FD-GAN: Pose-guided Feature Distilling GAN for Robust Person Re-identification

Yixiao Ge[1*], Zhuowan Li[2,3*], Haiyu Zhao[2], Guojun Yin[2,4], Shuai Yi[2], Xiaogang Wang[1], Hongsheng Li[1]

[1]The Chinese University of Hong Kong [2]SenseTime Research [3]Johns Hopkins University [4]University of Science and Technology of China

The Chinese University of Hong Kong

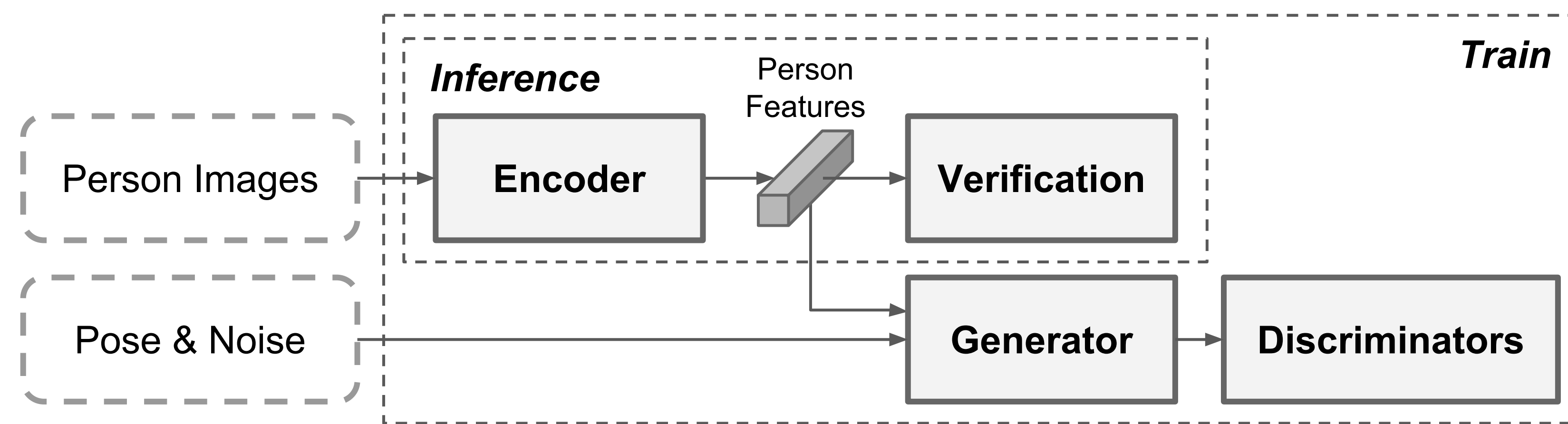Neural Information Processing Systems Foundation

