The 27th International Conference on Image Processing (ICIP 2020)

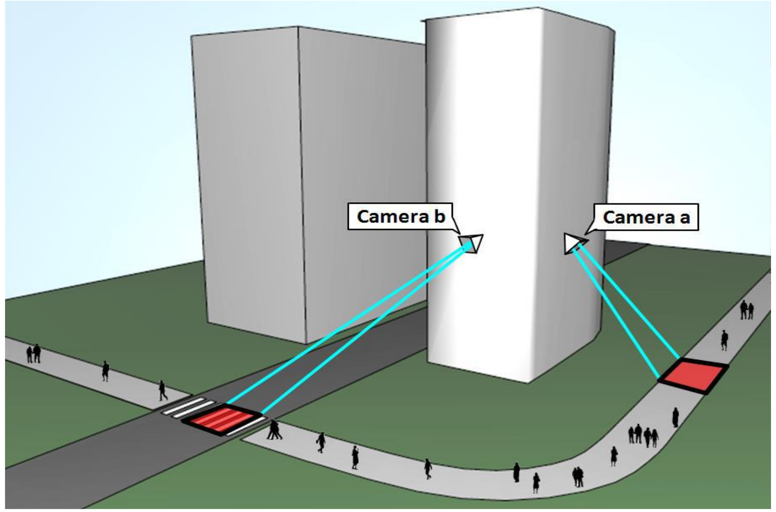
OPEN-SET METRIC LEARNING FOR PERSON RE-IDENTIFICATION IN THE WILD

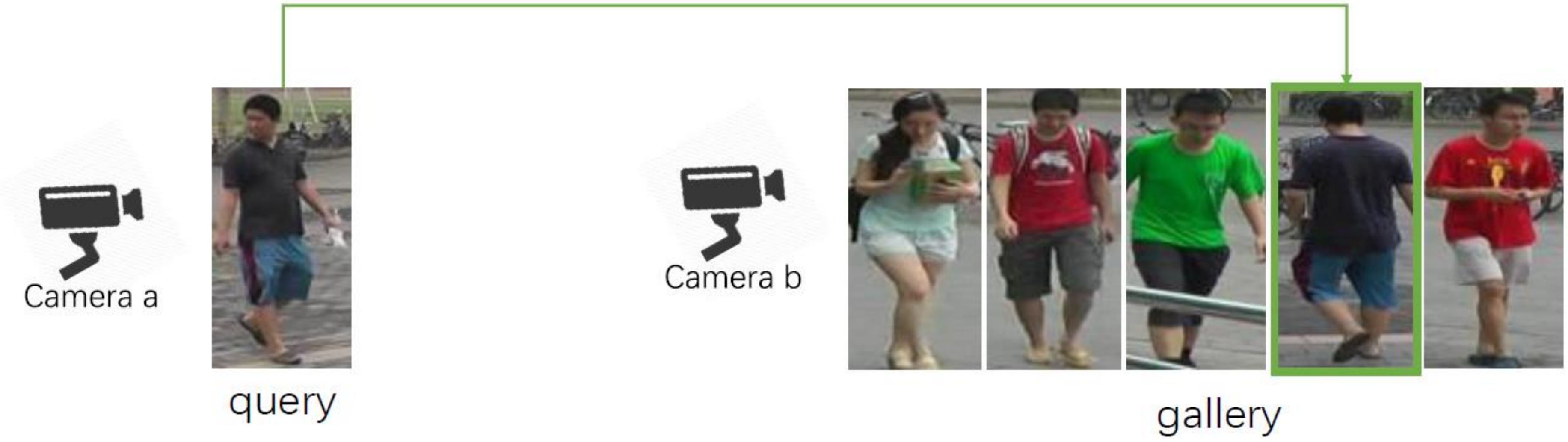
Arindam Sikdar, Dibyadip Chatterjee, Arpan Bhowmik, Ananda S. Chowdhury

Jadavpur University, India

Person Re-identification (Re-ID)?

- Associate same persons across two/multiple non-overlapping Field of views (FoVs)
- Sufficient temporal discontinuity between the visuals of same person

