**Mini Project Report on**



**IMAGE CLASSIFICATION USING DEEP LEARNING**



**Submitted in partial fulfillment of the requirement for the award of the degree of**

**BACHELOR OF TECHNOLOGY**

**IN**

**COMPUTER SCIENCE & ENGINEERING**

**Submitted by:**

**Student Name**  **University Roll No.**

**Anjali 2019532**

***Under the Mentorship of***

**Mr. Ashiwini Kumar**

**Assistant Professor**



**Department of Computer Science and Engineering**

**Graphic Era (Deemed to be University)**

**Dehradun, Uttarakhand**

**January 2024**



**CANDIDATE’S DECLARATION**

I hereby certify that the work which is being presented in the project report entitled **“Image Classification using Deep Learning”** in partial fulfillment of the requirements for the award of the Degree of Bachelor of Technology in Computer Science and Engineeringof the Graphic Era (Deemed to be University), Dehradun shall be carried out by the under the mentorship of **Mr. Ashwini Kumar**, **Assistant Professor**, Department of Computer Science and Engineering, Graphic Era (Deemed to be University), Dehradun.

Name : Anjali University Roll no : 2019532

**Table of Contents**

|  |  |  |
| --- | --- | --- |
| **Chapter No.** | **Description** | **Page No.** |
| Chapter 1 | Introduction | **1** |
| Chapter 2 | Literature Survey | **5** |
| Chapter 3 | Methodology | **7** |
| Chapter 4 | Result and Discussion | **10** |
| Chapter 5 | Conclusion and Future Work | **13** |
|  | References | **15** |

**CHAPTER 1**

**INTRODUCTION**

In the proceeding sections, an overview of the project and the problem statement at hand will be presented. The introduction will provide a general background on the topic and the problem statement will clearly define the objective and scope of the work.

* 1. **What constitutes an Image?**

Images are composed of pixels, serving as the fundamental digital representation. Each pixel encapsulates details such as color, shading, or opacity. These representations are commonly expressed in three primary formats:

* **Grayscale:** Utilizing a single channel to represent shades from black to white.
* **RGB (Red, Green, Blue):** Employing three channels to generate a wide spectrum of colors by blending red, green, and blue.
* **RGBA (Red, Green, Blue, Alpha):** Similar to RGB but includes an alpha channel to manage transparency or opacity.
  1. **Understanding Image Classification**

Image classification is the process of sorting images into predetermined classes or categories. This task primarily focuses on training a model to differentiate and recognize diverse objects, scenes, or unique patterns depicted within images. The fundamental aim is to equip the model with the ability to precisely allocate a particular label or category to each encountered image.

* 1. **What is Deep Learning?**

Deep learning is a subset of machine learning that involves the use of artificial neural networks to learn and make sense of data. It utilizes multiple layers of interconnected nodes, known as neurons, in these neural networks to extract high-level features and patterns from raw input data. These networks are structured in a hierarchical manner, enabling them to automatically learn representations of data through a process of feature extraction and transformation.

* 1. **Overview of MNIST dataset**

The fashion MNIST dataset holds a pivotal role in both machine learning and computer vision. It comprises an extensive repository of 70,000 grayscale images, each portraying clothing items in 10 different classes. These images are uniformly structured as 28x28 pixel squares, ensuring consistent and standardized analysis.

To enable effective model development and assessment, this dataset is segregated into two key subsets:

* **Training Set:** Consisting of 60,000 images, this subset serves as the cornerstone for training machine learning models. Models undergo exposure to this set, learning to discern and classify items by interpreting intricate patterns embedded within the pixel values of these images.
* **Test Set:** With 10,000 distinct images meticulously separated from the training set, this subset operates as an impartial evaluation platform. Models trained on the training set undergo rigorous assessment using this unseen data to gauge their performance, accuracy, and capacity to generalize to novel, previously unseen instances.

The fashion MNIST dataset holds profound significance as a fundamental asset in the realm of image classification. Its simplistic nature, coupled with its accessibility, positions it as a universal benchmark for refining and validating diverse machine learning and deep learning models. Specifically, it serves as an invaluable resource for refining algorithms focused on the intricate task of recognizing and interpreting clothing items—a crucial aspect in the evolution of image classification methodologies.

