

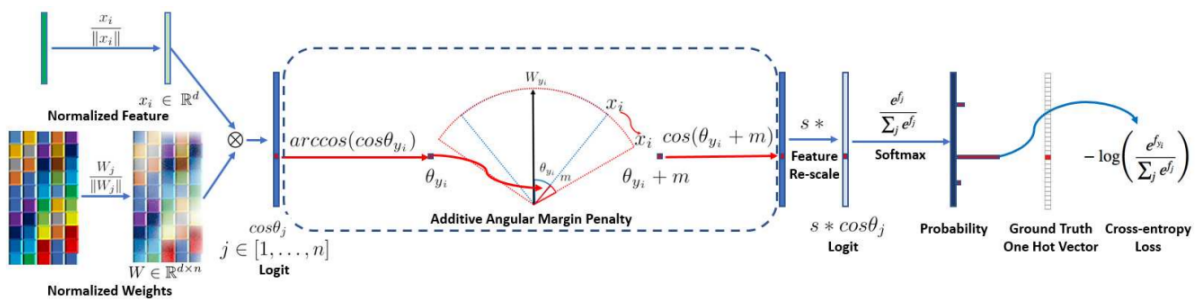
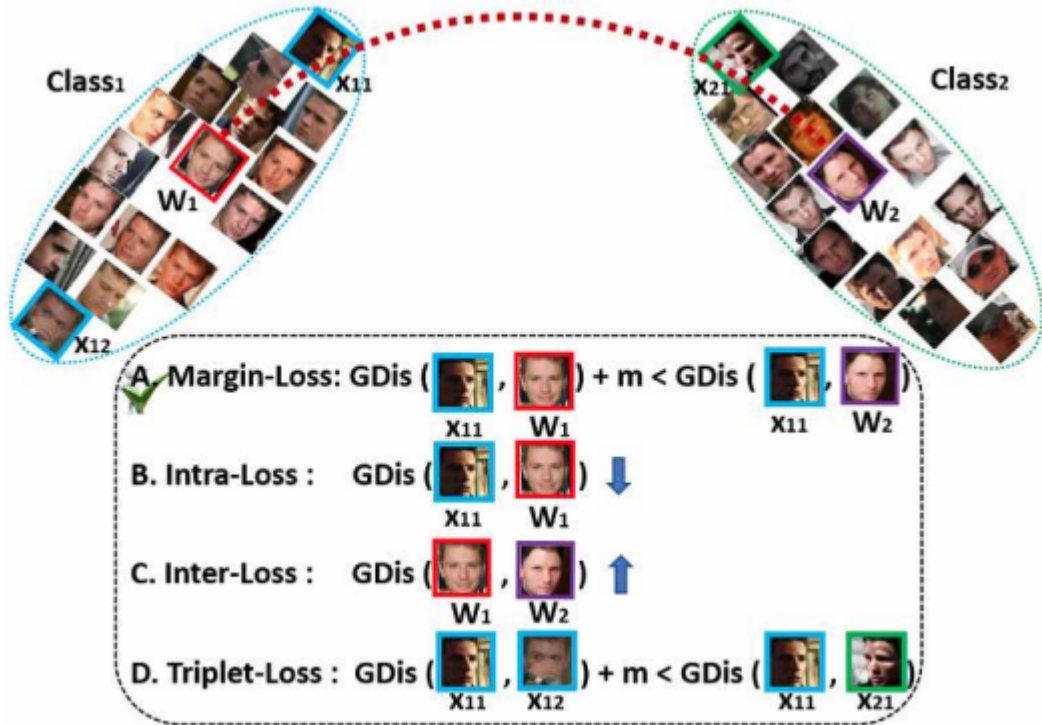


