



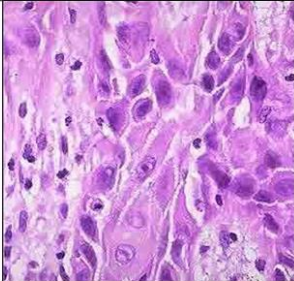
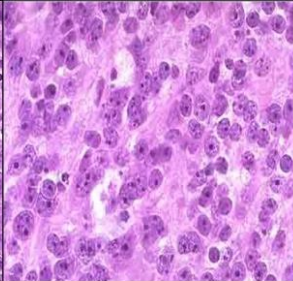


北京大学
PEKING UNIVERSITY

数字图像处理 大作业

2018.4.9

图像细粒度分类

- 图像分类的问题
 - 粗粒度的图像分类没有重点考虑内容相似的图像之间的精细差异，很难对指定类别的图像做进一步区分，无法很好地满足多层次分类需求
 - 粗粒度图像分类已不能满足人类的需求
- 细粒度图像分类
 - 细粒度图像分类指的是对属于相同类别下的不同子类别的物体进行分类
 - 常见应用场景：自动驾驶、生物保护、癌细胞检测等

输入图像				
粗粒度图像分类	癌细胞	癌细胞	车	车
细粒度图像分类	鳞状上皮癌细胞	腺癌细胞	现代伊兰特轿车 2007	丰田红杉越野车 2012

图像细粒度分类

- 细粒度图像分类在学术界和工业界都受到了广泛的关注



CVPR Workshop on FGVC



斯坦福Cars-196数据集



图像细粒度分类

- 细粒度图像分类面临着一些问题：
 - 类内差异大
 - 物体姿势、形状等的变化使得同一子类别内差异大，如鸟类的运动姿势和静止姿势差异很大
 - 拍摄角度、光照等的变化也使得同一子类别内差异大



黑脚信天翁



Smart Fortwo敞篷版2012

图像细粒度分类

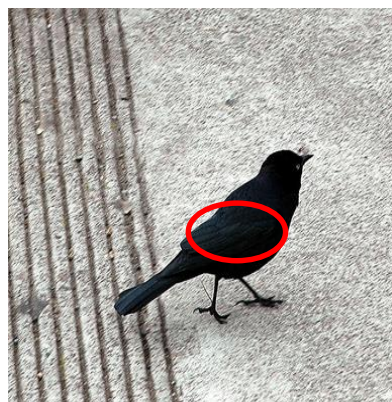
- 细粒度图像分类面临着一些问题：

- 类间差异小

- 粗粒度图像分类的对象属性差异大，而细粒度图像分类通常属于一个大类，差异很小
 - 不同子类别间的差异主要表现在细微的局部区域上，差异很小，如翅膀或者嘴的颜色等



红翅乌鸫



蓝头黑鹇



北极燕鸥



福斯特燕鸥

图像细粒度分类

- 细粒度图像分类面临着一些问题：
 - 子类别数量更多，分类难度更大
 - 据国际鸟类联盟（**BirdLife**）统计目前鸟类有**10,426**种
 - 据英国癌症研究（**Cancer Research UK**）统计目前有**200**多种癌细胞
 - 标注难度大，需要专业知识
 - 由于不同子类别间难以判别，需要具有专业知识的人进行标注
 - 不同子类间区别主要表现于局部区域上，局部区域的标注更加困难
- 如何有效提取不同子类别的具有判别性的特征？

图像细粒度分类

- 我们希望利用判别性部件来研究细粒度图像分类，缓解上述问题，**提取不同子类别的具有判别性的特征**
 - 不同子类别间差异主要是局部部件，人类视觉注意力系统处理图像时优先选择具有判别性的区域，判别性部件适用于细粒度图像分类
 - 利用视觉注意力模型获取判别性部件，使得不同子类更易区分，从而支持更多类别的分类
 - 弱监督（只使用图像类别的标注，不使用部件级的标注）得获取判别性部件，可以减少所需的标注信息

