CS588 Image Search

Variational Prototype Learning for Deep Face Recognition

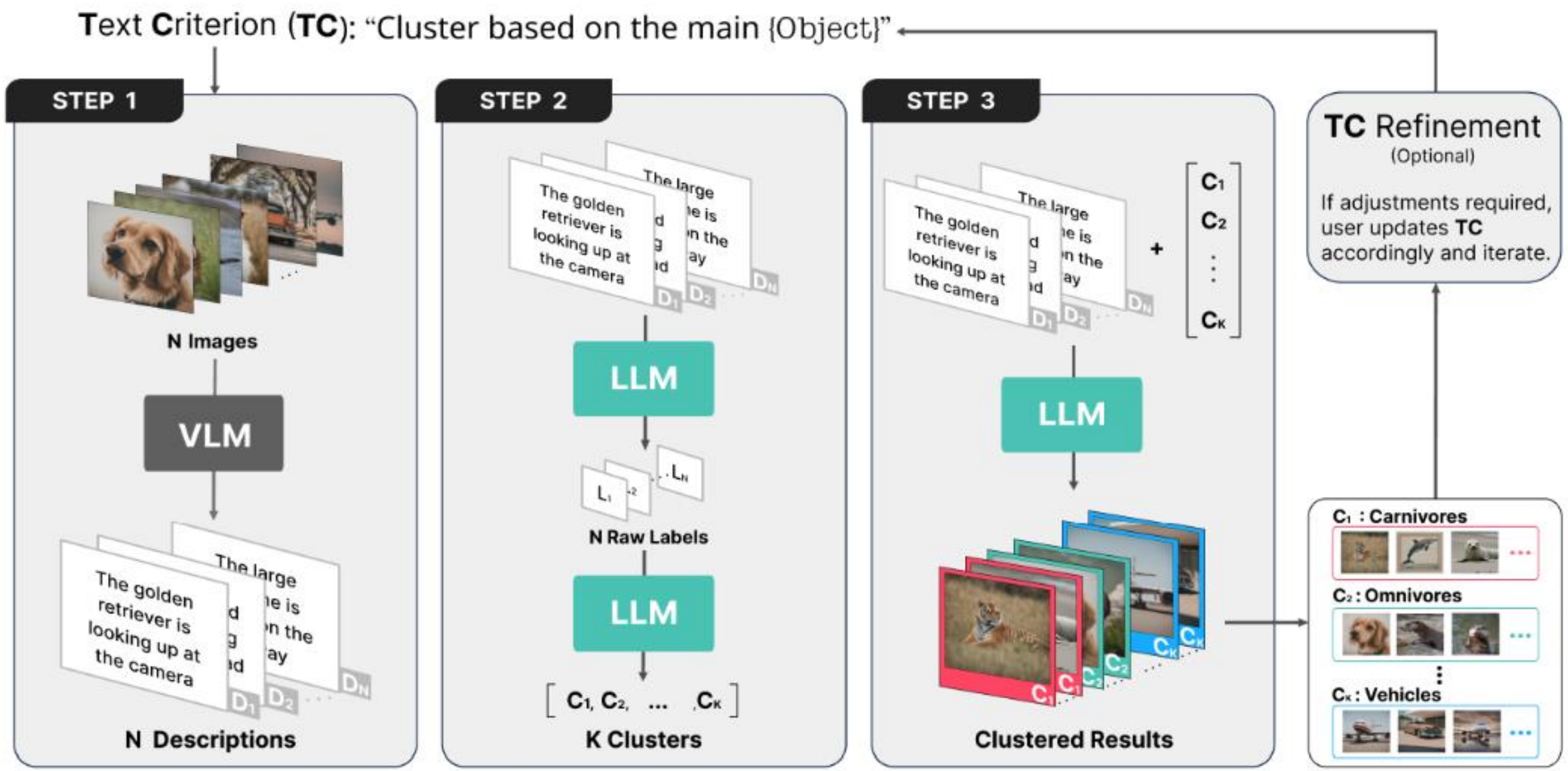
Suhyeon Ha

2024.04.24

KAIST

Contents

Review



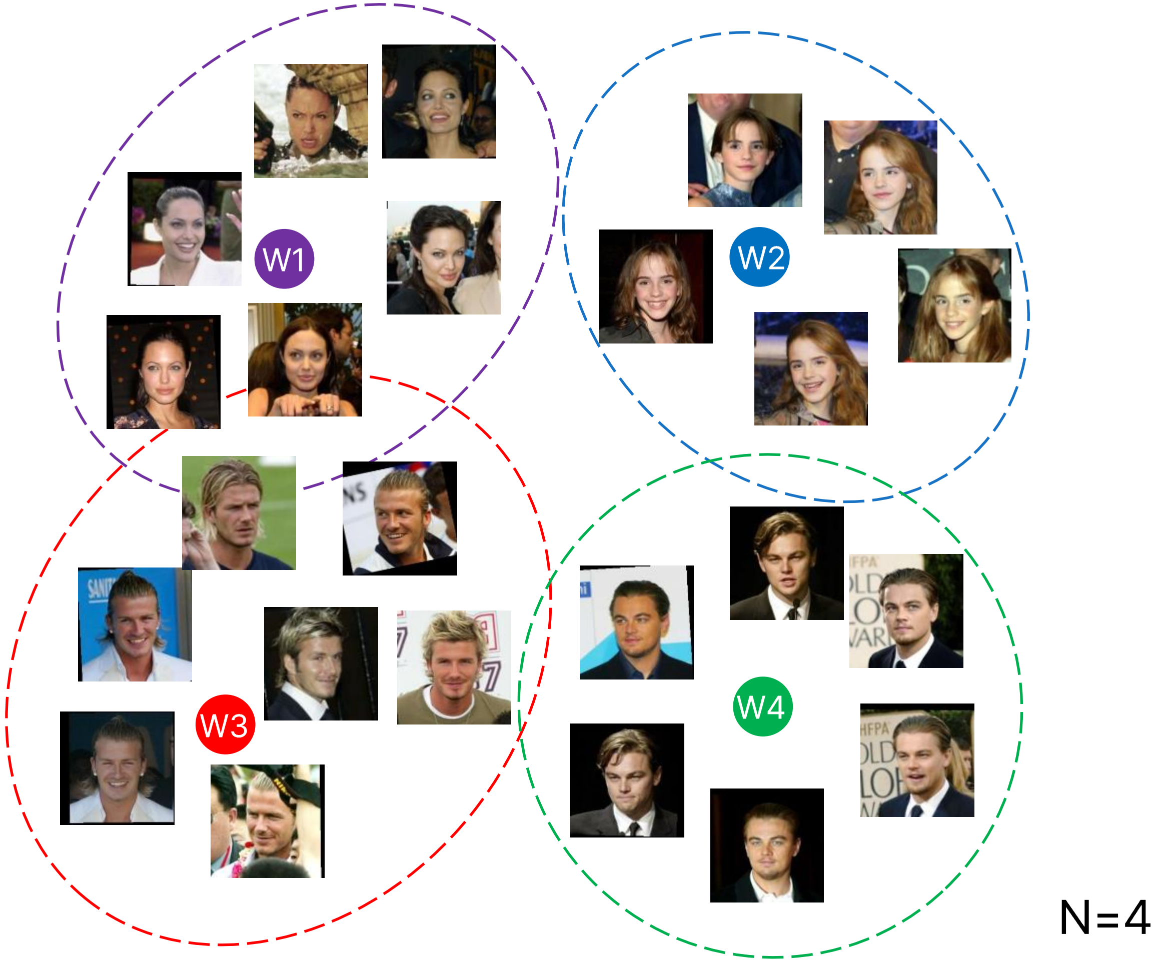
- Step 1 extracts textual descriptions of images from VLM + TC
- Step 2 identifies the names of the K clusters via LLM + TC
- Step 3 conducts clustering by assigning each description to the corresponding cluster + TC