# [3주차] ArcFace: Additive Angular Margin Loss for Deep Face Recognition

## 0. Abstract

- 주요 개념 및 기술
  - **센터 손실**: 유클리드 공간에서 깊은 특징과 해당 클래스 중심 간의 거리를 처벌하여 클래스 내의 밀도를 높임
  - **SphereFace**: 마지막 완전 연결 계층에서의 선형 변환 행렬을 각 공간에서 클래스 중심의 표현으로 사용할 수 있다고 가정하고, 깊은 특징과 해당 가중치 간의 각도를 곱셈 방식으로 처벌함
  - **ArcFace**: 얼굴 클래스의 분리 가능성을 극대화하기 위해 잘 확립된 손실 함수에 마진을 포함시키는 최근의 연구 라인을 따름. ArcFace(가산 각 마진 손실)을 제안하여 얼굴 인식을 위한 매우 차별적인 특징을 얻음.
- 실험 및 결과
  - **실험 평가**: ArcFace는 열 개의 얼굴 인식 벤치마크에서 최근의 모든 최신 얼굴 인식 방법에 대한 가장 광범위한 실험 평가를 제시함
  - **성능**: ArcFace는 일관되게 최신 기술을 능가하며, 미미한 계산 오버헤드로 쉽게 구현될 수 있음

## 1. Introduction



### • 기존 얼굴 인식 기술: DCNN

- 일반적으로 포즈 정규화 단계 이후에 얼굴 이미지를 특징으로 매핑하며, 이 특징은 클래스 내 거리는 작고 클래스 간 거리는 커야 함
- 다중 클래스 분류기(ex. 소프트맥스): 훈련 세트의 다른 정체성을 구분
- 직접 임베딩을 학습(ex. 트리플렛 손실)
- 대규모 훈련 데이터와 정교한 DCNN 아키텍처를 기반으로, 소프트맥스 손실 기반 방법과 트리플렛 손실 기반 방법 모두 얼굴 인식에서 있어서 우수한 성능을 달성할 수 있음
- 그러나 소프트맥스 손실과 트리플렛 손실 모두 단점이 존재

### • 소프트맥스 손실의 단점

- 선형 변환 행렬의 크기가 정체성 수와 선형적으로 증가함

- 학습된 특징은 폐쇄 집합 분류 문제에 대해 분리 가능하지만, 개방 집합 얼굴 인식 문제에 대해서는 충분히 차별적이지 않음
- 트리플렛 손실의 단점
  - 특히 대규모 데이터셋에서 얼굴 트리플렛의 수가 조합적으로 폭발하여 반복 단계 수가 크게 증가
- 소프트맥스 손실의 차별화 능력 강화
  - Wen 등은 센터 손실을 개척하여 각 특징 벡터와 그 클래스 중심 간의 유클리드 거리를 통해 클래스 내 밀도를 얻고 소프트맥스 손실의 공동 처벌을 통해 클래스 간 분산을 보장함
  - 그러나 훈련 중 실제 중심을 업데이트하는 것은 훈련 가능한 얼굴 클래스의 수가 최근 급격히 증가함에 따라 매우 어려움
- ArcFace
  - 얼굴 인식 모델의 차별화 능력을 더욱 향상시키고 훈련 과정을 안정화하기 위해 가산 각 마진 손실(ArcFace)을 제안함
  - ArcFace는 정규화된 초구체에서 각도와 호 사이의 정확한 대응을 통해 지오데식 거리 마진을 직접 최적화하며 512-D 공간에서 특징과 가중치 간의 각도 통계를 분석하여 직관적으로 설명
  - ArcFace에서는 몇 줄의 코드만 필요하며, MxNet, Pytorch, Tensorflow와 같은 계산 그래프 기반 딥러닝 프레임워크에서 매우 쉽게 구현할 수 있음
  - ArcFace는 다른 손실 함수와 결합할 필요가 없으며, 어떤 훈련 데이터셋에서도 쉽게 수렴할 수 있음
  - ArcFace는 훈련 중에 미미한 계산 복잡성만 추가하며, 현재 GPU는 훈련을 위한 수백만의 정체성을 쉽게 지원할 수 있음



### ArcFace 얼굴 인식 모델의 원리

- 얼굴 인식 분야에서 사용되는 고급 기술 중 하나로, 얼굴 사이의 구별 가능성을 높이기 위해 고안되었음

- 이 기술의 핵심은

**얼굴 특징을 초구체(고차원 공간에서의 구) 상에 표현**하고, 이 공간에서의 각도를 사용하여 서로 다른 얼굴을 구별하는 것

#### <ArcFace의 기본 원리>

- **정규화된 초구체:** ArcFace는 얼굴의 특징을 고차원 공간의 구(초구체) 상에 매핑, 지구상의 다양한 지점을 위도와 경도로 표현하는 것과 유사함