细粒度图像的标注

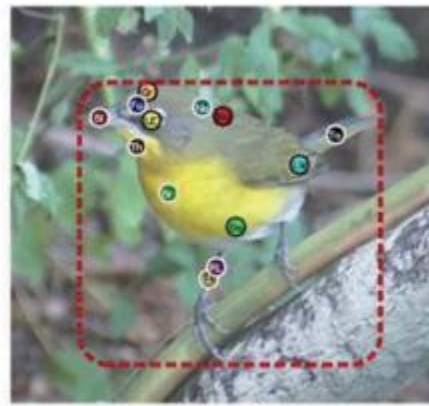
- 部件（parts）是细粒度图像分析的重要成分
 - 但是其标注复杂、耗时、昂贵、易错
 - 聘请专家是昂贵的，专家也会出现错误



北极燕鸥



里海燕鸥



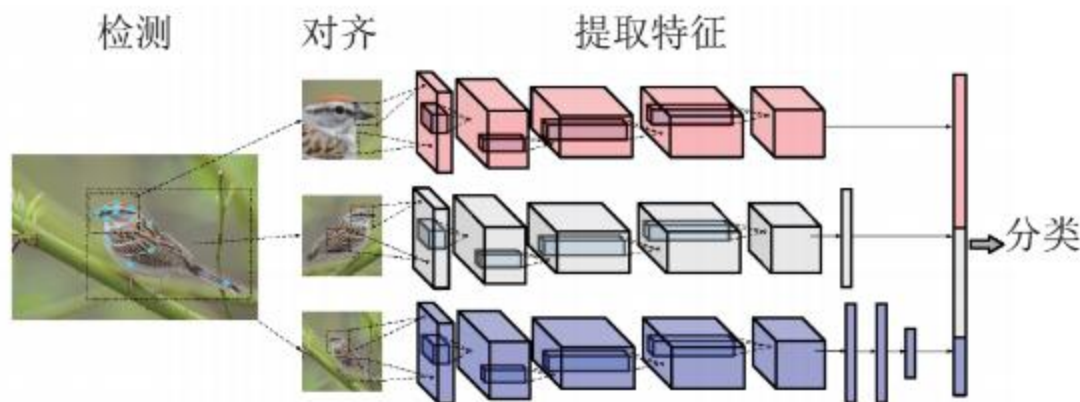
Caltech-UCSD CUB200-2011

早期工作

- Wah et al., 颜色直方图+Bag of words, 准确率10.3%
 - 使用细粒度标注, 17.3%
 - 是发布该数据集的技术报告
- Berg et al., 基于局部区域的特征编码POOF, 56.8%
 - 使用细粒度标注, 73.3%
 - CVPR 2013
- 物体部件parts对细粒度图像分析至关重要!