* 1. **Overview**

This project focuses on implementing an image classification system utilizing deep learning techniques, specifically with Keras—a high-level neural networks API, and the Fashion MNIST dataset. The primary goal is to train a model capable of accurately recognizing pictures of clothing items from grayscale images. Through meticulous data preprocessing, model building, training, evaluation, and potentially deployment, it provides a comprehensive demonstration of leveraging deep learning for image classification tasks.

**Following are the pre-requisites for the project:**

* Proficiency in Python programming.
* Familiarity with Keras, TensorFlow, or other deep learning frameworks.
* Understanding of CNN architecture and its application in image classification.
* Knowledge of data preprocessing techniques for image datasets.

**1.6 Applications of Image Classification**

Image classification find diverse applications in numerous fields. These applications underscore image classification's versatility across agriculture, forestry, retail, astronomy, and environmental conservation. They showcase its crucial role in solving real-world challenges and enhancing decision-making processes in various domains.

* **Agriculture(Crop Disease Detection):** Detecting diseases in crops by analyzing images of leaves or plants. Trained models identify anomalies in crop leaves, enabling early disease detection for timely intervention, supporting farmers in disease prevention and optimizing crop yield.
* **Forestry(Deforestation Monitoring):** Monitoring deforestation and forest health via satellite or drone images. Algorithms analyze images to track deforestation, detect illegal logging, and assess forest health, aiding conservation efforts.
* **Retail(Visual Search & Product Recommendation):** Empowering visual search and product recommendations in e-commerce. Categorizing product images enables systems to suggest visually similar items to users based on preferences, improving the shopping experience.
* **Astronomy(Celestial Object Classification):** Classifying celestial objects captured by telescopes and satellites. Machine learning models categorize stars, galaxies, etc., aiding astronomers in studying the universe and discovering new objects.
* **Environmental Conservation(Wildlife Monitoring):** Supporting wildlife monitoring and conservation efforts. ML algorithms classify species, track movements, and monitor endangered populations from images, assisting in conservation planning.
  1. **Problem Statement**

This project revolves around the core objective of categorizing grayscale images depicting clothing items into 10 distinct classes using the Fashion MNIST dataset. The primary focus is on achieving precise classification by accurately assigning each image to its corresponding category.

**In the context of this project:**

* The main task entails categorizing clothing items into their respective numeric classes.
* The goal is to architect and train a Keras-based Convolutional Neural Network (CNN) model specifically tailored to recognize and classify these images portraying items.
* The learning objective of the model encompasses the discernment and identification of varied patterns and inherent features within the images. This enables the model to effectively associate these distinctive features with their corresponding labels, ensuring accurate classification.

**CHAPTER 2**

**LITERATURE SURVEY**

Image recognition has witnessed transformative advancements, primarily driven by pioneering research and influential publications. The following survey highlights seminal papers and influential books that have significantly contributed to the evolution of image recognition techniques:

1. **"ImageNet Classification with Deep Convolutional Neural Networks" by Alex Krizhevsky et al.**

Published in 2012, this groundbreaking paper introduced AlexNet, a deep convolutional neural network (CNN) architecture. AlexNet's utilization of multiple convolutional layers reshaped image classification by achieving unprecedented performance on the ImageNet dataset, surpassing traditional methodologies.

1. **"You Only Look Once: Unified, Real-Time Object Detection" by Joseph Redmon et al.**

YOLO (You Only Look Once) pioneered real-time object detection by introducing an efficient system that treats object detection as a regression problem. While its primary focus lies in object detection, YOLO's innovative methodology has exerted a considerable influence on various approaches within the realm of image recognition tasks.

1. **"Digital Image Processing" by Rafael C. Gonzalez and Richard E. Woods**

This foundational book extensively delves into digital image processing techniques. Covering a wide spectrum of methods and algorithms pertinent to image recognition, the book offers comprehensive insights into various methodologies, including feature extraction and pattern recognition techniques essential for comprehending the fundamental principles of image processing.

1. **The Faster R-CNN paper by Shaoqing Ren et al.**

This predominantly revolves around object detection. It introduced the Faster R-CNN framework, which notably incorporates region proposal networks (RPNs). This integration brought about a substantial transformation in methodologies linked to region-based classification within image recognition, signifying significant progress and advancements in this particular domain.

