Víctor Ponce-López, Tilo Burdghart, Sion Hannuna, Dima Damen, Alessandro Masullo, Majid Mirmehdi

Dept. of Computer Science, Electric & Electronic Engineering, University of Bristol.
Merchant Venturers Building, Woodland Road, BS81UB, UK
vponcelop@gmail.com , {v.poncelopez, tb2935, sh1670, dima.damen, a.masullo, m.mirmehdi}@bristol.ac.uk

Semantically Controlled Augmentation for Deep Compact Person Re-Identification

ABSTRACT

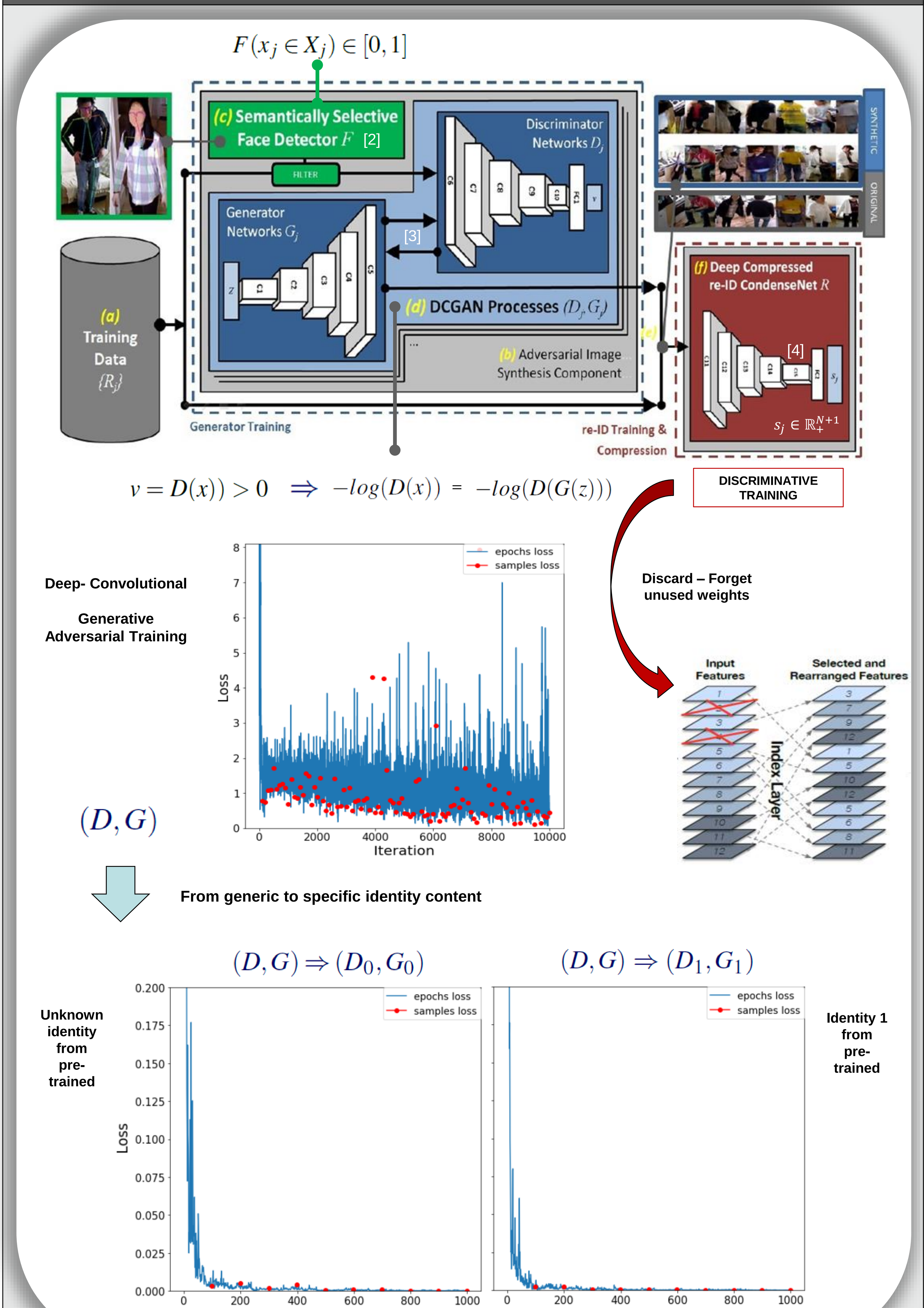
Deep person Re-ID approach that combines semantically selective, deep data augmentation with clustering-based network compression to generate high performance, light and fast inference networks. Augmented limited training data via sampling from DCGAN, whose discriminator is constrained by a semantic classifier to steer domain-specific adversarial synthesis and feed the ID-CondenseNet training. Variants on a number of datasets, outperforming SOTA for indoor long-term monitoring.

