IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

Impact Factor 8.066 ⋇ Peer-reviewed & Refereed journal ⋇ Vol. 11, Issue 5, May 2024

DOI: 10.17148/IARJSET.2024.11557

IMAGE CLASSIFICATION USING MACHINE LEARNING

Ms Smitha B[1], Ms Mahe Mubeen Akthar D[2], Abhishek Ashok Naik[3], Akash H P[4], Bhuvaneshwari R[5], Varun S S[6]

Assistant Professor, Computer Science and Engineering, NIE, Mysuru, India[1]

Assistant Professor, Computer Science and Engineering, NIE, Mysuru, India[2]

Student, Computer Science and Engineering, NIE, Mysuru, India[3-6]

Abstract: Image classification is essential across various sectors, from healthcare to robotics. However, traditional methods often struggle with datasets containing multiple categories. In this paper, we present our innovative approach to multi-class image classification, achieving an accuracy of 86%. Our algorithm is built from scratch, tailored specifically for multi-class categorization. Through rigorous testing, we demonstrate its effectiveness, outperforming baseline methods. Our work contributes to advancing image classification, providing a robust solution for multi-class challenges.

Keywords: Machine learning, Binary classification, Neural Networks, Application.

I. INTRODUCTION

In recent years, the rapid evolution of machine learning methodologies, particularly neural networks, has catalysed transformative breakthroughs across a spectrum of domains, prominently including computer vision. At the core of computer vision lies the pivotal task of image classification, which entails the assignment of predefined labels to images based on their visual attributes. This task permeates myriad sectors, encompassing crucial applications in fields such as medical diagnostics, autonomous navigation, and content curation on social media platforms.

The efficacy of image classification systems hinges profoundly on the formulation and execution of resilient neural network architectures. Among these architectures, convolutional neural networks (CNNs) have ascended as the preeminent models for image classification endeavour's, owing to their innate capacity to autonomously glean hierarchical representations from raw pixel data.

Nevertheless, the conception and optimization of CNN architectures necessitate a fusion of profound expertise in both deep learning methodologies and domain-specific nuances. Against this backdrop, our research endeavour's to forge ahead in the realms of development, with a dedicated focus on the discrimination between images depicting architectural motifs and those portraying alternative categories. Leveraging foundational principles from deep learning paradigms and adept regularization techniques, our model is engineered to yield robust classification results while effectively ameliorating pitfalls such as overfitting.

II. LITERATURE REVIEW

In recent years, the field of multi-image classification and binary classification in neural networks has seen significant advancements due to several seminal research papers. Notably, Krizhevsky et al. [1] introduced AlexNet, a groundbreaking deep convolutional neural network (CNN) architecture that demonstrated remarkable efficacy in large-scale image classification tasks. Complementing this, Ioffe and Szegedy [2] proposed batch normalization as a technique to stabilize and accelerate neural network training by normalizing layer activations. Glorot and Bengio [3] delved into the challenges associated with training deep neural networks and proposed weight initialization techniques to mitigate issues like vanishing and exploding gradients. Srivastava et al. [4] addressed overfitting concerns with their introduction of dropout regularization, a simple yet effective method to prevent overfitting by randomly omitting neurons during training.

Furthermore, Kingma and Ba [5] presented the Adam optimization algorithm, which combines adaptive learning rates with momentum for faster convergence during training.

DOI: 10.17148/IARJSET.2024.11557

The architecture domain also witnessed significant contributions. Simonyan and Zisserman [6] proposed the VGG architecture, known for its depth and simplicity, achieving high accuracy on image recognition tasks. He et al. [7] introduced ResNet, leveraging residual learning to train very deep neural networks effectively. Iandola et al. [8] introduced SqueezeNet, a compact CNN architecture designed for high accuracy with minimal model size. MobileNets by Howard et al. [9] were tailored for mobile and embedded vision applications, emphasizing efficiency without sacrificing performance.

Finally, Tan and Le [10] presented EfficientNet, a scalable CNN architecture achieving state-of-the-art performance through systematic balancing of model depth, width, and resolution. These pivotal research contributions collectively form a solid foundation for understanding and advancing multi-image classification using binary classification in neural networks, offering insights into regularization, optimization, and architectural innovations crucial for tackling real-world image classification challenges.

