Person Re-Identification in Surveillance Videos using Deep Learning based Body Part Partition and Gaussian Filtering

Fatih Aksu Cem Direkoğlu

Middle East Technical University

Northern Cyprus Campus

Outline

- Definition
- Literature review
- Our system
- Body part segmentation
- Gaussian filtering
- Classifier
- Dataset
- Experiments
- Conclusion

Definition

- Searching a person that previously observed over a camera network.

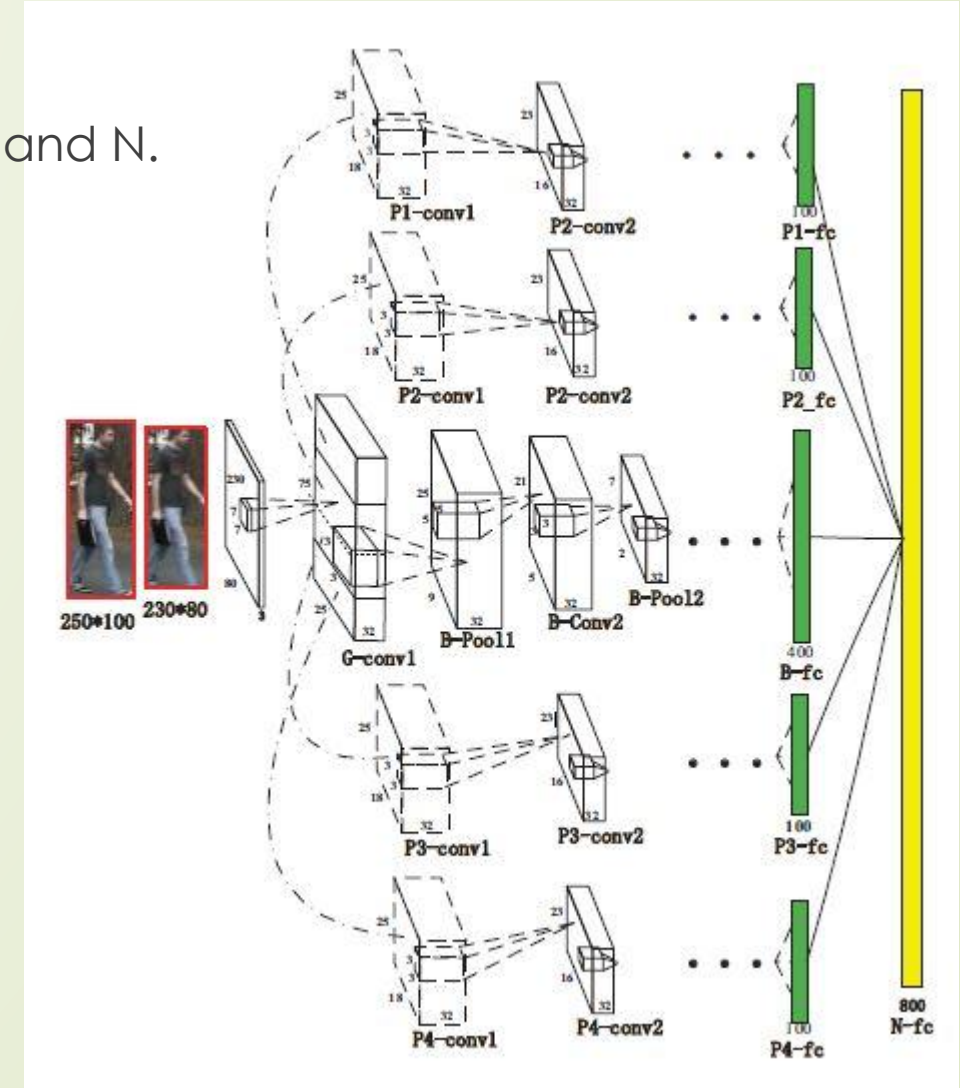
Wu et al., 2019

Earlier Works

- Normalized color and salient edgel histograms (Ghessari et al., 2006)
- 3 color and 19 texture channels chosen by AdaBoost (Gray & Tao, 2008)
- Support vector machines with color and texture features (Prosser et al., 2010)
- Siamese convolutional neural network (Li et al., 2014)

