RESEARCH ON NEW LOSS FUNCTION FOR FACE RECOGNITION

By:
Group 6
Ayush Sharma(22075017)
and Shankar Nath(22075082)

Introduction

CosFace: Large Margin Cosine Loss

Large Margin Cosine Loss (LMCL) which is referred as CosFace, reformulates the traditional softmax loss as a cosine loss by L2 normalizing both features and weight vectors to remove radial variations, based on this a cosine margin term is introduced to further maximize the decision margin in the angular space. As a result of which we get minimum intra-class margin and maximum inter-class margin for accurate face verification.

$$L_A = -\frac{1}{N}\sum_{i=1}^{N}\log\frac{e^{s(\cos(\theta_{y_i})-m)}}{e^{s(\cos(\theta_{y_i})-m)}+\sum_{j=1,j\neq y_i}^{N} e^{s\cos\theta_j}}$$

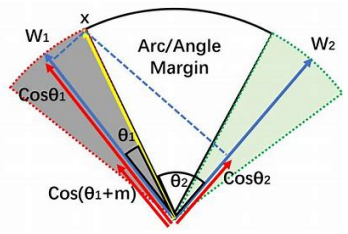
Drawbacks

- Handling Class Imbalance: CosFace loss may not handle imbalanced datasets well, and newer loss functions might include mechanisms to alleviate the impact of class imbalance on model training.

- Adaptability to Varied Data Distributions: Newer loss functions might be more adaptable to diverse data distributions and variations in pose, illumination, and other factors that can affect the quality of face recognition models.

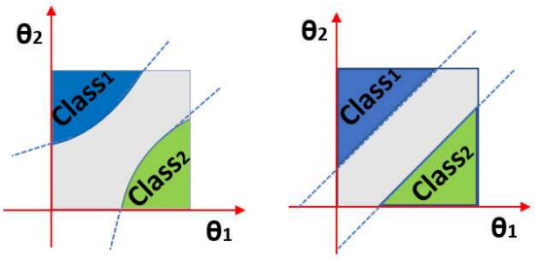
New loss function

Reduced_cosine_loss



Angular margin

In order to overcome the short comings of present cosface loss function , we devised a new loss function. In this new loss function , we define a new hyper parameter known as reducing factor (t).The new loss function also focuses more on angular margin thereby increasing the performance and leading to better classification.



Class distinction using:

1.(Cosface)Margin
2.(Reduced_cosine_loss)angular margin

New hyperparameter(t)

In the cosface loss, we introduce a wide interclass margin which separate data points belonging to different class, but in doing so we decide that the data points which belong to the margin area will have an ambiguous class .This loss function focuses on this aspect of cosface loss function. I.e. for instance if an object with y_i as an actual class has probability (p) of belonging to class y_i greater than belonging to other classes a little in comparison. But with the introduction of margin , its probability(p) of belonging to class y_i falls down as compared to other classes thus making the model declare that it belongs to some other class which is obviously incorrect. The new loss function tries to deal with this fault by decreasing the value of p of that object for classes other than y_i where it surpasses the probabilistic value of actual class by multiplying it with a reducing factor(t) where t<1.

Mathematical formula

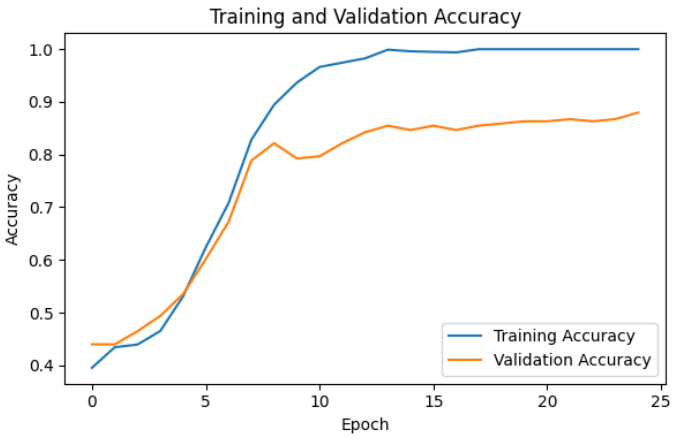
$$N(t) = cos\theta_j, \quad cos\theta_j < cos\left(\theta_{y_i} + m\right)$$

