ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

Nonlinear Ranking Loss on Riemannian Potato Embedding

Byung Hyung Kim (KAIST)
Yoon-Je Suh (KAIST)
Honggu Lee (Looxid Labs)
Sungho Jo (KAIST)

KAIST
School of Computing

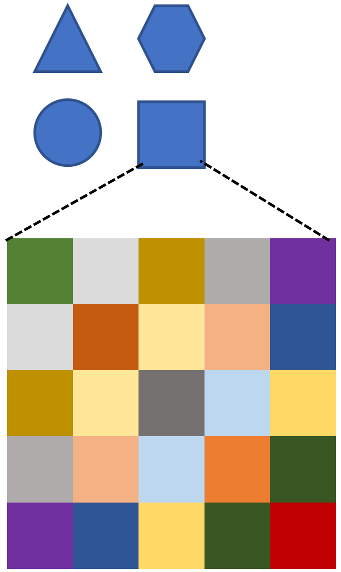
Introduction

- Goal : Aims to pull positive points closer than the potato-shaped region of acceptability (z-score) and push negative points out of the boundary.

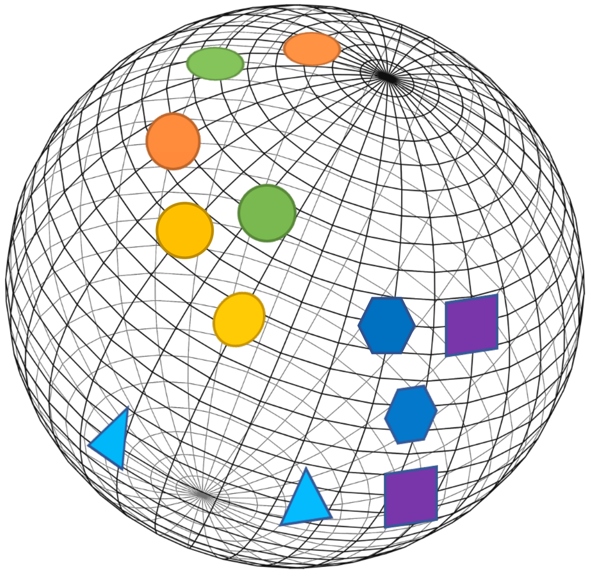
- Contribution
  - We propose a new ranking loss to learn discriminative embeddings on non-Euclidean spaces exploiting the structure of non-linear embedding spaces.
  - We achieve new state-of-the-art performance on three popular benchmarks, reducing the intra-class distances and enlarging the inter-class distances for the learned features.