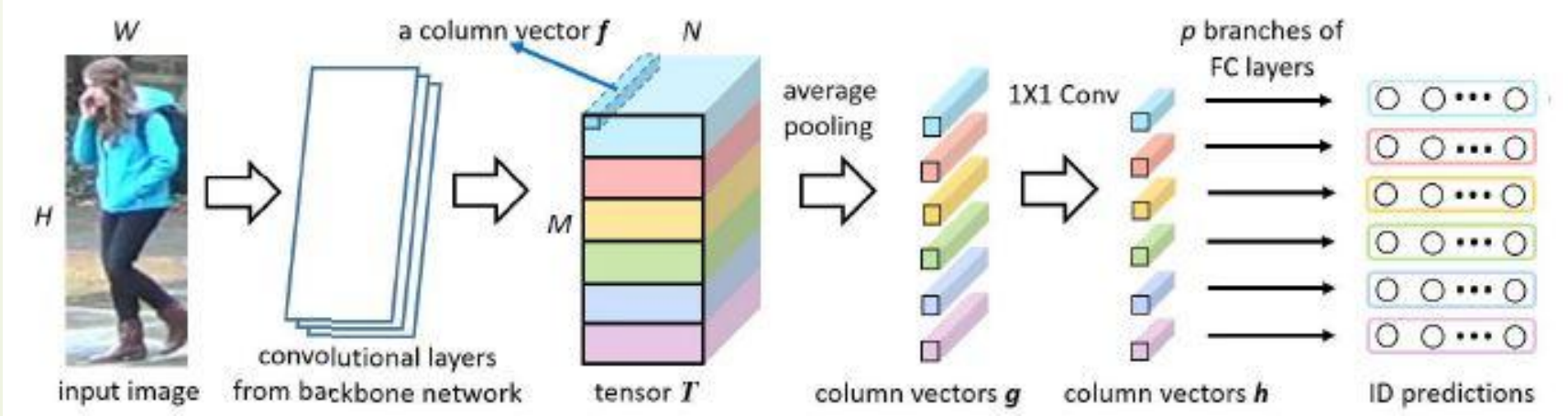
Recent Works

- D. Cheng, Y. Gong, S. Zhou, J. Wang and N. Zheng, "Person Re-identification by Multi-Channel Parts-Based CNN with Improved Triplet Loss Function," *2016*
  - 1 global, 4 local networks
  - Horizontal partition at the beginning
  - Concatanated at the end

