Final Report: Comparative Analysis of Feature Extraction Methods for MNIST Classification

1. Introduction

Image classification is a fundamental task in computer vision, with applications ranging from optical character recognition to medical imaging. The MNIST dataset, consisting of 70,000 handwritten digits (0-9), serves as a benchmark for evaluating different feature extraction techniques. This report compares traditional methods (HOG, LBP, Edge Detection) with deep learning-based features (MobileNetV2) when used with a Random Forest classifier.

Objectives

1. Evaluate performance of different feature extraction techniques.
2. Compare accuracy, training time, and computational efficiency.
3. Analyze why certain methods outperform others on MNIST.

---

2. Methodology

2.1 Dataset & Preprocessing

- **Dataset:** MNIST (28×28 grayscale images).
- **Train/Test Split:** 80% training (56,000 samples), 20% testing (14,000 samples).
- **Normalization:** Pixel values scaled to [0, 1].

2.2 Feature Extraction Techniques

Method	Description	Parameters
HOG	Captures gradient orientations	orientations=8, pixels_per_cell=(4,4)
LBP	Encodes local texture patterns	radius=3, n_points=24
Edge Detection	Extracts contours using Canny	sigma=1.0, resized to 14×14
Deep Features (MobileNetV2)	Pre-trained CNN features	Input resized to 32×32, RGB conversion

2.3 Classification

- **Classifier:** Random Forest (n_estimators=100).
- **Evaluation Metrics:** Accuracy, Precision, Recall, F1-Score.

---

3. Results & Analysis

3.1 Performance Comparison

Method	Accuracy	Precision	Recall	F1-Score	Training Time (s)
HOG	97.09%	97.07%	97.07%	97.07%	66.34
Edge Detection	87.30%	87.23%	87.02%	87.00%	17.04
Deep Features	78.71%	78.42%	78.39%	78.33%	64.89
LBP	60.77%	59.19%	59.96%	59.26%	16.23

3.2 Key Findings

1. **HOG is the Best Performer (97.09% Accuracy)**
   - Captures gradient structures effectively, ideal for digit recognition.
   - Confirmed by prior work (Dalal & Triggs, 2005).
2. **Edge Detection Performs Moderately (87.30%)**
   - Preserves shape but loses internal details.
3. **Deep Features Underperform (78.71%)**
   - Pre-trained on natural images (ImageNet), less suitable for MNIST.
   - Resizing grayscale to RGB adds no useful information.
4. **LBP Performs Worst (60.77%)**
   - Texture-based features are less discriminative for digits.

3.3 Computational Efficiency

Method	Feature Extraction Time (s)	Training Time (s)
HOG	75.67	66.34
LBP	37.19	16.23
Edge Detection	-	17.04
Deep Features	-	64.89

- **Fastest:** LBP (16.23s training).
- **Slowest:** HOG (66.34s training), but justified by high accuracy.

---

4. Discussion

Why HOG Outperforms Deep Learning?

- **MNIST Simplicity:** Handwritten digits have clear edges, making gradient-based methods highly effective.
- **Overkill for Deep Learning:** Pre-trained CNNs are optimized for complex natural images, not simple digits.

Limitations

1. **Deep Features:** Transfer learning may require fine-tuning.
2. **LBP:** Not suitable for shape-based recognition.
3. **Edge Detection:** Loses stroke thickness information.

---

5. Conclusion & Recommendations

5.1 Conclusions

- **HOG is the best choice** for MNIST digit recognition.
- **Deep learning is unnecessary** for simple datasets like MNIST.
- **LBP is unsuitable** for digit classification.

5.2 Recommendations

1. **For MNIST:** Use HOG + Random Forest for best results.
2. **For Complex Datasets:** Consider fine-tuned CNNs.
3. **Future Work:**
    - Combine HOG with CNNs for hybrid features.
    - Experiment with SVM for better classification.

---

6. References

1. Dalal, N., & Triggs, B. (2005). *Histograms of Oriented Gradients for Human Detection.* CVPR.
2. Ojala, T., et al. (2002). *Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns.* TPAMI.
3. LeCun, Y., et al. (1998). *Gradient-Based Learning Applied to Document Recognition.* Proceedings of the IEEE.