Background

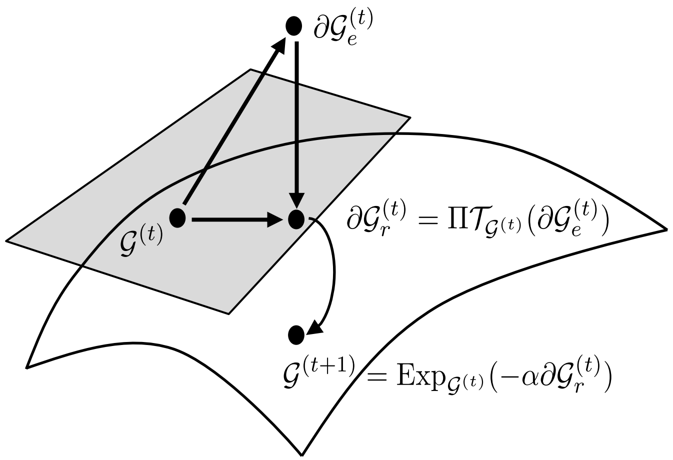
Covariance matrix representation



Riemannian Geometry over a SPD manifold



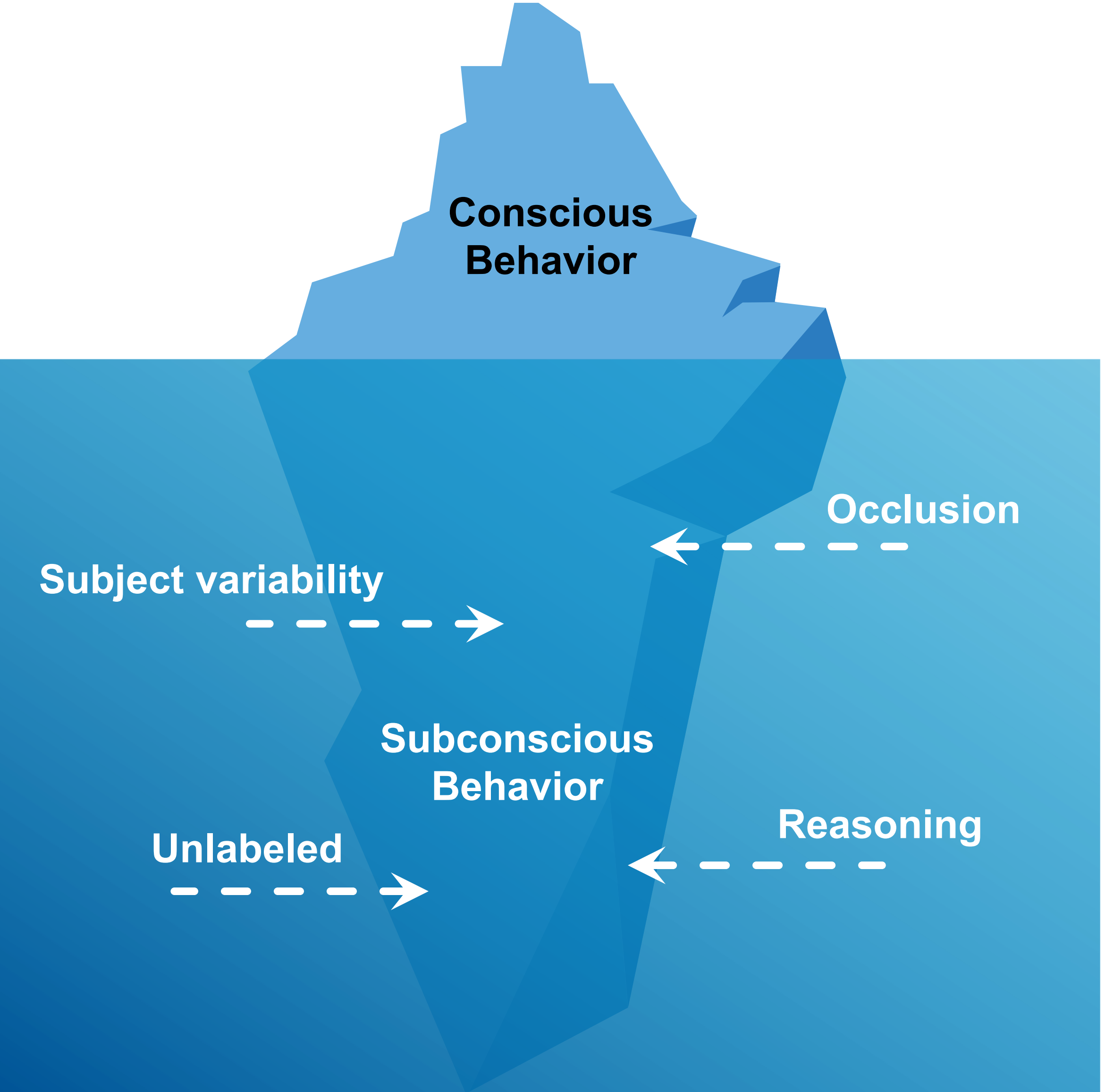
Tangent Approximation

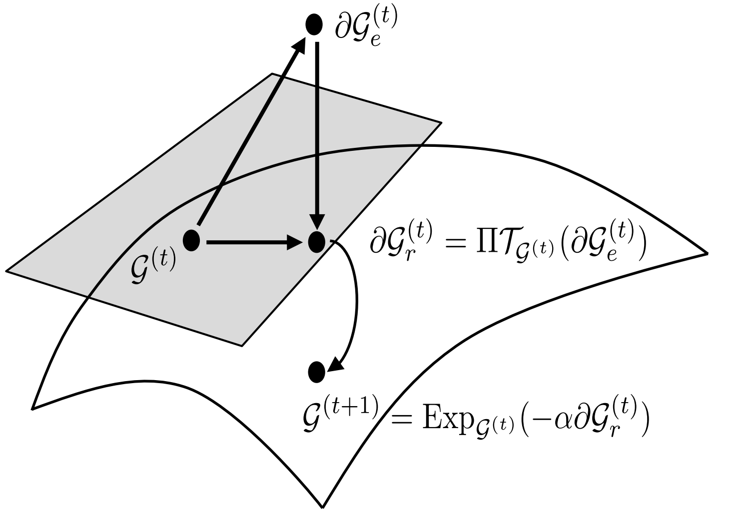


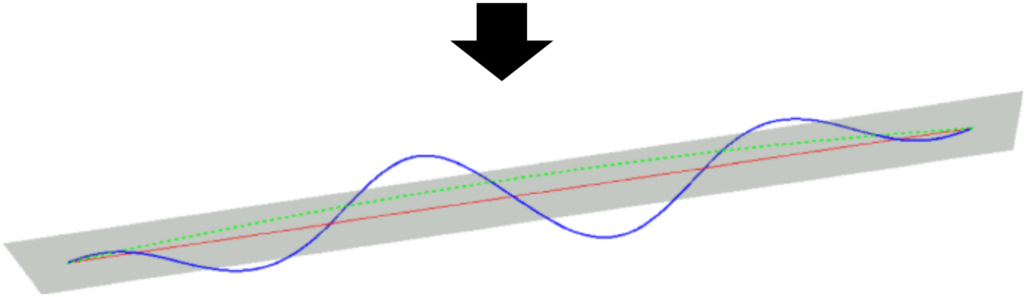
Background

ICPR 2020
25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION
Milan, Italy 10 | 15 January 2021

KAIST
School of Computing



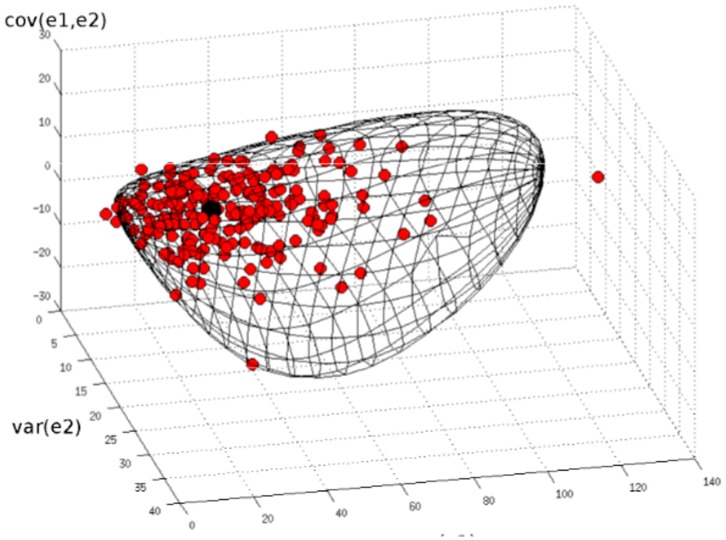




$$L(X, y) = \frac{1}{\mathcal{T}} \sum_{(i,j,k)\in\mathcal{T}} [d^2_{(i,j)} - d^2_{(i,k)} + \alpha]_+$$

Background

- Motivation
  - Riemannian Potato : Provides a measure of dispersion using the distribution of distances between covariance matrices and a reference matrix and rejects epochs whose covariance matrices lie out of a region of acceptability defined by a z-score threshold.



$$z_t = \frac{\log(d_t/\mu_t)}{\log(\sigma_t)} \qquad d_t = \delta_R(\Sigma_t, \bar{\Sigma}_{t-1})$$