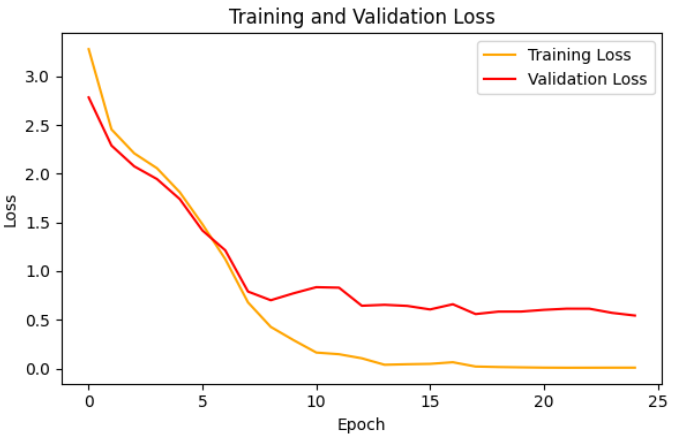
$$t * cos\theta_j, \quad cos\theta_j \geq cos\left(\theta_{y_i} + m\right)$$

$$\text{Loss} = -\frac{1}{N}\sum_{i=1}^{N} log\left(\frac{e^{s\,cos\left(\theta_{y_i}+m\right)}}{e^{s\,cos\left(\theta_{y_i}+m\right)} + \sum_{j\neq y_i} e^{s\,N\left(t,\theta_j\right)}}\right)$$

Conclusion / Performance

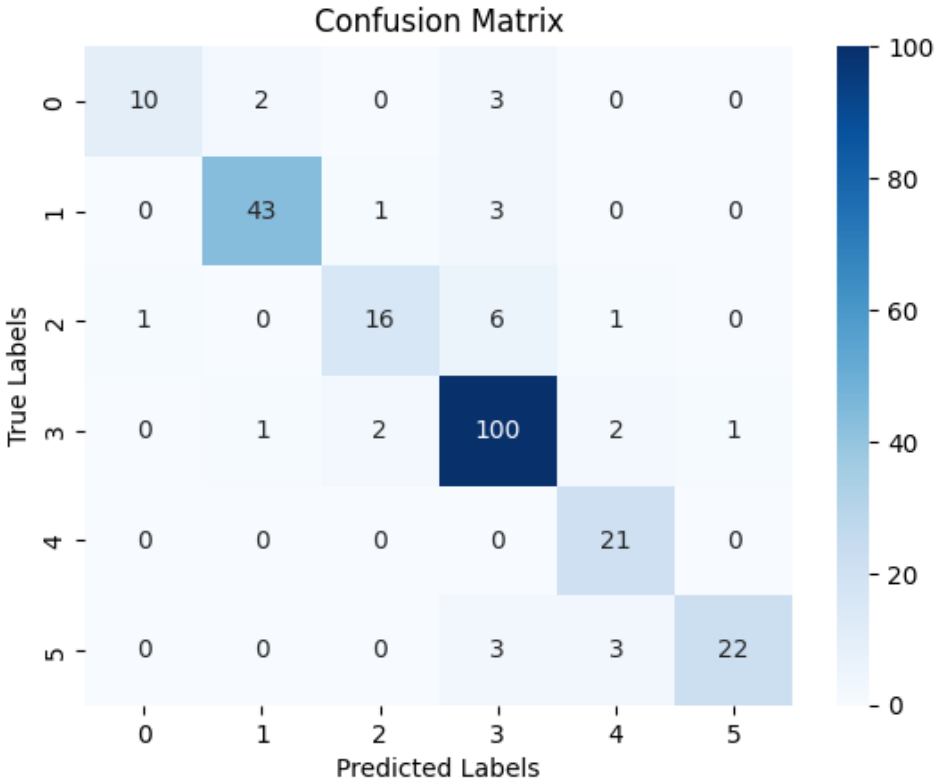
CNN face recognition model using reduced_cosine_loss loss function trained on lfw_dataset (here, no. of persons= 6)performance:





Model predictions on test data (lfw_dataset) :

Loss function used	Accuracy
Cosface	85%
Reduced_cosine_loss	88%



Classification report on Test data using new loss function:

```
Classification Report:
                   precision    recall  f1-score   support

     Ariel_Sharon       0.91      0.67      0.77        15
     Colin_Powell       0.93      0.91      0.92        47
  Donald_Rumsfeld       0.84      0.67      0.74        24
    George_W_Bush       0.87      0.94      0.90       106
Gerhard_Schroeder       0.78      1.00      0.88        21
       Tony_Blair       0.96      0.79      0.86        28

         accuracy                           0.88       241
        macro avg       0.88      0.83      0.85       241
     weighted avg       0.88      0.88      0.88       241
```