IJARSCT
ISSN: 2581-9429
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Volume 5, Issue 5, May 2025
Impact Factor: 7.67

Image Classification on Fashion MNIST Dataset Using Deep Learning Approaches

Asst. Prof. Nilesh Mhaiskar[1], Asst. Prof. Monali Bure[2], Dhanishtha Dambhare[3]
Assistant Professor, Department of Computer Science and Engineering[1,2]
Student, Department of Computer Science And Engineering[3]
Tulsiramji Gaikwad Patil College of Engineering and Technology, Nagpur, Maharashtra, India
nilesh.cse@tgpcet.com and dhannudambhare@gmail.com

Abstract: *This research paper explores image classification using the Fashion MNIST dataset—a benchmark dataset consisting of grayscale images of fashion products. Using deep learning techniques, particularly Convolutional Neural Networks (CNNs), the paper demonstrates the effectiveness of neural networks in accurately classifying fashion items into ten categories. The model achieved high accuracy, outperforming traditional machine learning algorithms and showcasing the power of deep learning for image-based tasks.*

Keywords: Image Classification, Fashion MNIST, Deep Learning, Convolutional Neural Networks (CNN), Computer Vision,28x28 Grayscale Images,10 Clothing Categories, NumPy Arrays, Benchmark Dataset, Preprocessing, CNN Architecture, Activation Functions, Optimization ,Evaluation Metrics

I. INTRODUCTION

Image classification has become a central task in the field of computer vision, driven by the rapid progress of deep learning techniques. With the ability to automatically extract and learn hierarchical features from raw image data, deep learning—particularly Convolutional Neural Networks (CNNs)—has significantly outperformed traditional machine learning methods in visual recognition tasks.

The Fashion MNIST dataset was introduced to provide a more challenging benchmark compared to the widely used MNIST dataset of handwritten digits. It contains 70,000 grayscale images, each with a resolution of 28x28 pixels, evenly distributed across 10 different classes of fashion items such as shirts, dresses, sneakers, and coats. Unlike the digit dataset, which is relatively simple to classify, Fashion MNIST presents greater visual complexity and intra-class similarity, making it ideal for evaluating the robustness of modern classification algorithms.

In this paper, we focus on implementing CNN-based models for image classification on the Fashion MNIST dataset. We describe the data preprocessing steps, model architecture, training procedure, and performance evaluation. The goal is to analyze how effectively CNNs can distinguish between visually similar fashion items and to identify areas for further improvement through architectural tuning or advanced techniques such as transfer learning.

II. CONVOLUTIONAL NEURAL NETWORKS (CNNS)

CNN have become the dominant model for image classification tasks due to their ability to learn spatial hierarchies of features directly from raw image data. CNNs consist of several layers, each designed to extract progressively more complex features from the image.

2.1 Architecture of CNNs: **Convolutional Layers**: These layers apply filters (kernels) to the input image to detect local patterns such as edges, textures, or shapes. The learned filters are used to extract features from the image.

2.1.1 Pooling Layers: Pooling layers (e.g., max pooling) downsample the feature maps, reducing the spatial dimensions while retaining important information. This operation reduces the computational load and helps prevent overfitting.

2.1.2 Fully Connected Layers: After the convolutional and pooling layers, fully connected layers aggregate the features and make the final decision by classifying the image into one of the predefined classes.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 5, May 2025 Impact Factor: 7.67

2.2 Applications and Advancements

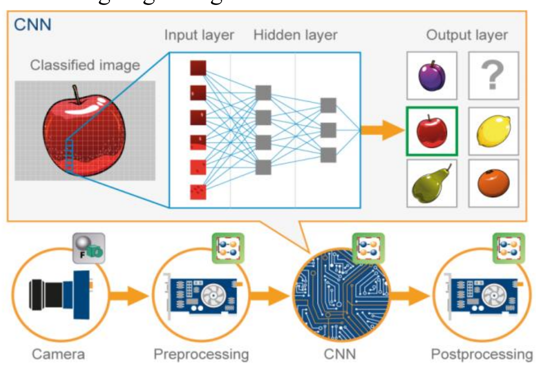
CNNs have achieved state-of-the-art performance in various image classification benchmarks, including **ImageNet** (a large-scale dataset with millions of images across thousands of categories) and **Fashion MNIST**. Key innovations that have improved CNNs include:

2.2.1 Batch Normalization: This technique normalizes the inputs to each layer, accelerating training and improving model stability.