- **각도와 호의 대응:** 초구체 상에서 서로 다른 얼굴(정체성)은 서로 다른 점으로 표현되며 이 점들 사이의 각도를 측정하여 얼마나 다른지를 판단함(각도가 클수록 더 다른 얼굴)

- **지오데식 거리 마진 최적화:** ArcFace는 이 각도를 기반으로 얼굴 사이의 '거리'를 최적화함. 즉, 같은 사람의 얼굴은 더 가깝게, 다른 사람의 얼굴은 더 멀게 배치하여 구별력을 높임

#### <ArcFace>

- **차별화 능력 향상:** 각도를 사용하여 얼굴을 구별함으로써 모델이 더 정확하게 얼굴을 인식하고 구별할 수 있음

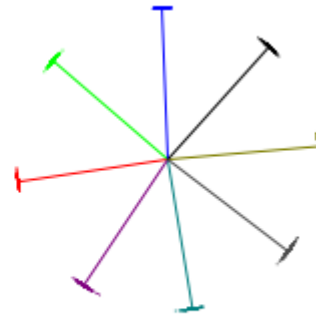
- **훈련 과정 안정화:** 이 방법은 훈련 과정을 더 안정적으로 만들어 모델이 더 빠르게 학습하고 더 좋은 성능을 낼 수 있게 도움

## 2. Proposed Approach

### 2.1 ArcFace



(a) Softmax



(b) ArcFace

**ArcFace(가산 각 마진 손실):** 분류 손실 함수인 소프트맥스 손실을 기반으로 하며 이를 개선한 새로운 손실 함수

- 소프트맥스 손실
  - 수식

$$L_1 = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{W_{y_i}^T x_i + b_{y_i}}}{\sum_{j=1}^n e^{W_j^T x_i + b_j}},$$

- (x<sub>i</sub>)는 i번째 샘플의 깊은 특징을 나타내며, (y<sub>i</sub>)번째 클래스에 속함
- 임베딩 특징의 차원 (d)는 512로 설정됨
- (W<sub>j</sub>)는 가중치 (W)의 j번째 열을 나타내고, (b<sub>j</sub>)는 편향
- 배치 크기와 클래스 수는 각각 (N)과 (n)
- 전통적인 소프트맥스 손실은 깊은 얼굴 인식에서 널리 사용되지만, **클래스 내 샘플 간의 유사성을 높이고 클래스 간 샘플의 다양성을 강제**하는 것을 명시적으로 최적화하지 않음
- 가중치와 특징에 대한 정규화 단계



**예측이 특징과 가중치 사이의 각도에만 의존**하게 만들어주며 따라서 학습된 임베딩 특징은 반지름이 (s)인 초구체에 분포하게 됨

$$L_2 = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{s \cos \theta_{y_i}}}{e^{s \cos \theta_{y_i}} + \sum_{j=1, j \neq y_i}^n e^{s \cos \theta_j}}.$$

- 임베딩 특징이 초구체의 각 특징 중심 주변에 분포함에 따라,  $(x_i)$ 와  $(W_{\{y_i\}})$  사이에 가산 각 마진 페널티 (m)을 추가하여 **클래스 내 밀도와 클래스 간 차이를 동시에 강화**함
- 제안된 가산 각 마진 페널티는 정규화된 초구체에서 지오데식 거리 마진 페널티와 동일하므로, 이 방법을 **ArcFace**라고 명명함

$$L_3 = -\frac{1}{N} \sum_{i=1}^N \log \frac{e^{s(\cos(\theta_{y_i} + m))}}{e^{s(\cos(\theta_{y_i} + m))} + \sum_{j=1, j \neq y_i}^n e^{s \cos \theta_j}}.$$

- 소프트맥스 손실과 ArcFace 손실을 각각 사용하여 2D 특징 임베딩 네트워크를 훈련시키기 위해, 8개의 다른 정체성에서 충분한 샘플(클래스당 약 1,500개 이미지)을 포함하는 얼굴 이미지를 선택
- 소프트맥스 손실은 대략적으로 분리 가능한 특징 임베딩을 제공하지만 **결정 경계에서 눈에 띄는 모호성을 생성**하는 반면, 제안된 ArcFace 손실은 **가장 가까운 클래스 사이에 더 명확한 간격을 강제**할 수 있음

## 2.2 Comparison with SphereFace and CosFace

### • ArcFace VS SphereFace VS CosFace

- ArcFace, SphereFace, 그리고 CosFace는 모두 얼굴 인식을 위한 딥러닝 모델에서 사용되는 기술들
- 얼굴의 특징을 추출하고 구별하는 방법에 있어서 각각 독특한 접근 방식을 취함
- 주요 차이점은 각 모델이 얼굴 특징 사이의 '거리'를 어떻게 측정하고 최적화하는지에 있음

### • 마진 유형의 비교

1 **SphereFace**: 곱셈 각도 마진(m1)을 사용, 클래스 내부의 밀집도를 높이고 클래스 간 다양성을 증가시키는 데 도움을 줌.

