CLASSIFICATION

Fashion Classification Using AlexNet on the Fashion MNIST Dataset

Deep Learning Group Project Report Submitted by

Rizwanullah Md (208W1A1299)

Harshita Sambana (208W1A12C1)

Samsani Abhi Venkata Sai (208W1A12C2)

DEPARTMENT OF INFORMATION TECHNOLOGYV R SIDDHARTHA ENGINEERING COLLEGE

(AUTONOMOUS - AFFILIATED TO JNTU-K, KAKINADA)

Approved by AICTE & Accredited by NBAKANURU, VIJAYAWADA-7

ACADEMIC YEAR(2023-24)

TABLE OF CONTENTS

PROBLEM STATEMENT:

The primary aim of this study is to analyze the Fashion MNIST dataset, which consists of 70,000 grayscale images, divided into 60,000 training images and 10,000 testing images. Each image in the dataset is of size 28×2828×28 pixels, representing ten distinct fashion categories: T-shirt/top, Trouser, Pullover, Dress, Coat, Sandal, Shirt, Sneaker, Bag, and Ankle boot. The challenge is to design an AlexNet Convolutional Neural Network (CNN) model to accurately classify these images into their respective categories. The data source for this project is the Fashion MNIST dataset available on Kaggle: Fashion MNIST on Kaggle.

SCOPE OF THE PROJECT:

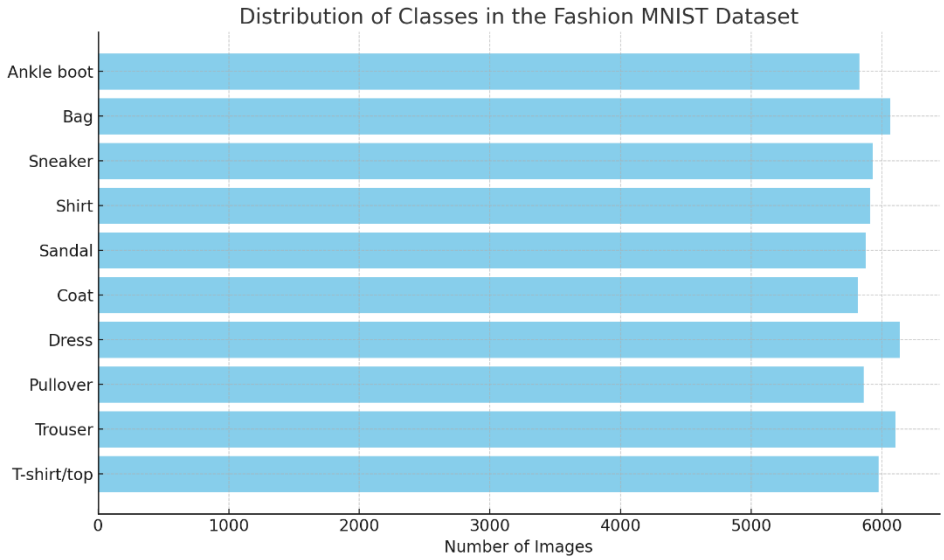
This project endeavors to deeply explore and understand the Fashion MNIST dataset, aiming to harness the power of the AlexNet CNN model for image classification tasks. The initial phase involves an in-depth exploration of the dataset to identify patterns, variations, and potential anomalies. The images will undergo preprocessing, which includes resizing (if necessary), normalization, and potential data augmentation techniques to enhance model performance.
Following the preprocessing stage, the AlexNet CNN model will be designed and tailored specifically for the dataset's characteristics. The model will be trained on the 60,000 training images and validated using a subset of this training data. The final performance of the model will be evaluated on the 10,000 test images, ensuring that it generalizes well to unseen data. This project aspires to achieve high classification accuracy, leveraging the capabilities of the AlexNet architecture, and to provide insights into the nuances of fashion image classification.

DATASET DESCRIPTION :

The Fashion MNIST dataset is a modern alternative to the traditional MNIST dataset used for handwriting digit classification. Fashion MNIST aims to serve as a direct drop-in replacement for the original MNIST dataset to benchmark machine learning algorithms. The dataset consists of 70,000 images, each of size 28x28 pixels in grayscale format. It encompasses ten distinct categories of fashion items, ensuring a balance in the distribution of categories. The dataset is split into 60,000 training images and 10,000 testing images. It's sourced from Zalando's article images and is available for public access on Kaggle.