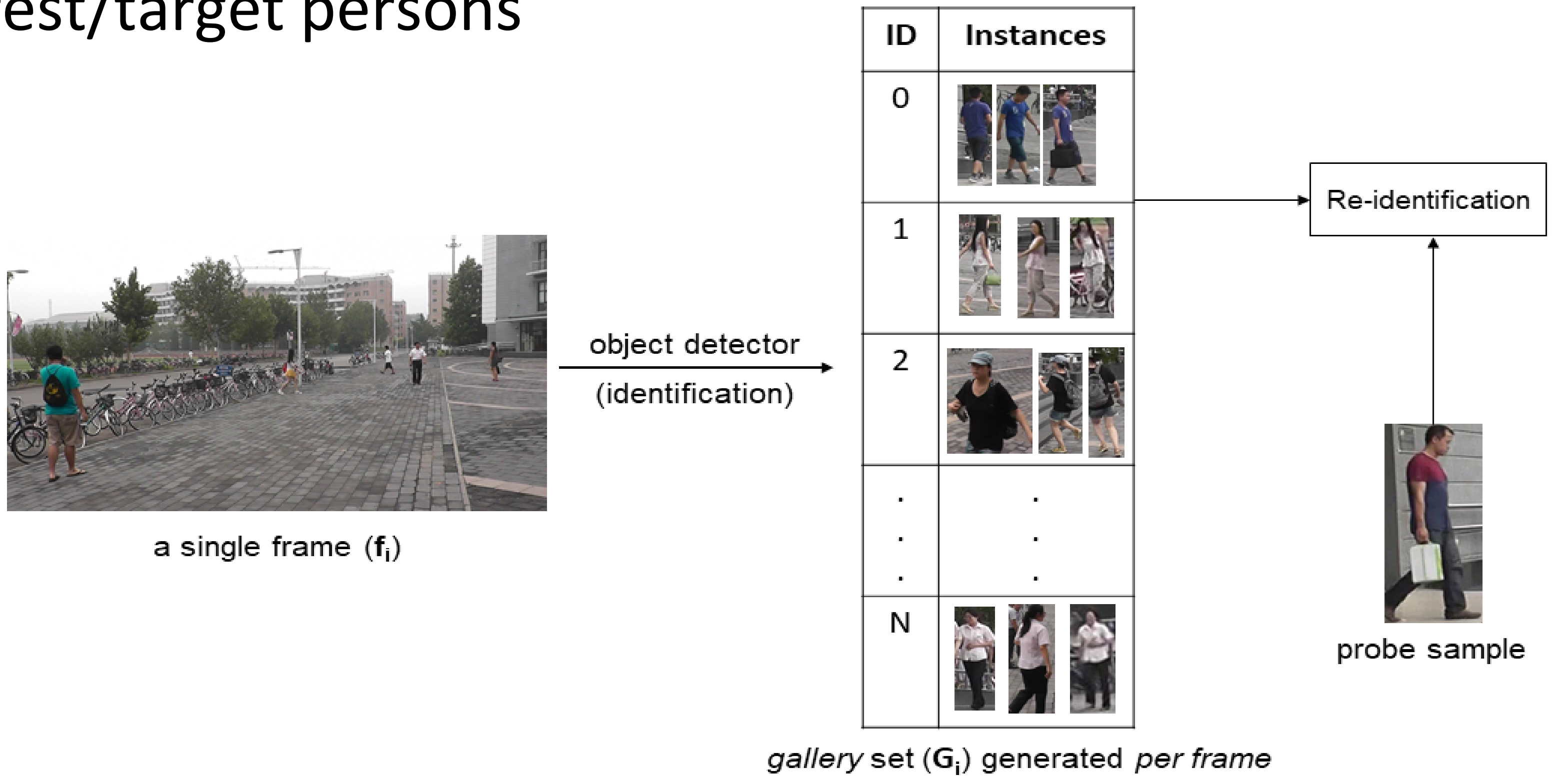




Introduction

Our Goal:

- To simultaneously detect and re-identify (re-id in the **wild**) interest/target persons



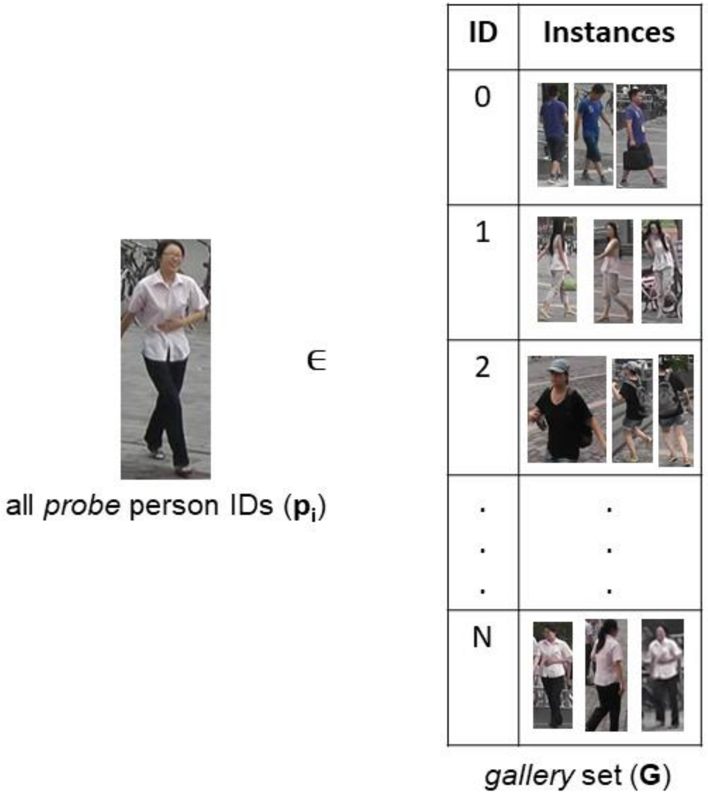
Challenges

- Frame-wise re-identification problem
- Gallery set is dynamically varying with each frame
- Probe set can be larger than gallery set
- From **closed-set** to **open-set re-ID** problem

