Deep Joint Discriminative Learning for Vehicle Re-identification and Retrieval

Yuqi Li, Yanghao Li,
Hongfei Yan, Jiaying Liu
Peking University

2017.09.18

Outline

Background

Deep Joint Discriminative Learning

Experimental Results

Conclusion

Outline

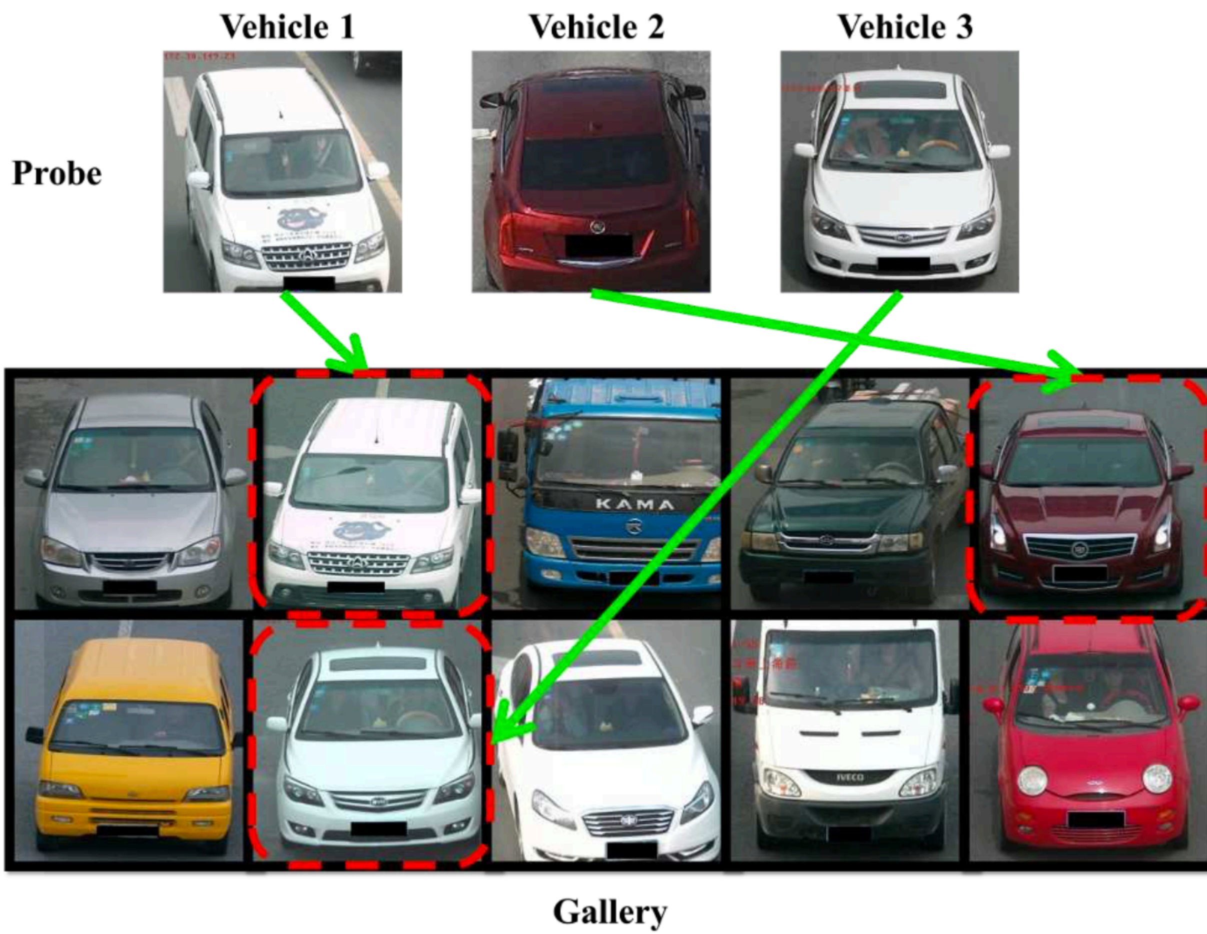
Background

Deep Joint Discriminative Learning

Experimental Results

Conclusion

Vehicle search and re-identification



Vehicle search and re-identification

- Practical applications in video surveillance systems
- Challenge
  - License plate is not clear
    - Low-resolution
    - Occluded or removed
  - →Vehicle ReID based on **appearance information**

Vehicle search and re-identification

- VehicleID dataset
  - Labeled in identity level
  - Remove license plate

Hongye Liu et al. “Deep relative distance learning: Tell the difference between similar vehicles,” in *CVPR*, 2016.

