SoftMaxout Transformation-Permutation Network

For Facial Template Protection

Hakyoung Lee, Cheng Yaw Low

2019311654

ADVICING PROFESSOR : ANDREW BENG-JIN TEOH

YONSEI UNIVERSITY

INTRODUCTION

BACKGROUND

- **Biometric template** is required in many applications, facilities, and personal systems.
- Biometric template has many advantages:
  i. Can identify the individual in high accuracy
  ii. Less time consuming
  iii. Easy to use and inexpensive.
- However, once the biometric template is taken by the attacker, the security will be lost forever because we cannot change the biometric template
- Thus, powerful protection scheme is required to ensure the security.

INTRODUCTION
MOTIVATION

- **Cancellable biometrics** is one of the proposed method to address the security issue.
- It refers to the non-invertible transformation that changes the original biometric templates
- Cancellable biometric must follow 4 criteria:
  i. Non-invertibility : It must not be recovered to the original template.
  ii. Revocability : It should be easily replaced whenever it needed.
  iii. Unlinkability : Two or more templates generated from the same identity must have no linkage.
  iv. Performance Preservation : It should preserve the performance of the original template.

- Motivated by Random Permutation Maxout (RPM):

  [20] Teoh, Andrew Beng Jin, Sejung Cho, and Jihyeon Kim. "Random permutation Maxout transform for cancellable facial template protection." Multimedia Tools and Applications 77.21 (2018): 27733-27759.
- The RPM transforms a biometric feature vector into a discrete hash code by localizing the maximal entries of the truncated and permuted original template
- However, the accuracy performance on RPM transform is unlikely to attain decent performance

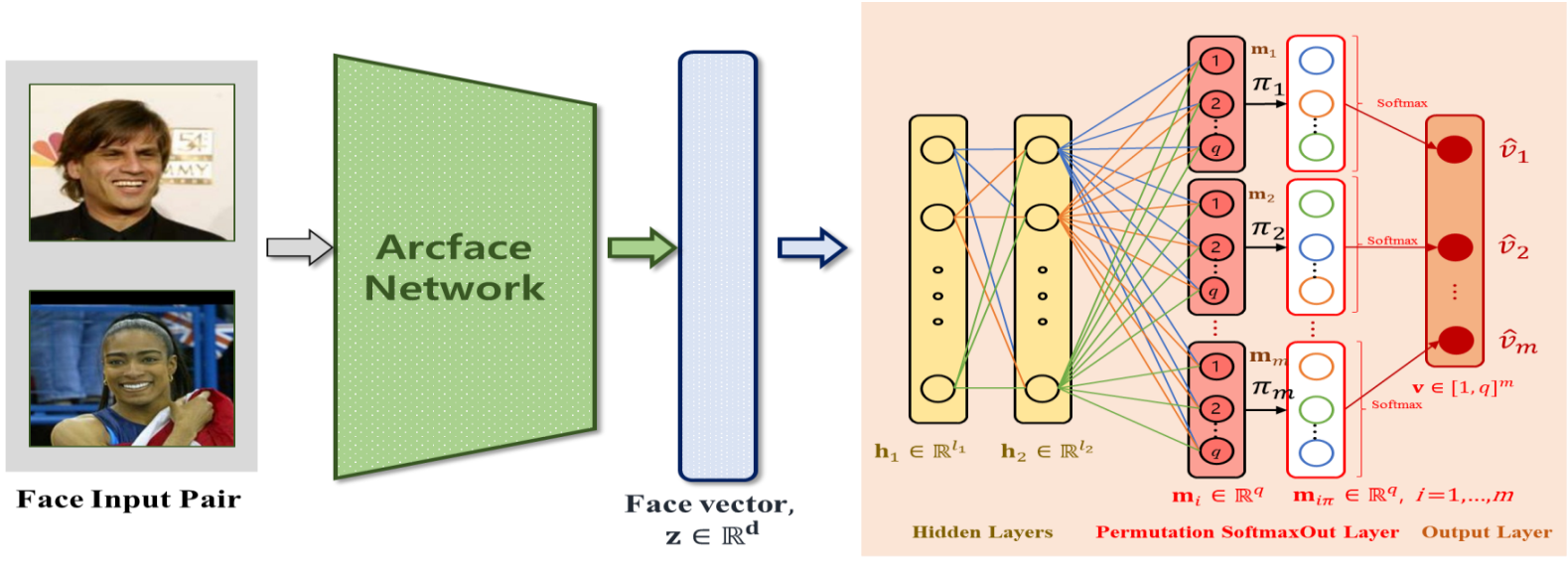
INTRODUCTION

CONTRIBUTIONS

- Our contributions are three-fold:
  i. Softmaxout Permutation-Transformation Network(SOTPN) inspired by RPM transformation
  ii. Satisfy four design criteria of cancellable biometrics, i.e. noninvertible, renewability, unlinkability and accuracy performance
  iii. Pairwise arcface loss function which is inspired by the Arcface loss function [2] and code-balancing loss that triggers hash code to have high entropy.

