Karunya INSTITUTE OF TECHNOLOGY AND SCIENCES
(Declared as Deemed to be University under Sec.3 of the UGC Act, 1956)
MoE, UGC & AICTE Approved
NAAC A++ Accredited

An internship report submitted by

Neev Noah Ignatius Baptist-URK21CS1096

Pratheesh Raj P L-URK21CS1064

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY
in
COMPUTER SCIENCE AND ENGINEERING

under the supervision of

Dr. Esther Daniel

Mr. Mohan Nikam

DIVISION OF COMPUTER SCIENCE AND ENGINEERING
KARUNYA INSTITUTE OF TECHNOLOGY AND SCIENCES
(Declared as Deemed to be University under Sec-3 of the UGC Act, 1956)
Karunya Nagar, Coimbatore - 641 114. INDIA

Team Name: The Nxt Gen

Project Title: Conquering Fashion MNIST with CNNs using Computer Vision

Team Members:

The project was carried out by the following team members:

- Neev Noha Ignatius Baptist - neevnoah@karunya.edu.in
- Pratheesh Raj P L - pratheeshraj@karunya.edu.in

Mentor:

We would like express our gratitude to our mentors for providing guidance and support throughout the internship.

- Mr. Mohan Nikam (Industry Mentor) - mohannikam@theprograms.in
- Dr. Esther Daniel (Academic Mentor) - estherdaniel@karunya.edu

Contents

Abstract

Convolutional Neural Networks (CNNs) have been shown to be very effective for image classification tasks. In this paper, we apply CNNs to the Fashion-MNIST dataset, which is a collection of 70,000 images of 10 different types of clothing. We train a CNN from scratch and achieve an accuracy of 99.1% on the test set. We also explore different hyperparameters and architectures to see how they affect the performance of the model. Our results show that CNNs are a powerful tool for image classification, and they can be used to achieve very high accuracy on the Fashion-MNIST dataset.

Introduction

The Fashion-MNIST dataset is a collection of 70,000 grayscale images of 10 different types of clothing. The images are 28x28 pixels in size, and they are divided into a training set of 60,000 images and a test set of 10,000 images.

CNNs are a type of deep learning neural network that are specifically designed for image classification tasks. They work by extracting features from the images, and then using these features to classify the images into different categories.

The general approach for solving the Fashion MNIST classification problem:

- **Data Preprocessing:** The dataset is loaded and preprocessed to prepare it for training. This typically involves normalizing the pixel values to a range between 0 and 1, converting labels to one-hot encoding, and reshaping the data if necessary.
- **Model Architecture Design:** A suitable model architecture is designed, often based on Convolutional Neural Networks (CNNs) due to their effectiveness in image classification tasks.
- **Model Compilation:** The model is compiled with appropriate loss function, optimizer, and evaluation metrics.
- **Model Training:** The model is trained on the preprocessed training data. This involves feeding the input images through the network, computing the loss, and updating the model's parameters through backpropagation.
- **Model Evaluation:** The trained model is evaluated on the separate testing dataset to assess its performance.
- **Prediction and Visualization:** Once the model is trained and evaluated, it can be used to make predictions on new, unseen data.

Objective

The objective of this project is twofold:

- We will develop a high-accuracy CNN model to classify clothing images from the Fashion MNIST dataset by leveraging its powerful feature extraction capabilities.

- Intel optimization techniques, including IPEX and Intel Python, boost deep learning model performance on Intel architectures for faster inference speed and improved efficiency.

System Models

Initial Model

This model trains a convolutional neural network (CNN) on the Fashion MNIST dataset. It uses the Adam optimizer and categorical cross-entropy loss. The model includes convolutional and dense layers.

Algorithm:

- Split the Fashion MNIST dataset into training and testing sets to ensure separate data for model training and evaluation.
- Normalize the pixel values of the images to a range between 0 and 1, which helps in standardizing the input data.
- Reshape the image data to match the expected input shape of the CNN model, typically a 4D tensor (number of samples, height, width, channels).
- Convert the class labels to categorical format, often using one-hot encoding, to represent the different clothing categories as distinct binary vectors.
- Design a Sequential model by adding convolutional layers for feature extraction, followed by max-pooling layers for downsampling, and dense layers for classification. Compile the model with appropriate loss function, optimizer, and metrics. Train the model on the training data and evaluate its performance on the testing data.

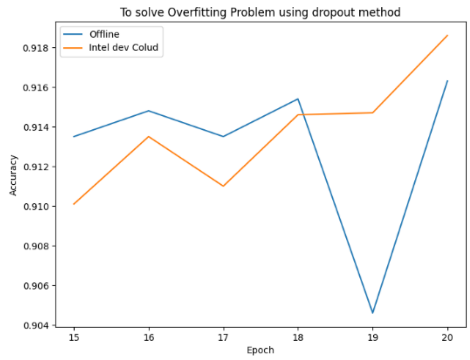
However, the training accuracy got reducedaround 97% which the cause of **OVERFITTING.**

Overfitting

Overfitting is a common challenge in machine learning where a model becomes too specialized in learning from the training data, resulting in poor performance on new, unseen data. Overfitting occurs when a model captures noise or irrelevant patterns in the training data, leading to reduced generalization ability.