1. **"Deep Learning," authored by Ian Goodfellow, Yoshua Bengio, and Aaron Courville.**

This book serves as a crucial handbook that simplifies the core concepts of deep learning. It encompasses discussions on neural networks—their structures, techniques to boost efficiency, and their role in identifying objects within images. The book underscores the importance of deep neural networks in augmenting image recognition abilities.

The reviewed literature signifies pivotal contributions to image recognition, introducing transformative architectures, methodologies, and frameworks that have significantly elevated the accuracy, efficiency, and comprehension of image recognition systems leveraging deep learning and other cutting-edge techniques. In the coming section we will discuss the methodology which we have used for our model.

**CHAPTER 3**

**METHODOLOGY**

The research methodology embraced in this study encompasses several stages to formulate and implement an image classification model utilizing the CNN architecture.

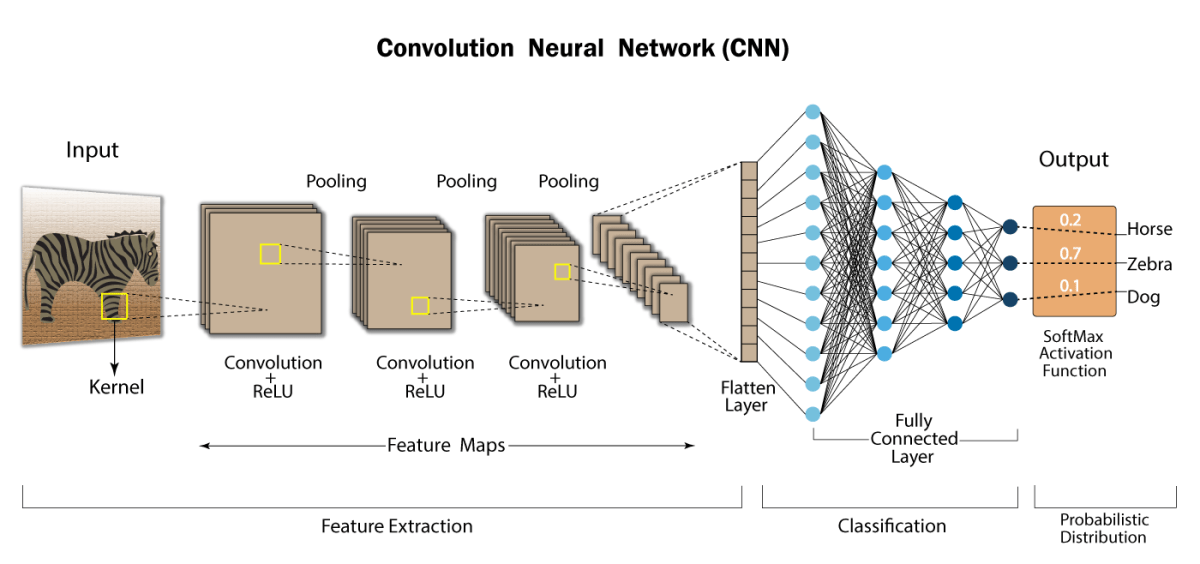
A specialized convolutional neural network (CNN) was trained on the Fashion MNIST dataset to effectively classify 28x28 grayscale images into ten specific categories representing diverse fashion items. These categories include T-shirts, trousers, pullovers, dresses, coats, sandals, shirts, sneakers, bags, and ankle boots. Fashion MNIST, comprising 60,000 training images and 10,000 test images, serves as a significant benchmark within the realm of fashion recognition in machine learning.

The model architecture incorporates CNN layers for feature extraction, max-pooling for down-sampling, and dense layers for classification. Rectified Linear Unit (ReLU) activation functions are applied within the CNN layers to handle pixel values, while the Softmax function performs the final classification.

To combat overfitting, a Dropout layer is utilized, and to normalize the inputs within the network layers and optimize training, a Batch Normalization layer is implemented**.**

Convolutional Neural Networks are essential for image recognition tasks due to their ability to reduce computational requirements. Instead of handling billions of weights and biases for RGB channels directly, CNNs employ convolution to reduce dimensions, leading to fewer calculations.