Introduction

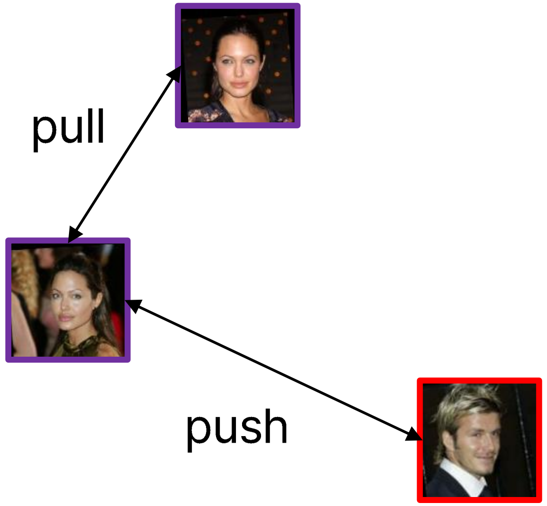
- To propose a novel **Variational Prototype Learning (VPL)** method which represents each class as a distribution instead of a point in the latent space
- To design **computationally efficient and memory-saving** way for the variational prototype sampling
- Extensive experimental results demonstrate the **superiority of proposed VPL over the SoTA competitors** in deep face recognition

KAIST

Introduction



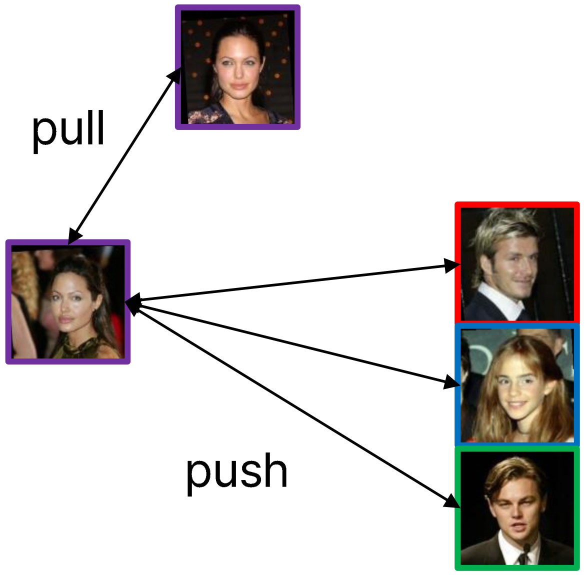
Introduction



Triplet

- Close to one positive sample
- Away from one negative sample

Introduction



Tuplet

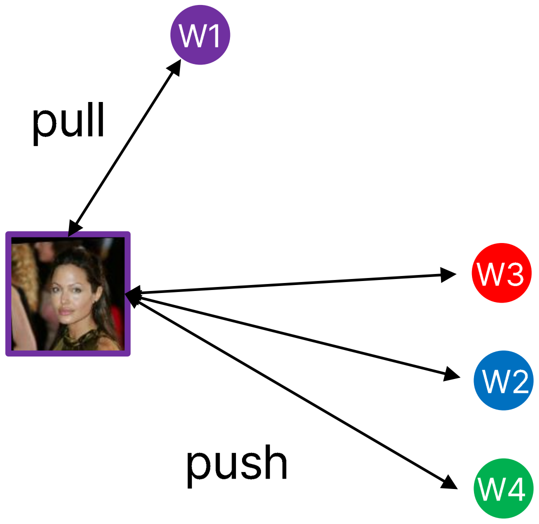
- Close to one positive sample
- Away from <u>multiple</u> negative samples

