Deep Joint Discriminative Learning for Vehicle Re-identification and Retrieval

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

2017.09.18
Deep Joint Discriminative Learning for Vehicle Re-identification and Retrieval

Outline

Background
Deep Joint Discriminative Learning
Experimental Results
Conclusion
Deep Joint Discriminative Learning for Vehicle Re-identification and Retrieval

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

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 $\rightarrow$ 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 $\rightarrow$ 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
Deep Joint Discriminative Learning for Vehicle Re-identification and Retrieval

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 $\rightarrow$ Identity label
  - Conventional recognition task
  - Softmax + cross-entropy loss

\[
L_{\text{identity}}(f_i) = -\sum_{j=1}^{n} p_j \log \hat{p}_j
\]

- Target label
- Predicted probability
Network Architecture

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

\[ L_{\text{attri}}(f_i) = - \sum_{k=1}^{n_{\text{attri}}} \sum_{j=1}^{n_k} a_j^k \log \hat{a}_j^k \]
## Network Architecture

- **Verification subnetwork**
  - Pair-wise siamese network
  - Use Euclidean distance after normalization
  - Distance \( \rightarrow \) small if same identity
  - Distance \( \rightarrow \) large if different identity

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

Margin parameter enforce distance \( > \alpha \)
Network Architecture

- Triplet subnetwork
  - Anchor + positive + negative

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

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

Positive pairs: [Diagram of positive pairs]
Random samples: [Diagram of random samples]
Training and Optimization

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

Positive pairs

Random samples
**Training and Optimization**

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

Positive pairs

Random samples
Vehicle Retrieval

- Discriminative features $\rightarrow$ Build index
- Vehicle Retrieval
- Nearest neighbor search

Deep Joint Discriminative Learning for Vehicle Re-identification and Retrieval

Outline

Background

Deep Joint Discriminative Learning

Experimental Results

Conclusion
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 $\alpha$, $\beta$ as 0.9
**Vehicle Retrieval**

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

<table>
<thead>
<tr>
<th>Method</th>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identi</td>
<td>0.712</td>
<td>0.684</td>
<td>0.670</td>
</tr>
<tr>
<td>Identi+Attri</td>
<td>0.718</td>
<td>0.686</td>
<td>0.672</td>
</tr>
<tr>
<td>Identi+Attri+Verifi</td>
<td>0.731</td>
<td>0.705</td>
<td>0.689</td>
</tr>
<tr>
<td>Identi+Attri+Verifi+Triplet</td>
<td><strong>0.786</strong></td>
<td><strong>0.747</strong></td>
<td><strong>0.720</strong></td>
</tr>
</tbody>
</table>
Vehicle Retrieval

- Compare with state-of-art

<table>
<thead>
<tr>
<th>Method</th>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGG+CCL [1]</td>
<td>0.492</td>
<td>0.448</td>
<td>0.386</td>
</tr>
<tr>
<td>Mixed Diff+CCL [1]</td>
<td>0.546</td>
<td>0.481</td>
<td>0.455</td>
</tr>
<tr>
<td>HDC + Contrastive [16]</td>
<td>0.655</td>
<td>0.631</td>
<td>0.575</td>
</tr>
<tr>
<td>Identi+Attri+Verifi+Triplet</td>
<td><strong>0.786</strong></td>
<td><strong>0.747</strong></td>
<td><strong>0.720</strong></td>
</tr>
</tbody>
</table>
Vehicle Re-identification

- Evaluation protocols
  - CMC curve

![CMC Curve Image]
## Vehicle Re-identification

**Evaluation protocols**

- Top1 and Top 5 match rates

<table>
<thead>
<tr>
<th>Method</th>
<th>Protocol</th>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGG+CCL [1]</td>
<td></td>
<td>0.436</td>
<td>0.370</td>
<td>0.329</td>
</tr>
<tr>
<td>Mixed Diff+CCL [1]</td>
<td></td>
<td>0.490</td>
<td>0.428</td>
<td>0.382</td>
</tr>
<tr>
<td>Identi+Attr</td>
<td>Top 1</td>
<td>0.670</td>
<td>0.667</td>
<td>0.651</td>
</tr>
<tr>
<td>Identi+Attr+Verifi</td>
<td></td>
<td>0.689</td>
<td>0.687</td>
<td>0.661</td>
</tr>
<tr>
<td>Identi+Attr+Verifi+Triplet</td>
<td></td>
<td>0.723</td>
<td>0.708</td>
<td>0.680</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Method</th>
<th>Protocol</th>
<th>Small</th>
<th>Medium</th>
<th>Large</th>
</tr>
</thead>
<tbody>
<tr>
<td>VGG+CCL [1]</td>
<td></td>
<td>0.642</td>
<td>0.571</td>
<td>0.533</td>
</tr>
<tr>
<td>Mixed Diff+CCL [1]</td>
<td></td>
<td>0.735</td>
<td>0.668</td>
<td>0.616</td>
</tr>
<tr>
<td>Identi+Attr</td>
<td>Top 5</td>
<td>0.735</td>
<td>0.729</td>
<td>0.716</td>
</tr>
<tr>
<td>Identi+Attr+Verifi</td>
<td></td>
<td>0.781</td>
<td>0.765</td>
<td>0.737</td>
</tr>
<tr>
<td>Identi+Attr+Verifi+Triplet</td>
<td></td>
<td>0.857</td>
<td>0.818</td>
<td>0.789</td>
</tr>
</tbody>
</table>
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
Thank you

lyttonhao@pku.edu.cn

Project Page: http://www.icst.pku.edu.cn/struct/Projects/djd1.html