PROPOSED METHOD

SOFTMAXOUT TRANSFORMATION-PERMUTATION NETWORK



SoftmaxOut Transformation-Permutation Network

- Major component of the SDTN is **Softmaxout Transformation-Permutation Network (SOTPN).**
- Pipeline of the Softmaxout Transformation-Permutation Network:
  i. Feature extraction network (Arcface network).
  ii. Two hidden layer $\mathbf{h}_1 \in \mathbb{R}^{l_1}$ and $\mathbf{h}_2 \in \mathbb{R}^{l_2}$ is configured with l_1 and l_2 neurons, respectively.
  iii. Permutable Softmaxout layer that approximates index of the maximum value.
  iv. Discrete m-dimensional hash code.

PROPOSED METHOD

SOFTMAXOUT TRANSFORMATION-PERMUTATION NETWORK

- Softmaxout layer is designed to approximate the index value of the maximum entry of a q-dimensional permuted vector. This process can be summarized as follows:

$$v_i = \arg\max_q \mathbf{m}_i \in \{1, \dots, q\},\ i = 1, \dots, m$$

- This process is not differentiable and not trainable with backpropagation
- In order to makes the network trainable, we suggest the differentiable Softmaxout layer with the following function:

$$v_i \approx \tilde{v}_i = \sum_{j=1}^{q} j\, \sigma_\beta(\mathbf{m}_i)\ \in \{1, \dots, q\}$$

- $\sigma_\beta()$ is the Softmax function parameterized with $\beta > 1$:

$$\sigma_\beta(v) = \frac{e^{\beta v}}{\sum_{i=1}^{q} e^{\beta v_i}}$$

PROPOSED METHOD

PAIRWISE ARCFACE LOSS

- Arcface loss is the classification-based loss that requires an explicit classification layer, of which its capacity and computational cost increases linearly with respect to the number of identities
- Thus, we modified the Arcface loss to fit to pair-wise loss function as:

$$PA_{ij} = -\log\left[s_{ij}\left(\frac{e^{\gamma\cos(\theta+\alpha)}}{e^{\gamma\cos(\theta+\alpha)}+e^{\gamma\sin\theta}}\right)\right] - \log\left[(1-s_{ij})\left(\frac{e^{\gamma\sin(\theta+\alpha)}}{e^{\gamma\sin(\theta+\alpha)}+e^{\gamma\cos\theta}}\right)\right]$$

- positive pairs : s_{ij} = 1, negative pairs : s_{ij} = 0
- α : angular margin, γ : scaling factor
- θ is the angle between $\hat{\mathbf{v}}_i$ and $\hat{\mathbf{v}}_j$ or $\theta = \cos^{-1}\left(\hat{\mathbf{v}}_i^T\hat{\mathbf{v}}_j\right)$, given that $\hat{\mathbf{v}}$ is the L2 normalized vector to be re-scaled with respect to γ. L2-normalization makes the similarity measure only relies on the angular between two vector as the original arcface loss.

PROPOSED METHOD

PAIRWISE ARCFACE LOSS

- Code balancing loss is applied to make the entropy of the hash code high.
- It is accomplished by imposing the occurrence probability be 1/q for each code
- We define the code balancing loss function as:

$$CB_{ij} = \sum_{b=1}^{B} \sum_{k=1}^{m} \{ \left| \, mean\left(v_i^{bk}\right) - \frac{q+1}{2} \right| + \left| \, mean\left(v_j^{bk}\right) - \frac{q+1}{2} \right| \}$$

- mean() : average operator

EXPERIMENTS AND ANALYSIS

DATASETS

- Three datasets are used in the following experiments

  i. LFW (labeled Faces in the wild) is a widely used face dataset collected from the web. It is mainly used to study face recognition and verification in unconstrained condition. The LFW face dataset has 13,233 images with 5,749 identities.

  ii. YTF (YouTube Faces) dataset contains video frames of 1,595 different identities. Every video frame is collected from YouTube and it has 3,245 videos in total.

  iii. FS (Facescrub) dataset contains 106,863 face images of 530 celebrities with approximately 200 images per person.