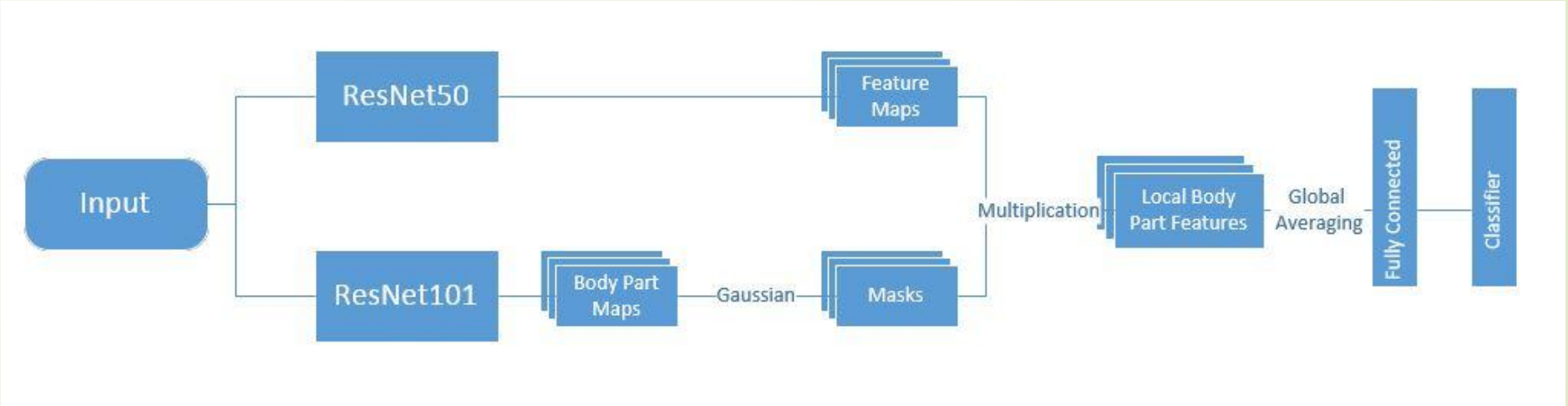


Recent Works

- Yifan Sun, Liang Zheng, Yi Yang, Qi Tian, Shengjin Wang, " Beyond Part Models: Person Retrieval with Refined Part Pooling (And A Strong Convolutional Baseline), " 2018
  - Horizontally divided into 6 parts at the end
  - Different classifier for each part
  - Max voting

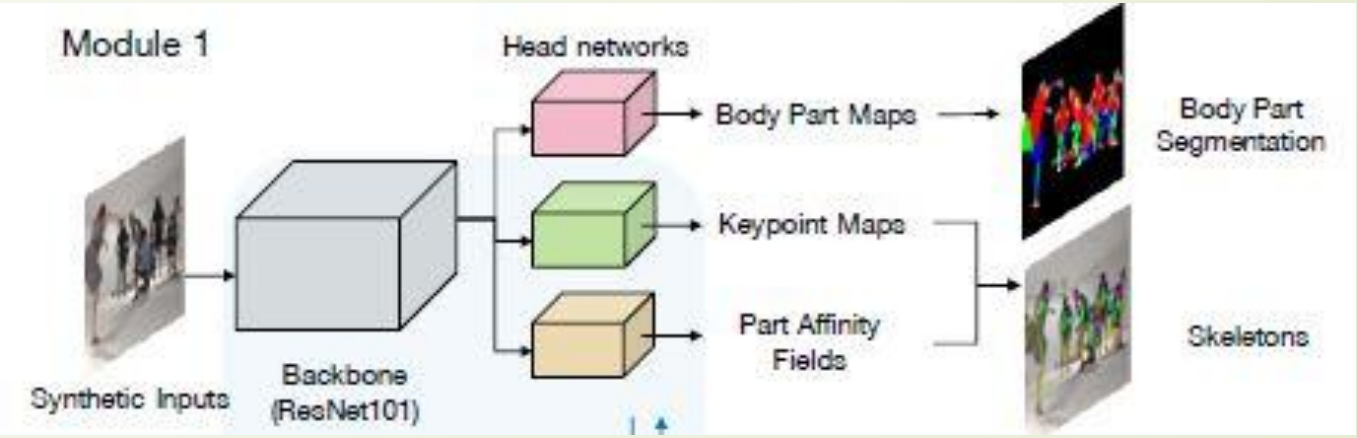


Our System


Input
ResNet50
ResNet101
Feature Maps
Body Part Maps
Gaussian
Masks
Multiplication
Local Body Part Features
Global Averaging
Fully Connected
Classifier