学习物体、部件检测器

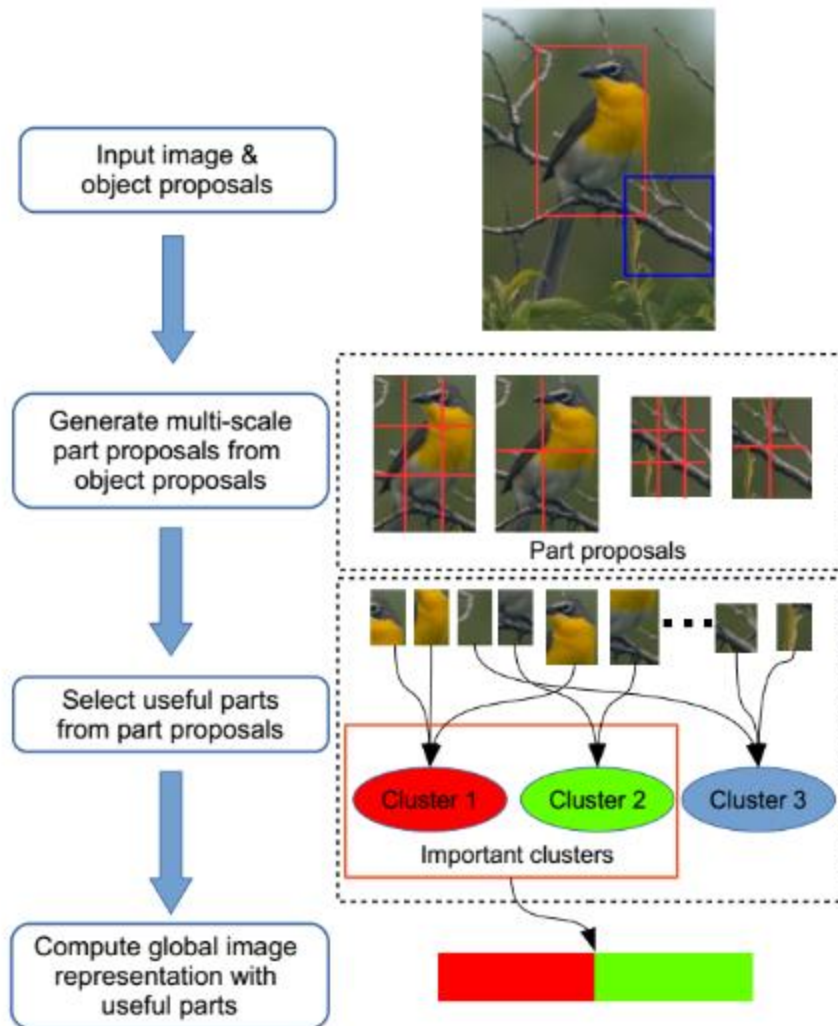
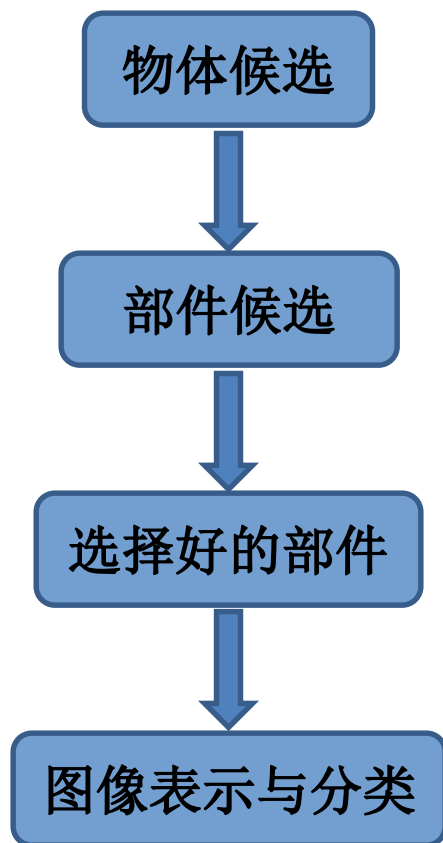
- Bird Species Categorization Using Pose Normalized Deep Convolutional Nets, BMVC2014
 - 使用细粒度标注+DPM训练检测器, CNN进行分类
 - 测试时: 用细粒度标注, 85.4%; 用检测器, 75.7%



- 训练检测器可能比分类更难!
- 直接全图像fine-tune VGG16可以有73.99%!