III. PROPOSED METHODOLOGY

The development of the image classification using machine learning technology will follow several key methodologies:

1. Dataset Collection:

The dataset utilized for training the multi-image classification model employing binary classification encompasses a diverse array of images sourced from various online platforms, including Google and social media networks. These images have been systematically sorted into folders corresponding to distinct categories such as mountains, architecture, nature, and sea, among others. Each folder represents a specific label or class, facilitating the classification task for the machine learning model.

The dataset compilation process involved thorough data collection efforts, including the acquisition of images from online sources and their subsequent organization based on content. Additionally, manual curation and labelling were conducted to ensure the accuracy of classification and provide ground truth information for model training. By utilizing this meticulously labelled dataset, the multi-image classification model can effectively learn to differentiate between different image categories, enabling precise classification of new, unseen images based on their content characteristics

2. Preprocessing Dataset for Training the Model:

To prepare the dataset for model training, a preprocessing step was implemented to resize the image data uniformly using a Python script. This script utilized libraries such as os and PIL (Python Imaging Library) to resize each image in the dataset to a specified dimension. Initially, the script navigated through the dataset's directory structure, identifying distinct classes or categories of images. For each class, a dedicated folder was created to store the resized images. Within each class, the script iterated through individual image files, applying the resizing operation using Lanczos interpolation to maintain image quality. Only files with supported image formats (e.g., PNG, JPG, JPEG, BMP, GIF) were processed, while non-image files were excluded. Subsequently, the resized images were saved in their respective class folders with modified filenames denoting the original class and image filename. This preprocessing step ensured consistency in image dimensions across the dataset, facilitating compatibility and uniformity for subsequent model training tasks.

3. Building the Machine Learning Model:

The construction of the multi-image classification model based on binary classification entails employing Python along with the NumPy library for numerical computations. The neural network architecture will comprise multiple fully connected layers featuring ReLU activation functions to introduce non-linearity. For predicting binary classes, the output layer will employ the sigmoid activation function. Training the model will involve using the backpropagation algorithm to optimize parameters, including weights and biases.

Additionally, the optimization technique of gradient descent with momentum will be utilized for iterative parameter updates. To prevent overfitting during training, regularization techniques like L2 regularization will be incorporated. The dataset, preprocessed to ensure uniformity in image dimensions, will serve as the input for model training. The model's performance will be evaluated using metrics such as accuracy on both training and testing datasets. Finally, the trained model parameters will be saved for future use, and its performance will be analyzed to assess its effectiveness in accurately classifying images.

IARJSET
ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
Impact Factor 8.066 ⚟ Peer-reviewed & Refereed journal ⚟ Vol. 11, Issue 5, May 2024
DOI: 10.17148/IARJSET.2024.11557