Related work

- Most identification works focus on face or person
  - Face recognition
  - Person re-identification
- Target: learn discriminative representations
  - State-of-art → Deep CNN based
  - DeepID [Sun et. al, 2014]
    - Directly classify identities (~1w)
  - DeepID2 [Sun et. al, 2014]
    - Pairwise verification loss
  - Triplet loss [Schroff et. al, 2015, Ding et. al 2015]
    - Triplet relationship between positive and negative pairs

Related work

- Difference of vehicle identification
  - Previous works focus on **model** classification
    - Recognize models instead of identities
  - Vehicles of same model → similar visual appearance
  - Capture special marks

Related work

- Difference of vehicle identification
  - Previous works focus on **model** classification
    - Recognize model instead of identities
  - Vehicles of same model → similar visual appearance
  - Capture special marks
- Large scale vehicle identification dataset
  - VehicleID [Liu et al. 2016]
  - Facilitate deep learning models

Related work

- Difference of vehicle identification
  - Previous works focus on **model** classification
    - Recognize model instead of identities
  - Vehicles of same model → similar visual appearance
  - Capture special marks
- Large scale vehicle identification dataset
  - VehicleID [Liu et al. 2016]
  - Facilitate deep learning models

- Deep Joint Discriminative Learning (DJDL) model
  - A unified framework to extract discriminative features