1. Motivation

- Person Re-Identification has potential impact applications:
  - CCVT Surveillance.
  - e-health for **indoor living environments**.

- CV & Practical Challenges:
  - **Unobserved** intervals, missing relevant **images**.
  - Varying **appearance, clothes**, low-**resolution**, non-**frontal** shot.
  - **Limited labelled** data.
  - Vast training **data pools** for **deep learning**,
  - **Computational demand** for network **inference**.

2. Methodology & Framework Overview



3. Experiments & Results

DCGAN – Selective Image Synthesis

(a) No Face Filtering (b) Face Filtering

INDOOR Re-ID RESULTS IN LIMA DATASET [5] - TEST ALL IDENTITIES

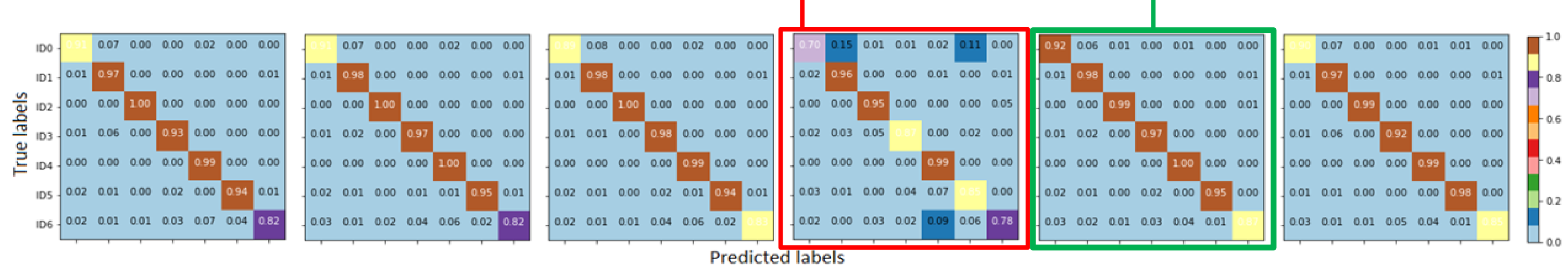
No Semantic Control	ALL prec@1	p-ID prec@1	ALL mAP	p-ID mAP
1: Baseline (M2&ME) [14]	89.1	-	-	-
2: No Augmentation (R)	91.98	93.49	90.90	96.28
3: Augmentation 24k$G \to R$	*92.43*	*94.27*	*91*	*96.95*
4: Augmentation 48k$G \to R$	91.74	93.48	90.61	96.54
Semantic Control via F	**ALL prec@1**	**p-ID prec@1**	**ALL mAP**	**p-ID mAP**
5: No Augmentation (FR)	82.02	92.14	72.90	95.48
6: Augmentation F322k$G \to R$	**92.58**	**94.57**	**91.14**	97.02
7: (24kG_0+F24kG_j) $\to R$	92.44	94.37	90.96	**97.04**