2 **ArcFace**: 덧셈 각도 마진(m2)을 도입, 지오데식 거리(구면 상의 최단 거리)에 정확히 대응되어 기하학적 특성이 우수함(해당 알고리즘이 얼굴 인식 분야에서 뛰어난 성능을 발휘하는 핵심 요소 중 하나임)

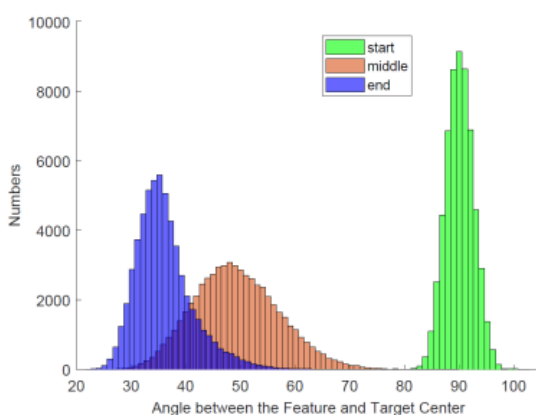
### 👉 덧셈 각도 마진의 역할

- **각도 마진 도입**: ArcFace는 각 클래스(얼굴)의 특징을 대표하는 중심과 각 데이터 포인트(얼굴 이미지) 사이의 각도에 마진을 추가함으로써, 같은 클래스 내의 데이터 포인트들이 더 가깝게, 다른 클래스의 데이터 포인트들과는 더 멀게 배치되도록 함

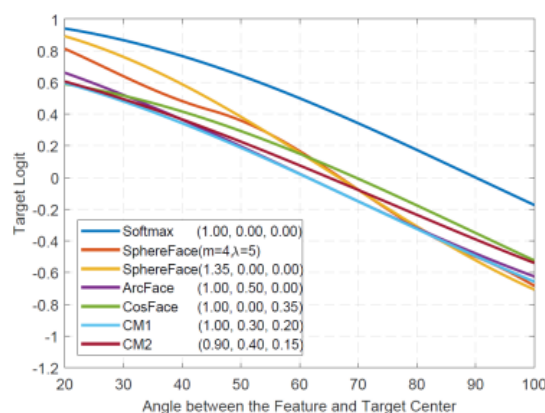
- **성능 향상**: 이 마진은 모델이 더 강력한 판별력을 갖도록 하며, 결과적으로 얼굴 인식 작업에서 더 높은 정확도를 달성할 수 있게 함