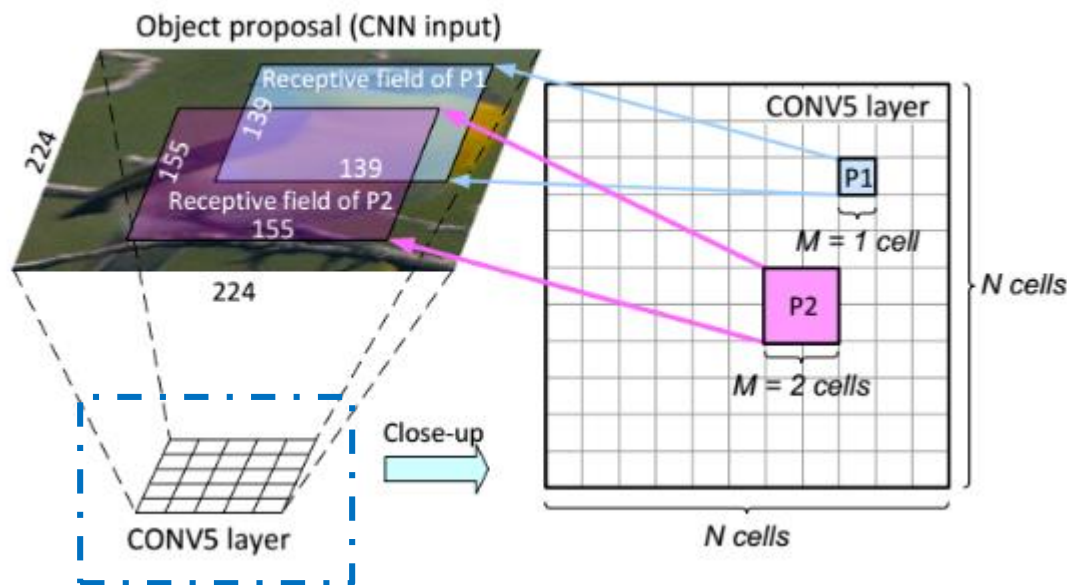
检测 vs. 候选+选择

- 弱监督：只使用图像标签（label）



物体和部件的候选

- 物体候选: Selective search
- 部件候选:
 - 需要物体中不同大小、位置的候选



$$z_{i,j,k}^M = \max_{\substack{i \leq p < i+M, \\ j \leq q < j+M}} X_{p,q,k},$$

s.t. $1 \leq M \leq N, 1 \leq k \leq d,$

研究现状

2.1 基于深度网络集成的方法

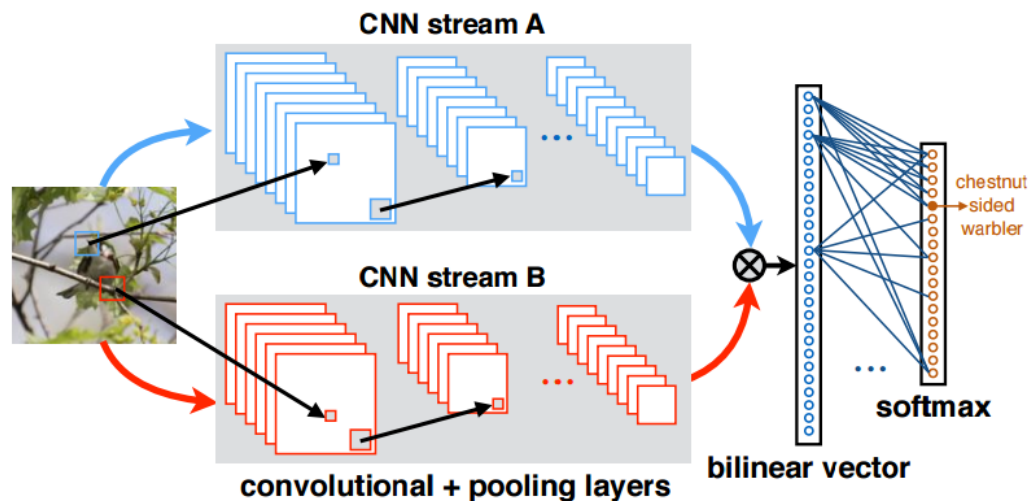
2.2 基于部件检测的方法