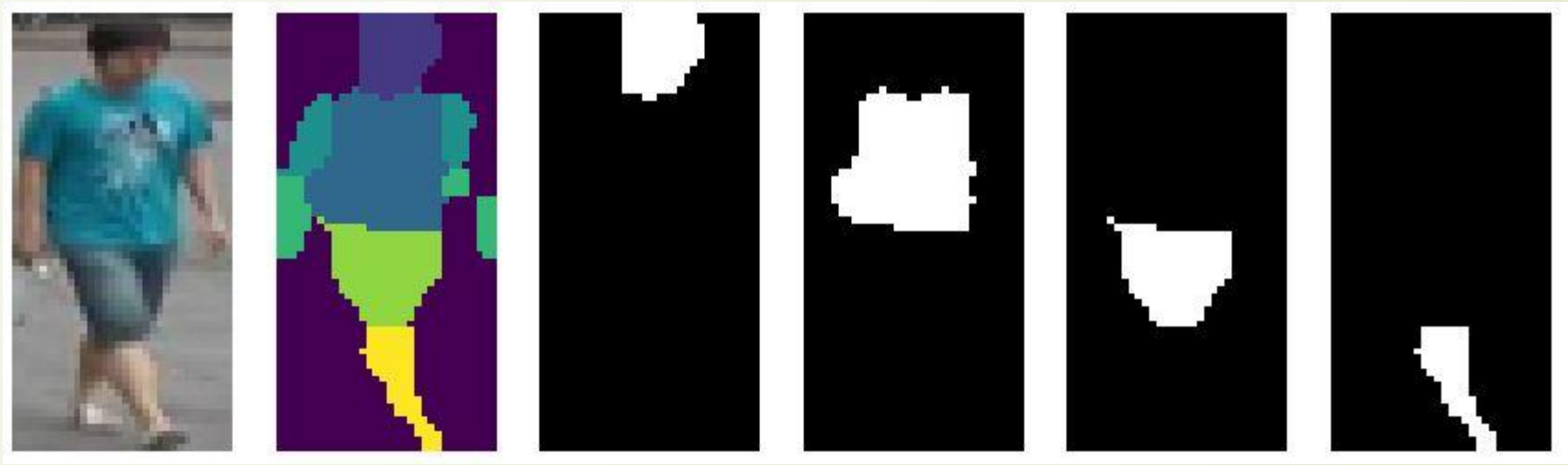
Body Part Segmentation

- K. Lin, L. Wang, K. Luo, Y. Chen, Z. Liu and M. Sun, "Cross-Domain Complementary Learning Using Pose for Multi-Person Part Segmentation, " 2020



