Comparative Study of Loss Functions in Face Verification and Variational Autoencoders

Fouzia Zilani

July 26, 2025

Contents

1 Introduction

This report presents a comparative study of different loss functions in two deep learning tasks:

- **Task 1:** Face verification using BCE, contrastive, and triplet loss.
- **Task 2:** Variational Autoencoder (VAE) training with BCE and MSE as reconstruction losses.

2 Task 1: Face Verification Loss Comparison

2.1 Dataset and Preprocessing

We use the Labeled Faces in the Wild (LFW) dataset. Pairs and triplets are constructed for training and evaluation. Images are resized to 128×128 and normalized.

2.2 Model Architecture

A simple convolutional neural network is used as the embedding network for all loss functions.

2.3 Loss Functions

- **Binary Cross-Entropy (BCE) Loss:** Used for binary classification of pairs.
- **Contrastive Loss:** Encourages similar pairs to be close and dissimilar pairs to be apart in embedding space.
- **Triplet Loss:** Uses anchor, positive, and negative samples to enforce a margin between positive and negative pairs.

2.4 Training and Evaluation

Each model is trained for 25 epochs. Performance is evaluated using ROC curves and AUC scores.

2.5 Results

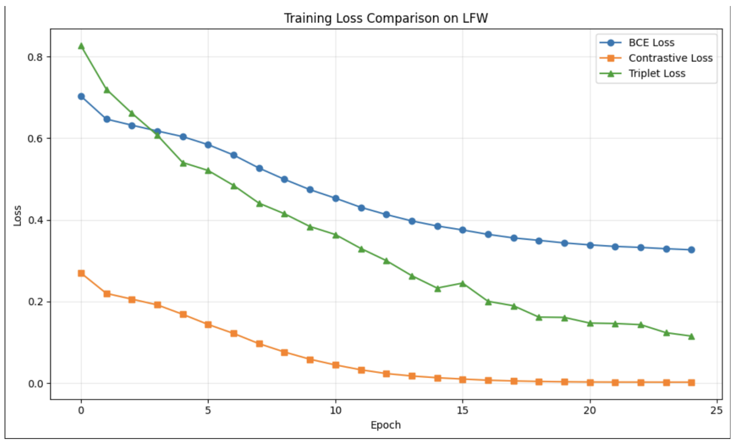

Figure 1: Training loss curves for BCE, Contrastive, and Triplet loss.

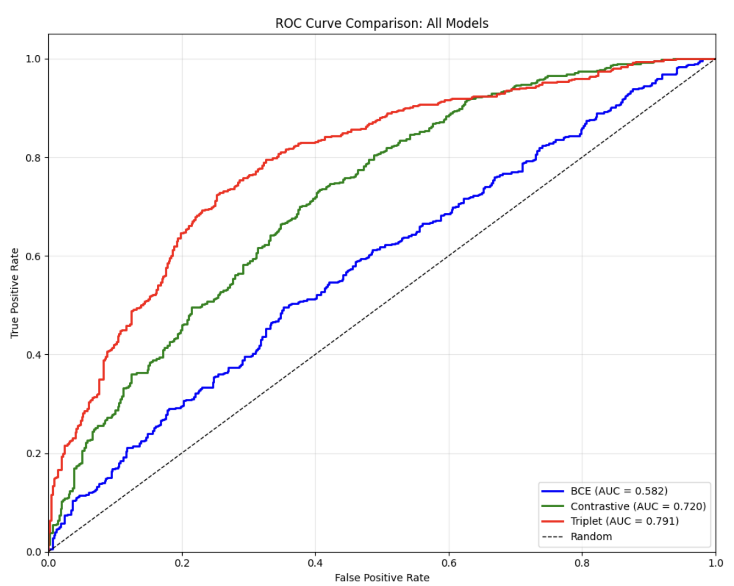

Figure 2: ROC curves for all models.

2.6 Discussion

The triplet loss model achieved the highest AUC, indicating better verification performance. BCE is easy to implement but may not structure the embedding space as well as contrastive or triplet loss.

3 Task 2: VAE Reconstruction Loss Comparison

3.1 Dataset and Preprocessing

We use the MNIST dataset. Images are normalized to $[0, 1]$.

3.2 Model Architecture

A fully-connected VAE with a latent dimension of 20.

3.3 Loss Functions

- **BCE:** Suitable for binary or normalized data.
- **MSE:** Suitable for real-valued data.

3.4 Training and Evaluation

Both models are trained for 50 epochs. We compare total loss, reconstruction loss, KL divergence, and visual quality of reconstructions.

