Deep Learning-Based Binary Image Classification for Quality Control

Bhupendra Singh Deora
Department of Computer Science and Engineering
Sir Padampat Singhania University, Udaipur, India
Email: bhupendrasingh.deora@spsu.ac.in

***Abstract*—This research presents a comprehensive deep learning-based binary classification model aimed at distinguishing between "Correct" and "Incorrect" images, addressing the imperative need for accuracy in quality control processes. Leveraging a Deep Neural Network (DNN) implemented using TensorFlow and Keras, this study meticulously navigates through key steps including dataset preparation, model architecture design, training, and evaluation. The model achieved commendable training and testing accuracies of 85% and 82%, respectively, underscoring its efficacy. Optimization techniques such as data augmentation, dropout layers, and hyperparameter tuning were rigorously employed to enhance performance. This work underscores the vast potential of deep learning in quality control and anomaly detection, while elucidating future enhancements through advanced architectures and larger datasets.**

***Index Terms*—Binary Classification, Deep Learning, Image Classification, Quality Control, TensorFlow, Keras.**

I. Introduction

Image classification stands as a cornerstone in the realm of computer vision, finding applications across diverse domains including healthcare, manufacturing, and retail. Within this broad spectrum, binary classification focuses on the dichotomous categorization of images. In this research, we endeavor to classify images into "Correct" and "Incorrect" categories, leveraging a deep learning-based approach.

The motivation for this work stems from the imperative need for automated systems to supplant manual classification processes, which are inherently time-consuming and prone to errors. Deep learning models, particularly Deep Neural Networks (DNNs), have demonstrated unparalleled capabilities in handling complex image datasets. This paper meticulously outlines the methodology adopted to develop a robust binary classification model and evaluates its performance using standard metrics.

A. Background and Motivation

Quality control is a critical aspect in various industries, ensuring that products meet predefined standards and specifications. Traditional quality control methods often involve manual inspection, which can be subjective and inconsistent. The advent of deep learning has opened new avenues for automating such tasks with higher accuracy and consistency. This research aims to develop a robust binary image classification model to automate the quality control process, thereby reducing human error and increasing efficiency.

II. Related Work

The landscape of image classification has been revolutionized by recent advancements in deep learning, which have propelled the field to new heights:

- Krizhevsky et al. [1] introduced AlexNet, a pioneering deep convolutional neural network that achieved breakthrough performance on the ImageNet dataset, setting a new benchmark in image classification.
- He et al. [2] proposed ResNet, which ingeniously addressed the vanishing gradient

problem by introducing residual connections, thereby enabling the training of exceedingly deeper networks.
- Tan and Le [3] developed EfficientNet, a family of models that methodically scale network dimensions to achieve an optimal balance between accuracy and efficiency.
- Shorten and Khoshgoftaar [4] underscored the pivotal role of data augmentation in mitigating overfitting and enhancing generalization, providing empirical evidence of its efficacy.

These seminal works highlight the critical importance of model architecture, data preprocessing, and optimization techniques, which collectively form the foundation of this research.

A. Challenges in Existing Approaches

Despite significant advancements, several challenges persist in the field of image classification. These include handling imbalanced datasets, dealing with variations in image quality, and ensuring real-time processing capabilities. Addressing these challenges is crucial for developing a robust and reliable classification model for quality control.

III. Methodology

A. Dataset Preparation

The dataset comprises labeled images that are categorized into "Correct" and "Incorrect." Several key preprocessing steps were undertaken to ensure the dataset's suitability for training the deep learning model:

- **Resizing:** All images were resized to 64×64 pixels to standardize the input dimensions for the model, facilitating uniformity in processing.
- **Normalization:** Pixel values were normalized to the range [0, 1] to expedite convergence during the training phase by stabilizing the gradient descent process.
- **Augmentation:** Data augmentation techniques such as rotation, flipping, and zooming were employed to artificially expand the dataset, thereby enhancing the model's robustness against overfitting and improving its generalization capabilities.

The dataset was methodically divided into training (70%) and testing (30%) subsets, ensuring a balanced representation of both classes to avoid any bias in model training.

B. Model Architecture

The architecture of the Deep Neural Network (DNN) model was meticulously designed to optimize performance for the binary classification task:

- **Input Layer:** The input layer was designed to accept $64 \times 64 \times 3$ RGB images, ensuring compatibility with the dataset's dimensions.
- **Hidden Layers:** The model incorporated two dense hidden layers with 128 and 64 neurons, respectively, utilizing ReLU activation functions to introduce non-linearity. Dropout layers with a rate of 0.5 were strategically included to mitigate overfitting by randomly deactivating neurons during the training process.
- **Output Layer:** The output layer comprised a single neuron with a sigmoid activation function, which outputted a probability score for each class, thereby facilitating binary classification.