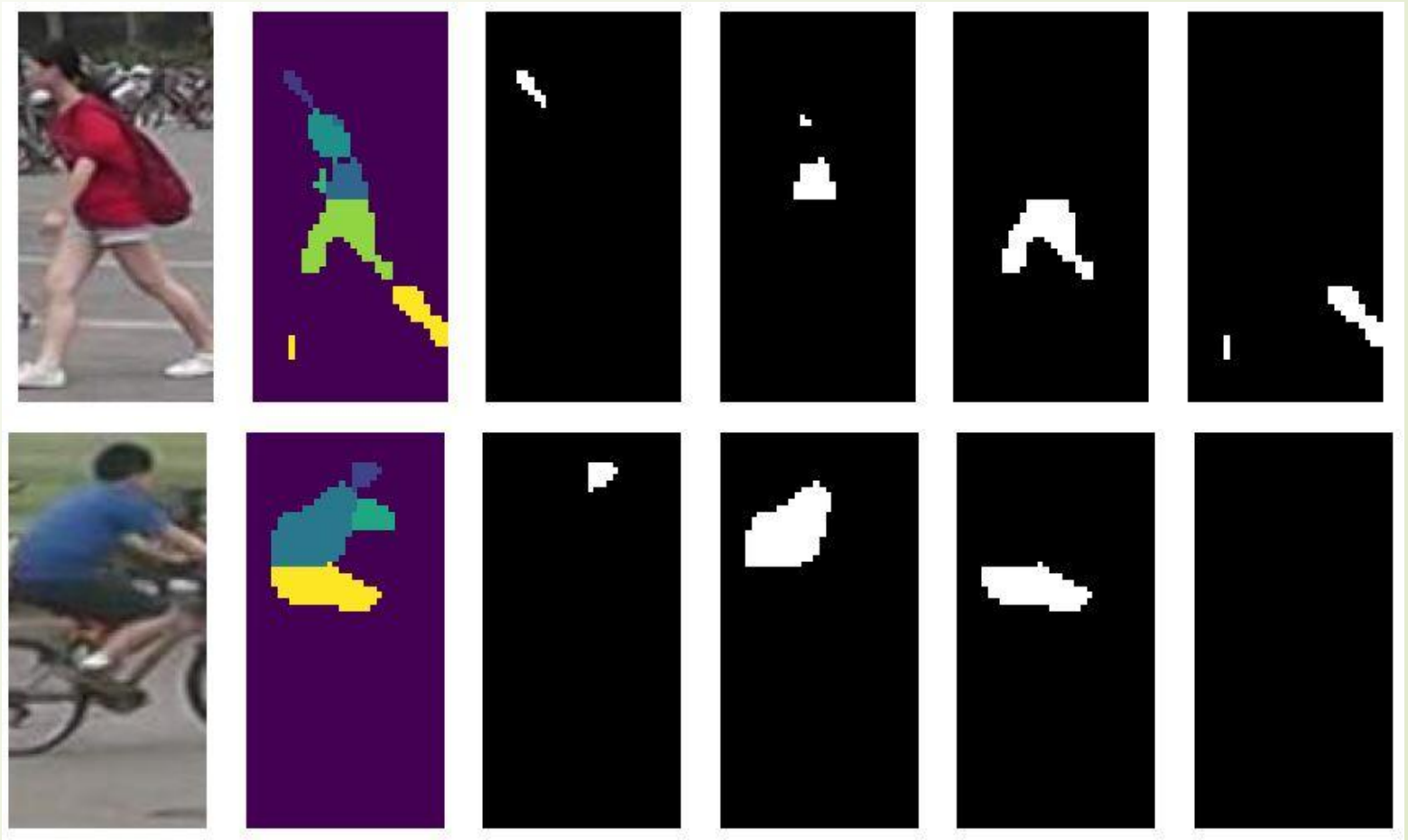
Mask Creation

- Extract binary masks from body part map
- 7 different masks corresponce to 7 different body parts
  - Head, upper body, lower body, arms, legs