3 **CosFace**: 덧셈 코사인 마진(m3)을 적용, 코사인 공간에서의 거리를 조정하여 클래스 간 및 클래스 내 분리를 강화함

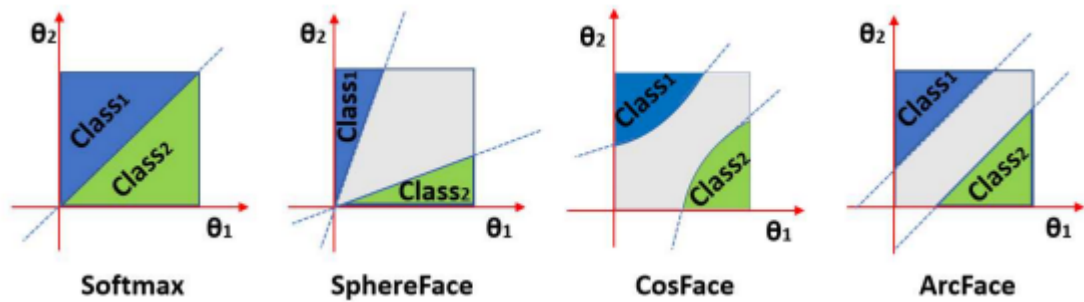
### • 수치적 유사성과 기하학적 차이



(a)  $\theta_j$  Distributions



(b) Target Logits Curves



- **수치적 유사성:** 세 모델 모두 클래스 내 밀집도와 클래스 간 다양성을 증가시키기 위해 로짓에 마진을 부과, 모델의 성능을 향상시키는 데 중요한 역할을 함
- **기하학적 차이:** ArcFace는 전체 구간에 걸쳐 일정한 선형 각도 마진을 가지는 반면, SphereFace와 CosFace는 비선형 각도 마진을 가짐. 이러한 작은 차이가 모델 훈련에 큰 영향을 미칠 수 있음
- 모델 훈련의 나비 효과
  - **SphereFace:** 복잡한 이중 각 공식 대신 아크코사인 함수를 사용하여 마진에 정수 요구 사항 없이 새로운 버전을 구현함, 이를 통해 훈련 초기의 발산을 피하고  $m=1.35$ 를 사용하여 원래 SphereFace와 유사한 성능을 달성할 수 있음

## 2.3 Comparison with Other Losses

- ArcFace 손실과 다른 손실 함수들의 비교 정리
  - 특징과 가중치 벡터의 각도 표현을 기반으로 한 다른 손실 함수들과의 비교
  - 특히 클래스 내의 밀집도를 높이고 클래스 간의 차이를 증가시키는 데 초점을 맞춘 손실 함수들을 다룸
- Intra-Loss(클래스 내 손실)
  - 목적: 클래스 내의 밀집도를 높이기 위해 설계되었음. 즉 샘플(얼굴 이미지)과 그 샘플의 실제 중심(해당 얼굴이 속한 클래스의 중심) 사이의 각도/호를 줄이는 것을 목표로 함

$$L_5 = L_2 + \frac{1}{\pi N} \sum_{i=1}^N \theta_{y_i}.$$



- $\theta_{yi}$ 는 샘플과 그 샘플의 실제 중심 사이의 각도를 나타냄
- Inter-Loss(클래스 간 손실)
  - 목적: **클래스 간의 차이를 증가시키기 위해** 설계되었으며, 다른 중심 사이의 각도/호를 증가시키는 것을 목표로 함

$$L_6 = L_2 - \frac{1}{\pi N (n - 1)} \sum_{i=1}^N \sum_{j=1, j \neq y_i}^n \arccos(W_{y_i}^T W_j).$$

- $W_{y_i}^T W_j$ 는 다른 중심 사이의 각도를 나타냄
- Triplet-Loss(트리플렛 손실)
  - 목적: 트리플렛 샘플 사이의 각도/호 마진을 확대하는 것을 목표로 하며, FaceNet에서는 **정규화된 특징에 유클리드 마진을 적용했지만** 여기서는 **특징의 각도 표현을 사용하여 트리플렛 손실을 적용함**
  - 기준:  $\arccos(x_{\text{pos}_i} \cdot x_i) + m \leq \arccos(x_{\text{neg}_i} \cdot x_i)$
  - $x_{\text{pos}_i}$ 와  $x_{\text{neg}_i}$ 는 각각 긍정적 샘플과 부정적 샘플을 나타냄



### 트리플렛 손실

- 트리플렛: 트리플렛은 세 개의 샘플로 구성되며, '

**기준 샘플(Anchor)**', '**긍정 샘플(Positive)**', '**부정 샘플(Negative)**'로 이루어짐. 여기서 기준 샘플과 긍정 샘플은 같은 클래스(예: 같은 사람의 얼굴)에 속하고, 부정 샘플은 다른 클래스에 속함.

<각도/호 마진 확대 목적>

트리플렛 샘플 사이의 각도/호 마진을 확대하는 것은

**기준 샘플과 긍정 샘플 사이의 거리를 줄이고, 기준 샘플과 부정 샘플 사이의 거리를 늘리는 것을 목표로 함.** 이를 통해 **같은 클래스의 샘플들은 서로 가까워지고, 다른 클래스의 샘플들은 서로 멀어지게 됨.**