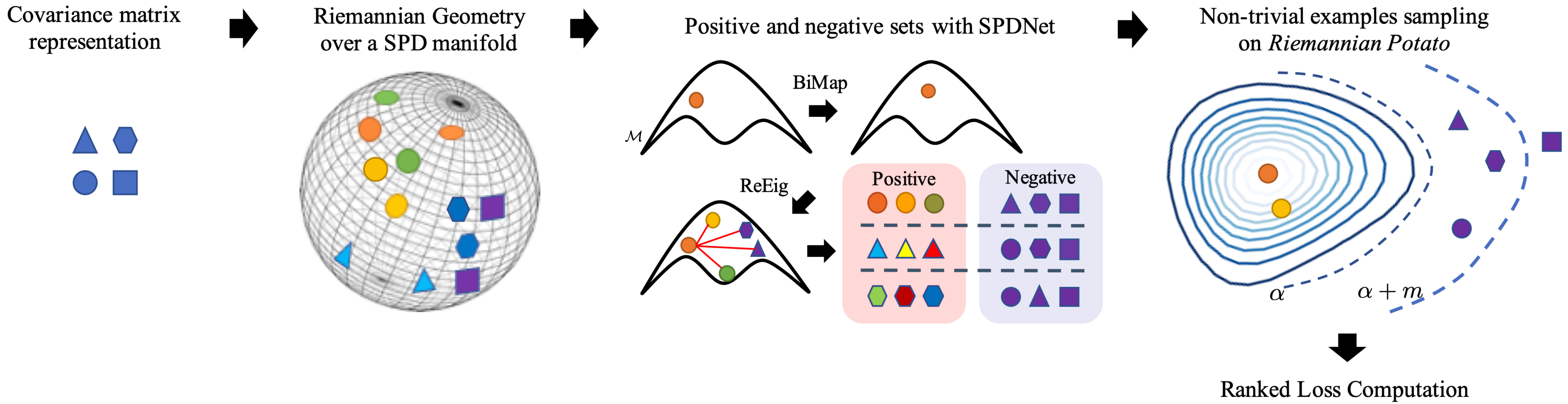
$$\bar{\Sigma} = \underset{\Sigma \in \mathcal{M}}{\arg\min} \sum_{i=1}^{N_I} \delta_R^2(\Sigma_i, \Sigma)$$

$$\mu_t = \exp\left(\frac{1}{t}\sum_{i=1}^{t}\log(d_i)\right),$$

$$\sigma_t = \exp\left(\sqrt{\frac{1}{t}\sum_{i=1}^{t}(\log(d_i/\mu_i))^2}\right)$$

Methodology

- Riemannian Potato-based Ranking Loss (RPL)
  - A rank-based metric learning method for learning discriminative embeddings