In practical terms, Convolutional Neural Networks (CNNs) are composed of layers of artificial neurons responsible for computing weighted sums of inputs and producing activation values. In the context of image recognition, each layer specializes in extracting distinct features. Early layers typically identify fundamental edges, while deeper layers progress to detecting more complex patterns such as corners or combinations. As the network gets deeper, it becomes adept at recognizing increasingly intricate features, including specific objects or fashion items.

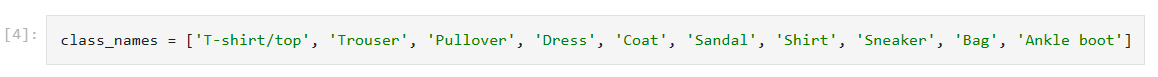


At the final classification layer, the model generates confidence scores (ranging from 0 to 1) for each class, denoting the probability or likelihood of an image belonging to a particular category.

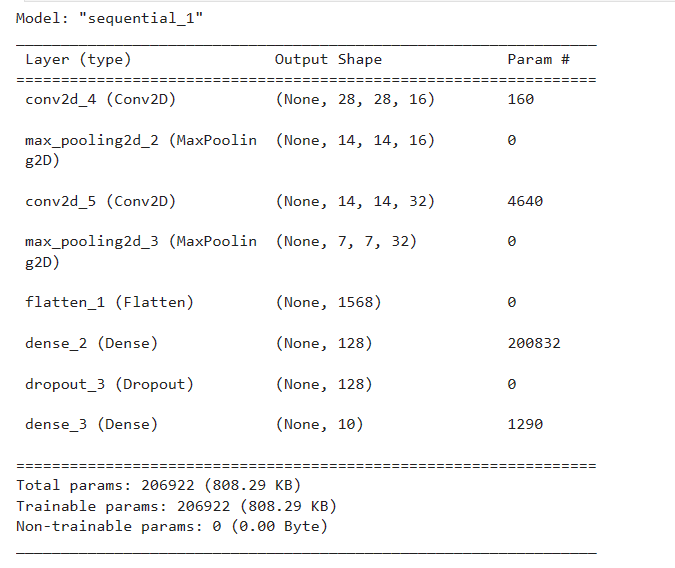
The programming approach for this model includes importing necessary libraries such as pandas, numpy, tkinter, matplotlib, and Keras. Further steps encompass reading the Fashion MNIST dataset, exploratory data analysis, data preprocessing, model construction using Keras, evaluation visualization, result prediction, and deployment.

This model, comprising 206,922 trainable parameters, attains an accuracy of 94.53% during training and a test accuracy of 91.53%. The code for building the model is executed in Google Colab, while the graphical user interface (GUI) is developed using Visual Studio Code, utilizing the trained model's JSON file.

There are in total 10 classes defined in the fashion\_MNIST dataset on which our model is trained for the classification of images.



Following is the model summary for our project:



**CHAPTER 4**

**RESULT AND DISCUSSION**

**4.1 Training data for the model**



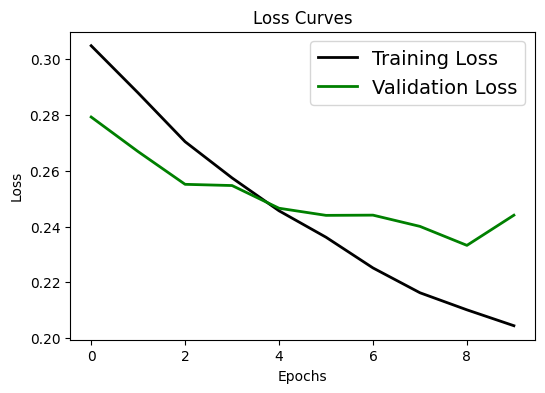
**4.2 Testing data for the model**

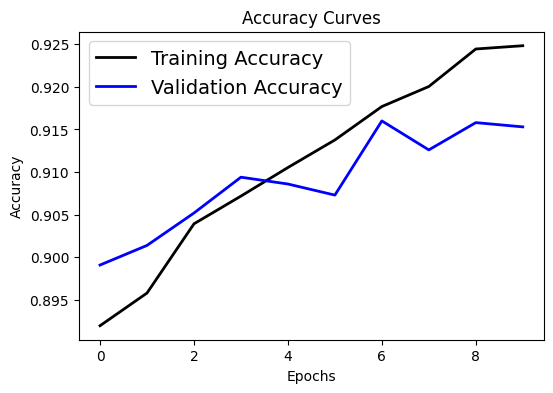


**4.3 Visualization and Evaluation**

We have the following two curves which help in the evaluation of our classification model:

* **Loss Curve** - Comparing the Training Loss with the Testing Loss over increasing epochs.
* **Accuracy Curve** - Comparing the Training Accuracy with the Testing Accuracy over increasing epochs.