Introduction

Sample-to-sample comparisons

Combinational explosions on large-scale datasets

Introduction



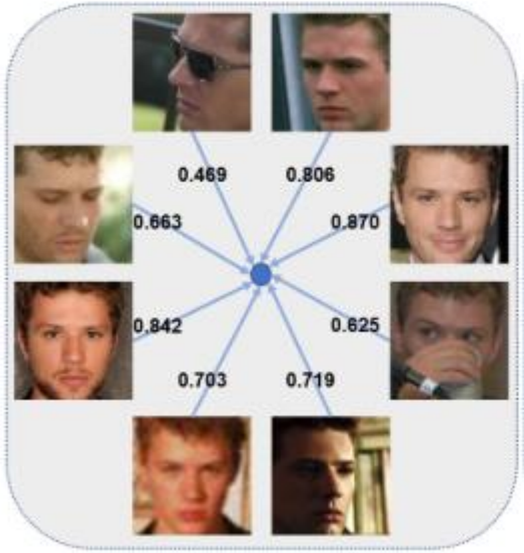
Sample-to-prototype comparison

- Close to one positive prototype
- Away from multiple negative prototypes

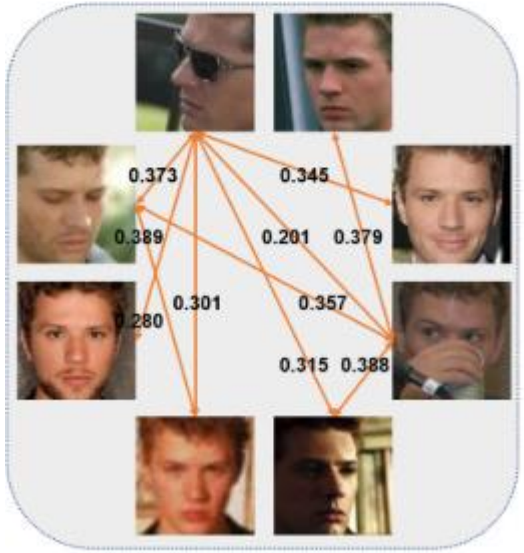
- More efficient and stable