<FaceNet과 각도 표현의 사용>

-

**FaceNet의 접근:** FaceNet에서는 정규화된 특징에 유클리드 마진 (Euclidean margin)을 적용, 이는 샘플들의 특징을 정규화한 후 유클리드 거리(직선 거리)를 사용하여 마진을 계산하는 방식

-

**각도 표현의 사용:** '특징의 각도 표현을 사용하여 트리플렛 손실을 적용함'은 **샘플들의 특징을 각도로 표현하고, 이 각도를 기반으로 트리플렛 손실을 계산하는 방식을 의미함.** 이는 샘플들 사이의 관계를 각도로 표현함으로써 더 직관적이고 효과적인 학습이 가능하게 합니다

## 3. Experiments

### 3.1 Implementation Details

- 사용된 데이터셋
  - **훈련 데이터:** CAISA, VGGFace2, MS1MV2, DeepGlint-Face(MS1M-DeepGlint 및 Asian-DeepGlint 포함)를 훈련 데이터로 사용하여 다른 방법들과 공정한 비교를 진행함. 특히, MS1MV2는 MS-Celeb-1M 데이터셋의 반자동으로 정제된 버전임
  - **테스트 데이터:** 얼굴 검증을 위해 LFW, CFP-FP, AgeDB-30 등의 데이터셋을 사용했으며 LFW와 YTF 외에도 CPLFW, CALFW 등 **최근의 대규모 포즈 및 연령 데이터셋**에서 ArcFace의 성능을 보고함. 대규모 이미지 데이터셋(MegaFace, IJB-

B, IJB-C, Trillion-Pairs)과 비디오 데이터셋(iQIYI-VID)에서도 ArcFace를 광범위하게 테스트함

- 실험 설정
  - **데이터 전처리:** 최근 논문들을 따라 얼굴 포인트 5개를 이용해 정규화된 얼굴 크롭 ( $112 \times 112$ )을 생성함
  - **임베딩 네트워크:** 널리 사용되는 CNN 아키텍처인 ResNet50과 ResNet100을 사용했으며, 마지막 컨볼루션 레이어 후에는 BN-Dropout-FC-BN 구조를 통해 최종 512-D 임베딩 특징을 얻음
  - **학습 설정:** 특징 스케일(s)을 64로 설정하고, ArcFace의 각도 마진(m)을 0.5로 선택함. 모든 실험은 MXNet을 사용하여 구현되었으며 NVIDIA Tesla P40 GPU 4개에서 모델을 훈련시킴
- 테스트 중 주의 사항
  - 훈련 세트와 테스트 세트 사이의 중복 신원은 엄격한 평가를 위해 제거됨
  - 테스트 시에는 **완전 연결 계층 없이** 특징 임베딩 네트워크만을 유지함

## 3.2 Ablation Study on Losses

- 각도 마진 설정 탐색
  - **ArcFace 최적 마진:** CASIA 데이터셋과 ResNet50을 사용한 실험에서 0.5가 최적의 마진으로 관찰되었음
  - **SphereFace와 CosFace 마진:** SphereFace는 1.35, CosFace는 0.35로 설정했을 때 최적 성능을 보였음. 이 두 방법 모두 우수한 성능을 보이며 수렴에 어려움이 없었음
- 결합 마진 프레임워크
  - **CM1과 CM2 성능:** CM1(1, 0.3, 0.2)과 CM2(0.9, 0.4, 0.15) 설정에서 가장 좋은 성능을 관찰했으며, 이 프레임워크는 SphereFace와 CosFace보다 더 나은 성능을 보였지만 ArcFace의 성능에는 미치지 못함
- 다른 손실 함수와의 비교
  - **기본 손실 함수:** 소프트맥스 손실을 기본으로 사용했을 때, CFP-FP와 AgeDB-30에서 성능 저하가 관찰되었음

- **내부 클래스 및 외부 클래스 손실:** 소프트맥스와 내부 클래스 손실을 결합하면 성능이 향상되었지만, 외부 클래스 손실과의 결합은 정확도를 약간만 향상시킴.
- **트리플렛 손실:** 트리플렛 손실이 Norm-Softmax 손실보다 우수한 성능을 보였으며, 이는 마진의 중요성을 시사함. 그러나 ArcFace처럼 **샘플과 중심 사이에 마진을 삽입하는 것이** 트리플렛 샘플 내에서 마진 벌칙을 사용하는 것보다 더 효과적인 것으로 드러났음
- ArcFace의 우수성 분석
  - **각도 통계:** 훈련 데이터(CASIA)와 테스트 데이터(LFW)에서 다양한 손실 하에 상세한 각도 통계를 제공함. ArcFace는 이미 내부 클래스의 밀집도와 외부 클래스 간의 차이, 그리고 분류 마진을 강화하고 있음을 보여줌