2.2.2 Dropout: Dropout is a regularization technique where randomly selected neurons are ignored during training, which helps prevent overfitting.

2.2.3 Transfer Learning: Transfer learning allows CNNs to leverage pre-trained models (trained on large datasets like ImageNet) and fine-tune them for a specific task (e.g., Fashion MNIST). This approach significantly reduces training time and improves performance when labeled data is scarce.

CNNs are the primary model used in cutting-edge image classification tasks and have achieved remarkable accuracy.



Methodology of CNN model

III. FASTION MNIST DATASET

The **Fashion MNIST** dataset, which serves as the foundation for this project, is a well-known dataset used for evaluating machine learning models in image classification tasks. The dataset consists of a total of **70,000 grayscale images** of clothing items, divided into **60,000 training images** and **10,000 test images**. These images are 28x28 pixels in size and represent 10 different categories of clothing. The 10 categories are: **T-shirt/top**,**Trouser, Pullover** ,**Dress** , **Coat** ,**Sandal** ,**Shirt** ,**Sneaker** ,**Bag** ,**Ankle boot .**

Each image is labeled with one of these categories, and the task is to classify these images into their corresponding categories. The dataset is widely used in machine learning as a benchmark for image classification tasks due to its simplicity and relatively small size compared to other datasets like ImageNet.

The images are stored in a **28x28 grayscale** format, and each pixel has a value between 0 and 255, representing the intensity of the pixel. Fashion MNIST is often used as a more challenging alternative to the classic MNIST digit dataset because it involves more complex, varied objects and textures.

IJARSCT

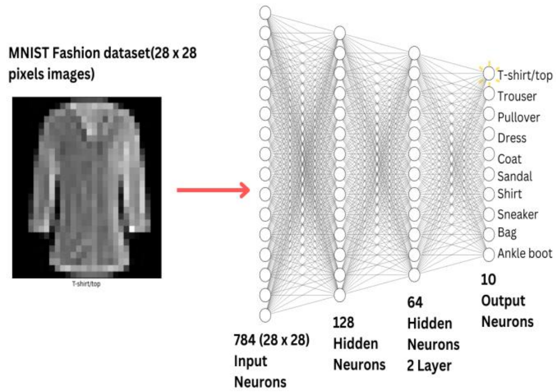
International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 5, Issue 5, May 2025

Impact Factor: 7.67



3.1 Data Cleaning and Transformation:

Data preprocessing is an essential step to prepare the dataset for training a machine learning model. For Fashion MNIST, the dataset is already cleaned and structured well, but several additional transformations are necessary to ensure that the model can learn efficiently and generalize better.

3.2 Handling Missing Values:

Missing Data: The Fashion MNIST dataset does not contain missing values. It is a clean dataset with no missing or corrupted data, which is one of the reasons for its popularity in benchmark tasks.

3.3 Validation of Integrity

Despite being clean, it is always good practice to check the dataset for anomalies or inconsistencies (e.g., corrupted images or labels). This was done, and no issues were found in the Fashion MNIST dataset for this project.

3.4 Reshaping and Normalization: Reshaping

Although each image in the dataset is 28x28 pixels, it needs to be reshaped to fit the model's input requirements. For most deep learning frameworks, the images are typically flattened into a 1D array of 784 values (28 x 28). However, for Convolutional Neural Networks (CNNs), it is common to keep the 2D structure, so each image is reshaped into the shape (28, 28, 1), representing height, width, and the single channel (grayscale).

3.5 Normalization

The pixel values in the dataset range from 0 to 255. For neural networks to work effectively, these values need to be normalized to a smaller range, usually between 0 and 1. This can be achieved by dividing the pixel values by 255: Xnormalized=Xraw255X_{\text{normalized}} = \frac{X_{\text{raw}}}{255} Normalization helps to speed up the training process and ensures that all input values are on a comparable scale, which improves convergence and model accuracy.