Introduction

- Limitation of Prototype Learning



Sample-to-prototype
High similarities

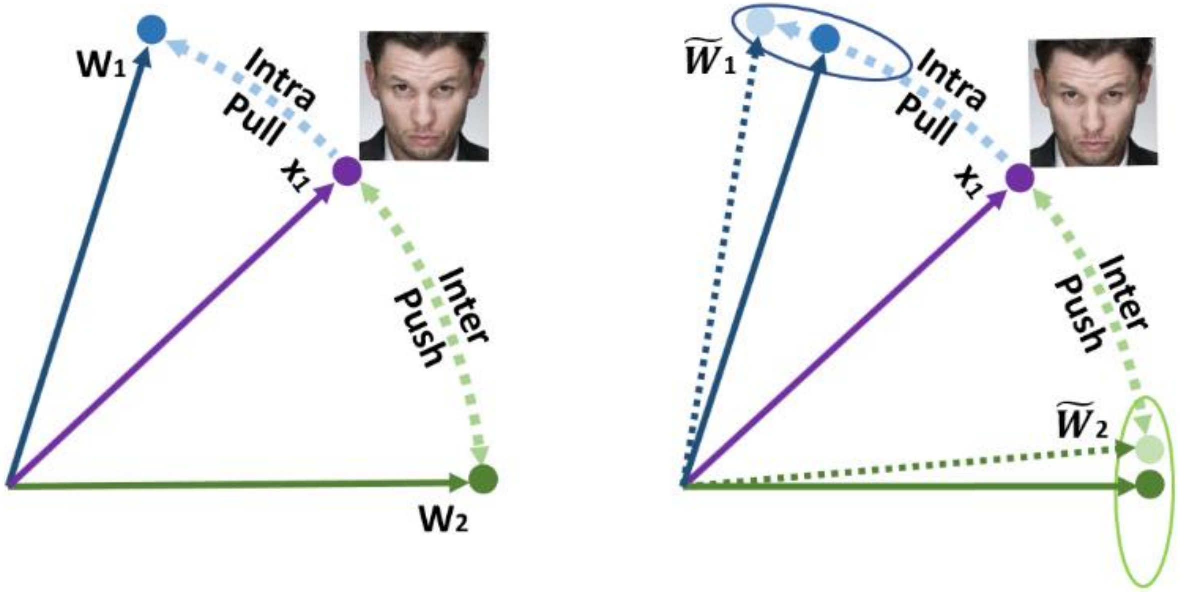


Sample-to-sample
Low similarities

Motivation

- Variational Prototype Learning

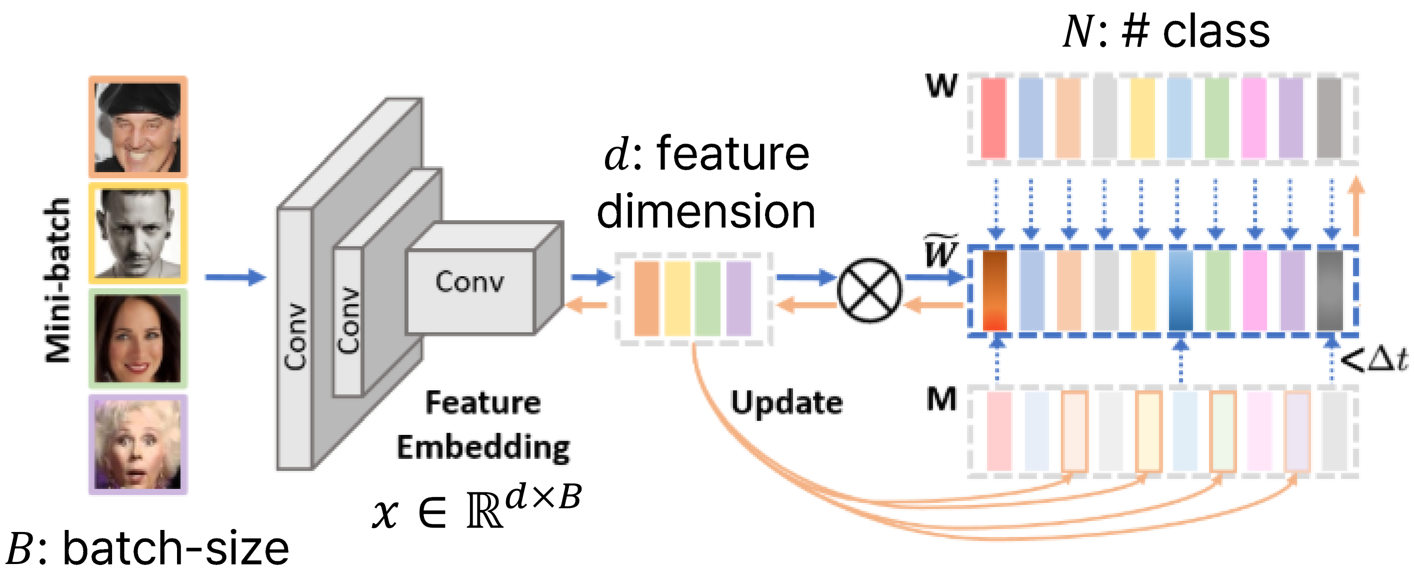
Representation of every class as a <u>distribution</u> instead of a point



