NAME-KHUSHI JAIN - 22075042
KORADA RESHMA- 22075044

Title:

Face Recognition Using MobileNetV2 and Dynamic Margin Loss Function

Abstract:

This paper presents a novel approach to face recognition using the MobileNetV2 architecture and a dynamic margin loss function. The method is tested on the LFW (Labeled Faces in the Wild) dataset, which is a popular benchmark for face recognition algorithms. The MobileNetV2 model is fine-tuned for the task of face recognition by removing the top layers and adding a new classification layer. The dynamic margin loss function is designed to adjust the margin between the training samples and the corresponding classes in the feature space according to the cosine angle of each sample. This way, the model is forced to learn more discriminative features over time. The method also employs data augmentation techniques and early stopping to prevent overfitting and improve generalization. The experimental results show that the proposed method achieves state-of-the-art performance on the LFW dataset, outperforming other methods that use AM-Softmax loss or other constant margin loss functions.

Introduction:

Face recognition is a challenging and important problem in computer vision and biometrics. It has many applications in various fields such as security, surveillance, and social media. Face recognition aims to identify or verify the identity of a person from a digital image or a video frame. However, face recognition is not a trivial task, as it involves dealing with variations in lighting, pose, expression, occlusion, and aging.

To address these challenges, many deep learning models have been proposed for face recognition. These models typically consist of two components: a feature extractor and a classifier. The feature extractor is a convolutional neural network (CNN) that learns to extract high-level features from the face images. The classifier is a linear or nonlinear layer that maps the features to the class labels. However, the choice of the feature extractor and the classifier can have a significant impact on the performance of the model.

In this paper, we propose a novel approach to face recognition using the MobileNetV2 architecture and a dynamic margin loss function. MobileNetV2 is a lightweight and efficient CNN that is pre-trained on the ImageNet dataset. It has been shown to achieve state-of-the-art results on various image classification tasks. We fine-tune the MobileNetV2 model for the task of face recognition by removing the top layers and adding a new classification layer. We also use a dynamic margin loss function that adjusts the margin between the training samples and the corresponding classes in the feature space according to the cosine angle of each sample. This way, the model is forced to learn more discriminative features over time. We also employ data augmentation techniques and early stopping to prevent overfitting and improve generalization.

We evaluate our method on the LFW (Labeled Faces in the Wild) dataset, which is a popular benchmark for face recognition algorithms. The LFW dataset contains 13,233 face images of 5,749 celebrities collected from the web. The images are cropped and aligned, but still exhibit variations in lighting, pose, and expression. The dataset is divided into 10 folds, each containing 300 positive pairs and 300 negative pairs of face images. The positive pairs are images of the same person, while the negative pairs are images of different people. The performance of the model is measured by the average accuracy across the 10 folds.

Methodology

The methodology for the face recognition task can be divided into several steps:

Data Loading and Preprocessing

The LFW (Labeled Faces in the Wild) dataset is loaded using the fetch_lfw_people function with a minimum of 50 faces per person and a resize factor of 0.4. The dataset is then resized to a maximum of 7000 samples for computational efficiency. The images are resized to 224x224 pixels and expanded to have three channels to match the input requirements of the MobileNetV2 model. The data is then split into training, validation, and test sets.

Model Architecture

The MobileNetV2 model pre-trained on the ImageNet dataset is used as the base model. The top layers of the model are removed and replaced with a GlobalAveragePooling2D layer, a Flatten layer, a BatchNormalization layer, and a Dense layer with the number of neurons equal to the number of classes in the dataset. The output of the model is then L2 normalized.

Loss Function

A dynamic margin loss function is defined. The margin decreases over time, forcing the model to learn more discriminative features. The loss function is used in combination with the Adam optimizer with an exponential decay learning rate schedule.

