

# RelatedWork

---

下面我们将从几个角度简要回顾与本文 Task 相关的 Related Work，包括 CNN 特征融合、细粒度特征学习和两种线性池化方法。

## CNN中的特征融合

深度学习的迅速发展，让CNN成为了目前计算机视觉任务的头号选手。目前CNN已经称为各种计算机视觉识别任务的通用的特征提取方法。经过一段时间的发展后，可以发现简单的进行特征图融合是不利于细粒度分类问题的。其缺陷有如下两点：

- 导致训练参数的增大
- 卷积层相结合时无法正确捕捉层间特征关系

## 特征学习对细粒度识别问题的影响

特征学习在细粒度识别问题中是十分重要的。由于细粒度分类问题中各个类别之间的差异十分细小且局部性较强，所以在训练过程中只使用传统方法获取整体语义分类信息是不利于网络学习的，这限制了网络的语义信息表达能力，从而使得最终分类的结果不理想。为了缓解这一问题，*双线性池化*（Bilinear-pooling）[[1] Lin T Y , Roychowdhury A , Maji S . Bilinear CNN Models for Fine-grained Visual Recognition[J]. 2015.] 方法于2015年被提出用于 Fine-grained 分类，到现今，此方法得到了充分的发展与补充。在本文中，我们使用了基于此方法与 Bilinear-Model 改进而来的 Hierarchical Bilinear Pooling [[1] Yu C , Zhao X , Zheng Q , et al. Hierarchical Bilinear Pooling for Fine-Grained Visual Recognition[J]. 2018.]

## 双线性池化（Bilinear-pooling）

*双线性池化*（Bilinear-pooling）主要用于特征融合，即对于同一个样本的同一位置提取出来的不同特征  $x$  和  $y$ ，通过双线性池化操作得到两个特征融合后的向量，进而将此向量用于分类。在*双线性池化*论文中，作者使用*双线性池化*对两种特征提取器提取到的特征进行融合，再用于 Fine-grained 分类，取得了不错的效果。与此同时，*双线性池化*的优点是形式简单、梯度计算简单、反向传播迅速。然而本方法仅考虑来自单个卷积层的特征融合，并不能从根本上解决上述问题。为此，Yu 等人提出了一种新的模型来克服这些问题。

## Hierarchical Bilinear Pooling

Hierarchical Bilinear Pooling（HBP）框架使用的是一种通过级联多个跨层的双线性池化模块来融合更多的卷积层特征。这相比传统的双线性池化有着如下优点：

- 增强了卷积层之间的特征融合
- 几乎没有增加训练参数

- 模型表达能力更强