Outline

Background

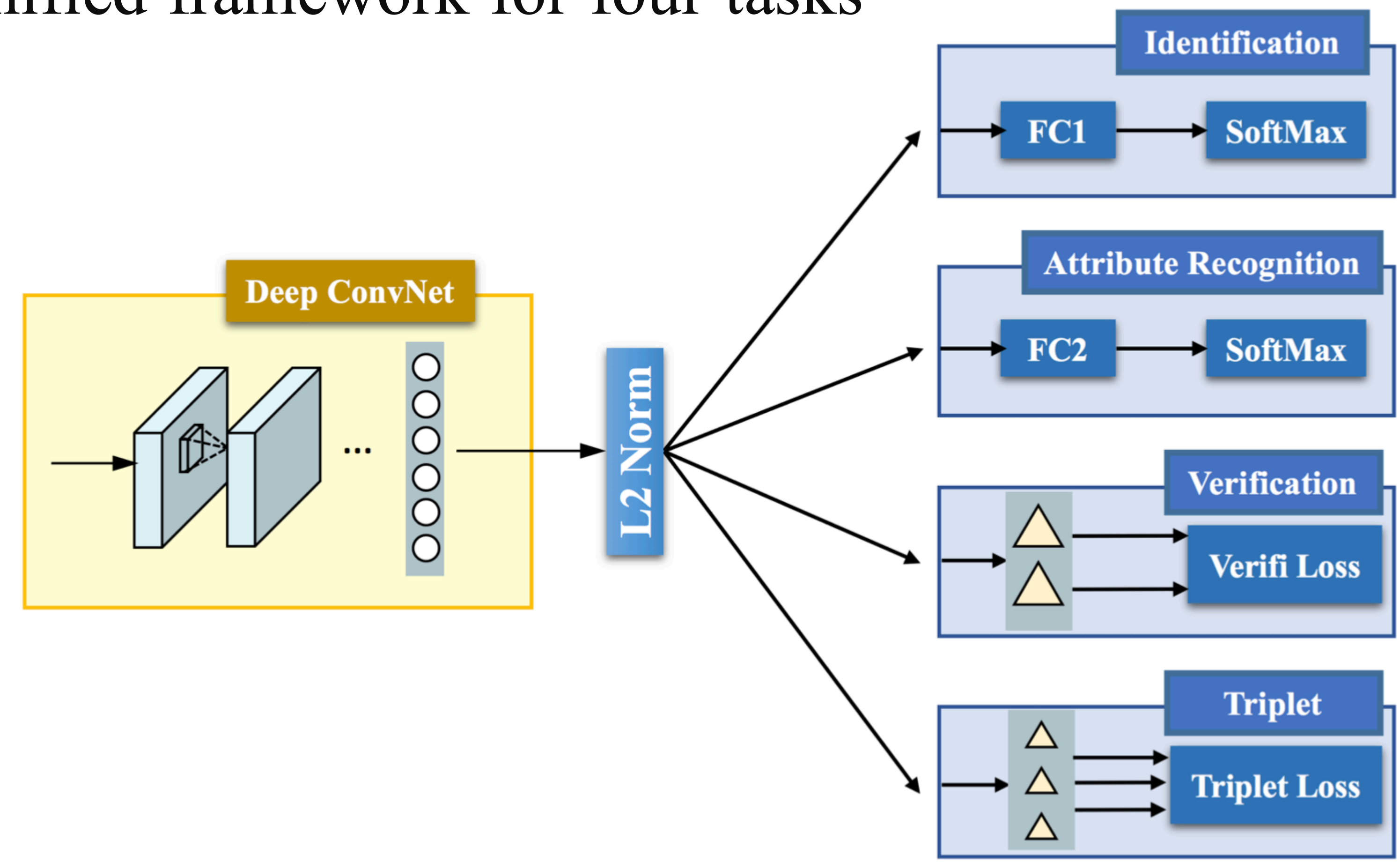
Deep Joint Discriminative Learning

Experimental Results

Conclusion

Architecture Overview

- Unified framework for four tasks



Network Architecture

- Unified framework for four tasks
  - Shared base convolution network
    - A common CNN pretrained on ImageNet
  - Classification tasks
    - Identification
    - Attribute recognition
  - Verification subnetwork
    - Two images
  - Triplet subnetwork
    - Three images

Network Architecture

- Identification subnetwork
  - Each input image → Identity label
  - Conventional recognition task
    - Softmax + cross-entropy loss

$$L_{identi}(f_i) = -\sum_{j=1}^{n} p_j \log \hat{p_j}$$

target label

Predicted probability

Network Architecture

- Attribute recognition subnetwork
  - Jointly recognize vehicle attributes
    - Such as color and vehicle model

$$L_{attri}(f_i) = -\sum_{k=1}^{n_{attri}} \sum_{j=1}^{n_k} a_j^k \log \hat{a}_j^k$$

Network Architecture

- Verification subnetwork
  - Pair-wise siamese network
    - Use Euclidean distance after normalization
      - Distance → small if same identity
      - Distance → large if different identity

$$L_{verif}(f_i, f_j) = \begin{cases} \frac{1}{2}\|f_i - f_j\|_2^2, & v_i == v_j, \\ \frac{1}{2} max(0, \alpha - \|f_i - f_j\|_2)^2, & v_i \neq v_j, \end{cases}$$

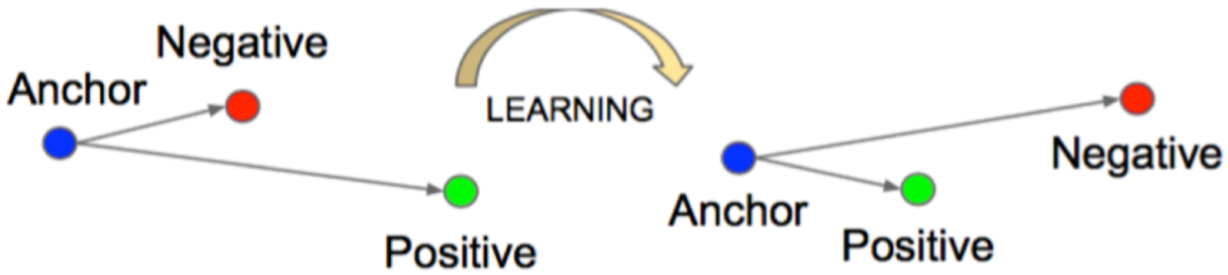
Margin parameter
enforce distance > α

Network Architecture

- Triplet subnetwork
  - Anchor + positive + negative

$$L_{triplet}(f_i, f_j, f_k) = max(0, \|f_i - f_j\|_2^2 - \|f_i - f_k\|_2^2 + \beta)$$