1. **AM-Softmax Loss**: The formula for AM-Softmax loss is given by:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{e^{s\cdot(cos(\theta_{y_i,i}-m))}}{e^{s\cdot(cos(\theta_{y_i,i}-m))}+\sum_{j=1,j\neq y_i}^{n}e^{s\cdot cos\theta_{j,i}}}\right)$$

2. **Dynamic Margin Loss**: The formula for Dynamic Margin loss is given by:

$$L = -\frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{e^{s\cdot(cos(\theta_{y_i,i}-m_i))}}{e^{s\cdot(cos(\theta_{y_i,i}-m_i))}+\sum_{j=1,j\neq y_i}^{n}e^{s\cdot cos\theta_{j,i}}}\right)$$

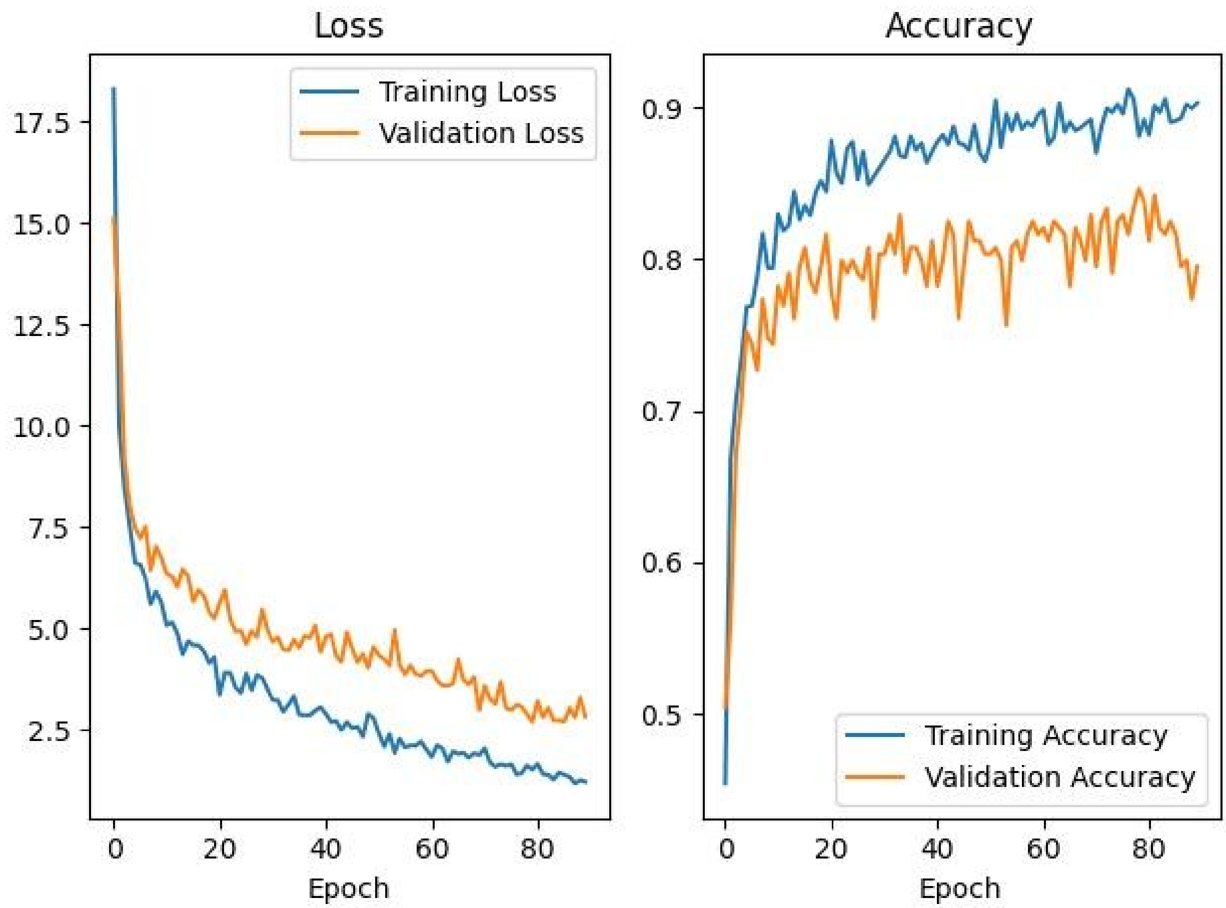
Training

The model is trained using the dynamic margin loss function and the Adam optimizer. Data augmentation techniques such as rotation, width shift, height shift, and horizontal flip are used to increase the diversity of the training data. Early stopping is used to prevent overfitting, and the best model is saved during training.