* 1. **Model Deployment**
* Load the Trained Model:
  + The trained model is loaded using the model architecture saved in model.json and model weights stored in model.h5 files.
  + The model\_from\_json function loads the model architecture, while load\_weights is used to load the model weights.
* Load Image Button:
  + A "Load Image" button (load\_button) is created to open a file dialog, enabling users to select an image file.
  + The selected image is then displayed within the image\_label.
* Classify Image Button:
  + The "Classify Image" Button triggers the execution of a function named "classify\_image."
  + The "classify\_image" function performs the following steps:

a. Resizes the loaded image to match the model's required input size (28x28).

b. Normalizes the pixel values of the image to a range between 0 and 1.

c. Expands the image dimensions to align with the expected input shape for the model.

d. Makes predictions using the model (model.predict).

e. Retrieves the predicted class index with the highest probability using np.argmax.

f. Displays a message box containing the predicted class name.

**CHAPTER 5**

**CONCLUSION AND FUTURE WORK**

In the realm of image classification, leveraging deep learning models on datasets like Fashion MNIST provides invaluable insights and practical applications across diverse industries. The exploration of convolutional neural networks (CNNs) on image datasets has demonstrated remarkable capabilities in categorizing and recognizing intricate visual patterns. Although our model for the Image Classification is working fine for the given data but there may have situations which when arise will lead to the wrong result production.

Following are some of the limitations for the model:

* **Limited Dataset Diversity:** Fashion MNIST might lack diversity, limiting the model's capability to generalize to real-world scenarios.
* **Resolution Constraints:** Low-resolution images might affect the model's ability to classify intricate details in certain clothing items.
* **Class Imbalance:** Unequal representation of classes can lead to biased predictions and affect model performance.
* **Overfitting Risks:** Models might become overly specialized to the training data, impacting their performance on new, unseen data.
* **Challenges in Fine-Grained Classification:** Distinguishing between visually similar categories (e.g., differentiating between various types of shirts) might pose challenges.
* **Domain Adaptability:** Adapting pre-trained models to different fashion domains or new datasets may require extensive retraining or fine-tuning.

The Fashion MNIST dataset, while beneficial for image classification, poses certain challenges concerning dataset diversity and resolution limitations. However, leveraging this dataset for fashion image classification applications can lead to significant advancements in e-commerce, retail, predictive analytics, and customer experiences. Future work involves addressing these limitations and expanding the model's applications to meet the demands of real-world fashion scenarios.

* **Fashion Recommendation Systems:** Utilizing the trained model to develop sophisticated recommendation systems for fashion e-commerce platforms.
* **Adaptation for Real-World Datasets:** Extending the model's applicability by fine-tuning or retraining it with larger and more diverse fashion datasets from real-world scenarios.
* **Ensemble Models:** Integrating the Fashion MNIST model into ensemble models or cascading it with other models for improved performance.
* **Personalized Styling Apps:** Creating apps that offer personalized styling advice based on the user's wardrobe or style preferences.
* **Retail Inventory Management:** Implementing the model to automate inventory management in retail stores by categorizing and organizing clothing items.
* **Visual Search Technologies**: Incorporating the model into visual search technologies for efficient exploration and discovery of fashion items.

**REFERENCES**

[1] https://www.analyticsvidhya.com/blog/2021/05/convolutional-neural-networks-cnn/#be44

[2] https://www.simplilearn.com/image-processing-article

[3] https://scholar.google.com/

[4]https://ieeexplore.ieee.org/Xplore/home.jsp

[5]https://arxiv.org/

[6] https://www.coursera.org/learn/cifar-10-image-classification/home