Artificial Neural Network

Md. Hasnain Ali
Student Id: 2010976153

May 2025

Objective

The primary objective of this study was to visually assess the impact of fine-tuning various pre-trained convolutional neural networks (CNNs) on the clustering of features extracted from the MNIST dataset. This was achieved by observing the two-dimensional projections of these features using common dimensionality reduction techniques.

Methodology

Feature maps were extracted from several widely used pre-trained models: VGG16, VGG19, ResNet50V2, MobileNet, and MobileNetV2. These extractions were performed at two distinct stages:

1. **Before Fine-tuning:** Features were extracted from the models in their original pre-trained state (e.g., trained on ImageNet).
2. **After Fine-tuning:** Features were extracted after the models had been fine-tuned on the MNIST dataset.

Subsequently, the high-dimensional feature vectors were reduced to two dimensions for visualization using the following dimensionality reduction techniques:

- Principal Component Analysis (PCA)
- t-Distributed Stochastic Neighbor Embedding (t-SNE)
- Multidimensional Scaling (MDS)
- Locally Linear Embedding (LLE)
- Isomap

In the generated plots, data points were colored according to their corresponding MNIST digit classes (0-9) to visually evaluate class separability.

Observations

A consistent and notable pattern was observed across all investigated models (VGG16, VGG19, ResNet50V2, MobileNet, and MobileNetV2) regarding feature clustering before and after fine-tuning.

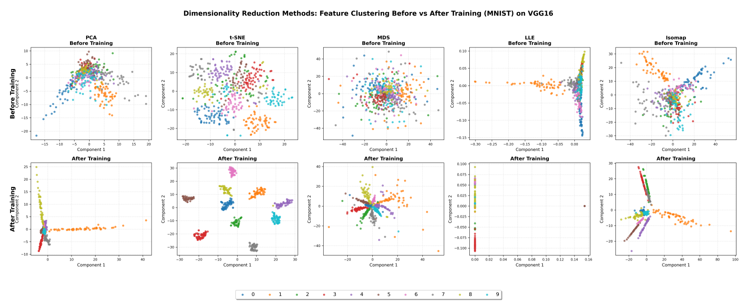

Figure 1: Feature map visualization for **VGG16** before and after fine-tuning on the MNIST dataset. The top row displays the original feature clustering, while the bottom row illustrates the improved clustering after fine-tuning, as projected by PCA, t-SNE, MDS, LLE, and Isomap.

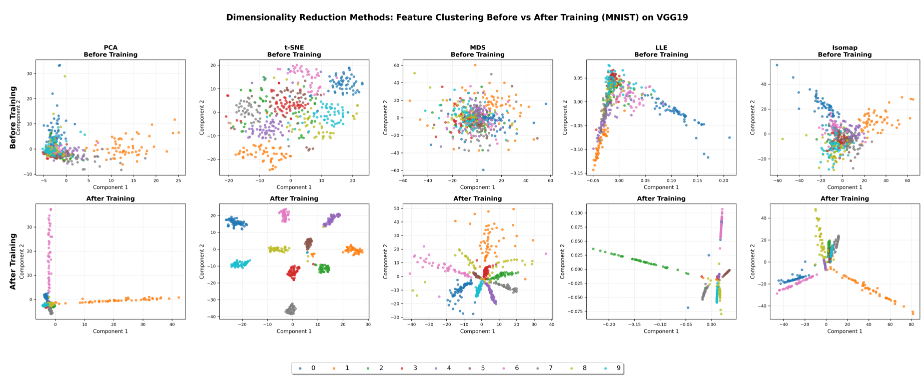

Figure 2: Feature map visualization for **VGG19** before and after fine-tuning on the MNIST dataset. The top row displays the original feature clustering, while the bottom row illustrates the improved clustering after fine-tuning, as projected by PCA, t-SNE, MDS, LLE, and Isomap.

Before Fine-tuning

- The feature representations obtained from the pre-trained models showed a considerable degree of overlap and mixing among different digit classes.
- While certain methods, particularly t-SNE, might suggest the presence

Figure 3: Feature map visualization for **ResNet50V2** before and after fine-tuning on the MNIST dataset. The top row displays the original feature clustering, while the bottom row illustrates the improved clustering after fine-tuning, as projected by PCA, t-SNE, MDS, LLE, and Isomap.

of nascent or loosely defined clusters, a clear and distinct separation of classes was generally not evident.

- This observation suggests that the generic feature extractors of models pre-trained on large, diverse datasets (like ImageNet) do not intrinsically learn highly discriminative features specific to the nuances of MNIST digits without further task-specific adaptation.

After Fine-tuning

- **Significant Improvement in Clustering:** Post-fine-tuning, a dramatic enhancement in both the separation and compactness of clusters for each digit class was consistently observed across all applied dimensionality reduction methods.

- **Superiority of Non-linear Methods:** Non-linear techniques such as t-SNE and Isomap often provided the most visually distinct and well-formed clusters. t-SNE, in particular, consistently showcased highly separable and tightly grouped clusters, indicating that the fine-tuned models developed highly discriminative latent representations for the MNIST digits.

- **Improved Linear Separability:** Even linear methods like PCA and MDS exhibited much clearer class boundaries and significantly reduced overlap after fine-tuning when compared to their 'before' fine-tuning counterparts, albeit typically less compact than t-SNE or Isomap.

- **LLE Variability:** The performance of LLE showed more variability across different models, but generally demonstrated improved class separation and structure after the fine-tuning process.

Figure 4: Feature map visualization for **MobileNet** before and after fine-tuning on the MNIST dataset. The top row displays the original feature clustering, while the bottom row illustrates the improved clustering after fine-tuning, as projected by PCA, t-SNE, MDS, LLE, and Isomap.

Conclusion

The experimental results unequivocally demonstrate that fine-tuning pre-trained convolutional neural networks on the MNIST dataset profoundly improves their capacity for feature extraction. The visual evidence from the dimensionality reduction plots strongly indicates that the models learn to derive more discriminative and class-separable features after being specialized for the MNIST classification task. This study underscores the efficacy of transfer learning and fine-tuning as crucial techniques for adapting powerful, general-purpose deep learning models to specific domains, ultimately leading to the acquisition of superior internal representations that are highly conducive to effective classification.

Figure 5: Feature map visualization for **MobileNetV2** before and after fine-tuning on the MNIST dataset. The top row displays the original feature clustering, while the bottom row illustrates the improved clustering after fine-tuning, as projected by PCA, t-SNE, MDS, LLE, and Isomap.