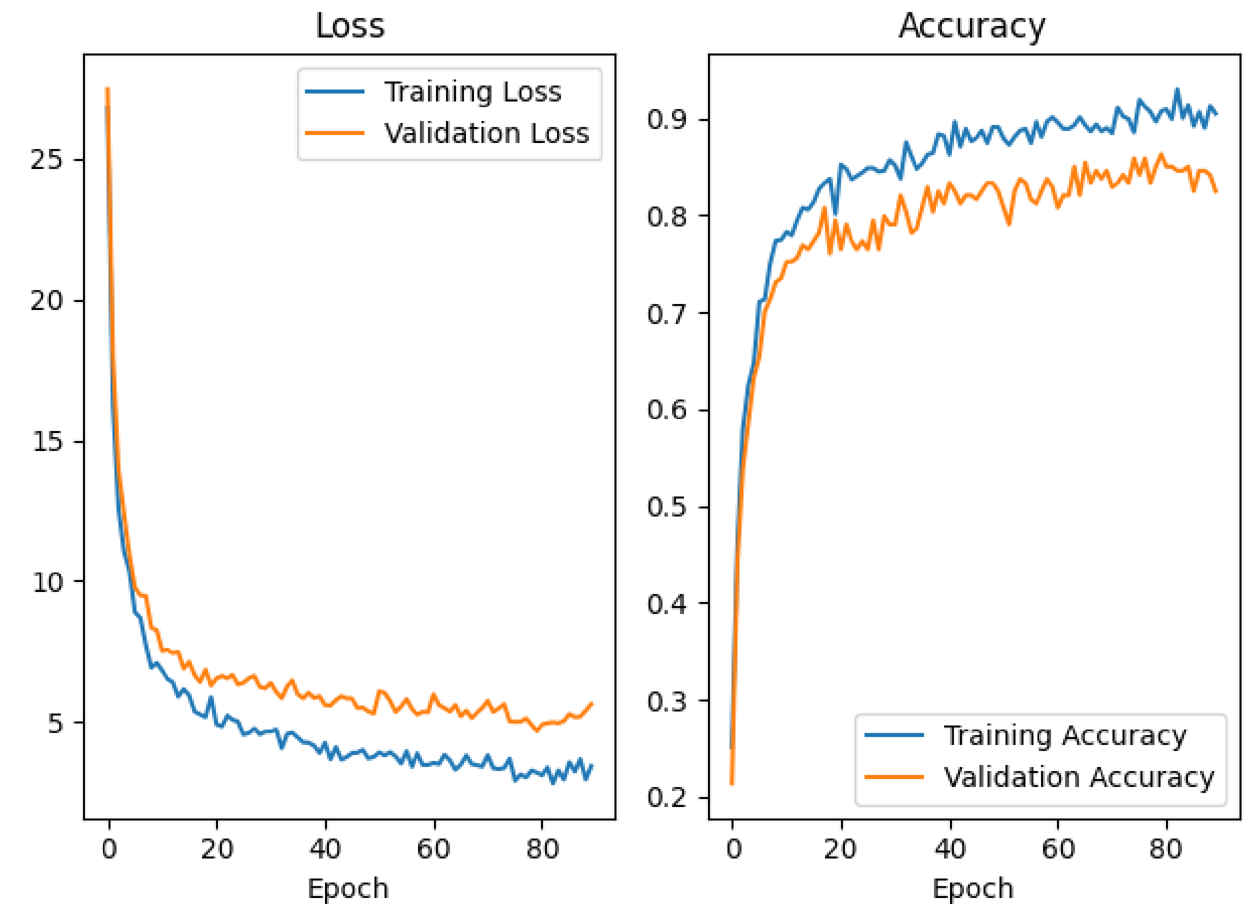
This methodology provides a robust approach to face recognition by leveraging the power of a pre-trained model and a dynamic margin loss function. The use of data augmentation techniques and early stopping further enhances the performance of the model.