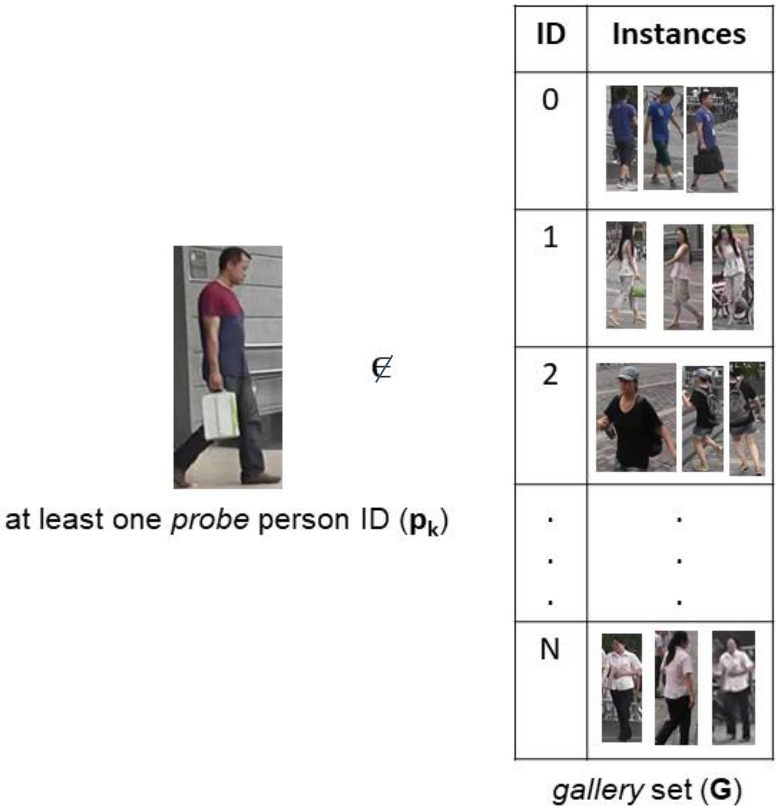
Challenges

- **Open-set problem**: A probe set is not guaranteed to be present in a gallery
- Increase of several false alarms.

closed-set Re-ID



open-set Re-ID



Previous Works

- Closed-set re-ID:
  - Most existing like [Zheng *et al.*, ICCV 2015], [Xiong *et al.*, ECCV 2014] [Liao *et al.*, CVPR 2015] re-ID solution are based on closed sets

- Open-set re-ID:
  - [Liao *et al.,* arXiv 2014] introduced the concept of open-set person re-ID
    - Performed on private dataset with poor performance
  - [Wang *et al.*, ICIP 2016] indicated method to address the open-set re-ID problem
    - Regularized Kernel Subspace Learning but experimented over fixed gallery

- Re-ID in wild:
  - [Zheng *et al.*, CVPR 2017] introduced large-scale Person Re-identification in the Wild (**PRW**) dataset
    - Facilitate End-to-end pedestrian detection and recognition over raw video frames
    - Analyzed only closed-set re-ID performance

Our Contributions

- Introduced **open-set metric learning** (**OSML**) for a more realistic open set re-ID setting

- Joint optimization of Weibull distribution with Mahalanobis metric [Weinberger et al., JMLR 2009] based on **OS-LMNN** loss

- Perform re-ID over dynamically (frame-wise) generated gallery sets

- Converting open-set to closed-set re-ID problem by *rejecting* probe samples over dynamic gallery set

[1] K. Weinberger et al., “Distance metric learning for large margin nearest neighbor classification,” JMLR 2009.

Our Re-ID Framework

- Different components of our proposed re-ID framework
  - ❑ **Pedestrian Detection** using Mask-RCNN followed by feature extraction
  - ❑ Open to closed-set conversion following **Weibull rejection** before similarity ranking