Motivation

- Posture variations, blur and occlusion pose great challenges for learning discriminative person features.
- Existing works which attempted to address the above issues require auxiliary pose information and more computational cost in the inference stage.
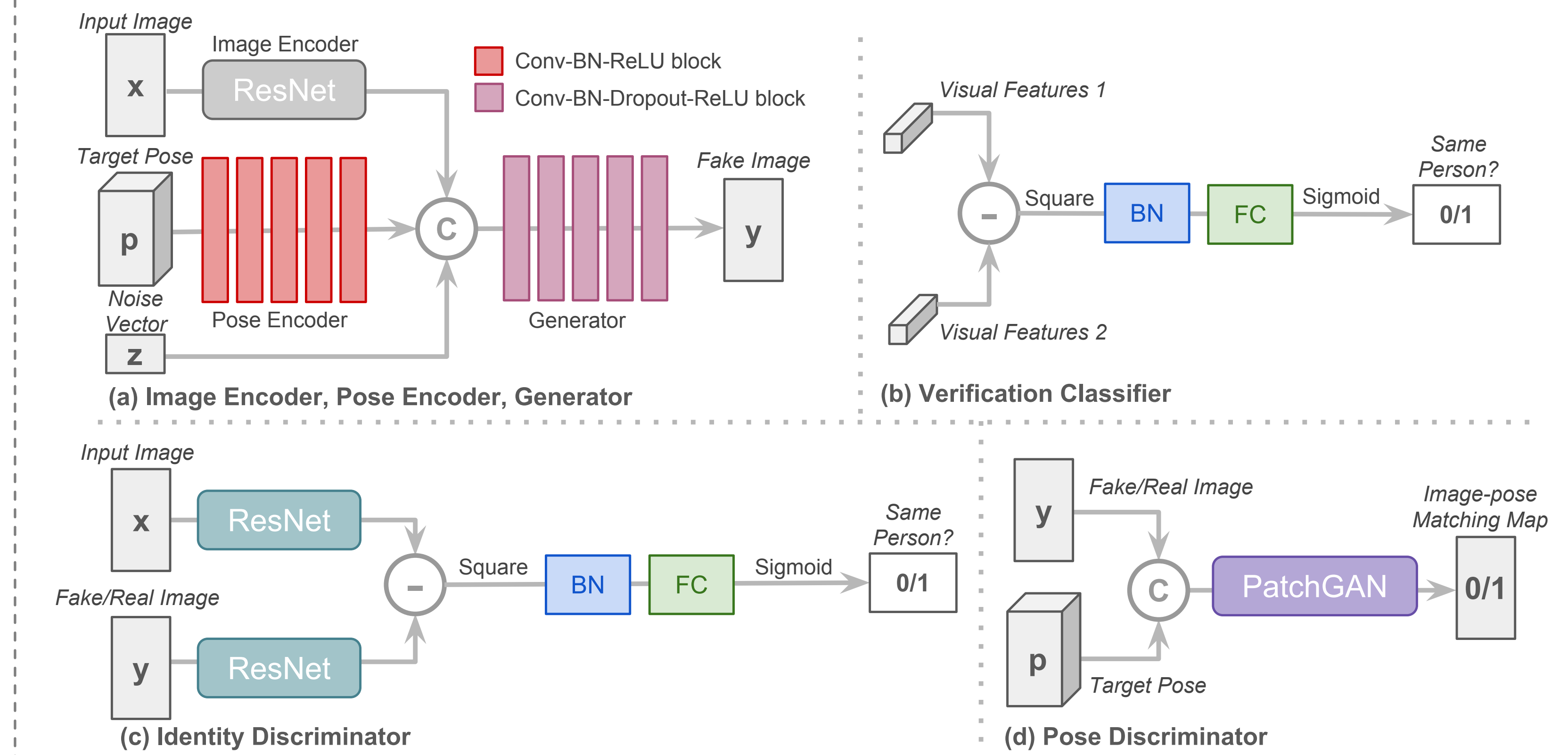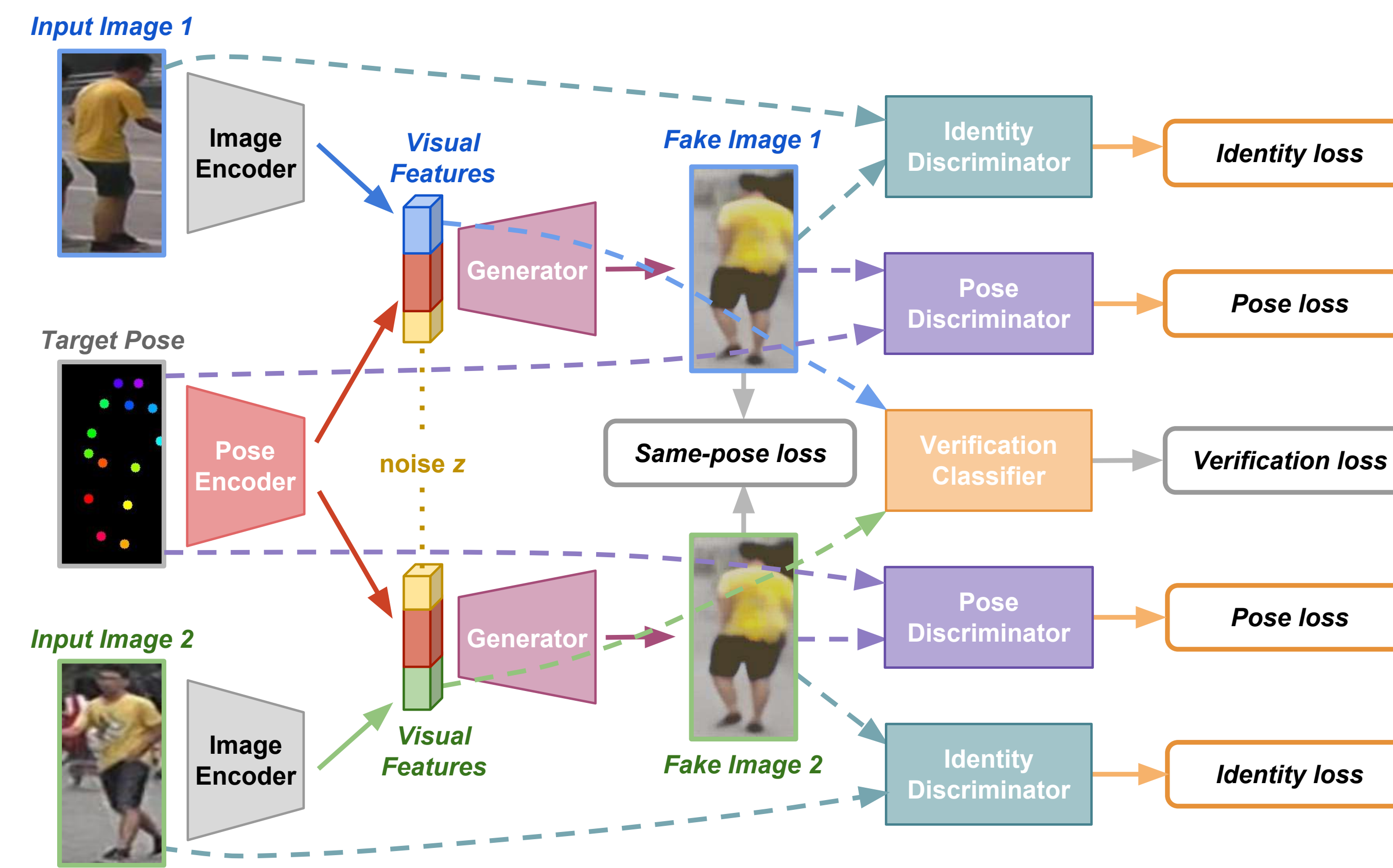

