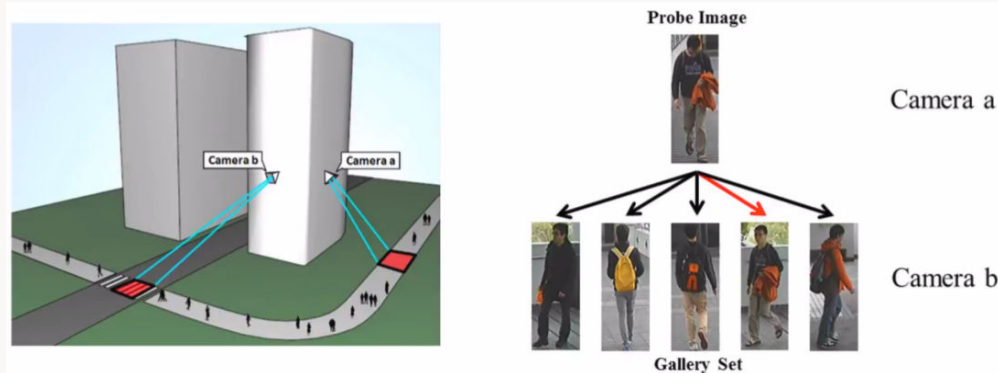


# Identity tracking in multi-view camera surveillance system

## Motivation

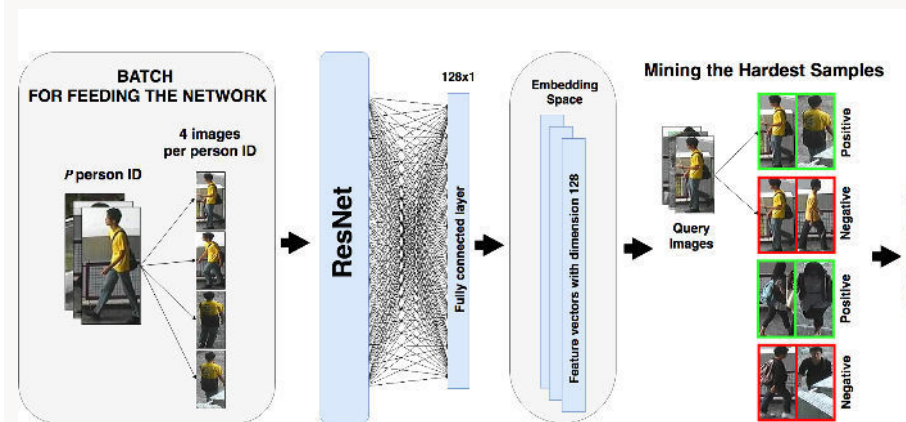
Crime in the cities is growing simultaneously with them. Our advantage is that cities have more and more cameras, unfortunately, they have usually low resolution. Therefore we proposed person and vehicle re-identification solution, where we do not need car plate numbers. The result of our work could be used for following a criminal in their car from the crime scene through the whole city or general surveillance. Our solution could be also used for optimization city traffic or calculating statistics in shopping centers.



## Contribution

- ▶ We proposed a solution based on ResNet50 neural network.
- ▶ We created new loss function, based on famous triplet loss function and we proposed new approach to choose the hardest samples for training the neural network.
- ▶ We trained the neural network for person and vehicle re-identification separately.
- ▶ We proposed a new re-identification method with spatio-temporal constraints for the datasets, where we know the distances between cameras.

## Network training flow



Each batch of images is processed with ResNet to obtain feature vectors. We then mine the hardest samples and feed them into our novel loss function to force the network to learn the hardest cases.

## Qualitative results



The **left column** - the query image, the **remaining ten images** are the image representations of the most similar images from gallery. A **green border** means correctly matched object according to ground truth while a **red border** indicates a different identity from the query.

## Experiments

- ▶ The proposed architecture was exhaustively evaluated on the most popular modern data sets: CUHK03, DukeMTMC4REID, Market1501, VeRi, and VehicleReid.
- ▶ The comparison with state of the art results are in the following tables:

Person dataset: CUHK03 Detected		
	mAP	rank-1
State of the art methods	88.4	84.6
<b>Ours + re-ranking</b>	<b>93.4</b>	<b>91.0</b>

Vehicle dataset: VeRi (image-to-track)			
	mAP	rank-1	rank-5
State of the art methods	59.47	81.56	95.11
<b>Ours (spatio-temporal)</b>	<b>77.09</b>	<b>92.79</b>	<b>97.97</b>

## Conclusion and future work

- ▶ We have suggested and implemented a method, which finds top-n likeliest matches for a given object throughout cameras of the system by visual matching and also with spatio-temporal constraints.
- ▶ Our framework obtains state-of-the-art results on both, person and vehicle re-identification tasks, with no change to the network architecture.
- ▶ **Future work:**
  - Training the same framework on data sets from both domains and performing vehicle and person re-ID simultaneously.
  - Suggesting a neural network, which would be able to learn spatio-temporal information by itself.