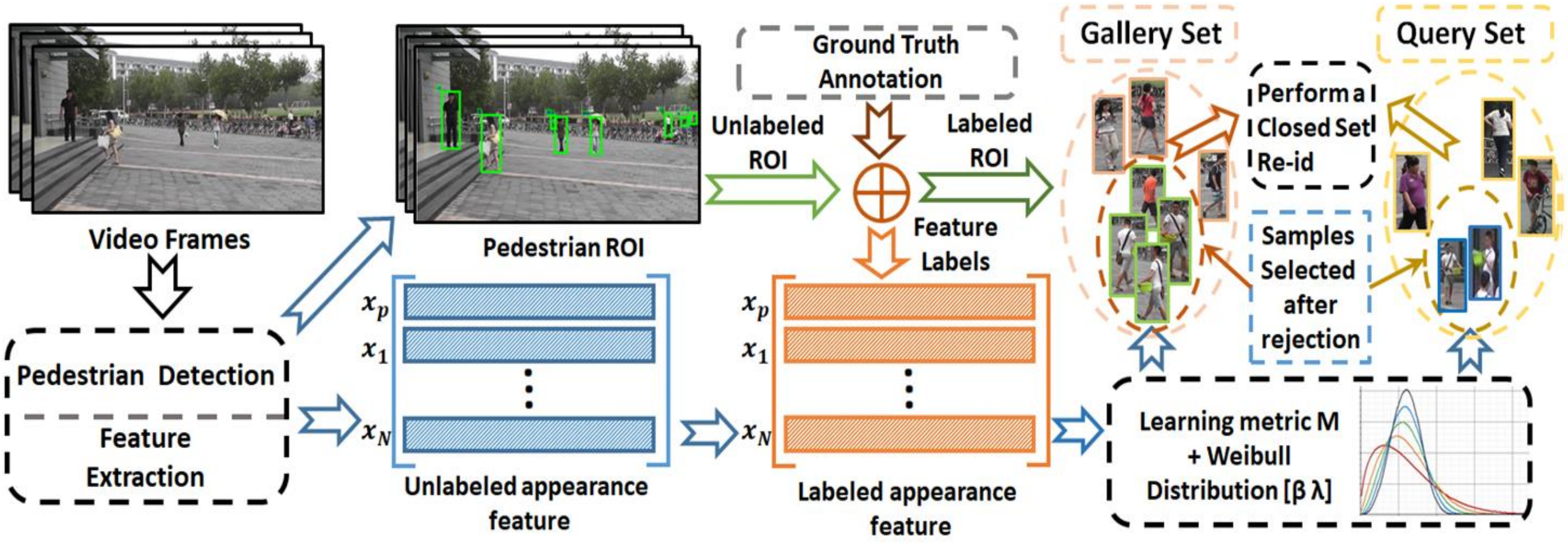


Fig1. An illustrative overview of the proposed re-ID framework

Proposed Method : Pedestrian Detection

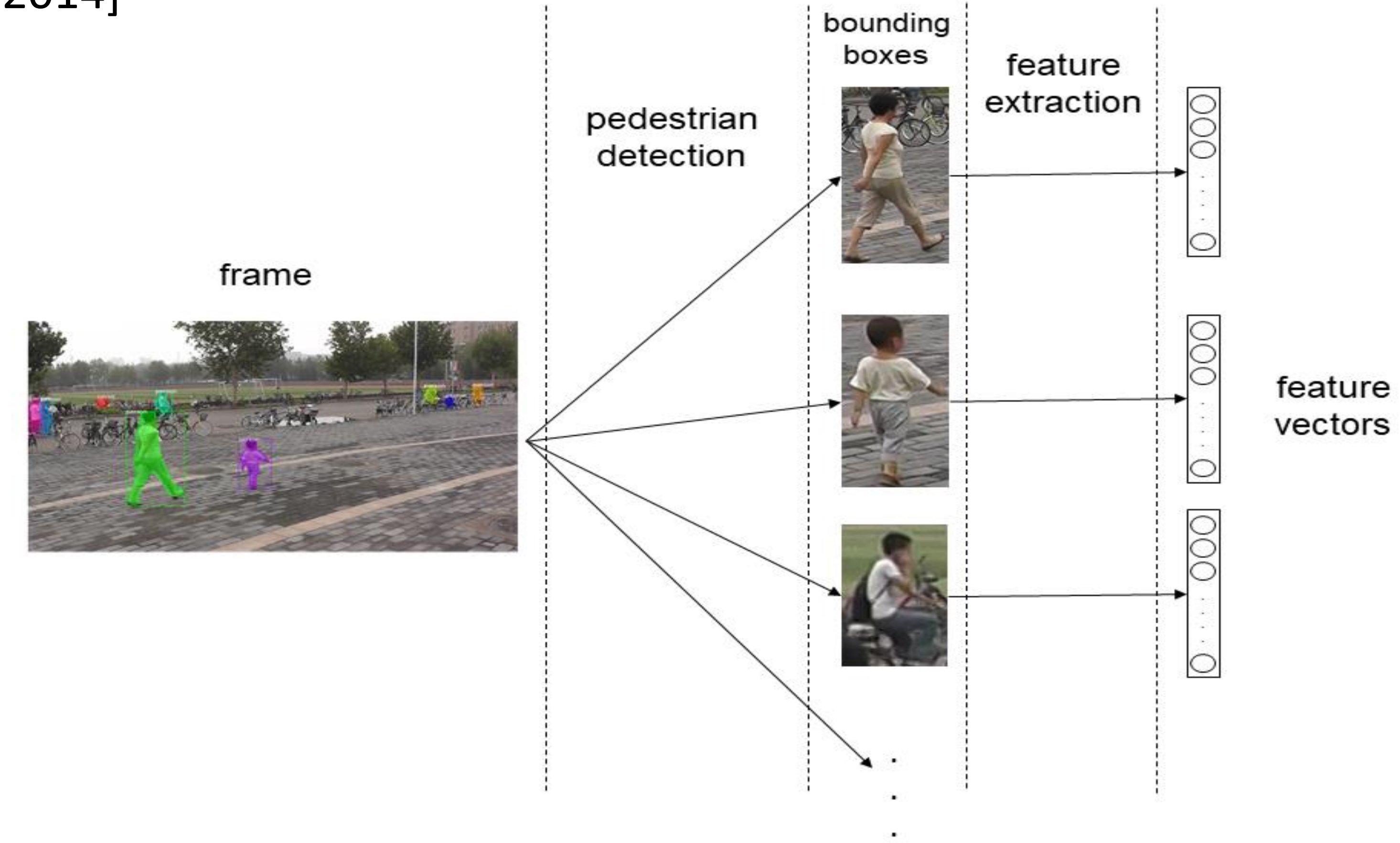
- Pedestrian Detection using Mask R-CNN [He *et al.*, ICCV2017]
  - IOU > 0.5
  - Detector Threshold > 0.9 (for accurate pedestrian localization)

