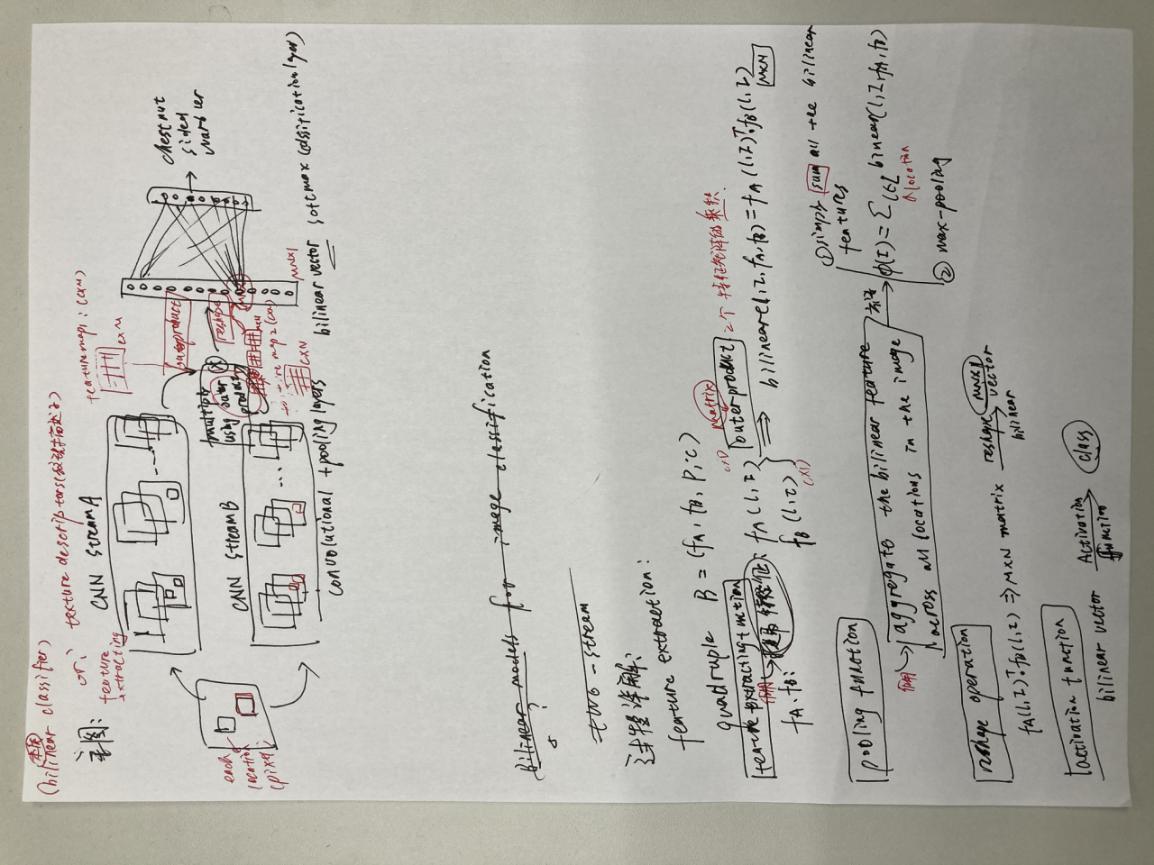
## Bilinear CNN Models for Fine-grained Visual Recognition

## 论文解读：



## 要查的几个问题：

1 bilinear vector

2 hand-crafted features

3 Bilinear classifiers for visual recognition

4 orderless pooling methods（无序池化即为不考虑特征的location信息）与bilinear model的关系

5 texture descriptors纹理描述子

6 SIFT

Bilinear的核心：

两个流stream：

1 腹流：models the temporal aspect 时间

2 背流： models the spatial aspect 空间

biinear模型的组成：

A bilinear model B for image classification consists of a quadruple B = (f A , f B , P, C). Here f A and f B are feature functions, P is a pooling function and C is a classification function.

A feature function is a mapping f : L ⇥ I ! R c⇥D that takes an image.

bilinear特征融合的表示形式：

the bilinear feature combination of f A and f B at a location l is given by

bilinear(l, I, f A , f B ) = f A (l, I) T （转置）f B (l, I).

下面是这3种函数的具体使用和形式：

fa fb：特征提取函数

P：池化函数

C：分类函数

## 1 自己的总结、评价以及应用

We propose bilinear models, a recognition architecture that consists of two feature extractors.