### 3.3 Evaluation Results

- **LFW와 YTF:** 이 두 데이터셋은 이미지와 비디오에서의 제약 없는 얼굴 검증을 위한 가장 널리 사용되는 벤치마크임. ArcFace는 MS1MV2 데이터셋과 ResNet100을 사용하여 훈련되었으며, SphereFace와 CosFace 같은 기존 모델들을 큰 차이로 능가함. 이는 추가 각 마진 패널티가 깊게 학습된 특징의 구별력을 현저히 향상시킨다는 것을 보여줌.
- **CPLFW와 CALFW:** 이 두 최신 데이터셋은 LFW에서 동일한 신원의 더 높은 포즈와 나이 변화를 보여줌. ArcFace는 이 데이터셋에서도 최고의 성능을 보여주며, 모든 오픈 소스 얼굴 인식 모델 중에서 최상위를 차지함.
- **MegaFace** 데이터셋의 성능
  - MegaFace는 1백만 개의 이미지와 69만 명의 다른 개인을 포함하는 갤러리 세트와 FaceScrub에서 53만 명의 고유 개인의 10만 개 사진을 포함하는 프로브 세트로 구성됨
  - ArcFace는 CASIA와 MS1MV2에서 각각 작은 프로토콜과 큰 프로토콜 하에 훈련되었으며, 식별 및 검증 모두에서 최고의 단일 모델 성능을 달성함
  - 이는 강력한 기준 모델들을 능가할 뿐만 아니라 다른 발표된 방법들보다도 우수한 성능을 보여줌
- **IJB-B 및 IJB-C** 데이터셋의 성능

- **IJB-B 데이터셋:** 1,845명의 대상자, 21.8K개의 정지 이미지, 55K개의 비디오 프레임 포함하며, 총 12,115개의 템플릿, 10,270개의 진짜 매치, 그리고 8M개의 가짜 매치가 있음
- **IJB-C 데이터셋:** IJB-B의 확장판으로, 3,531명의 대상자, 31.3K개의 정지 이미지, 117.5K개의 비디오 프레임 포함하며, 총 23,124개의 템플릿, 19,557개의 진짜 매치, 그리고 15,639K개의 가짜 매치가 있음
- 성능 평가 및 결과
  - ArcFace는 VGG2 데이터셋을 훈련 데이터로 사용하고 ResNet50을 임베딩 네트워크로 사용하여 IJB-B와 IJB-C 데이터셋에서 훈련되었음. 최신 방법들과의 공정한 비교를 위해 이러한 설정이 사용됨
  - **성능 향상:** ArcFace는 IJB-B와 IJB-C 데이터셋 모두에서 이전의 최고 성능 모델들보다 약 3~5%의 성능 향상을 보여주었음. 이는 오류율의 상당한 감소를 의미함
  - **더 나은 결과를 위한 추가 조치:** 더 많은 훈련 데이터(MS1MV2)와 더 깊은 신경망(ResNet100)을 사용함으로써, ArcFace는 IJB-B에서 94.2%, IJB-C에서 95.6%의 TAR(@FAR=1e-4)를 달성하여 더욱 향상된 성능을 보여줌
  - **새로운 기준 설정:** ArcFace는 FAR=1e-6에서도 인상적인 성능을 달성하여 새로운 기준선을 설정함

## • Trillion-Pairs 데이터셋의 성능

- 데이터셋 구성: Trillion-Pairs 데이터셋은 Flickr에서 가져온 1.58M개의 이미지 (갤러리 세트)와 5.7k LFW 신원에서 가져온 274K개의 이미지(프로브 세트)로 구성됨. 갤러리와 프로브 세트 간의 모든 쌍이 평가에 사용되며 총 0.4조 쌍이 있음
- 성능 비교: ArcFace는 CASIA 데이터셋과 비교하여 MS1MV2 데이터셋을 사용할 때 성능이 크게 향상되었으며 심지어 신원 수가 두 배인 DeepGlint-Face 데이터셋보다도 약간 더 나은 성능을 보였음. MS1MV2와 DeepGlint의 아시아 유명인사 신원을 모두 결합할 때, ArcFace는 최고의 식별 성능(84.840% @FPR=1e-3)을 달성했음

