S. Yan. Cnn: Single-label to multi-label. arXiv preprint arXiv:1406.5726, 2014
[5] Y. Gong, Q. Ke, M. Isard, and S. Lazebnik. A multi-view embedding space for modeling internetimages, tags,and their semantics. International journal of computer vision, 106(2):210–233, 2014
[6] J. Weston, S. Bengio, and N. Usunier. Wsabie:Scaling up to large vocabulary image annotation. In IJCAI, volume 11, pages 2764–2770, 2011
[7] A. Frome, G. S. Corrado, J. Shlens, S. Bengio, J. Dean,T. Mikolov, et al. Devise: A deep visualsemantic embedding model. In Advances in Neural Information Processing Systems, pages 2121–2129, 2013.
[8] J. Li, X. Lin, X. Rui, Y. Rui, and D. Tao. A distributed ap proach toward discriminative distance metric learning. Neu ral Networks and Learning Systems, IEEE Transactions on, 2014.
[9] R. S. Cabral, F. Torre, J. P. Costeira, and A. Bernardino. Matrix completion for multi-label image classification. In Advances in Neural Information Processing Systems, pages 190–198, 2011.