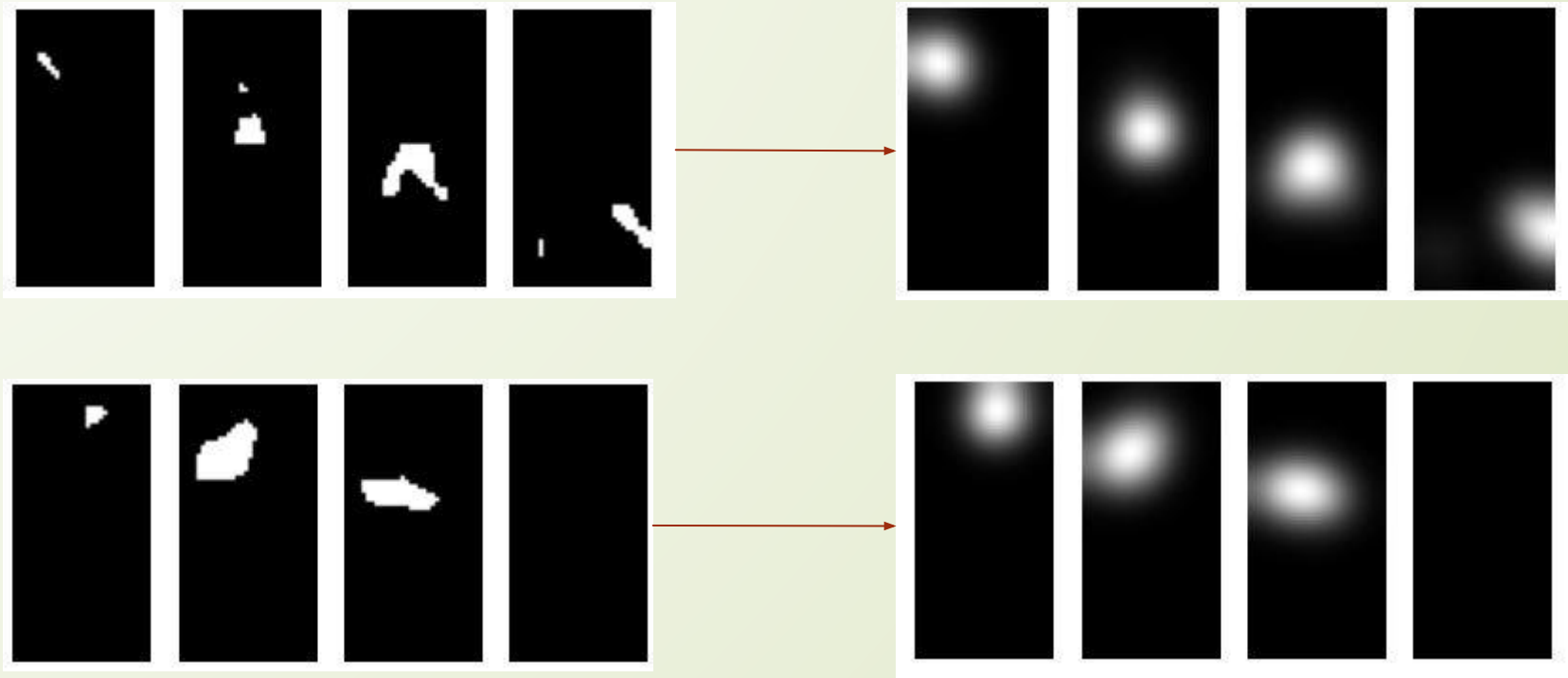
Mask Creation

- However, not all the masks are perfect.

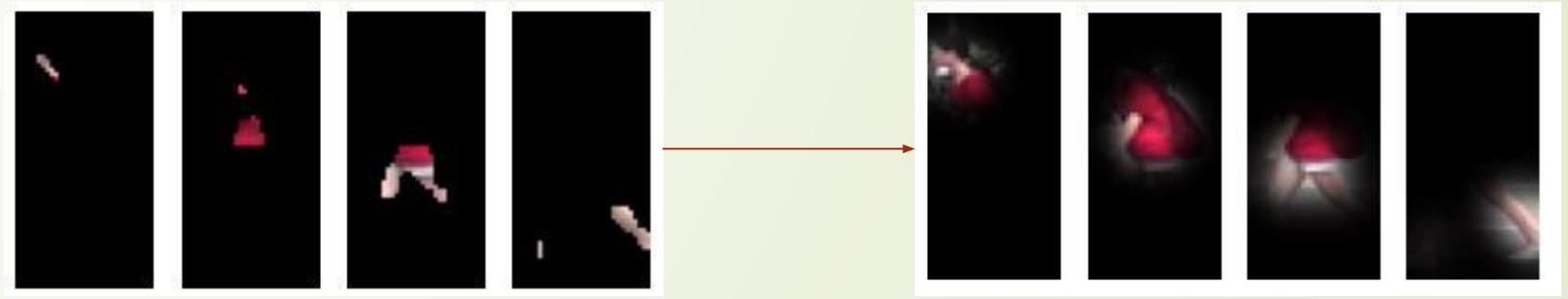


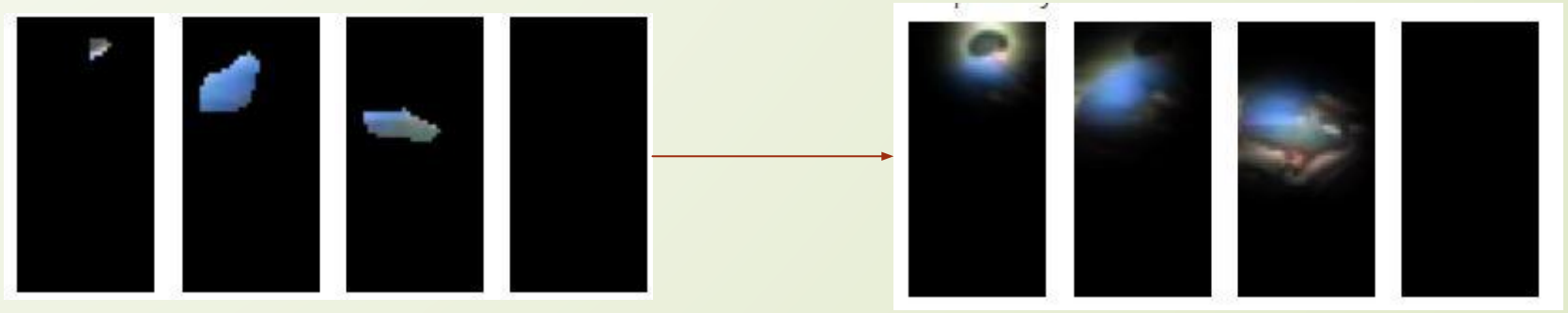
Gaussian Filtering

- Gaussian filter smooths the masks
- Undetected parts are taken into consideration
- Weight decreases as distance increase from the part



Gaussian Filtering





Classifier

- Seperate classifier for each part;
  - Lose correlation between parts
  - Similar clothes of different individuals create confusions
- Shared classifier;
  - Keeps correlation between parts
  - Creates more unique features of individuals
- Thus, we used shared classifier.

