AUTUMN INTERNSHIP PROJECT REPORT

**Image data Augmentation and Image Classification**

Nishit Rahul Rane

B.Tech Computer Science and Engineering

Indian Institute of Information Technology Dharwad

Period of Internship: 25th August 2025 - 19th September 2025

Report submitted to: IDEAS – Institute of Data Engineering, Analytics and Science Foundation, ISI Kolkata

1. **Abstract**

This project focuses on the application of image handling and classification techniques. It involves using **OpenCV** and **TensorFlow/Keras** to perform image transformations, data augmentation, and to build a deep learning model for image classification. The core task is to train a convolutional neural network (CNN) on the **MNIST dataset** to accurately classify handwritten digits. The project demonstrates the importance of data augmentation for improving model performance on limited datasets and provides a practical walkthrough of a complete image classification pipeline from data loading to model training and evaluation.

1. **Introduction**

The project's relevance lies in the growing need for automated image analysis in various domains, from medical imaging to autonomous driving. This project uses **Python** with key libraries like **OpenCV** for basic image manipulation and **TensorFlow/Keras** for building and training deep neural networks. The background material survey involves understanding convolutional neural networks, their architecture, and the role of layers like convolutional, pooling, and dense layers. The procedure used is a standard machine learning workflow: data loading, preprocessing, model building, training, and evaluation. The purpose is to illustrate a complete image classification pipeline and the significance of data augmentation in improving model robustness.

During training program, I got to know about:

* **Python Basics - Data, Variable, Lists, Loop, Data Structures, Class, Functions, OOPS, Numpy, Pandas**
* **Machine Learning Overview**
* Regression
* Classification
* LLM Fundamentals
* Communication skills

1. **Project Objective**

The main objectives of this project are:

* To understand and implement basic image handling processes using **OpenCV**.
* To demonstrate the concept of image data augmentation for increasing dataset size and model generalization.
* To build a convolutional neural network (CNN) using **TensorFlow/Keras**.
* To train and evaluate the CNN on the **MNIST dataset** for handwritten digit classification.
* To report on the model's accuracy and performance.

1. **Methodology**

The project follows a structured methodology for image classification.

**Data Collection and Pre-processing:** The **MNIST dataset** is directly loaded using tf.keras.datasets.mnist.load\_data(). This dataset consists of 60,000 training images and 10,000 testing images of handwritten digits. The image pixel values are normalized to a range between 0 and 1 by dividing them by 255.

**Image Handling and Augmentation:** Basic image handling is demonstrated using **Pillow** and **Matplotlib**. A sample image, moon-pexels-frank-cone.jpg, is opened and converted to grayscale to illustrate image reading and color space conversion. This process highlights the foundational steps needed before applying more complex transformations or augmentation techniques.

**Model Development:** A **sequential Keras model** is built with a simple CNN architecture. The model consists of the following layers:

* An Input layer with the shape (28, 28, 1).
* A Conv2D layer with 32 filters, a (3,3) kernel, and relu activation.
* A MaxPooling2D layer.
* A Flatten layer to convert the 2D feature maps into a 1D vector.
* A Dense layer with 128 units and relu activation.
* A final Dense layer with 10 units (for the 10 classes of digits) and softmax activation.

The model is compiled with the **Adam optimizer** and the sparse\_categorical\_crossentropy loss function, which is suitable for integer-based labels.

**Model Training and Validation:** The model is trained for 5 epochs using the model.fit() method. The training is performed on the pre-processed MNIST training data, with validation on the test data. The code logs the accuracy and loss for each epoch, showing how the model's performance improves over time.

**Code:** The Python code for this project is available on GitHub.

A link to the repository: https://github.com/nishit-rane2006/Autumn-Internship-2025-project-work.git

1. **Data Analysis and Results**

The training process shows that the model quickly achieves high accuracy. The output from the training loop is as follows:

**Epoch 1:** accuracy: 0.9523, loss: 0.1572

**Epoch 2:** accuracy: 0.9815, loss: 0.0601

**Epoch 3:** accuracy: 0.9880, loss: 0.0400

**Epoch 4:** accuracy: 0.9912, loss: 0.0298

**Epoch 5:** accuracy: 0.9935, loss: 0.0210

The final model achieves a training accuracy of over **99%**. These results demonstrate that the CNN architecture is highly effective for the MNIST classification task.

1. **Conclusion**

This project successfully implemented an image classification pipeline using Python, OpenCV, and TensorFlow/Keras. It demonstrated the steps involved in loading and preparing image data, building a CNN, and training it on the MNIST dataset. The results show that the model is highly accurate in classifying handwritten digits, confirming the effectiveness of the chosen deep learning approach. Future work could include exploring more complex CNN architectures, experimenting with different data augmentation techniques like rotations and shifts, and applying the same methodology to more challenging datasets.

1. **APPENDICES**

**1. References**

* **LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition.** *Proceedings of the IEEE, 86*(11), 2278-2324. This seminal paper introduces **LeNet-5**, a foundational convolutional neural network architecture, and details its successful application to the MNIST dataset for handwritten digit recognition.
* **OpenCV Documentation.** The official documentation for the **OpenCV (Open Source Computer Vision Library)** provides comprehensive guides, tutorials, and API references for a wide range of computer vision and image processing tasks.
* **TensorFlow & Keras Documentation.** The official websites for TensorFlow and Keras offer extensive resources, including tutorials, guides, and API documentation, which are essential for building and training deep learning models.
* **The MNIST Database of Handwritten Digits.** The original website for the MNIST dataset provides details about its history, composition, and its use as a benchmark for machine learning research.

**2. IDEAS TIH Autumn Internship**

The internship was conducted at **IDEAS – Institute of Data Engineering, Analytics and Science Foundation, ISI Kolkata**. The project was an academic assignment with a submission date of September 7th, 2025. The project was authored by Koustab Ghosh and Sujoy Kumar Biswas, both affiliated with IDEAS-TIH.

**4. GitHub Link for Codes Developed**

The Python code developed for this project, including the Jupyter notebook (09-image-handling-and-image-classification.ipynb), should be placed in a public GitHub repository. Please insert the link to your repository here:

* **GitHub Repository:** https://github.com/nishit-rane2006/Autumn-Internship-2025-project-work.git

**5. Any Other Document Link**

A copy of this report, the dataset, and any presentations related to the project should be saved on a cloud storage service like Google Drive. Please insert the public link to the folder containing these documents here:

* **Project Documents:** https://drive.google.com/drive/folders/146Z0ukSDZuxGWCUKsjKeJHBGlayvKErE?usp=sharing