(a) Prototype Learning (b) Variational Prototype Learning

Method

- Feature Injection with Memory Bank



Prototypes $W \in \mathbb{R}^{d \times N}$

$\lambda_1 \downarrow$

Variational Prototype $\widetilde{W} \in \mathbb{R}^{d \times N}$

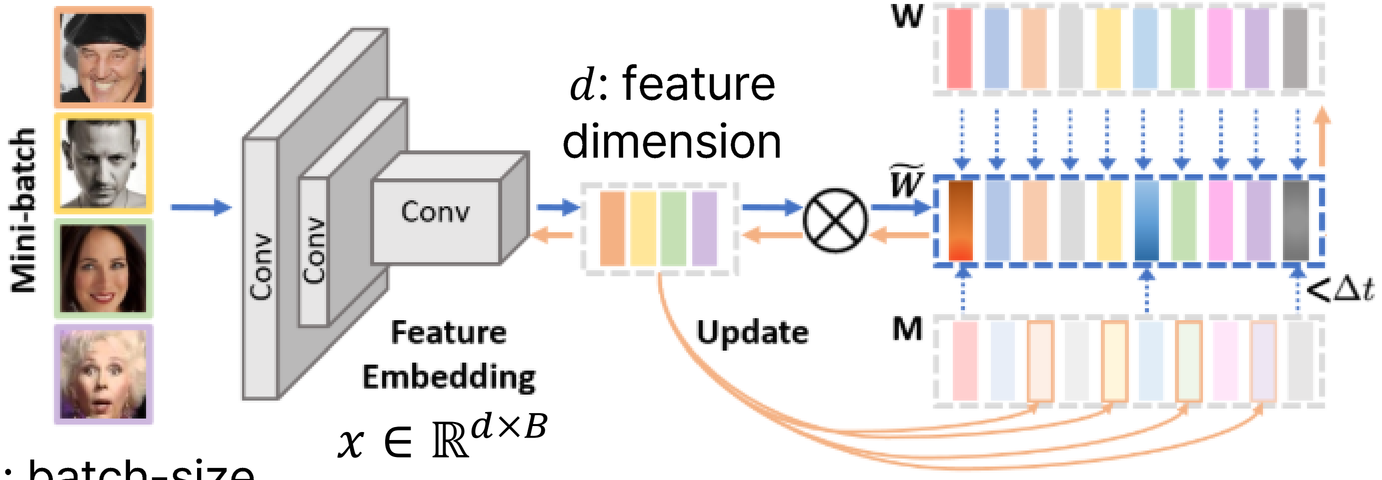
$\lambda_2 \uparrow$

Memorized samples $M \in \mathbb{R}^{d \times N}$

Method

- Feature Injection with Memory Bank

N: # class

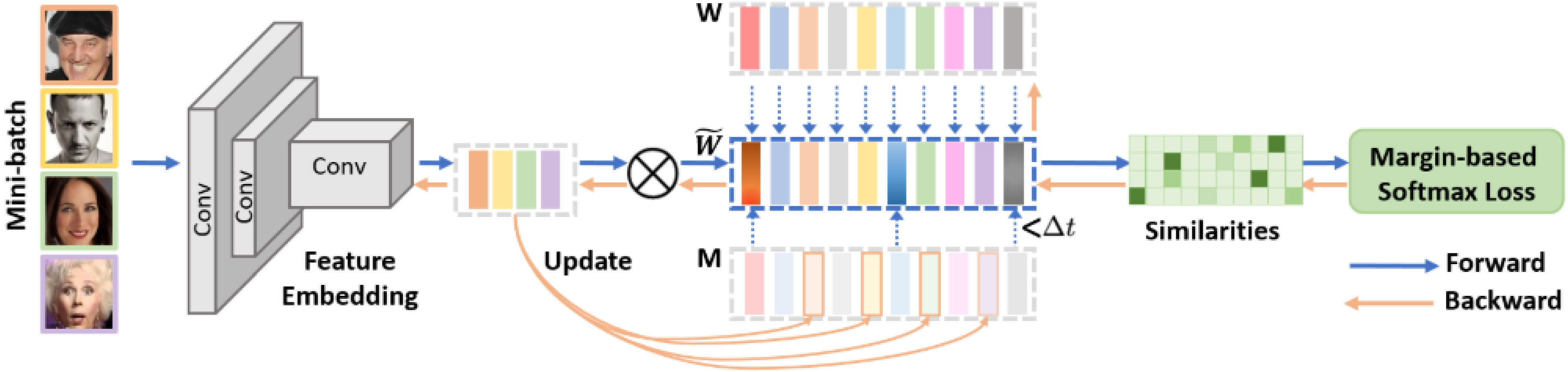


B: batch-size

Δt: # of iterations to stay in the memory bank for each sample

Method

- Variational Prototype Learning



Method

- Softmax Loss

$$\mathcal{L}_{VPL} = -\log \frac{e^{\widetilde{W}_{y_i}^{T} x_i}}{e^{\widetilde{W}_{y_i}^{T} x_i} + \sum_{j=1, j=y_i}^{N} e^{\widetilde{W}_{j}^{T} x_i}}$$

  - i-th sample belonging to the y_i-th class
  - Positive Prototype $\widetilde{W}_{y_i}^{T}$
  - Negative Prototype $\widetilde{W}_{j}^{T}$