R → Deep CondenseNet
FR → Direct Semantic Control
G → Augmentation via DCGANs
$G_0 + G_j$ → Individual-specific Augmentation

Training restriction down to only 39% withholds critical identity-relevant information

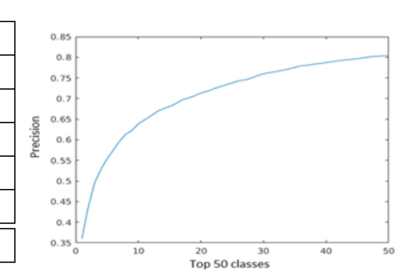
TEST ON FACE-FILTERED ONLY

Semantic Control via F	ALL prec@1	p-ID prec@1	ALL mAP	p-ID mAP
1. No Augmentation (R)	93.35	96.58	93.38	97.73
2. Augmentation F24kG → R	95.34	96.34	92.94	97.08
3. Augmentation F48kG → R	**96.17**	**97.38**	**93.58**	**97.88**



OUTDOOR Re-ID RESULTS IN DukeMTMC WITH Market1501 CROSS-SUPPORT

Method / No Semantic Control	prec@1	prec@5	mAP	CMC@1 S-Q	mAP S-Q
1: Baseline BoW + KISSME [31]	-	-	-	25.13	12.17
2: Baseline LOMO + XQDA [31]	-	-	-	30.75	17.04
3: No Augmentation (R)	87.70	95.54	87.79	*29.04*	*15.99*
4: Augmentation 24k$G \to R$	88.08	95.73	88.26	**36.45**	**21.11**
5: Transfer 24k(Market1501)$G \to R$	**88.84**	**95.82**	**88.64**	*35.95*	*20.6*
6: ResNet50+LSRO [33] (8x larger)	-	-	-	67.68	47.13



→ 8x more parameters and operations required: 600 million more operations to perform inference on a single image [4].

4. Conclusion & Future Hints

- Deep compact Re-ID approach with **semantically controlled data augmentation** and **compression** for **fast inference networks**.
- **Face-constrained DCGAN** sampling **outperform** SOTA on LIMA for **long-term monitoring**, and achieves **competitive Re-ID** performance.
- Explore **generic** semantic controllers for the **discriminator** networks.
- Learn **generators** from other **person-like representations**.

References

[1] V. Ponce-López, T. Burdghart, S. Hannuna, D. Damen, A. Masullo, M. Mirmehdi. **Semantically Selective Augmentation for Deep Contact Person Re-Identification**. arXiv:1806.04074 pre-print, 2018.
[2] T. Simon, H. Joo, I. Matthews, Y. Sheikh. **Hand Keypoint Detection in Single Images using Multiview Bootstrapping**. Computer Vision and Pattern Recognition Conference (CVPR), 2017.
[3] A. Radford, L. Metz, S. Chintala. **Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks**. International Conference on Learning Representations (ICLR), 2015.
[4] G. Huang, S. Liu, L.-v.-d. Maaten, K,-Q. Weinberger. **CondenseNet: An Efficient DenseNet using Learned Group Convolutions**. arXiv preprint arXiv:1711.09224, 2017.
[5] R. Layne, S. Hannuna, M. Camplani, J. Hall, T.-M. Hospedales, T. Xiang, M. Mirmehdi, D. Damen. **A dataset for persistent multi-target multi-camera tracking in RGB-D**. In CVPR Workshops, pages 1462–1470, 2017.
[31] L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang, Q. Tian. **Scalable person re-identification: A benchmark**. International Conference on Computer Vision (ICCV), pages 1116–1124, Dec 2015.
[33] Z. Zheng, L. Zheng, Y. Yang. **Unlabeled samples generated by GAN improve the person re-identification baseline *in vitro***. International Conference on Computer Vision (ICCV), pages 3754–3762, 2017.

@poncevi