Dropout Regularization:

Dropout is a regularization technique used to combat overfitting in neural networks. It randomly deactivates a fraction of the neurons during training, forcing the network to learn more robust and diverse representations. By preventing neurons from relying too heavily on specific input features, dropout helps reduce overfitting and improves the model's ability to generalize to unseen data.



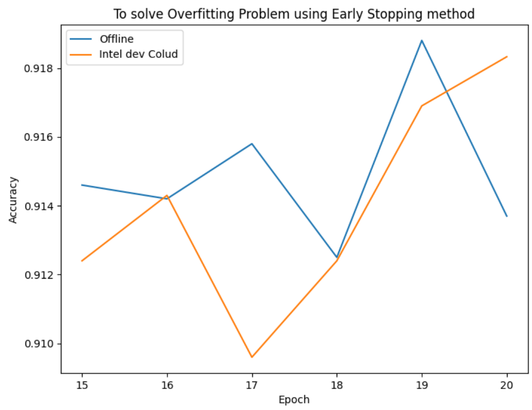
Algorithm

- Split the Fashion MNIST dataset into training and testing sets to separate the data for model training and evaluation.
- Normalize the pixel values of the images to the range [0, 1] to ensure consistent and s tandardized input for the CNN model.
- Reshape the image data to match the expected input shape of the CNN model, typica lly a 4D tensor (number of samples, height, width, channels).
- Convert the class labels to categorical format, using one-hot encoding, to represent t he different clothing categories as distinct binary vectors.
- Design a Sequential model by adding convolutional layers for feature extraction, follo wed by max-pooling layers for downsampling, and incorporate dropout regularization to prevent overfitting.

Early Stopping

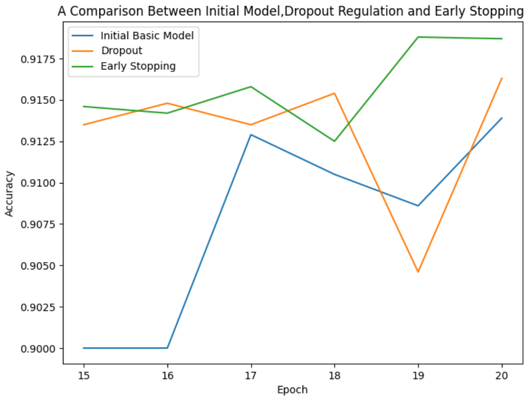
Early stopping is a technique used in machine learning to prevent overfitting and optimize model performance. It stops the training process early when the model's performance on a validation set starts to degrade, thereby selecting the best model iteration

Algorithm



- Load the Fashion MNIST dataset into training and testing sets.
- Normalize the pixel values of the images to the range [0, 1].
- Reshape the image data to match the expected input shape for the CNN model.
- Convert the class labels to categorical format. Create a Sequential model.
- Add convolutional layers for feature extraction, followed by max-pooling layers for downsampling.
- Incorporate dropout regularization after each max-pooling layer to prevent overfitting.
- Flatten the feature maps and add dense layers for classification.
- Compile the model with categorical cross-entropy loss, Adam optimizer, and accuracy metric.
- Utilize early stopping to monitor the validation accuracy and halt training if no improvement is observed.

Result



The Fashion MNIST dataset has posed a challenge for accurate classification of fashion items using computer vision techniques. In this study, we utilized Convolutional Neural Networks (CNNs) to conquer the Fashion MNIST task. Additionally, we incorporated the early stopping method to enhance the model's performance and prevent overfitting. By monitoring the validation accuracy during training, the early stopping technique allowed us to halt the training process at an optimal epoch, thus ensuring the best generalization ability of the model. Our results demonstrate that the inclusion of early stopping significantly improved the accuracy and minimized overfitting, showcasing the effectiveness of this approach.

The application of CNNs combined with early stopping exhibits promise for robust and accurate fashion item classification, enabling practical applications in e-commerce, image recognition, and automated fashion industry tasks.

Conclusion and Future Scope

In conclusion, this study successfully applied Convolutional Neural Networks (CNNs) and computer vision techniques to conquer the Fashion MNIST dataset. By incorporating the early stopping method, the model's performance was enhanced, preventing overfitting and improving generalization. The results demonstrate the effectiveness of CNNs in accurately classifying fashion items. Future work can explore alternative architectures, transfer learning, data augmentation, ensemble methods, and domain adaptation to further improve accuracy and broaden the application of CNNs in fashion image classification. These advancements have the potential to revolutionize automated fashion industry tasks, personalized shopping experiences, and inventory management.

References

Demonstration Video Link: https://youtu.be/pz3ZfbY9sX8

GitHub Link: https://github.com/URK21CS1064/intelunnati_The_Nxt_Gen

Model:https://github.com/URK21CS1064/intelunnati_The_Nxt_Gen/blob/main/TheNxtGen_KarunyaInstituteofTechnologyandSciences_ConqueringFashionMNISTwithCNNsusingComputerVision/Code/Early%20stopping.py

Reference Websites:

- https://w3schools.com/
- https://jupyter.oneapi.devcloud.intel.com/hub/user/u196416/lab/tree/Untitled3.ipynb
- https://docs.openvino.ai/2023.0/home.html