Margin parameter



Training and Optimization

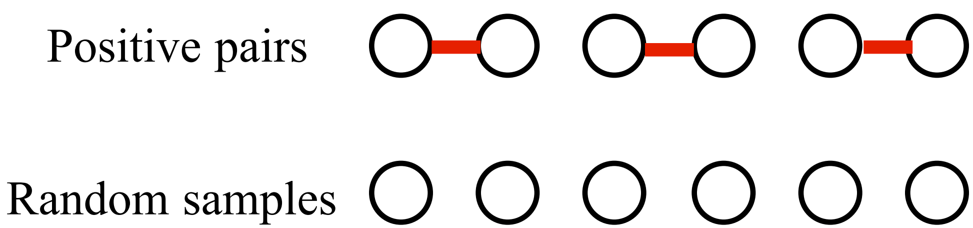
- Objective function

$$L = L_{identi} + L_{attri} + L_{verif} + L_{triplet}$$

  - SGD optimization
  - Jointly learning in a single batch
    - Specific batch composition design

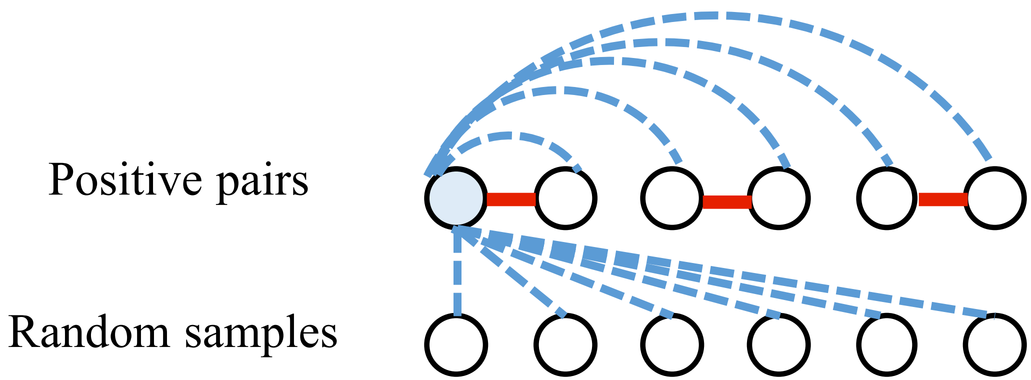
Training and Optimization

- Batch composition design
  - Satisfy four tasks at the same time
    - Half positive pairs + half random samples



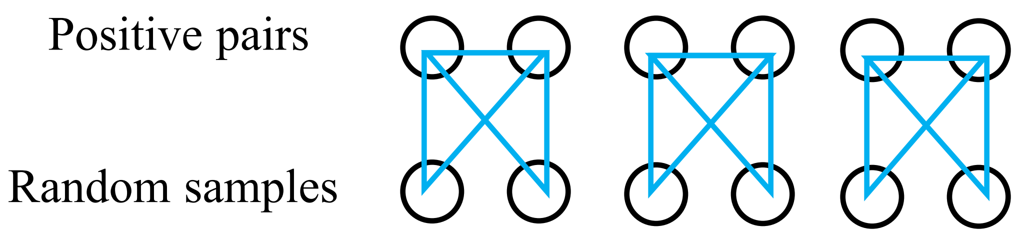
Training and Optimization

- Batch composition design
  - Satisfy four tasks at the same time
    - Verification samples