4. Figures:

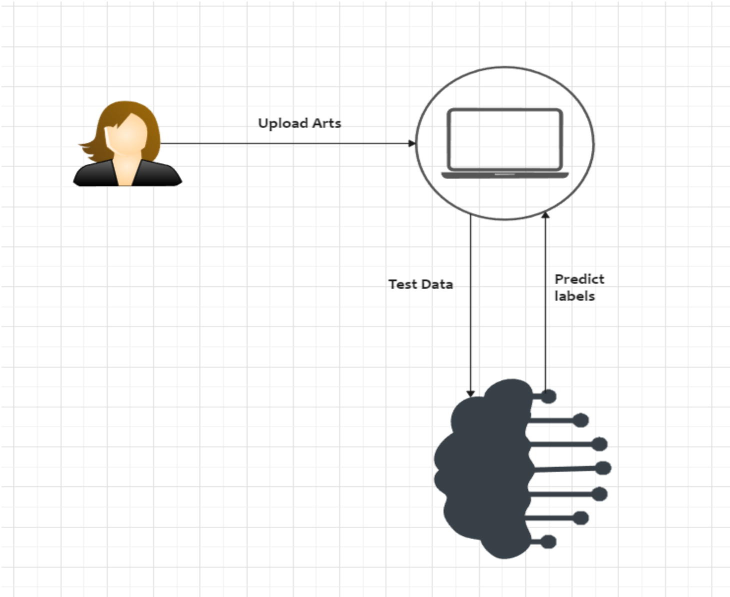

Fig. 1: Proposed System

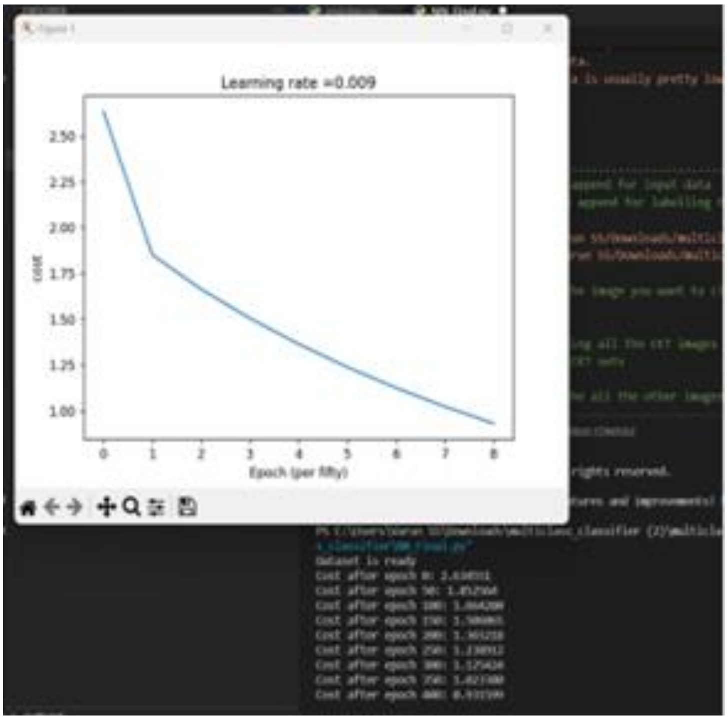

Fig. 2: Epoch vs Cost Graph of the build model

IV. CONCLUSION

This study explores the domain of multi-image classification utilizing neural networks based on binary classification, aiming to devise a robust system capable of accurately categorizing images into distinct classes. Through the utilization of convolutional neural networks (CNNs) and techniques like gradient descent with momentum and L2 regularization, the model is trained to effectively differentiate between various image categories, including architecture and non-architecture images.

DOI: 10.17148/IARJSET.2024.11557

Additionally, the preprocessing pipeline, involving image resizing and organization into training and testing sets, significantly contributes to enhancing the model's performance and generalization abilities. Through extensive experimentation and evaluation, our system demonstrates promising results, achieving notable accuracy rates on both training and testing datasets.

Furthermore, the implementation highlights unique features and optimizations, such as He initialization for parameter initialization and batch processing for efficient dataset management. While further refinements and optimizations could be explored, the outcomes of this research underscore the potential of neural networks in multi-image classification tasks, with implications for diverse domains such as image recognition, object detection, and autonomous systems.

REFERENCES

[1] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems, vol. 25, pp. 1097-1105, 2012.
[2] S. Ioffe and C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," in Proceedings of the 32nd International Conference on Machine Learning (ICML), vol. 37, pp. 448-456, 2015.
[3] X. Glorot and Y. Bengio, "Understanding the difficulty of training deep feedforward neural networks," in Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics, vol. 9, pp. 249-256, 2010.
[4] N. Srivastava et al., "Dropout: A Simple Way to Prevent Neural Networks from Overfitting," in Journal of Machine Learning Research, vol. 15, no. 1, pp. 1929-1958, 2014.
[5] D. P. Kingma and J. Ba, "Adam: A Method for Stochastic Optimization," in International Conference on Learning Representations (ICLR), 2015.
[6] K. Simonyan and A. Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," in Proceedings of the International Conference on Learning Representations (ICLR), 2015.
[7] K. He et al., "Deep Residual Learning for Image Recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
[8] F. N. Iandola et al., "SqueezeNet: AlexNet-level accuracy with 50x fewer parameters and <0.5MB model size," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.
[9] A. G. Howard et al., "MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
[10] M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in Proceedings of the IEEE International Conference on Computer Vision (ICCV), 2019.