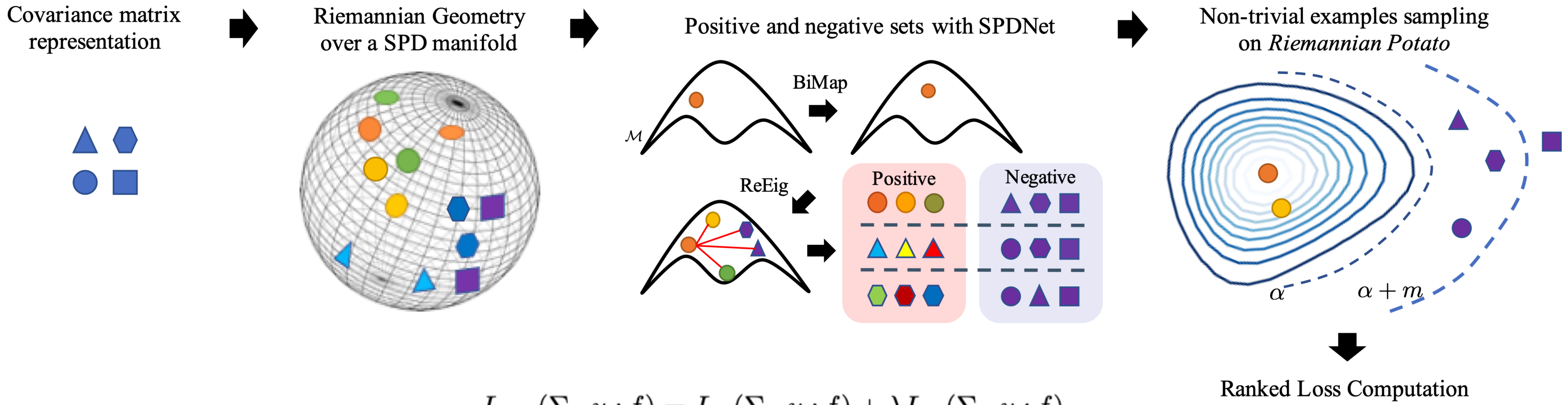


$$\mathcal{P}_i^c = \{\forall \Sigma_j | j \neq i \wedge c = y_i = y_j\} \quad \hat{\mathcal{P}}_i^c = \{\forall \Sigma_j | j \neq i \wedge c = y_i = y_j, z_j^c > z_{th}\} \quad L(\Sigma_i, \Sigma_j, y_i; f) = (1 - y_{ij})[z_{th} - z_j^c)]$$

$$\mathcal{N}_i^c = \{\forall \Sigma_j | c = y_i, y_i \neq y_j\} \quad \hat{\mathcal{N}}_i^c = \{\forall \Sigma_j | c = y_i, y_i \neq y_j, z_j^c < z_{th} + m\} \quad + y_{ij}[z_j^c - (z_{th} + m)]_+,$$

Methodology

- Riemannian Potato-based Ranking Loss (RPL)
  - A rank-based metric learning method for learning discriminative embeddings



$$L_{RP}(\Sigma_i, y_i; f) = L_P(\Sigma_i, y_i; f) + \lambda L_N(\Sigma_i, y_i; f).$$

$$L_P(\Sigma_i, y_i; f) = \frac{1}{|\hat{\mathcal{P}}_i^c|} \sum_{\Sigma_j \in \hat{\mathcal{P}}_i^c} L(\Sigma_i, \Sigma_j, y_i; f) \qquad L_N(\Sigma_i, y_i; f) = \frac{1}{|\hat{\mathcal{N}}_i^c|} \sum_{\Sigma_j \in \hat{\mathcal{N}}_i^c} L(\Sigma_i, \Sigma_j, y_i; f)$$

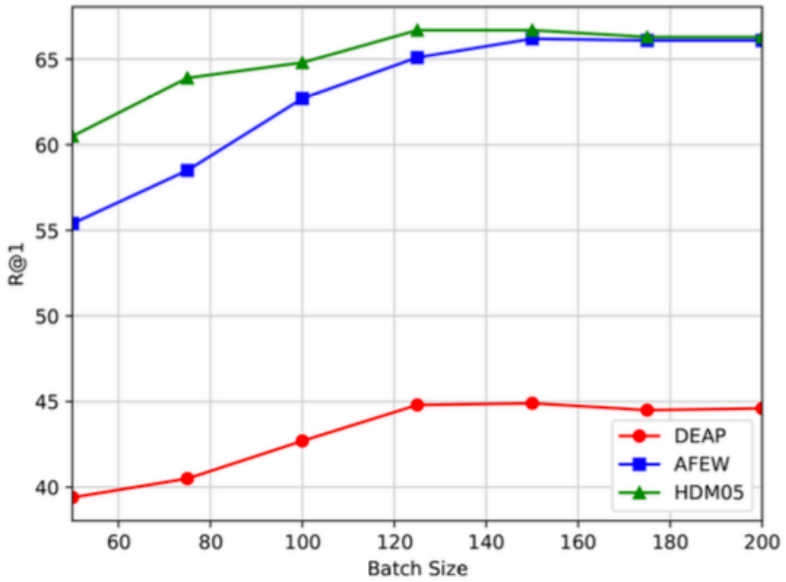
Results

- Public Datasets
  - Emotion Recognition (DEAP, AFEW), Skeleton-based Human Action Recognition (HDM05)