KAIST

Method

- Margin-based Softmax Loss

$$\mathcal{L}_{VPL-Arc} = -\log \frac{e^{s\cos(\tilde{\theta}_{y_i}+m)}}{e^{s\cos(\tilde{\theta}_{y_i}+m)} + \sum_{j=1,j=y_i}^{N} e^{s\cos\tilde{\theta}_j}}$$

- $\tilde{W}_j^T x_i = \|\tilde{W}_j\|\|x_i\| \cos\tilde{\theta}_j$
- $\tilde{\theta}_{y_i}$: Angle between the feature x_i and positive variational prototype $\tilde{W}_{y_i}$
- $\tilde{\theta}_j$: Angle between the feature x_i and negative variational prototype $\tilde{W}_j$
- m: additive angular margin
- s: scaling parameter

Experiments

Method	Verification Accuracy					IJB		MegaFace	
	LFW	CFP-FP	CPLFW	AgeDB	CALFW	IJB-B	IJB-C	Id	Ver
CosFace(0.35) [45] (CVPR18)	99.81	98.12	92.28	98.11	95.76	94.80	96.37	97.91	97.91
ArcFace(0.5) [8] (CVPR19)	99.83	98.27	92.08	98.28	95.45	94.25	96.03	98.35	98.48
AFRN [22] (ICCV19)	99.85	95.56	93.48	95.35	96.30	88.5	93.0	-	-
MV-Softmax [50] (AAAI20)	99.80	98.28	92.83	97.95	96.10	93.6	95.2	97.76	97.80
GroupFace [24] (CVPR20)	99.85	98.63	93.17	98.28	96.20	94.93	96.26	98.74	98.79
CircleLoss [41] (CVPR20)	99.73	96.02	-	-	-	-	93.95	98.50	98.73
DUL [4] (CVPR20)	99.83	98.78	-	-	-	-	94.61	98.60	-
CurricularFace [19] (CVPR20)	99.80	98.37	93.13	98.32	96.20	94.8	96.1	98.71	98.64
URFace [39] (CVPR20)	99.78	98.64	-	-	-	-	96.6	-	-
DB [2] (CVPR20)	99.78	-	92.63	97.90	96.08	-	-	96.35	96.56
Sub-center ArcFace [7](ECCV20)	99.80	98.80	-	98.31	-	94.94	96.28	98.16	98.36
BroadFace [25] (ECCV20)	99.85	98.63	93.17	98.38	96.20	94.97	96.38	98.70	98.95
SST [11](ECCV20)	99.75	95.10	88.35	97.20	94.62	-	-	96.27	96.96
MS1M, R100, VPL-ArcFace	99.83	99.11	93.45	98.60	96.12	95.56	96.76	98.80	98.97

KAIST

Experiments

Base Model	Diff	LFW	CFP-FP	AgeDB	IJB-C
Softmax-Norm	PL	99.48	96.99	95.70	91.32
	VPL	99.65	97.56	96.23	92.54
CosFace [45]	PL	99.80	98.51	97.96	96.18
	VPL	99.81	98.81	98.24	96.52
ArcFace [8]	PL	99.78	98.54	98.05	96.21
	VPL	99.83	98.96	98.38	96.61
AdaptiveFace [27]	PL	99.80	98.62	98.08	96.28
	VPL	99.83	98.98	98.36	96.62
CurricularFace [19]	PL	99.80	98.58	98.10	96.30
	VPL	99.83	99.01	98.38	96.65

The VPL improves the accuracy in all cases.

Conclusion

- To propose a novel **Variational Prototype Learning (VPL)** method which represents each class as a distribution instead of a point in the latent space

- To design **computationally efficient and memory-saving** way for the variational prototype sampling

- Extensive experimental results demonstrate the **superiority of proposed VPL over the SoTA competitors** in deep face recognition

Strength and Weakness

Strength

- Orthogonal improvement with negligible extra memory and computation cost.
- Significant improvement when long-tail prototypes are variational.

Weakness

- The explanation for why certain hyper-parameters that determine the quantity of features to inject are effective is not provided.

Thank you