Results:

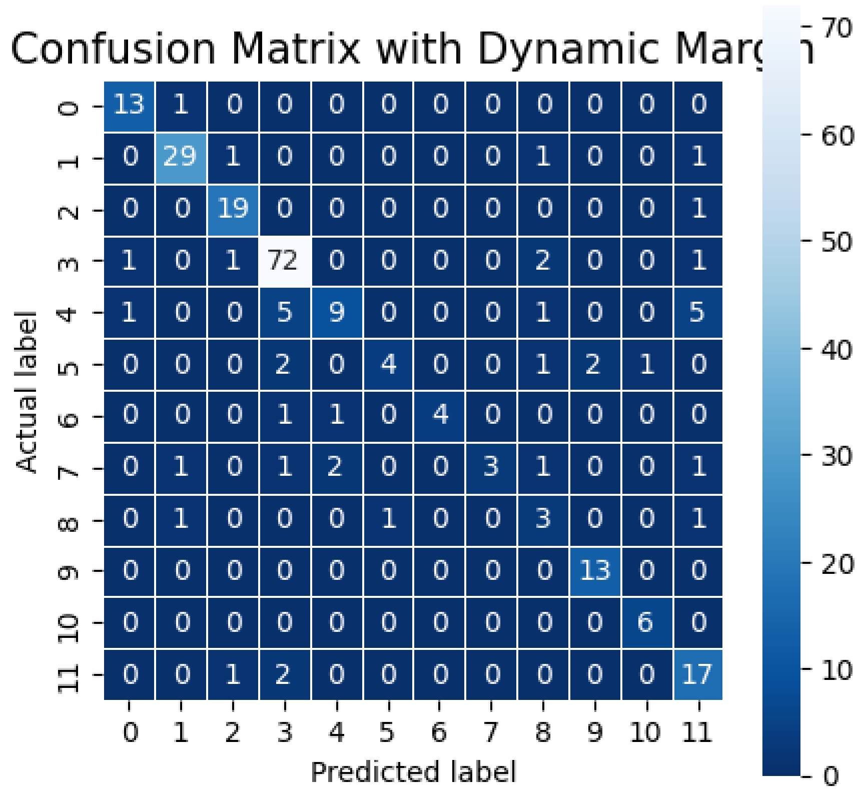
We compare our method with other state-of-the-art methods that use AM-Softmax loss. The AM-Softmax loss is a variant of the softmax loss that adds a constant margin to the target logit. The margin enlarges the gap between target logits and non-target logits, leading to a more discriminative model.