Fig 2. Detected pedestrians using Mask R-CNN

[2] Kaiming He et al., “Mask r-cnn,” ICCV 2017.

Proposed Method : Feature Extraction

- Feature Extraction using traditional descriptors like **BoW** [Zheng et al., ICCV 2015], **HistLBP** [Xiong et al., ECCV 2014], **LOMO** [Liao et al., CVPR 2015], **gBiCov** [Ma et al., IMAGE VISION COMPUT. 2014]



Proposed Method : LMNN Loss

- Mahalanobis Distance metric

$$D_{ij}^{M} = (x_i - x_j)^T M (x_i - x_j)$$

- LMNN loss

$$\varepsilon(M) = (1-\mu)\sum_{i \rightsquigarrow j} D_{ij}^{M} + \mu \sum_{i,j \rightsquigarrow i} \sum_{k} [\alpha + D_{ij}^{M} - D_{ik}^{M}]_+$$

- μ is a weighting parameter that balances the pull and push factors

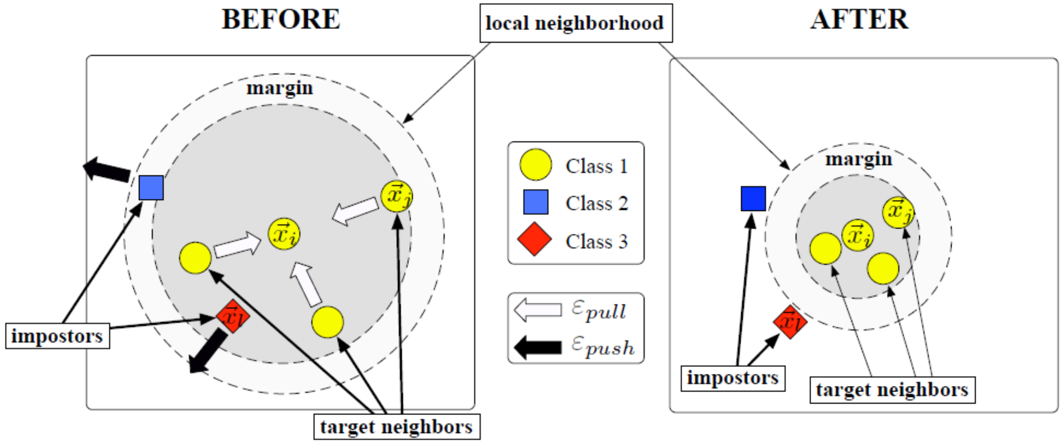


Fig 3. Schematic illustration of push pull concept based on LMNN [1] loss

[1] K. Weinberger et al., “Distance metric learning for large margin nearest neighbor classification,” JMLR 2009.

Proposed Method : Introducing OSML

- **Open Set Recognition** Problems : Ability to distinguish between known and unknown/uncertain samples

- Most Open-set Recognition models [Scheirer *et al.*, TPAMI 2014], [Rudd *et al.*, TPAMI 2017] are applicable over **fixed known classes**

- Open-set metric learning (**OSML**) is extends the concept of Open-set Recognition over similarity metric learning of **variable known samples**/IDs.

- Based on Extreme Value theorem (**EVT**) a learned Weibull distribution can represent unlikely samples at the tail of their distribution.

- **OS-LMNN** combines existing LMNN approach with Weibull distribution to reject unlikely samples.