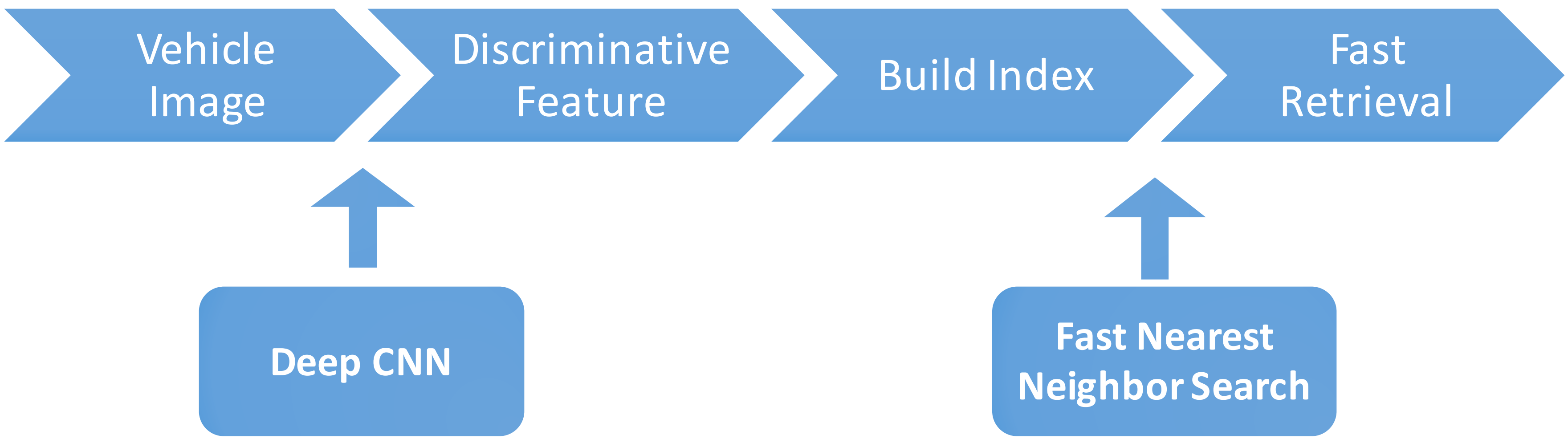
Training and Optimization

- Batch composition design
  - Satisfy four tasks at the same time
    - Triplet samples



Vehicle Retrieval

- Discriminative features → Build index
  - Vehicle Retrieval
    - Nearest neighbor search



Marius Muja and David G Lowe, "Fast approximate nearest neighbors with automatic algorithm configuration," in *VISAPP*, 2009.

Outline

Experimental settings

- VehicleID Dataset
  - 221763 images of 26267 vehicles
  - Three test sets
    - Small, medium, large size
- Two tasks
  - Vehicle retrieval
  - Vehicle re-identification

Experimental settings

- Implementation Details
  - MXNet platform
  - Base convolutional network
    - Inception-BN
  - Augmentation
    - Random crop
    - Random flip
  - Batch size: 64
  - Margin parameters α, β as 0.9

Vehicle Retrieval

- Evaluation protocol
  - Mean average precision (MAP)
- Ablation results

Method	Small	Medium	Large
Identi	0.712	0.684	0.670
Identi+Attri	0.718	0.686	0.672
Identi+Attri+Verifi	0.731	0.705	0.689
Identi+Attri+Verifi+Triplet	**0.786**	**0.747**	**0.720**

