PERSON RE-IDENTIFICATION WITH DEEP DENSE FEATURE REPRESENTATION AND JOIN T BAYESIAN

Shengke Wang, Lianghua Duan, Na Yang, Junyu Dong
Department of Computer Science and Technology, Ocean University of China, China

ABSTRACT
Person re-identification that aims at matching individuals across multiple camera views has become indispensable in intelligent video surveillance systems. It remains challenging due to the large variations of pose, illumination, occlusion and camera viewpoint. Feature representation and metric learning are the two fundamental components in person re-identification. In this paper, we present a Special Dense Convolutional Neural Network (SD-CNN) to extract the feature and apply Joint Bayesian to measure the similarity of pedestrian image pairs. The SD-CNN can preserve more horizontal information against viewpoint changes, maximize the feature reuse and ensure feature distributing discriminative. Joint Bayesian models the extracted feature representation as the sum of inter- and intra-personal variations, and the joint probability of two images being a same person can be obtained through log-likelihood ratio. Experiments show that our approach significantly outperforms state-of-the-art methods on several benchmarks of person re-identification.

RESULTS
We present a comprehensive evaluation of our framework by comparing it against the baseline SD-CNN feature with Euclidean distance as well as other state-of-art methods for person re-identification. All evaluations are based on the Cumulative Matching Characteristics (CMC).

CONCLUSIONS
This paper proposed a novel framework for person re-identification, it consists of a convolutional neural network extractor named SD-CNN and metric measure named Joint Bayesian. Experiments shown that our framework achieved state-of-the-art result in several dataset.

REFERENCES

CONTACT
Department of Computer Science and Technology, Ocean University of China, China
weimee@ouc.edu.cn
lianghua_duan@outlook.com

We propose a novel framework (See Fig 1) to solve person re-identification problem including a SD-CNN feature extractor and a Joint Bayesian model for distance metric.

The proposed Special Dense Convolutional Neural Network (SD-CNN) architecture, we present a Special Dense Convolutional Neural Network architecture we present a SD-CNN feature extractor and a Joint Bayesian model for distance metric.

For Market-1501 dataset, our final framework gains 72.6% rank-1 accuracy, both Euclidean distance and Joint Bayesian with our SD-CNN feature are greatly better than the previous methods.

We learn the Joint Bayesian model for distance metric based on the extracted SD-CNN feature. Follow the Joint Bayesian (2), the feature of a pedestrian image can be represent as the sum of inter- and intra-personal variations.

\[ X = \mu + \epsilon \]

where \( \mu \) and \( \epsilon \) follow two Gaussian distributions \( \mathcal{N}(0, \Sigma_L) \) and \( \mathcal{N}(0, \Sigma_I) \) can be estimated from the training data.

Given the features of image pairs \( x_1, x_2 \) extracted by SD-CNN from two images, let \( \mathcal{M} \) represents the inter-personal (same) hypothesis that two images belong to the same person, and \( \mathcal{I} \) is the extra-personal (not same) hypothesis, then the person re-id problem amounts to classifying the difference \( x_1 - x_2 \) as inter-personal variation or extra-personal variation. Based on the MAP/Maximization of the posterior rule, the distance is made by testing a log-likelihood ratio:

\[ d(x_1, x_2) = \log \frac{P(\mathcal{M} | H_1)}{P(\mathcal{I} | H_1)} = \mu_1^T \Delta \Sigma_1^{-1} \Delta \mu_2 + \frac{1}{2} \left( \mu_1^T \Delta \Sigma_1^{-1} \Delta \mu_2 + 2 \mu_1^T \Delta \Sigma_1^{-1} \Delta \epsilon_2^{T} + 2 \epsilon_2^T \epsilon_2 - 2 \mu_1^T \epsilon_2 \right) \]

where \( \Delta = \Sigma_L - \Sigma_I \) and \( \epsilon_1 \) can be estimated by the algorithm in Table 1

Table 1. The Joint Bayesian learning algorithm. Assume there are \( n \) identities and each identity has \( m_i \) images.

For CUKH03 dataset, our final framework gains up to 82.3%.

RESULTS
We present a comprehensive evaluation of our framework by comparing it against the baseline SD-CNN feature with Euclidean distance as well as other state-of-art methods for person re-identification. All evaluations are based on the Cumulative Matching Characteristics (CMC).