Proposed Method : Weibull PDF and CDF

- Weibull Distribution PDF

$$\rho(x;\beta,\lambda) = \begin{cases} \frac{\beta}{\lambda}\left(\frac{x}{\lambda}\right)^{\beta-1} e^{-(\frac{x}{\lambda})^{\beta})}, & x \geq 0 \\ 0 & , x < 0 \end{cases}$$

- Weibull Distribution CDF

$$\mathrm{F}(x;\beta,\lambda) = \left[1 - e^{-(\frac{x}{\lambda})^{\beta}}\right] \in \ [0,1]$$

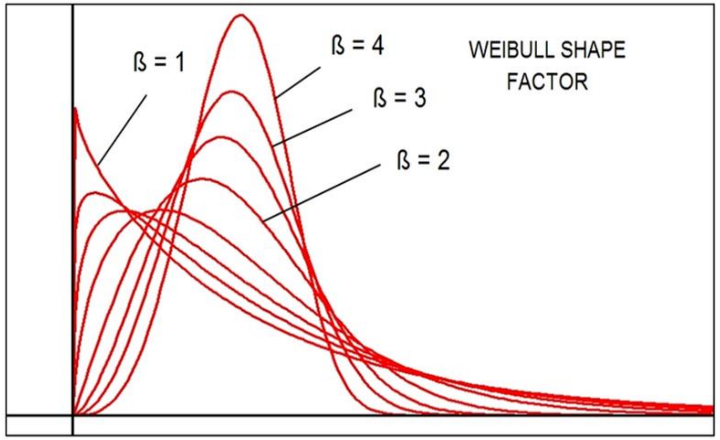


Fig 4. Variation of Weibull PDF based on shape factor β

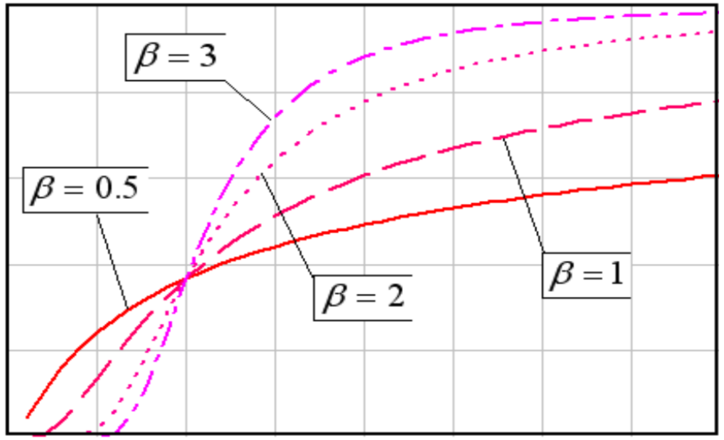


Fig 5. Variation of Weibull CDF based on shape factor β

Proposed Method : OS-LMNN Loss

- Dynamically adjust the push pull weights based on Weibull parameters at each iteration.

- Thus our proposed loss measure is:

$$\varepsilon(M,\beta,\lambda) = \sum_{i,j \rightsquigarrow i} \sum_{k} \left\{ \left(\frac{\omega_{ki}}{1+\omega_{ki}} \right) D_{ij}^{M} + \left(\frac{1}{1+\omega_{ki}} \right) \left[\alpha + D_{ij}^{M} - D_{ik}^{M} \right]_{+} \right\}$$

where, $\omega_{ki} = F(D_{k,\mu_i}^{M};\beta,\lambda)$ is monotonically increasing

and μ_i is mean of samples of person i belonging to same ID

Proposed Method : OS-LMNN Loss (Contd.)

$$\varepsilon(M,\beta,\lambda) = \sum_{i,j \rightsquigarrow i} \sum_{k} \left\{ \left(\frac{\omega_{ki}}{1+\omega_{ki}} \right) D_{ij}^{M} + \left(\frac{1}{1+\omega_{ki}} \right) \left[\alpha + D_{ij}^{M} - D_{ik}^{M} \right]_{+} \right\}$$

where, $\omega_{ki} = F(D_{k,\mu_i}^{M};\beta,\lambda)$ is monotonically increasing

- Based on property of Weibull CDF,

$$\omega_{ki} \Downarrow \qquad \left(\frac{1}{1+\omega_{ki}} \right) \Uparrow$$

- Smaller distance between dissimilar pair increases push factors weight w.r.t pull

Proposed Method : Optimization

- **Regularization**

$$M^*, \beta^*, \lambda^* = \underset{M,\beta,\lambda}{argmin}\left[\epsilon(M,\beta,\lambda) + \gamma R(\beta,\lambda)\right]$$

where, $R(\beta,\lambda) = \frac{1}{2}\mathbb{N}.\,(\beta + \lambda)$ is a regularization term and

$\mathbb{N}$ is the total no. of valid triplets

- We use L-BFGS-B [7] optimizer to solve the objective by alternatively fixing $\boldsymbol{M}$ and $\boldsymbol{w} = [\beta, \lambda]$

[7] Richard H Byrd et al., "A limited memory algorithm for bound constrained optimization," SIAM Journal on scientific computing, 1995.

