# 浏览式阅读

## 1 自己的总结、评价以及应用

本文就介绍了这样一个东西：a context learning framework with graph model.

In this work, we design a target-context graph and employ a pairwise GCN to learn visual relations in the scene.

## 2 文章的主要问题（abstract、疑问句中）

存在的问题：

when there are no overlapping fields of view between cameras.

提出的解决方案：

两点：１ （特征提取：network architecture）we present a novel multi-channel parts-based convolutional neural network (CNN) model 得到一个整体（全局）＋局部　的　good image features 换言之：more discriminative and robust feature representations

２ （损失计算：loss function.）under the triplet framework for person re-identification.　A suitable distance metrics　具体而言：that requires the intra-class feature distances to be less than not only the inter-class ones, but also a predefined threshold.

简而言之，就是一个基于多局部特征的多通道卷积神经网络＋triplet损失函数。

## 3 结论（abstract以及conclusion中）

## 4 思路脉络（小标题中的关键句）

**1. Introduction**

Multi-view存在challenge的原因：

①appearance包括两点：variations和similar②human pose③background

问题：in order to find the correct matches among a large set of candidate images captured by different cameras

需要解决的２个问题：

①good image features

First, good image features are required to represent both the query and the gallery images.

②suitable distance metrics

suitable distance metrics are indispensable to determine whether a gallery image contains the same as the query image.

具体措施：

1. To extract better features for raw person images, we propose a new, multi-channel CNN model that learns features for both the input person’s full body and the body parts.

②We also borrow the idea from Wang’s et al. [39] and the FaceNet work [34] to use triplet training examples and the improved triplet loss function to further enhance the discriminative power of the learned features.

很重要的一句话：

In contrast to the original triplet loss function that only requires the intra-class feature distances to be less than the inter-class ones.类内距离小于类间距离

最后，该文献的两大贡献：

The main contributions of this paper are twofold: 1) a novel, multi-channel CNN model that learns both the global full-body and the local parts features, and integrates them

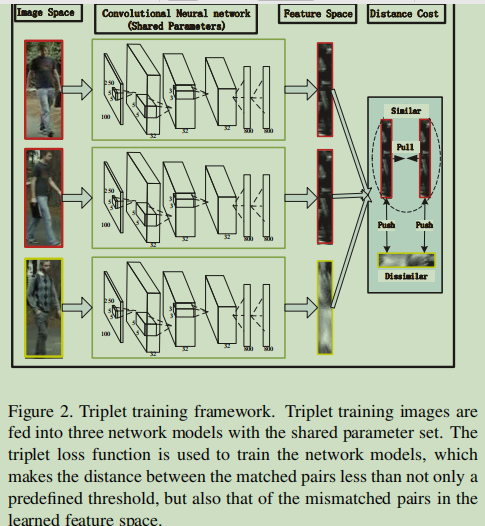
together to produce the final feature representation of the input person; 2) an improved triplet loss function that requires the intra-class feature distances to be less than not only the inter-class ones, but also a predefined threshold. Experimental evaluations results show that the proposed method achieves the state-of-the-art performances on several widely adopted person re-id benchmark test datasets.

1. Related Work

什么是better features？

better features that are at least partially invariant to lighting, pose, and viewpoint variations

1. The Proposed Person Re-Id Method
   1. The Overall Framework



3.2. Multi-Channel Parts-based CNN Model

1. **Conclusion**

In this work, we propose to employ contextual information to improve the robustness of person search results.（说白了就是一个上下文情境图）

## 5 难理解点

专业术语：

different illumination：照明

Occlusion：遮挡

context information：上下文信息

probe-gallery pair：探测库对

contextual instance expansion module：上下文实例拓展模块

multi-camera surveillance systems：多摄像机监控系统 surveillance：监视

intra-class variations：组内变异 background clutter：背景混乱

manually cropped image snapshots or video clips：手动裁剪的图像快照或视频剪辑

appearance cue：外观提示 manual annotations：手动注释 semantic group：语义群

spatial and temporal cue：空间和时间线索 texture：纹理

latent subspace：潜在子空间