This architecture can model local pairwise feature interactions in a translationally invariant manner which is particularly useful for fine-grained categorization. It also generalizes various orderless texture descriptors such as the Fisher vector, VLAD and O2P.

## 2 文章的主要问题（abstract、疑问句中）

CCA（典型相关性分析）方法在multi-view应用上存在问题：

（问题1：）it makes no use of class labels.

Recent CCA methods have started to address this

weakness but are limited in（问题2：又存在别的问题） that they do not simultaneously optimize the CCA

projection for discrimination and the CCA projection itself, or they are linear only

解决方法：

We address these deficiencies by simultaneously optimizing a CCA-based and a

task objective in an end-to-end manner.具体而言：We proposed a method to find a shared latent space that is also discriminative by adding a task-driven

component to deep CCA while enabling end-to-end training.再具体而言：This was accomplished by replacing the CCA projection with `2 distance minimization and orthogonality constraints on the activations, and was implemented in three different ways.

## 3 结论（abstract以及conclusion中）

## 4 思路脉络（小标题中的关键句）

1 Introduction

尽管CCA很多特征有助于判别任务，但是它依旧有很多缺点... ...non-linear extensions of CCA更是放大了这个问题

我们提出了一个方法：

Therefore, we present a new deep learning technique

to project the data from two views to a shared space that is also discriminative

我们的工作与之前的不同：

both the shared latent space and a task-driven objective

the CCA objective can equivalentlybe expressed as an `2 distance minimization in the shared space plus an orthogonality constraint.

we present three techniques to accomplish this.While our method is derived from CCA, by manipulating the orthogonality constraints, we obtaindeep CCA approaches that compute a shared latent space that is also discriminative.

2 Background

CCA. The objective is to maximize the correlation between a1 = w> 1 X1 and a2 = w> 2 X2

DCCA. Deep CCA adds non-linear projections to CCA by non-linearly mapping the input via a multilayer perceptron (MLP).

SoftCCA. While DCCA enforces orthogonality constraints on projections W>

1 A1 and W>2 A2,SoftCCA relaxes them using regularization [8].

总结：

we optimize the non-linear mapping into the shared space together with the CCA part

3 Task-Optimal CCA (TOCCA)

一开始

DCCA formulation + task-driven term

然后DCCA存在问题，解决办法：

We tackle this by focusing on the two components of DCCA: maximizing the sum correlation between activations A1 and A2 and enforcing orthonormality constraints within A1 and A2. We achieve both by transforming the CCA objective and present three methods that progressively relax the orthogonality constraints.以及 ：enabling mini-batch computations for improved flexibility and test performance

**Task-driven objective.**

**Orthogonality constraints.**The remaining complications for mini-batch optimization are the orthogonality constraints, for which we propose three solutions, each handling the orthogonality constraints of CCA in a different way: whitening, soft decorrelation, and no decorrelation.

1) Whitening (TOCCA-W) We use a Zero-phase Component Analysis (ZCA) whitening transform

2) Soft decorrelation (TOCCA-SD) In this second formulation we relax the orthogonality constraints using regularization

3) No decorrelation (TOCCA-ND) TOCCA-ND therefore removes the decorrelation term entirely

**Computational complexity**

In summary, all three variants are motivated by adding a task-driven component to deep CCA.

5 Discussion

提供了一种方法：

We proposed a method to find a shared latent space that is also discriminative by adding a task-driven component to deep CCA while enabling end-to-end training. This was accomplished by replacing the CCA projection with `2 distance minimization and orthogonality constraints on the activations, and was implemented in three different ways.

## 5 难理解点

专业术语：

i.e.=that is / in other，用来进一步解释前面所表明的观点，可以翻译为“亦即”

semi-supervised learning：半监督学习

clustering：聚类

orthogonality constraint：正交性约束

decorrelation ：解相关

covariance matrice：协方差矩阵 cross-covariance matrix：互协方差矩阵

singular value decomposition：奇异值分解

denotes：表示 regularization：正则化

matrix trace norm objective：矩阵跟踪规范目标

off-diagonal elements：非对角线元素

incorporate ：合并

non-linear mapping：非线性映射

schematic diagrams：示意图‘原理图

equivalent loss function：等效损失函数

stochastic gradient optimization： 随机梯度优化

identity matrix：单位矩阵

eigendecomposition ：本征合成