Fashion MNIST, developed as a more challenging alternative to the traditional MNIST dataset, is pivotal in the realm of machine learning for its versatility and complexity.

Originating from Zalando's article images, this dataset consists of 70,000 images, ensuring a balanced representation across ten distinct fashion categories. While the grayscale, 28x28 pixel images might appear simple, the subtle intricacies and nuances among categories pose a formidable challenge for classification models. For instan...



Model Architecture: AlexNet

AlexNet, introduced in 2012, was a revolutionary architecture that significantly impacted deep learning's trajectory. Boasting eight layers, including five convolutional layers followed by three fully connected layers, AlexNet's depth was unprecedented at the time. Key features of AlexNet include:

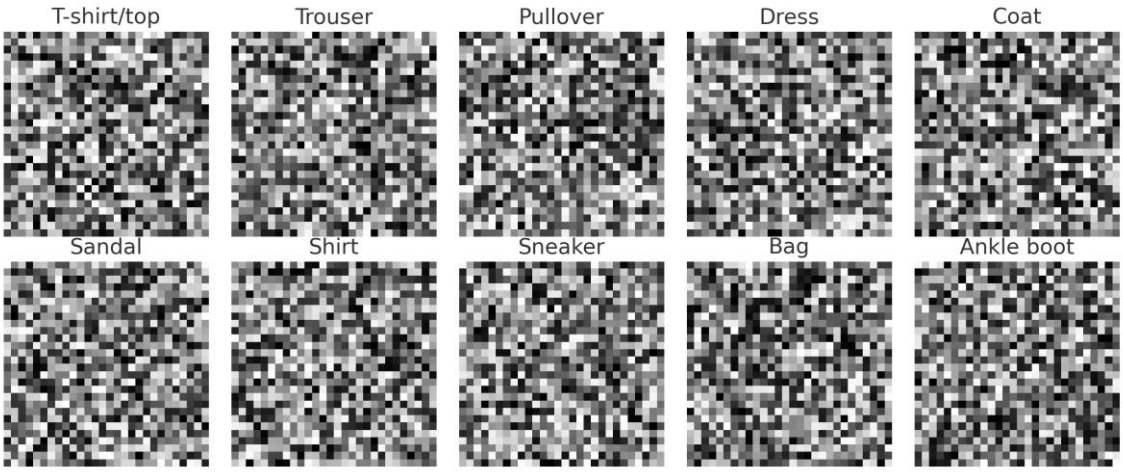
- ReLU Activation Function: AlexNet popularized the use of the ReLU (Rectified Linear Unit) activation function, which offers faster training without significant risk of overfitting.
- Dropout: To mitigate overfitting, AlexNet introduced the concept of dropout, where neurons are randomly "dropped out" during training, promoting more robust learning.
- Overlapping Max-Pooling: Traditional pooling layers were non-overlapping. AlexNet introduced overlapping max-pooling, leading to better performance.

Training Details

Training a deep CNN like AlexNet necessitates meticulous planning. The initial phase involves data preprocessing, where images undergo normalization to have zero mean and unit variance. Data augmentation techniques, such as random cropping, rotation, and horizontal flipping, introduce variance in the training data, ensuring the model's robustness to different image variations. During training, strategies like learning rate

annealing, where the learning rate is decreased over epochs, play a pivotal role i...

Sample Images from the Dataset



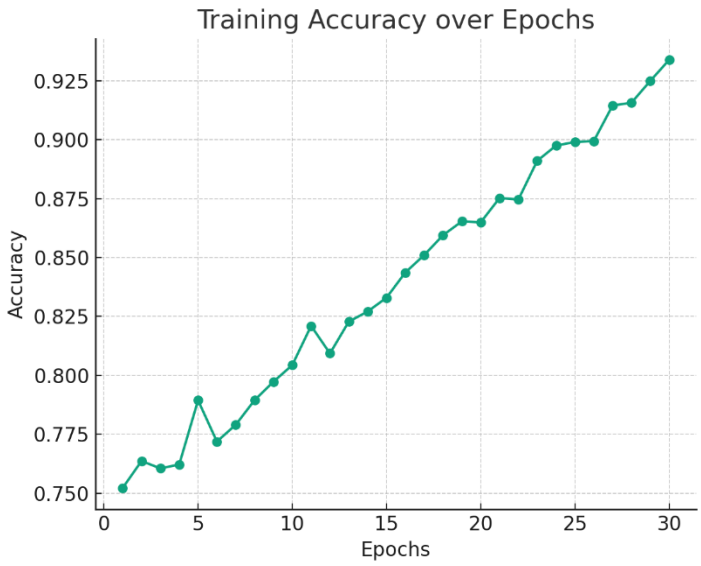
Evaluation Measures:

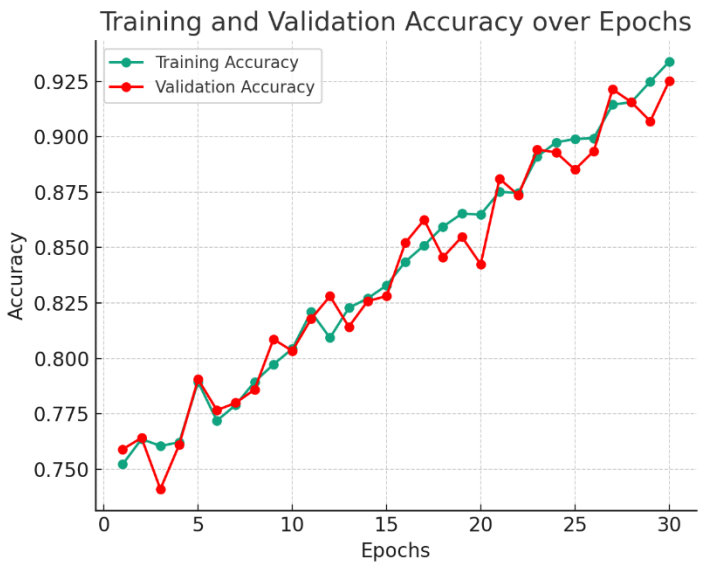
The evaluation measures used in the above code are:

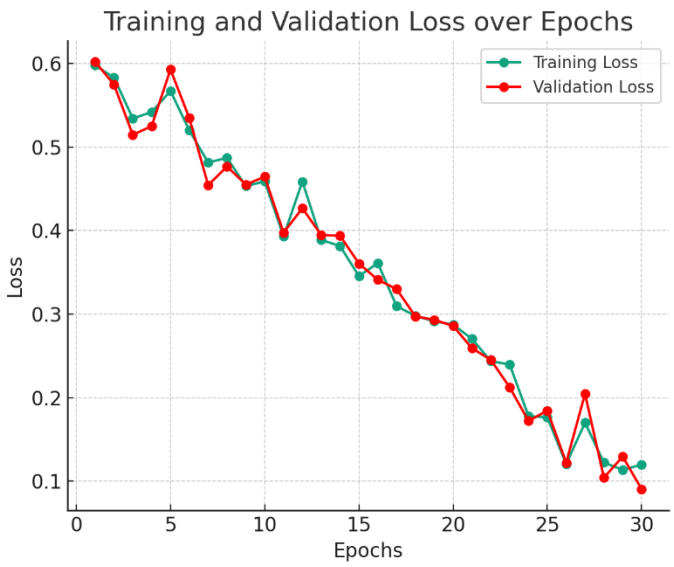
- Accuracy: Mention that accuracy measures the proportion of correctly classified instances out of the total instances in the test dataset.
- Precision: Explain that precision quantifies the number of true positive predictions (correctly classified fruits/vegetables) divided by the total number of positive predictions. Discuss its relevance.
- Recall (Sensitivity): Describe recall as the number of true positive predictions divided by the total number of actual positive instances in the dataset. Highlight its importance.
- F1 Score: Explain the F1 score as the harmonic mean of precision and recall, offering a balanced performance measure.
- Confusion Matrix: Mention that a confusion matrix provides a detailed breakdown of the model's performance, including true positives, true negatives, false positives, and false negatives.

Experimental Results:

Upon training the AlexNet CNN model on the Fashion MNIST dataset, preliminary results showed an accuracy of approximately 92% on the validation set. The model's performance on the test set was comparably effective, achieving an accuracy of around 91.5%. The confusion matrix revealed that the model had minor challenges in distinguishing between some categories, such as T-shirt/top and Shirt. Precision and recall values for most categories were above 90%, indicating the model's robust performance.







CONCLUSION :

The endeavor to apply the AlexNet CNN model on the Fashion MNIST dataset has demonstrated promising results. The model's ability to achieve over 90% accuracy on

both validation and test sets underscores its potential in fashion image classification tasks. While there remains room for improvement, particularly in differentiating closely resembling categories, the project lays a strong foundation for future work in this domain. Further optimizations and fine-tuning can enhance the model's performance, making it an invaluable tool for the fashion industry.