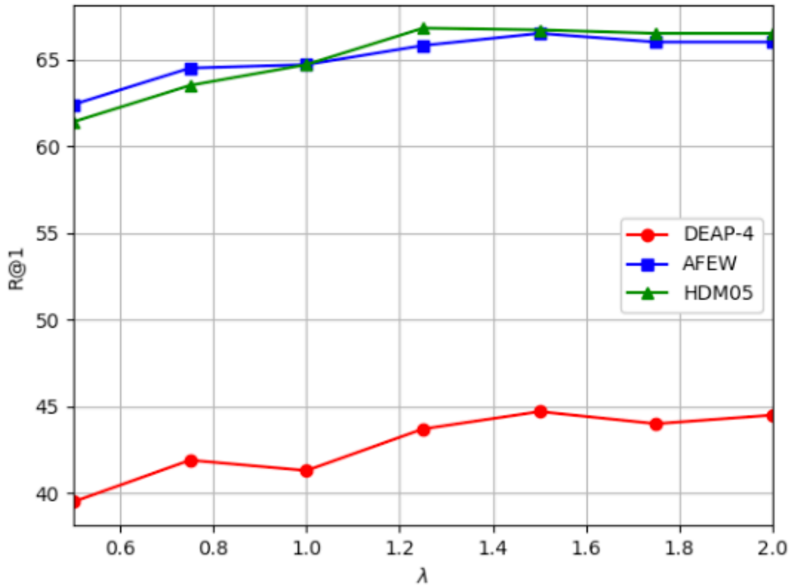
Method	DEAP-4				AFEW				HDM05			
	F1	R@1	R@3	NMI	F1	R@1	R@3	NMI	F1	R@1	R@3	NMI
DSML-Triplet	38.7	35.5	37.8	36.1	59.3	61.3	64.2	61.9	53	57.3	58.5	52.4
Triplet-Random	33.5	31.4	32.5	28.7	55.8	55.1	55.4	57.4	48.3	44.5	47.5	45.6
Triplet-Semihard	35.5	30.1	31.4	27.3	57.4	54.4	60.5	58	51.3	50.4	55.3	47.5
Lifted Struct	35.3	35.2	35.8	33.4	62.5	65.4	68.4	69.4	55.5	59.3	59.2	53.4
N-pair-mc	41.5	38.4	39.5	34.8	66.4	63.5	66.4	65.1	59.8	60	61	59.4
Proxy NCA	39.8	41.3	41.4	**38.1**	66.5	64.2	66	66.3	59	63.3	64.5	62
NRA	42.2	44.4	**46.2**	37.2	67.2	**66.5**	68.6	**66.8**	59.2	64.3	65.2	64.1
RPL	**43.3**	**44.7**	**46.2**	37.5	**67.4**	**66.5**	**69.4**	66.4	**59.4**	**66.7**	**68.8**	**65.4**

Results

- The effect of RPL
  - Batch size, Parameter λ



Recall@1 results of different batch size on the three datasets (DEAP, AFEW, and HDM05).



Recall@1 results of different λ on the three datasets (DEAP, AFEW, and HDM05).

Conclusion

ICPR 2020 25th INTERNATIONAL CONFERENCE ON PATTERN RECOGNITION Milan, Italy 10 | 15 January 2021

KAIST School of Computing

- A rank-based metric learning method for learning discriminative embeddings
  - Given a query covariance matrix, our RPL splits its positive and negative sets and forces a margin between them on a SPD manifold.
  - Non-trivial samples mining and negative examples weighting are exploited to make better use of all informative data points.
  - The proposed method achieves state-of-the-art performance, reducing the intra-class distances and enlarging the inter-class distances for learned features.
- Our next work will study the non-stationary nature of brain activity as revealed by EEG
  - subject to noises from various artifacts, low signal-to-noise ratio (SNR) of sensors, and inter- and intra- subject variability.

Thank you