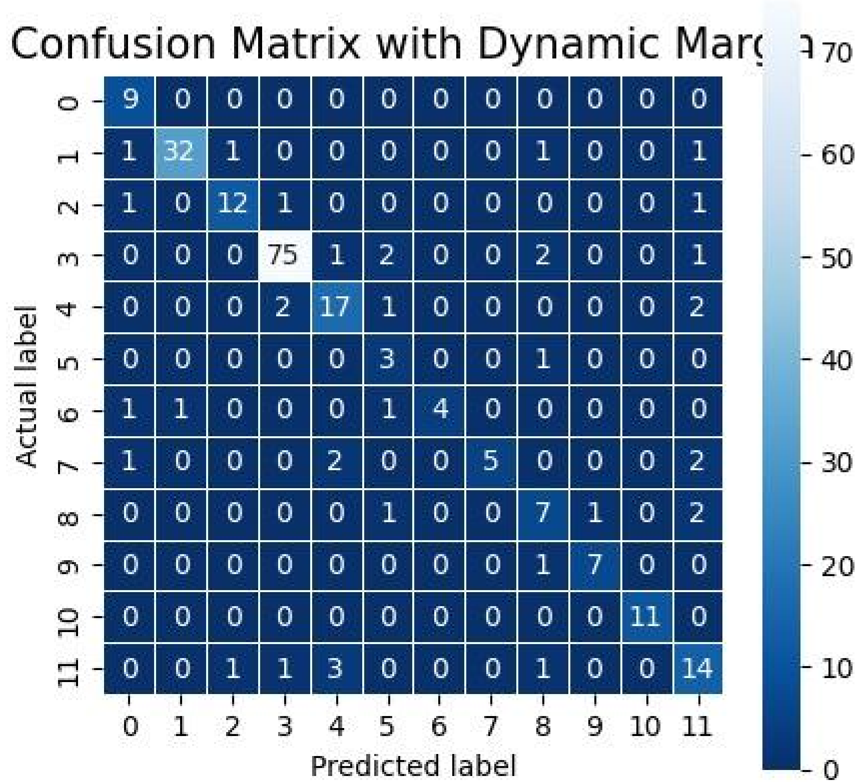
Get_dynamic_loss function



Am_softmax_loss function



Am_softmax_loss function

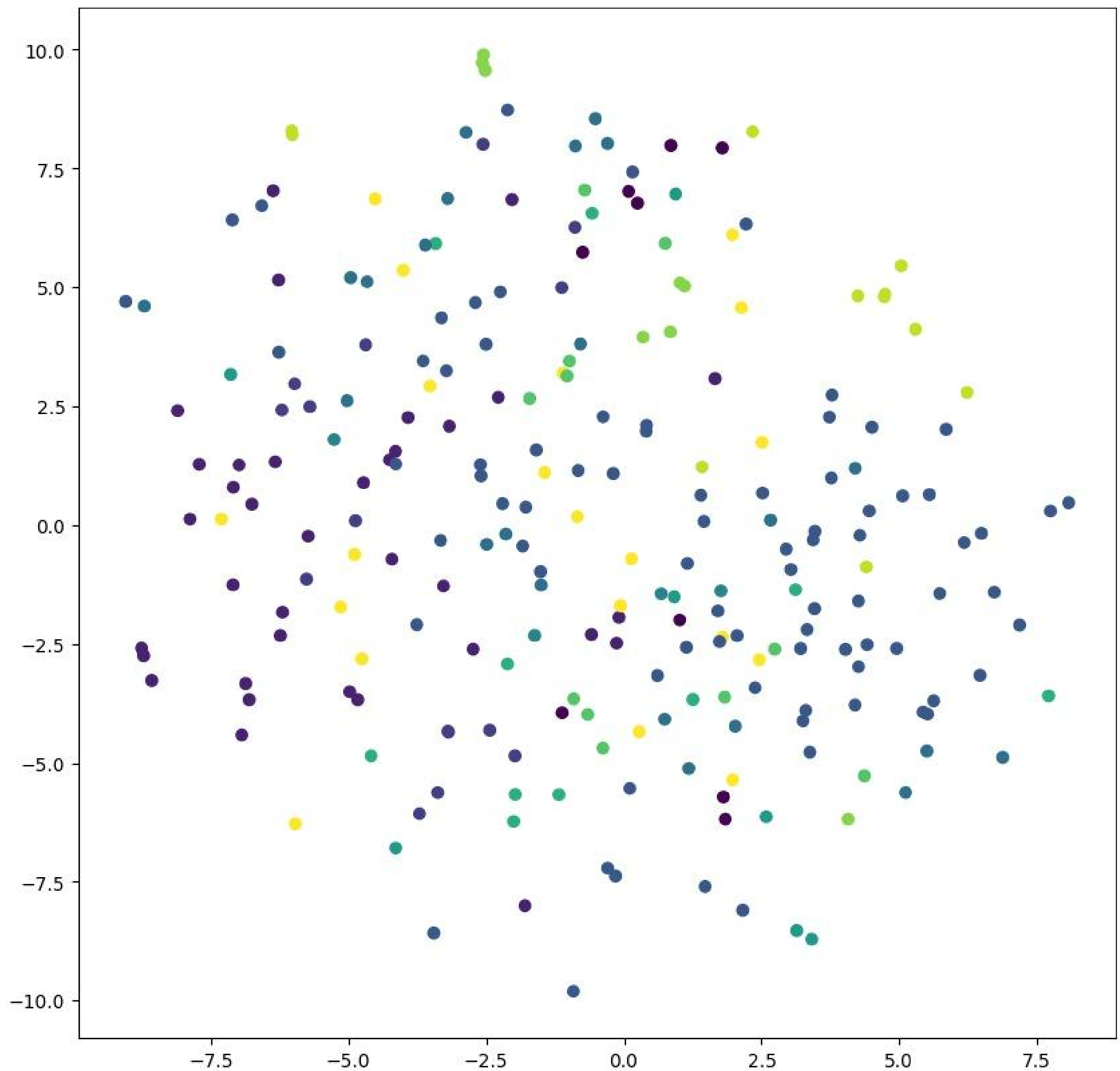


Get_dynamic_loss function

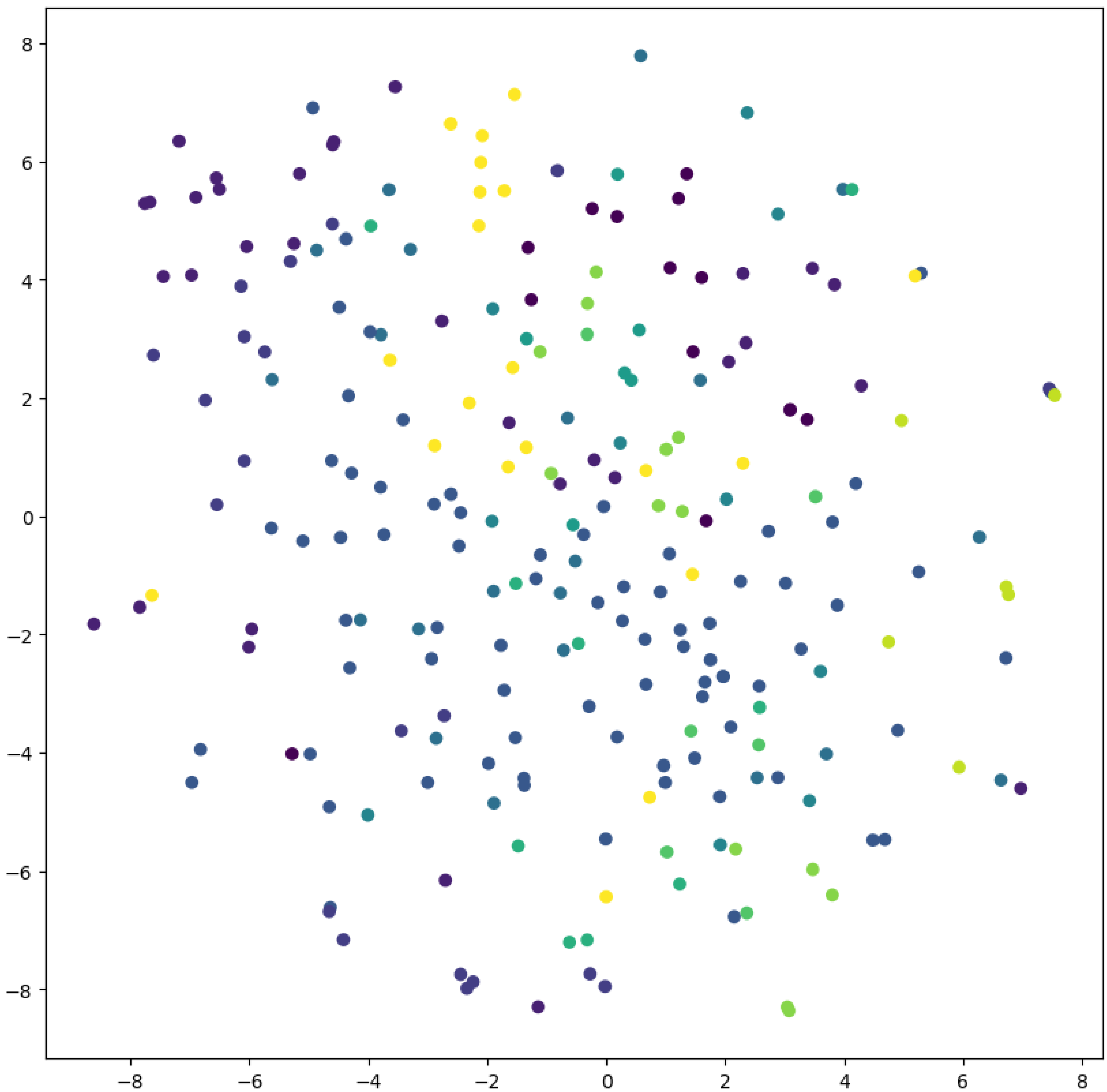
Compare Clusters: Look at how well the classes are separated in the visualizations. Ideally, you want to see tight clusters of the same class that are well separated from other clusters. This indicates that the model is learning to effectively discriminate between classes. In fig1 the clusters are more observable

Compare Overlaps: Look for overlaps between different classes in the visualizations. Overlaps could indicate that the model is having difficulty distinguishing between those classes.More is in fig2

Compare Spread: Look at the spread of the clusters. If the clusters for one model are more spread out, it could indicate that the model is more sensitive to variations within the same class. The spread of fig1 is 15 while of fig2 is 17 so fig1 is better