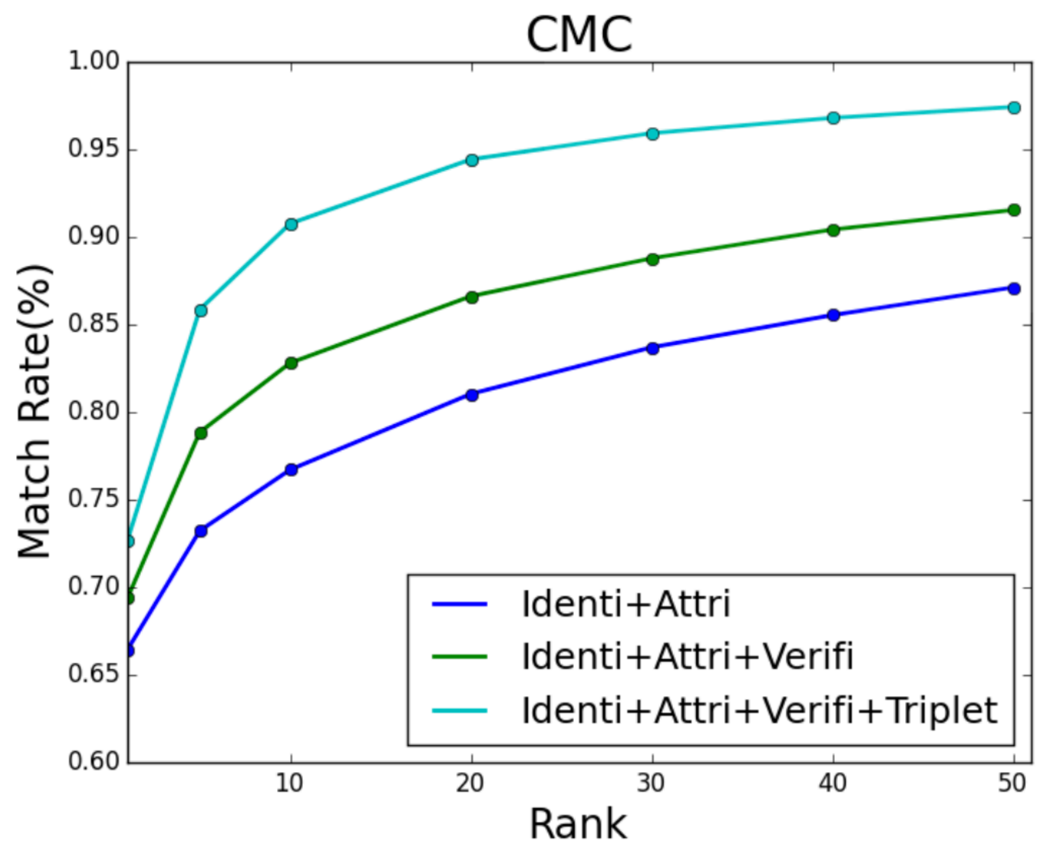
Vehicle Retrieval

- Compare with state-of-art

Method	Small	Medium	Large
VGG+CCL [1]	0.492	0.448	0.386
Mixed Diff+CCL [1]	0.546	0.481	0.455
HDC + Contrastive [16]	0.655	0.631	0.575
Identi+Attri+Verifi+Triplet	**0.786**	**0.747**	**0.720**

Vehicle Re-identification

- Evaluation protocols
  - CMC curve



Vehicle Re-identification

- Evaluation protocols
  - Top1 and Top 5 match rates

Method	Protocol	Small	Medium	Large
VGG+CCL [1]		0.436	0.370	0.329
Mixed Diff+CCL [1]		0.490	0.428	0.382
Identi+Attr	Top 1	0.670	0.667	0.651
Identi+Attr+Verifi		0.689	0.687	0.661
Identi+Attr+Verifi+Triplet		**0.723**	**0.708**	**0.680**
VGG+CCL [1]		0.642	0.571	0.533
Mixed Diff+CCL [1]		0.735	0.668	0.616
Identi+Attr	Top 5	0.735	0.729	0.716
Identi+Attr+Verifi		0.781	0.765	0.737
Identi+Attr+Verifi+Triplet		**0.857**	**0.818**	**0.789**

Conclusion

- A novel Deep Joint Discriminative Learning model
  - For vehicle re-identification and retrieval
  - A unified framework by incorporating four tasks
    - Different properties → benefit each other
    - Jointly optimize
      - specific designed batch composition
- Experiments validate the effectiveness of DJDL model
  - State-of-the-art results on two tasks

lyttonhao@pku.edu.cn

Project Page: http://www.icst.pku.edu.cn/struct/Projects/djdl.html