A Comparative Analysis of Conventional and Deep Learning-Based Feature Extraction for MNIST Classification

1. Introduction

Handwritten digit classification has been a core challenge in computer vision and machine learning. Feature extraction techniques like Histogram of Oriented Gradients (HOG), Local Binary Patterns (LBP), and edge detection yield interpretable and compact representations, whereas deep learning-based techniques like ResNet50 take advantage of hierarchical feature learning. In this research, traditional and deep learning-based feature extraction techniques are compared on the MNIST dataset.

2. Literature Review

Feature extraction is the backbone of machine learning models. Conventional techniques such as HOG and LBP have been extensively employed because they perform well in analyzing texture and shape. Edge orientations are captured by HOG, while texture patterns are encoded by LBP, and edge detection highlights gradient-based features. Deep architectures, specifically Convolutional Neural Networks (CNNs), learn hierarchical representations automatically, diminishing the necessity of handcrafted features. Pretrained models such as ResNet50 offer a feature extraction mechanism that is robust and generalizable across domains.

3. Experimentation

3.1 Dataset

The MNIST dataset of 60,000 training and 10,000 testing handwritten digit images (0-9) is utilized for benchmarking. Images are resized and normalized for deep learning models.

3.2 Feature Extraction Methods

3.2.1 Conventional Approaches:

- **HOG + SVM:** Extracts features based on gradients and classifies with an RBF kernel-based SVM.
- **LBP + Gradient Boosting**: Features based on texture are encoded and classified with a Gradient Boosting model.
- **Edge Detection + Random Forest**: Applies Laplacian edge detection followed by a Random Forest classifier.

3.2.2 Deep Learning-Based Approach:

- **ResNet50 + MLP**: Employs ResNet50 as a feature extractor, with subsequent Multi-Layer Perceptron (MLP) classification.

3.3 Performance Metrics

All methods are tested against classification accuracy and a comprehensive classification report.

4. Results and Analysis

Model	Accuracy
HOG + SVM	98.26%
LBP + Gradient Boosting	96.08%
Edge Detection + Random Forest	92.34%
ResNet50 + MLP	91.59%

4.1 Discussion

- **HOG + SVM** has the maximum accuracy (98.26%) because of strong edge detection along with structured feature extraction.
- **LBP + Gradient Boosting** yields good performance (96.08%) but just falls behind HOG because LBP is biased towards local texture information.
- **Edge Detection + Random Forest** is not as accurate (92.34%) because mere edge features will not offer as much discriminative power.
- **ResNet50 + MLP** underperforms compared to **HOG + SVM** (91.59%), which is probably because the limited size of the dataset hinders the efficiency of deep feature learning.

5. Trade-offs Between Conventional and Deep Learning Feature Extraction

Aspect	Conventional Methods	Deep Learning Methods
Feature Interpretability	High	Low
Computational Cost	Low	High
Training Data Requirement	Low	High
Robustness	Moderate	High
Generalization	Task-Specific	More Generalizable

5.1 Feature Representation and Model Performance

Feature representation has a strong effect on classification performance:

- Traditional features are designed in an explicit manner and perform well with small datasets.
- Deep learning models learn hierarchical features but need large datasets to perform at their best.
- Hybrid methods (e.g., deep learning + HOG) might increase accuracy with less computational overhead.

6. Conclusion

This research shows that traditional techniques such as HOG + SVM can be made more accurate on MNIST than deep learning feature extraction with minimal data. Nonetheless, deep models provide better generalization and flexibility. Future studies may investigate hybrid feature extraction algorithms and optimized deep architectures for small datasets.