Proposed Method : Sample Rejection

- **Weibull Rejection** mechanism:
  - Given a dynamic gallery set a G a likelihood value is assigned to every probe sample P based on Weibull PDF
  - Pairwise computation
  - Assigns a low probability value to dissimilar pairs in new metric space
  - Reject pairs with likelihood less than a threshold $\boldsymbol{\tau}$ (a user parameter)
  - A probe samples rejected by all samples ϵ G are inferred absent in dynamic set G
  - Similarity ranking are performed with remaining probe samples (closed set comparison)
  - The gallery ID attaining highest similarity with a probe are inferred same person

Implementation Details

- Our model performance is evaluated and compared over **PRW** dataset [Zheng et al., CVPR 2017]
- Mask R-CNN [He et al., ICCV 2017] detector (pre-trained on ImageNet dataset) was fine-tuned on the PRW dataset
- Our model has two hyper-parameters which are set experimentally
  - margin α = 25
  - regularization constant γ = 0.5

[2] Kaiming He et al., “Mask r-cnn,” ICCV 2017.

[8] Liang Zheng et al., “Person reidentification in the wild,” CVPR 2017.

Evaluation Metrics

- Detection and Identification Rate (DIR) [Liao *et al.*, arXiv 2014]
- False Acceptance Rate (FAR)

$$DIR(\tau, k) \;\; \frac{|\{p: p \in P_G, rank(p) \leq k,\; \rho(D_{pg}^{M}) \geq \tau\}|}{P_G}$$

$$FAR(\tau, k) = \frac{|\{p: p \in P_N \; and \; \rho(D_{pg}^{M}) \geq \tau\}|}{P_N}$$

where P_G and P_N are the two probe sets and G is the gallery set with g ∈ G

- Rank-1 recognition rate and Area under ROC (AUC) curve

Results : DPM Detector

Table 1. DIR vs. varying FAR for Rank-1 scores with the DPM detector. Best values are shown in **bold**.

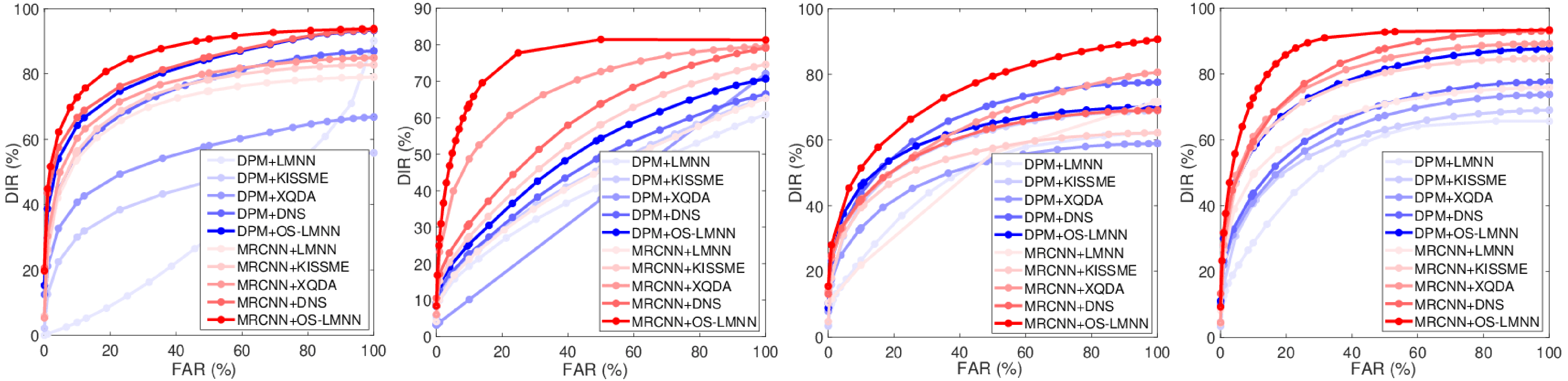
Detector	Feature	Recognizer	FAR (%)				AUC (%)
			1	10	50	100	
DPM [Felzenszwalb *et al.*, TPAMI 2009]	HistLBP	LMNN	9.89	19.34	41.21	60.92	39.64
		KISSME	11.17	21.83	46.01	65.45	43.88
		DNS	12.79	22.70	49.34	66.50	45.98
		XQDA	3.92	10.12	37.65	**72.06**	37.65
		OS-LMNN (ours)	**14.31**	**25.40**	**54.37**	70.71	**50.26**
	LOMO	LMNN	12.64	28.69	61.56	65.58	53.89
		KISSME	15.95	40.60	63.11	68.98	58.34
		DNS	23.12	43.72	70.87	77.58	65.22
		XQDA	21.97	41.54	67.33	73.70	61.96
		OS-LMNN (ours)	**30.04**	**57.63**	**81.42**	**87.61**	**76.30**
	BoW	LMNN	0.43	3.96	29.84	90.11	32.85
		KISSME	12.56	30.04	47.07	55.82	44.35
		DNS	30.03	55.30	78.71	87.03	74.19
		XQDA	21.17	40.72	58.02	66.82	55.16
		OS-LMNN (ours)	**38.69**	**64.23**	**84.62**	**93.34**	**80.83**
	gBiCov	LMNN	10.12	23.41	54.70	62.31	48.24
		KISSME	15.95	40.60	63.11	68.98	58.34
		DNS	23.12	43.72	70.87	**77.58**	**65.22**
		XQDA	17.57	33.22	53.86	58.96	49.57
		OS-LMNN (ours)	**24.03**	**46.10**	**65.14**	70.09	61.40

Results : Mask R-CNN Detector

Table 2. DIR vs. varying FAR for Rank-1 scores with Mask-RCNN detector. Best values are shown in **bold**.

