EXECUTIVE SUMMARY – COMPARATIVE STUDY OF ANN,CNN AND ENSEMBLE MODELS ON THE FASHION- MNIST DATASET

Abstract

The current study examines deep learning techniques for image classification utilizing the Fashion-MNIST dataset focused on three models: an **Artificial Neural Network (ANN), Convolutional Neural Network (CNN)**, and an **Adaptive Deep Neural Ensemble Integration (ADNEI)** model. The models are evaluated on accuracy, stability and generalisation. The novelty of the proposed ADNEI ensemble is that it combines the probabilistic outputs of the ANN and CNN through weighted soft-voting to give balanced performance with an average accuracy of **89.58%.** The results showed the CNN model achieved the best accuracy (90.01%) due to better spatial feature extraction capabilities, while ADNEI provided improved generalisation and better class balance for accuracyThe results clearly show that ensemble deep learning enhances robustness and accuracy of visually similar classes, which may have important consequences for fashion analytics, medical imaging and fraud detection.

1. Introduction

Fashion-MNIST is a newer benchmark for quantitatively evaluating image classifiers based on deep neural networks. Fashion-MNIST consists of **70,000** different grayscale images, evenly distributed among ten classes, such as coats, t-shirts, pants, and sneakers. The goal of this project is to design, train, and evaluate deep learning models based on three different architectures; (ANN), (CNN), and a new hybrid model based on an ensemble deep learning approach named ADNEI (Adaptive Deep Neural Ensemble Integration) to classify these images.

The ANN can be viewed as a reference point for fully connected networks, while the CNN uses convolutional layers to account for spatially linked features of the images. ADNEI is an ensemble model that combines the strengths of output from both ANN and CNN models.The objective is to achieve high accuracy, improved generalisation, and reduced misclassification between visually similar classes, such as shirt, coat, and pullover. This work demonstrates the potential of **hybrid ensemble models** in enhancing reliability and decision accuracy in automated visual recognition tasks used in real-world applications like retail tagging and recommendation systems.

2. Methodology

The implementation was carried out in **Python (TensorFlow/Keras)**, ensuring reproducibility and consistent evaluation across all models.

Data_Pre-processing:
 The dataset was loaded from **keras.datasets.fashion_mnist** and normalised to a [0, 1] scale to improve convergence. The data was split into 50,000 training, 10,000 validation, and 10,000 testing samples. All images were reshaped to maintain consistent input dimensions for both ANN and CNN architectures.

Model Design:

- **ANN:** Three dense layers (512–256–128 units) with ReLU activation and 0.3 dropout to prevent overfitting.

- **CNN Model**: Three convolutional layers with 32-64-128 filters and max-pooling and batch normalization layers, finally flattening to dense layers with softmax as the output.

- **ADNEI ensemble of ANN and CNN models:** Softmax probability outputs were combined through weighted soft voting. The weights were tuned with validation accuracy to achieve a balance learning global (ANN) and spatial (CNN) features.

Training Configurations:

All models were trained using **TensorFlow/Keras** with the Adam optimiser and Categorical Cross-Entropy loss for **15** epochs and a batch size of **128**.EarlyStopping and ModelCheckpoint were applied to prevent overfitting and save the best weights.The ADNEI Ensemble combined ANN and CNN outputs using weighted soft-voting based on validation accuracy.

3. Results and Analysis

Model	Accuracy(%)	Macro Precision	Macro Recall	Macro F1-Score
ANN	86.79	86.88	86.79	86.63
CNN	90.01	89.98	90.01	89.96
ADNEI Ensemble	89.58	89.54	89.58	89.46

Compared to the ANN, the CNN model demonstrated improved feature extraction and spatial awareness, with the highest test accuracy of **90.01%.** The accuracy of the ADNEI Ensemble method was slightly lower than that achieved with CNN at **89.58%,** which utilized a weighted soft-voting strategy to combine the softmax probability outputs from ANN and CNN, but it showed a better balance of precision, recall, and F1-score values.

The training and validation curves showed a consistent convergence with no evidence of overfitting, reflecting the effective use of dropout and early stopping. Both the CNN and ADNEI confusion matrices indicate the modules reduce major misclassifications between visually similar classes(for example, shirt-coal-pullover). Overall, both approaches best for class stability.

4. Innovation and Technical Contribution

Our project presents the **ADNEI framework**, which uses an adaptive weighting strategy to form a unified ensemble prediction based on the outputs of an **ANN and a CNN**. This adaptive fusion of models improves classification robustness without the need to retrain the component base models. Reproducibility is achieved using fixed random seeds within the models, consistent preprocessing, and standard evaluation metrics. The **ADNEI ensemble** shows that it is possible to simultaneously learn global characteristics and spatial features to significantly improve image classification tasks. The work is relevant outside of image classification tasks, and could be transferable to tasks in other fields such as **medical imaging**, **fraud detection** and **retail analytics**.

5. Conclusion and Future Scope

The study demonstrates that **ensemble deep learning** improves reliability and generalisation in image classification tasks. While CNN performed best overall, the **ADNEI Ensemble** achieved competitive accuracy with enhanced class balance and stability, validating the benefits of adaptive integration. Future work will focus on **data augmentation**, **learning rate scheduling**, and **Explainable AI (Grad-CAM)** to improve interpretability and expand ADNEI to **transformer-based hybrid models** for broader applications in **fashion analytics**, **healthcare imaging**, and **fraud detection**.