We learn the Joint Bayesian model for distance metric based on the extracted SD-CNN feature. Follow the Joint Bayesian (2), the feature of a pedestrian image can be represent as the sum of inter- and intra-personal variations.

\[ X = \mu + \epsilon \]

where \( \mu \) and \( \epsilon \) follow two Gaussian distributions \( \mathcal{N}(0, \Sigma_L) \) and \( \mathcal{N}(0, \Sigma_I) \) can be estimated from the training data.

Given the features of image pairs \( x_1, x_2 \) extracted by SD-CNN from two images, let \( \mathcal{M} \) represents the inter-personal (same) hypothesis that two images belong to the same person, and \( \mathcal{I} \) is the extra-personal (not same) hypothesis, then the person re-id problem amounts to classifying the difference \( x_1 - x_2 \) as inter-personal variation or extra-personal variation. Based on the MAP/Maximization of the posterior rule, the distance is made by testing a log-likelihood ratio:

\[ d(x_1, x_2) = \log \frac{P(\mathcal{M} | H_1)}{P(\mathcal{I} | H_1)} = \mu_1^T \Delta \Sigma_1^{-1} \Delta \mu_2 + \frac{1}{2} \left( \mu_1^T \Delta \Sigma_1^{-1} \Delta \mu_2 + 2 \mu_1^T \Delta \Sigma_1^{-1} \Delta \epsilon_2^{T} + 2 \epsilon_2^T \epsilon_2 - 2 \mu_1^T \epsilon_2 \right) \]

where \( \Delta = \Sigma_L - \Sigma_I \) and \( \epsilon_1 \) can be estimated by the algorithm in Table 1

Table 1. The Joint Bayesian learning algorithm. Assume there are \( n \) identities and each identity has \( m_i \) images.

For CUKH03 dataset, our final framework gains up to 82.3%.

RESULTS
We present a comprehensive evaluation of our framework by comparing it against the baseline SD-CNN feature with Euclidean distance as well as other state-of-art methods for person re-identification. All evaluations are based on the Cumulative Matching Characteristics (CMC).

We learn the Joint Bayesian model for distance metric based on the extracted SD-CNN feature. Follow the Joint Bayesian (2), the feature of a pedestrian image can be represent as the sum of inter- and intra-personal variations.

\[ X = \mu + \epsilon \]

where \( \mu \) and \( \epsilon \) follow two Gaussian distributions \( \mathcal{N}(0, \Sigma_L) \) and \( \mathcal{N}(0, \Sigma_I) \) can be estimated from the training data.

Given the features of image pairs \( x_1, x_2 \) extracted by SD-CNN from two images, let \( \mathcal{M} \) represents the inter-personal (same) hypothesis that two images belong to the same person, and \( \mathcal{I} \) is the extra-personal (not same) hypothesis, then the person re-id problem amounts to classifying the difference \( x_1 - x_2 \) as inter-personal variation or extra-personal variation. Based on the MAP/Maximization of the posterior rule, the distance is made by testing a log-likelihood ratio:

\[ d(x_1, x_2) = \log \frac{P(\mathcal{M} | H_1)}{P(\mathcal{I} | H_1)} = \mu_1^T \Delta \Sigma_1^{-1} \Delta \mu_2 + \frac{1}{2} \left( \mu_1^T \Delta \Sigma_1^{-1} \Delta \mu_2 + 2 \mu_1^T \Delta \Sigma_1^{-1} \Delta \epsilon_2^{T} + 2 \epsilon_2^T \epsilon_2 - 2 \mu_1^T \epsilon_2 \right) \]

where \( \Delta = \Sigma_L - \Sigma_I \) and \( \epsilon_1 \) can be estimated by the algorithm in Table 1

Table 1. The Joint Bayesian learning algorithm. Assume there are \( n \) identities and each identity has \( m_i \) images.

For CUKH03 dataset, our final framework gains 72.6% rank-1 accuracy, both Euclidean distance and Joint Bayesian with our SD-CNN feature are greatly better than the previous methods.

TABLE 1.

<table>
<thead>
<tr>
<th>Method</th>
<th>Rank-1 Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>SD-CNN</td>
<td>67.8%</td>
</tr>
<tr>
<td>Joint Bayesian</td>
<td>72.6%</td>
</tr>
</tbody>
</table>

However, for small dataset CUKH01 [5], it would be insufficient to learn such a large capacity network from scratch, our framework with SD-CNN and Joint Bayesian fully achieve the best result compared with others (see Table 2).

CONTACT
Department of Computer Science and Technology, Ocean University of China, China
weimee@ouc.edu.cn
lianghua_duan@outlook.com

Code is available at https://github.com/duanLH/