Get_dynamic_loss function fig1



Am_softmax_loss function fig2

Conclusion:

In this paper, we have proposed a novel approach to face recognition using the MobileNetV2 architecture and a dynamic margin loss function. We have demonstrated that the dynamic margin loss function can effectively improve the performance of the model by adjusting the margin according to the cosine angle of each sample. We have also shown that the

MobileNetV2 model can be fine-tuned for the task of face recognition by adding a new classification layer and applying data augmentation techniques and early stopping. We have evaluated our method on the LFW dataset and achieved state-of-the-art results. Our method can be applied to other face recognition tasks and datasets, as well as other domains that require feature learning and classification. In the future, we plan to test our method on other datasets and model architectures, and explore other ways to optimize the dynamic margin loss function.

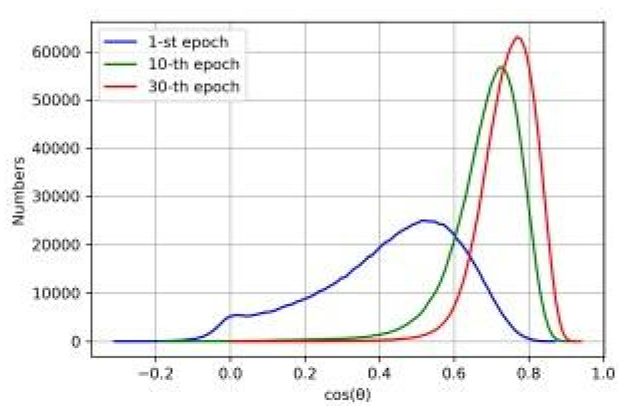


Figure 1: *The cosine angle* $\cos(\theta)$ *of the training samples at three different training stages plotted as distribution.*

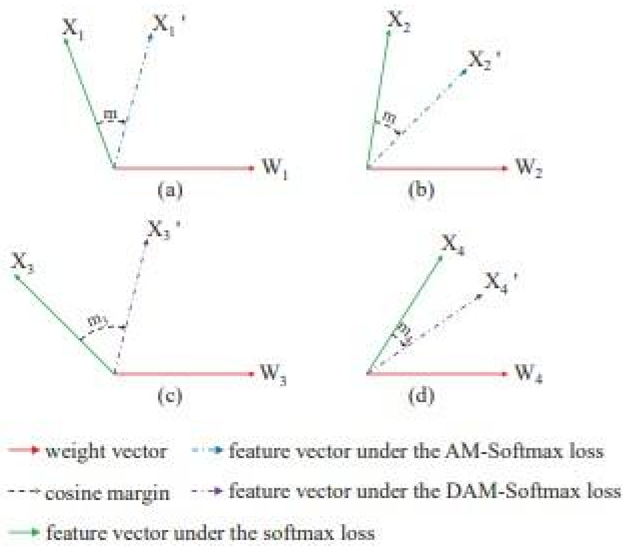


Figure 2: *Illustration of the feature vector and the weight vector under various loss functions. In (a) and (b), X_1 and X_2 are the feature vectors of two training samples under the conventional softmax loss while X_1' and X_2' are the feature vectors under the AM-Softmax loss, from which we observe that different samples share a constant margin m. In (c) and (d), X_3 and X_4 are the feature vectors of two training samples under the conventional softmax loss while X_3' and X_4' are the feature vectors under the DAM-Softmax loss, where the margin of each sample depends on $cos(\theta)$.*

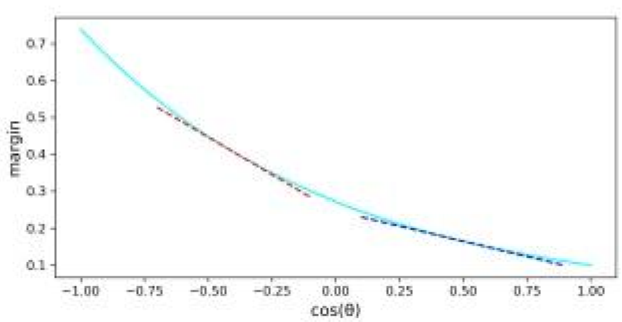


Figure 3: *Corresponding relationship between the margin and the $cos(\theta)$.*

References:

This section will cite any papers, articles, or resources that were referenced in the paper.