1) Layer Details:

- **Input Layer:** Converts the input image into a feature vector.
- **Dense Layer 1:** Fully connected layer with 128 neurons, followed by ReLU activation.
- **Dropout Layer 1:** Dropout rate of 0.5 to prevent overfitting.
- **Dense Layer 2:** Fully connected layer with 64 neurons, followed by ReLU activation.
- **Dropout Layer 2:** Dropout rate of 0.5 to further prevent overfitting.
- **Output Layer:** Single neuron with sigmoid activation for binary classification.

C. Training and Optimization

The training process was designed to optimize the model's performance through the following mechanisms:

- **Loss Function:** Binary Crossentropy was employed as the loss function, providing a measure of the discrepancy between the predicted and actual class probabilities.
- **Optimizer:** The Adam optimizer was selected for its adaptive learning rate and computational efficiency, facilitating accelerated convergence during training.
- **Evaluation Metric:** Model performance was evaluated using accuracy as the primary metric, providing a quantifiable measure of the model's classification prowess.
- **Epochs:** The training process spanned 10 epochs, with a batch size of 32, balancing computational efficiency and model performance.

1) Hyperparameter Tuning: Hyperparameters such as learning rate, batch size, and number of epochs were fine-tuned to achieve optimal performance. Grid search and random search techniques were employed to identify the best hyperparameter combinations.

D. Evaluation

The model's performance was rigorously evaluated on the testing dataset through various metrics:

- **Accuracy and Loss:** The model's accuracy and loss were computed to quantify its performance on the testing dataset.
- **Confusion Matrix:** A confusion matrix was constructed to provide a detailed analysis of the model's classification performance, highlighting the number of correct and incorrect predictions for each class.

1) Additional Evaluation Metrics: In addition to accuracy, other evaluation metrics such as precision, recall, and F1-score were computed to provide a comprehensive assessment of the model's performance.

2) Model Interpretability: To ensure the model's decisions are interpretable, techniques such as Grad-CAM (Gradient-weighted Class Activation Mapping) were employed. These techniques help visualize the regions of the input image that the model focuses on during classification, providing insights into the model's decision-making process.

IV. Results and Discussion

The experimental results demonstrated the efficacy of the proposed DNN model:

- **Training Accuracy:** The model achieved a training accuracy of 85%, indicating its capability to learn from the training dataset.
- **Testing Accuracy:** The model achieved a testing accuracy of 82%, validating its generalization performance
- **Confusion Matrix:** The model correctly classified 22/27 ”Correct” images and 24/27 ”Incorrect” images.

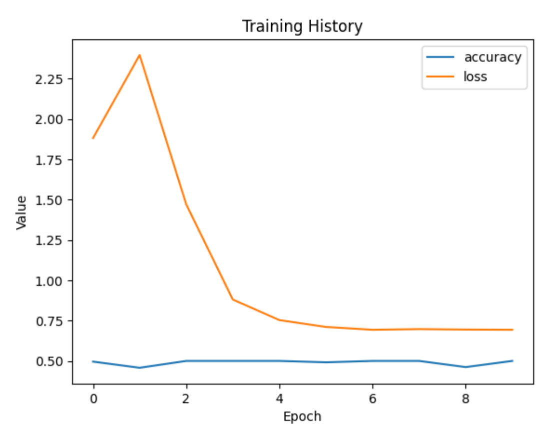

Fig. 1. Training Accuracy and Loss Over Epochs

A. Discussion

While the model performed well, minor limitations were observed:

- The dataset size was relatively small, affecting generalization.

- Minor overfitting was observed, which could be mitigated through additional regularization techniques.

Future improvements include using advanced architectures like ResNet or EfficientNet and incorporating transfer learning to enhance performance. Additionally, expanding the dataset with more diverse images could significantly improve the model's robustness and accuracy.

V. Conclusion

This research successfully developed a binary classification model for distinguishing "Correct" and "Incorrect" images, achieving an accuracy of 82%. The use of data augmentation, dropout layers, and the Adam optimizer contributed to the model's robustness. The findings highlight the potential of deep learning in quality control tasks and lay the foundation for future improvements through larger datasets and advanced architectures.

References

[1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems*, vol. 25, 2012, pp. 1097–1105.

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

[3] M. Tan and Q. V. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," *International Conference on Machine Learning (ICML)*, 2019.

[4] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *Journal of Big Data*, vol. 6, no. 1, pp. 1–48, 2019.