Dataset

- Market-1501 dataset;
  - 32,668 images of 1,501 persons
  - Each individual is captured by at least two different cameras
  - 6 different camera views in total

Zheng et al. 2015

Evaluation

- CMC (Cumulative Matching Characteristics)
  - probability of a person appears in different-sized candidate lists
  - rank-1, rank-5 and rank-10 lists are used for first, first five and first ten guesses
- mAP (Mean Average Precision)
  - Mean of the average precision (AP)
  - AP = area under the Precision-Recall

Experiments

- Implemented in Pytorch
- Pre-trained ResNet-50 network used
  - Batch size : 32
  - Epoch : 30
  - Learning rate : 0.01
    - with a decay by a factor of 0.01 in every 10 epoch
  - Optimizer : stochastic gradient descent
- Gaussian parameters;
  - 41x41 kernel
  - Standard deviation : 5
  - Applied in spatial domain

Results

Model	Rank-1	Rank-5	Rank-10	mAP
Baseline Method #1 (ResNet-50) [22]	0.858	0.949	0.969	0.670
Baseline Method #2 (Horizontal partition with individual classifiers) [28]	0.796	0.912	0.941	0.590
Semantic partition with individual classifiers (no Gaussian filtering) (Proposed Method)	0.817	0.912	0.940	0.468
Semantic partition with Gaussian filtering and individual classifiers (Proposed Method)	0.845	0.940	0.961	0.634
Semantic partition with Gaussian filtering and shared classifier (Proposed Method)	**0.870**	**0.952**	**0.971**	0.668

Conclusion

- Resnet-50 to extract global features
- Masking global features to obtain part features
- Gaussian smoothing to prevent feature lose
- Shared classifier to keep correlation between parts
- Future work;
  - More precise tuning of parameters
  - Applying the algorithm to more specifik networks
  - Using more beneficial lost functions
  - Adjust our functions to make them be able to use augmentation techniques

References

- B. Prosser, W.-S. Zheng, S. Gong, T. Xiang, and Q. Mary, “Person re-identification by support vector ranking.” in BMVC, 2010
- D. Cheng, Y. Gong, S. Zhou, J. Wang and N. Zheng, "Person Re-identification by Multi-Channel Parts-Based CNN with Improved Triplet Loss Function," IEEE Conference on Computer Vision and Pattern Recognition, 2016.
- D. Gray and H. Tao., " Viewpoint invariant pedestrian recognition with an ensamble of localized features, " ECCV, 2008
- D. Wu et al., "Deep learning-based methods for person reidentification: A comprehensive review," Neurocomputing, 2019
- K. Lin, L. Wang, K. Luo, Y. Chen, Z. Liu and M. Sun, "Cross-Domain Complementary Learning Using Pose for Multi-Person Part Segmentation, " 2020
- L. Zheng, L. Shen, L. Tian, S. Wang, J. Wang and Q. Tian, "Scalable Person Re-identification: A Benchmark," *IEEE International Conference on Computer Vision*, 2015
- N. Gheissari, T. B. Sebastian and R. Hartley, "Person Reidentification Using Spatiotemporal Appearance,« Computer Vision and Pattern Recognition, 2006
- W. Li, R. Zhao, T. Xiao, and X. Wang, “Deepreid: Deep filter pairing neural network for person re-identification,” Computer Vision and Pattern Recognition, 2014
- Yifan Sun, Liang Zheng, Yi Yang, Qi Tian, Shengjin Wang, " Beyond Part Models: Person Retrieval with Refined Part Pooling (And A Strong Convolutional Baseline), " 2018

Thank You