Detector	Feature	Recognizer	FAR (%)				AUC (%)
			1	10	50	100	
Mask R-CNN [He *et al.*, ICCV 2017]	HistLBP	LMNN	10.08	21.23	45.54	65.36	43.23
		KISSME	15.21	27.42	58.14	74.59	53.68
		DNS	17.19	30.92	63.87	79.13	58.54
		XQDA	23.23	48.72	72.64	79.54	67.61
		OS-LMNN (ours)	**26.92**	**63.71**	**81.45**	**81.51**	**76.11**
	LOMO	LMNN	24.81	49.45	70.79	75.90	65.99
		KISSME	26.63	57.95	80.30	84.73	74.86
		DNS	**31.96**	59.41	87.68	93.04	81.29
		XQDA	28.03	61.00	84.53	89.19	78.80
		OS-LMNN (ours)	31.65	**72.70**	**92.65**	**93.31**	**87.15**
	BoW	LMNN	25.3	54.00	74.74	79.02	69.83
		KISSME	26.25	56.41	78.27	82.78	73.10
		DNS	43.51	66.68	85.25	93.48	81.85
		XQDA	30.41	60.29	80.28	84.95	75.49
		OS-LMNN (ours)	**44.82**	**72.82**	**90.66**	**93.80**	**86.23**
	gBiCov	LMNN	11.62	21.93	56.01	71.54	49.17
		KISSME	18.05	39.39	57.48	62.18	53.58
		DNS	24.88	41.90	63.69	69.06	59.13
		XQDA	24.06	41.72	67.57	80.60	63.50
		OS-LMNN (ours)	**28.04**	**51.49**	**79.40**	**90.61**	**74.24**

Results



(a) With BoW descriptor (b) With HistLBP descriptor (c) With gBiCov descriptor (d) With LOMO descriptor

Fig 5. ROC Curve with DIR vs FAR comparison at rank-1 recognition rate for different feature descriptors

Conclusion

- A new metric learning model has been proposed especially for performing open-set re-ID in the wild
- The concept of Weibull rejection has been introduced convert an open-set re-ID problem to a closed-set.
- The proposed model can be further improved by introducing non-linearity through kernels that can better represent the metric space for complex sets
- We plan to extend our open-set metric learning framework to end-to-end trainable deep architectures

References

[1] K. Weinberger et al., “Distance metric learning for large margin nearest neighbor classification,” JMLR 2009.

[2] Kaiming He et al., “Mask r-cnn,” ICCV 2017.

[3] Liang Zheng et al., “Scalable person reidentification: A benchmark,” ICCV 2015.

[4] Fei Xiong et al., “Person re-identification using kernel-based metric learning methods,” ECCV 2014.

[5] Shengcai Liao et al., “Person re-identification by local maximal occurrence representation and metric learning,” CVPR 2015.

[6] Bingpeng Ma et al., “Covariance descriptor based on bio-inspired features for person reidentification and face verification,” Image Vision Computing, 2014.

[7] Richard H Byrd et al., “A limited memory algorithm for bound constrained optimization,” SIAM Journal on scientific computing, 1995.

References

[8] Liang Zheng et al., “Person reidentification in the wild,” CVPR 2017.

[9] P. F. Felzenszwalb et al.,“Object detection with discriminatively trained part-based models,“ IEEE Transactions on Pattern Analysis and Machine Intelligence, 2009.

[10] Shengcai Liao, Zhipeng Mo, Jianqing Zhu, Yang Hu, and Stan Z Li, “Open-set person re-identification,” arXiv:1408.0872, 2014.

[11] Hanxiao Wang, Xiatian Zhu, Tao Xiang, and Shaogang Gong, “Towards unsupervised open-set person reidentification,” in ICIP, 2016, pp. 769–773.

[12] Walter J Scheirer, Lalit P Jain, and Terrance E Boult, “Probability models for open set recognition,” IEEE Trans. Pattern Anal. Mach. Intell., pp. 2317–2324, 2014.

[13] Ethan M Rudd, Lalit P Jain, Walter J Scheirer, and Terrance E Boult, “The extreme value machine,” IEEE Trans. Pattern Anal. Mach. Intell., pp. 762–768, 2017.

THANK YOU!

For more information, please visit:

https://sites.google.com/site/ivprgroup/