Contributions

- **Representation learning.** The proposed FD-GAN learns identity-related and pose-unrelated representations for person re-identification with pose-variation, not requiring extra auxiliary pose information or increase the computational complexity during inference.
- **Person image generation.** Although it is an auxiliary task for our framework, the generated person images by our proposed method show better quality than existing specific person-generation methods.

Feature Distilling Generative Adversarial Network

(a) Image Encoder, Pose Encoder, Generator
(b) Verification Classifier
(c) Identity Discriminator
(d) Pose Discriminator

The proposed FD-GAN adopts a Siamese structure, taking a person image and a target pose landmark map as inputs for each branch. The assumption is that, if the learned person features are pose-unrelated and identity-related, then it can be used to accurately generate the same person's image but with different target poses.

- **Image encoder** transforms the input person image into feature representations.
- **Image generator** generates desired person images conditioned on the identity-related features from the image encoder and a target pose map.
- **Identity verification classifier** determines whether the two images belong to the same person, given the visual features of the two input images from the image encoder.
- **Identity discriminator** is trained to distinguish whether the generated person image and the input person image of the same branch belong to the same person.
- **Pose discriminator** is proposed to distinguish whether the generated person image matches the given target pose.

Person reID performance

Proposed Networks	Market-1501		DukeMTMC-reID		CUHK03	
	mAP (%)	top-1 (%)	mAP (%)	top-1 (%)	mAP (%)	top-1 (%)
Baseline (Siamese)	72.5	88.2	61.3	78.2	88.5	90.1
Our FD-GAN	**77.7**	**90.5**	**64.5**	**80.0**	**91.3**	**92.6**

Generation results