- Following the LFW standard evaluation protocol, we apply the pre-determined 3,000 matched pairs and 3,000 non-matched pairs for verification tasks.

EXPERIMENTS AND ANALYSIS

PARAMETER ANALYSIS

- The EER of the SOTPN with respect to different setting of m and q

Table 5.1. EER (%) of SoftmaxOut layer with various $\boldsymbol{m}$ by fixing $\boldsymbol{q}$ =32

Training/Testing	m				
	32	64	128	256	512
FS/LFW	9.47	6.20	5.00	4.23	3.97
FS/YTF	17.44	12.80	11.40	9.72	9.12
LFW/FS	10.06	5.60	4.04	3.30	3.22

Table 5.2. EER (%) of SoftmaxOut layer with various $\boldsymbol{q}$ by fixing $\boldsymbol{m}$ =512

Training/Testing	q				
	8	16	32	64	128
FS/LFW	3.90	3.83	3.97	3.73	4.77
FS/YTF	11.60	11.76	9.12	9.86	9.50
LFW/FS	3.12	3.00	3.22	3.24	3.48

EXPERIMENTS AND ANALYSIS

PARAMETER ANALYSIS

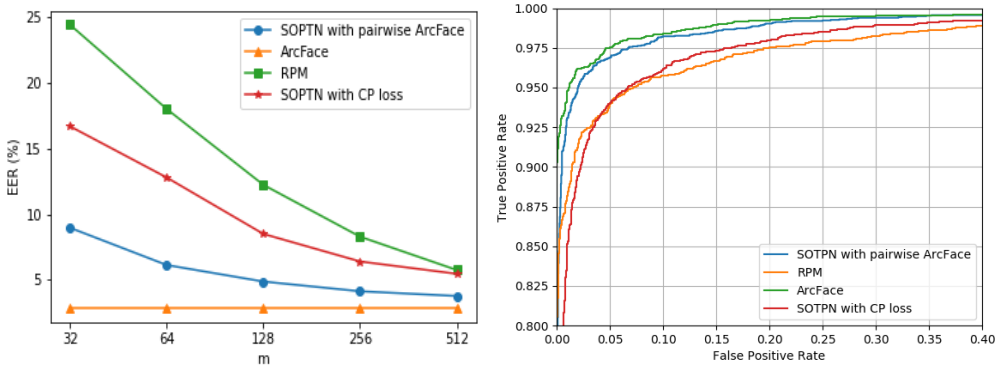


Fig. 5.1 (a) EER vs m; (b) ROC curves (trained with FS and tested with LFW)

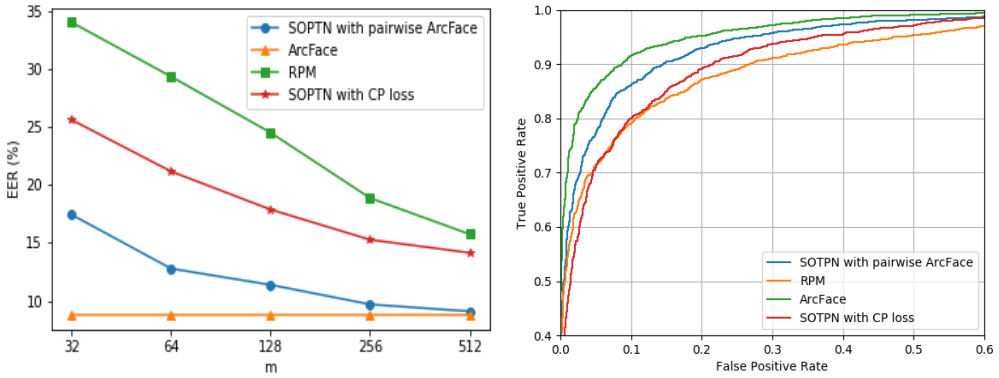


Fig. 5.2 (a) EER vs m; (b) ROC curves (trained with FS and tested with YTF)

EXPERIMENTS AND ANALYSIS

ABLATION STUDY

- Ablation study for the proposed SDTN trained with respect to the FS dataset for the best hyperparameter from previous experiment $m = 512$ and $q = 32$, and our evaluation is performed on the LFW dataset.
- The effect of the code balancing loss, quantization and random weight activation can be found on here.

SOTPN Configuration	EER (%)
PA loss with $\beta = 9$ + CB Loss	**3.53**
PA loss with $\beta = 9$, without CB loss	4.07
PA loss with $\beta = 1$ + CB loss	4.50
PA loss with $\beta = 1$, without CB loss	5.70