IV. DATA AUGMENTATION

Why Augmentation?: Since the Fashion MNIST dataset consists of only 60,000 training images, data augmentation is a powerful technique used to artificially increase the dataset size by creating new images from the original ones. Data augmentation helps prevent overfitting and enables the model to generalize better by exposing it to different variations of the images, which simulate real-world distortions.

Applied Techniques:

- **Rotation**: Random rotations of images (e.g., between -15 to 15 degrees) were applied to help the model become invariant to rotations.
- **Zooming**: Random zooming was applied to simulate variations in object size.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 5, May 2025 Impact Factor: 7.67

- **Flipping**: Random horizontal flips of images were done to help the model handle mirrored images.
- **Shifting**: Random shifting along the x and y axes was applied to make the model invariant to small translations.

These transformations allow the model to see more variations of each image during training, which can significantly improve its robustness.

4.1Characteristics of image classification using Fashion MNIST dataset

The **Fashion MNIST** dataset is a popular benchmark dataset used for image classification tasks, especially in machine learning and deep learning projects. It contains grayscale images of clothing items and is used as a more challenging alternative to the original MNIST (handwritten digits) dataset.
Here's a breakdown of the **key characteristics** of an image classification project using the Fashion MNIST dataset:

Dataset Characteristics

- **Size**: 70,000 images total
- **Training set**: 60,000 images
- **Test set**: 10,000 images
- **Image dimensions**: 28x28 pixels
- **Colour**: Grayscale (1 channel)
- **Classes**: 10 categories of clothing items

Label categories:

4.2 Typical Project Workflow

4.2.1. Data Loading and Preprocessing

- Normalize pixel values (0-255 scaled to 0-1)
- Optionally reshape or flatten images depending on model

4.2.2. Model Selection

- Traditional ML models: Logistic Regression, SVM, k-NN
- Deep learning models: Convolutional Neural Networks (CNNs)

Training and Validation

- Split training set further into training and validation
- Use techniques like early stopping, dropout to prevent overfitting

DOI: 10.48175/IJARSCT-26634

- **Evaluation:** Accuracy, confusion matrix, precision/recall per class.
- **Prediction & Visualization:** Visualize predictions with actual vs. predicted labels. Use softmax to interpret class probabilities

4.2.3. Technical Aspects

- **Input shape**: (28, 28, 1) for CNNs
- **Output**: One-hot encoded or integer class (0–9)
- **Loss Function**: Categorical Crossentropy or Sparse Categorical Crossentropy
- **Optimizer**: Adam, SGD, etc.
- **Performance Benchmark**: Simple CNN: ~90% accuracy

Advanced techniques (data augmentation, deeper networks): >93% accuracy

4.3 Challenges

1. Similar looking classes (e.g., shirt vs. T-shirt/top)
2. Small image size limits detail
3. Imbalanced difficulty per class

V. DISCUSSION

The model exhibited strong generalization on the test set. Misclassifications were mostly between similar categories (e.g., shirt vs. T-shirt). Data augmentation could further improve results. More complex architectures like ResNet or transfer learning with pre-trained models may yield even higher accuracy.

ACKNOWLEDGEMENT

I would like to thank **to the Director and HR of NU Intelligence pvt.ltd. , Nagpur 440015**

VI. CONCLUSION

CNN-based models are highly effective for image classification tasks, especially on structured datasets like Fashion MNIST. The research highlights the importance of deep learning in pattern recognition and lays the foundation for more advanced applications in real-world image classification.

REFERENCES

[1]. Xiao, H., Rasul, K., & Vollgraf, R. (2017). Fashion-MNIST: a Novel Image Dataset for Benchmarking Machine Learning Algorithms.

[2]. LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition.

[3]. Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012).

[4]. Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.

[5]. Dhanishtha Rukesh Dambhare:- B.tech student final year Computer Science Engineering Department (TGPCET college Nagpur) .