## • iQIYI-VID 데이터셋의 성능

- 데이터셋 구성: iQIYI-VID 챌린지는 iQIYI의 다양한 쇼, 영화, TV 드라마에서 가져온 4934개의 신원에 대한 565,372개의 비디오 클립으로 구성됨. 각 비디오의 길

이는 1초에서 30초 사이이며, 얼굴, 옷, 목소리, 보행 및 자막을 포함한 다중 모달 단서를 제공함

- 성능 평가: ArcFace는 MS1MV2와 아시아 데이터셋을 결합하여 훈련되었고, ResNet100을 사용하여 높은 기준선(MAP=79.80%)을 설정했음. iQIYI-VID 데이터셋에서 맞춤형 특징 설명자를 얻기 위해 추가적으로 **3층 완전 연결 네트워크**를 분류 손실로 훈련시켰으며 이 MLP는 MAP를 6.60% 크게 향상시킴. 모델 앙상블과 현장에서 사용 가능한 객체 및 장면 분류기의 컨텍스트 특징을 활용하여, 최종 결과는 준우승자보다 0.99% 높은 성능을 달성했음



### 3층 완전 연결 네트워크(MLP)

- **목적:** 맞춤형 특징 설명자를 통해 추출된 특징을 기반으로 비디오나 이미지 내의 객체를 더 정확하게 분류하기 위해 설계됨
- **작동 방식:** 입력된 특징을 받아 각 층을 통과시키면서 분류 작업을 수행, 이 과정에서 네트워크는 가중치를 조정하며 학습하게 됨
- **분류 손실:** 네트워크의 출력과 실제 레이블 간의 차이를 측정하는 것으로, 이 손실을 최소화하는 방향으로 네트워크를 훈련시키면 분류 성능이 향상됨

- 성능 향상의 핵심
  - ArcFace의 성공은 **추가 각 마진 손실(ArcFace)**에 기인하며, 이는 **특징 임베딩을 명시적으로 최적화**하여 동일 클래스 내 샘플의 유사성을 높이고 다른 클래스 간 샘플의 다양성을 증가시킴. 결과적으로 ArcFace는 얼굴 인식 작업에서 더 높은 성능을 달성할 수 있었음

## 4. Conclusions

- 부가 각 마진 손실 함수
  - 원래 얼굴 인식 작업에는 소프트맥스 손실 함수가 사용되었으나, 소프트맥스 손실 함수는 **클래스 내 샘플 간의 유사성을 높이고 클래스 간 샘플의 다양성을 강제하는 특징 임베딩을 명시적으로 최적화하지 않아** 큰 클래스 내 외형 변화 하에서의 깊은 얼굴 인식 성능에 격차가 있었음.

- **ArcFace의 접근:** ArcFace 손실은 로짓(logit)을 변환하여, 가중치와 특징 간의 각도에 기반하여 계산됨. 이는 클래스 간 경계를 더 명확하게 하여 모델이 각 클래스를 더 잘 구분할 수 있도록 함
- 핵심 이점
  - **판발력 향상:** 얼굴 인식을 위한 특징 임베딩의 판별력을 크게 향상시켰으며 이는 얼굴 인식 작업에서 더 높은 정확도와 신뢰성을 달성하는 데 도움이 되었음
  - **재현성 지원:** 연구자들이 이 논문에서 보고된 결과를 쉽게 재현할 수 있도록 자세한 설명과 함께 코드를 공개함으로써 연구 커뮤니티에 기여함

## 논문에 대한 의견 및 의문점(꼭지)

➡ 해당 논문을 읽어보았을 때 ArcFace는 얼굴 이미지를 분류하는 데 있어서 다른 모델들에 비해 월등히 높은 성능을 보이는 것으로 나타나지만, 이는 연구자들에 의한 실험 결과일 뿐이지 실제 사용자가 이용하는 일반적인 경우에도 똑같이 높은 성능을 보일 것인지 검증할 필요가 있다고 생각함. 또한 이러한 고성능의 얼굴 인식 기술이 무단으로 사용되지 않도록 접근 제어를 하는 보안 시스템이 함께 개발되어야 한다고 생각함.