3.5 Results

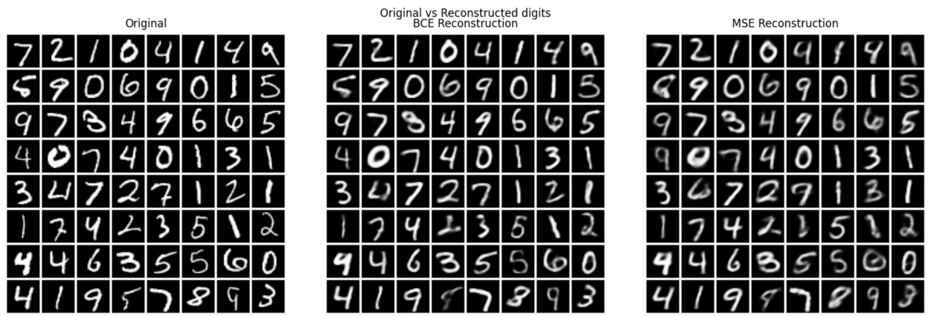

Figure 3: Original and reconstructed images (BCE and MSE).

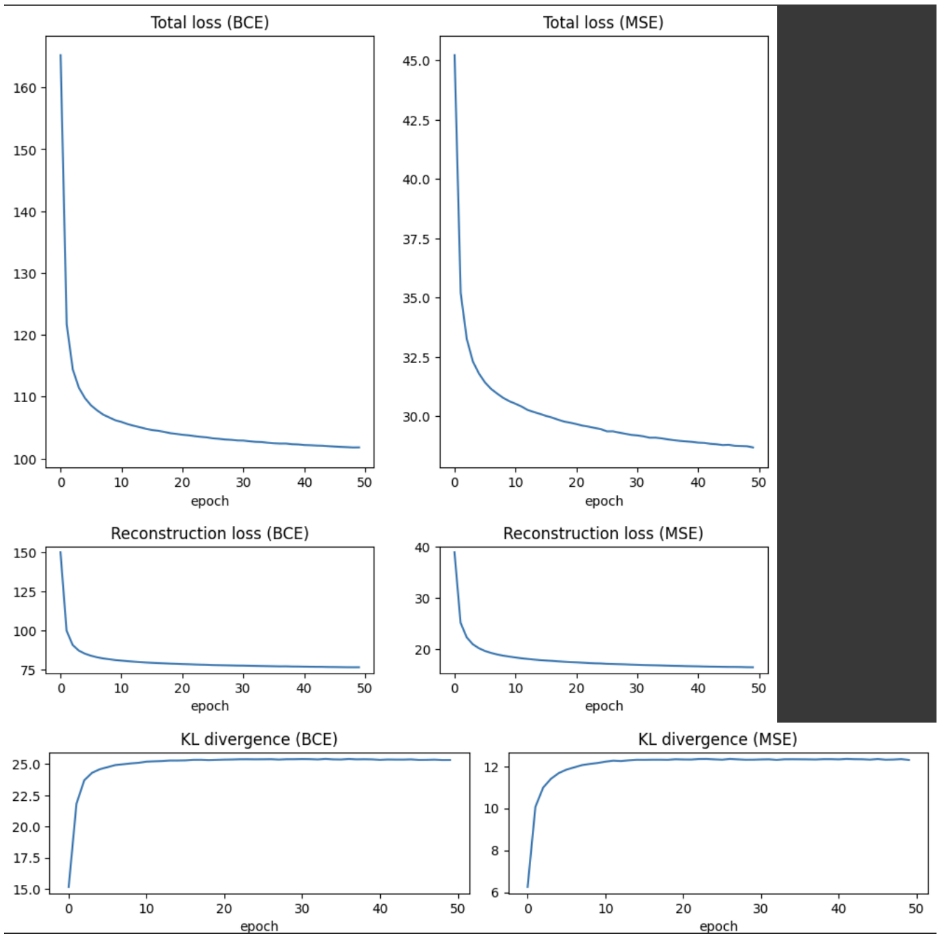

Figure 4: Training loss curves for BCE and MSE VAEs.

3.6 Discussion

BCE loss produced sharper reconstructions on MNIST, while MSE led to slightly blurrier images. BCE is preferred for binary/normalized data, while MSE is more general for real-valued data.

4 Conclusion

This study demonstrates that the choice of loss function significantly impacts the performance and output quality in both face verification and VAE training. For face verification, margin-based losses (contrastive, triplet) structure the embedding space better than BCE. For VAEs, BCE is preferable for binary data, while MSE is more general but may yield blurrier reconstructions.

Code is in this link: Link