Literature Review: Feature Extraction Methods for Image Classification on MNIST

1. Introduction

Image classification is a fundamental task in computer vision, and feature extraction plays a crucial role in determining classification performance. The MNIST dataset, consisting of handwritten digits, has been widely used to benchmark different feature extraction techniques. This literature review explores traditional and deep learning-based feature extraction methods, their effectiveness on MNIST, and their comparative performance.

2. Traditional Feature Extraction Methods

2.1 Histogram of Oriented Gradients (HOG)

HOG (Dalal & Triggs, 2005) is a widely used feature descriptor that captures local gradient information. It works by:

- Dividing the image into small cells.
- Computing gradient magnitudes and orientations.
- Creating histograms of gradient directions.

Relevance to MNIST:

- HOG performs exceptionally well on MNIST (as seen in our results: **97.09% accuracy**) because digits have well-defined edges and gradients.
- Studies (Majeed & Al-Khalid, 2016) confirm that HOG is highly effective for digit recognition due to its ability to capture stroke orientations.

2.2 Local Binary Patterns (LBP)

LBP (Ojala et al., 1996) is a texture descriptor that encodes local pixel intensity variations.

- It compares each pixel with its neighbours, generating binary patterns.
- These patterns are used to construct histograms representing texture features.

Relevance to MNIST:

- Our results show poor performance (**60.77% accuracy**), likely because:
  - MNIST digits are better characterized by shape rather than texture.
  - LBP loses structural information, making it less discriminative for digits (Zhao et al., 2012).

2.3 Edge Detection (Canny)

Canny edge detection (Canny, 1986) extracts boundaries by:

- Smoothing the image to reduce noise.
- Computing gradients to detect edges.
- Applying non-max suppression and hysteresis thresholding.

Relevance to MNIST:

- Achieved **87.30% accuracy** in our experiments.
- Edge maps preserve digit structure but lose internal details, limiting performance compared to HOG.
- Works better than LBP but lacks the robustness of gradient-based methods (LeCun et al., 1998).

3. Deep Learning-Based Feature Extraction

3.1 CNN-Based Feature Extraction (MobileNetV2)

Deep learning models, particularly CNNs, automatically learn hierarchical features. MobileNetV2 (Sandler et al., 2018) is a lightweight CNN designed for efficiency.

Relevance to MNIST:

- Surprisingly, our results show **78.71% accuracy**, worse than HOG.
- Possible reasons:
  - **Pre-training on ImageNet (natural images) vs. MNIST (digits):** The features learned from natural images may not transfer optimally.
  - **Input size mismatch:** MNIST digits (28×28) were upscaled to 32×32, potentially introducing artifacts.
  - **Grayscale vs. RGB:** Converting grayscale to RGB provides no additional discriminative information.

Comparison with Literature:

- Traditional CNNs (LeNet-5, LeCun et al., 1998) achieve >99% on MNIST, but require training from scratch.
- Transfer learning from large models (e.g., ResNet, VGG) may not always help for simple datasets (Yosinski et al., 2014).

4. Comparative Analysis of Methods

Method	Accuracy (Our Results)	Strengths	Weaknesses
HOG	97.09%	Captures gradient structure well	Computationally expensive
LBP	60.77%	Fast, texture-based	Poor for shape recognition
Edge Detection	87.30%	Preserves contours	Loses internal details
Deep Features (MobileNetV2)	78.71%	Automatic feature learning	Suboptimal transfer learning

5. Conclusion and Insights

- **HOG is the best-performing traditional method** for MNIST, confirming prior studies (Dalal & Triggs, 2005).
- **LBP performs poorly** due to its texture-based nature, which is less discriminative for digits.
- **Edge detection is moderate** but lacks fine-grained discriminative power.
- **Deep learning features underperform** when used as fixed feature extractors without fine-tuning.

Future Research Directions

1. **Hybrid Approaches:** Combining HOG with CNN features could improve robustness.
2. **Domain-Specific Fine-Tuning:** Adapting pre-trained CNNs to MNIST may enhance performance.
3. **Alternative Traditional Features:** Exploring SIFT or SURF for comparison.

References

- Dalal, N., & Triggs, B. (2005). *Histograms of Oriented Gradients for Human Detection.* CVPR.
- Ojala, T., Pietikäinen, M., & Mäenpää, T. (2002). *Multiresolution Gray-Scale and Rotation Invariant Texture Classification with Local Binary Patterns.* TPAMI.
- Canny, J. (1986). *A Computational Approach to Edge Detection.* TPAMI.
- Sandler, M., et al. (2018). *MobileNetV2: Inverted Residuals and Linear Bottlenecks.* CVPR.
- LeCun, Y., et al. (1998). *Gradient-Based Learning Applied to Document Recognition.* Proceedings of the IEEE.