2.3 基于视觉注意力模型的方法

研究现状

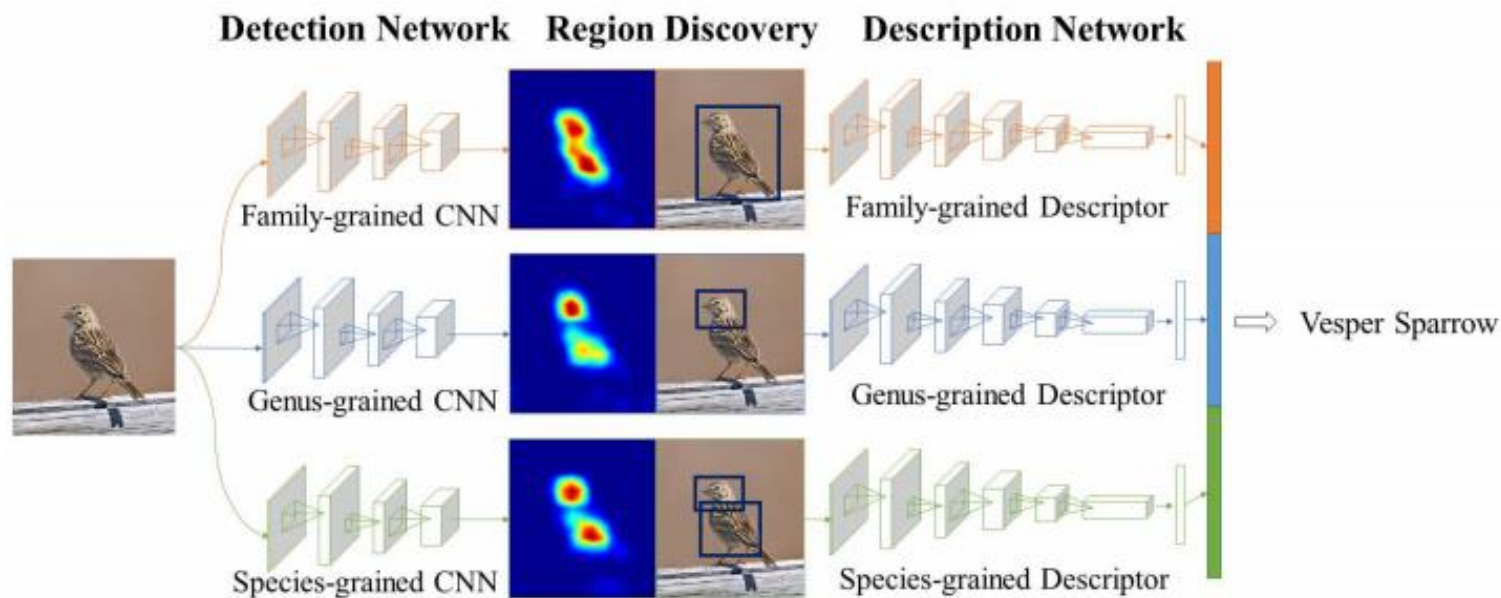
2.1 基于深度网络集成的方法

- 基于深度网络集成（ensemble of networks）的方法
 - 集成不同的深度网络模型，学习多种图像的特征
- [Lin, et al. ICCV 2015]提出一个端到端的集成两个CNN网络的方法，网络A的作用是对物体进行定位，即完成传统算法的物体与局部区域检测工作，而网络B则是用来对网络A检测到的物体位置进行特征提取



研究现状

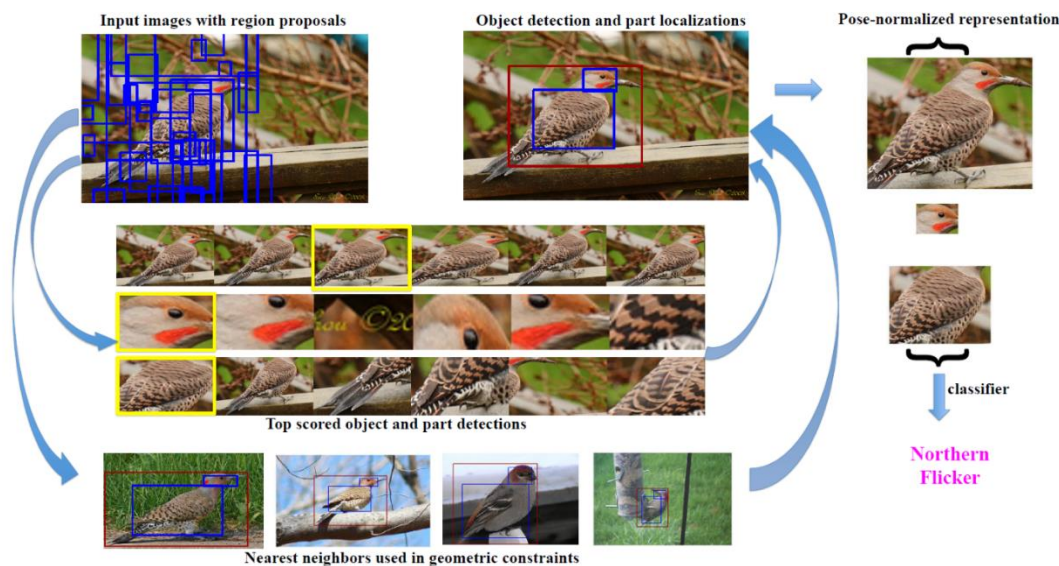
- [Wang, et al. ICCV 2015] 提出一个结合分类层次和特征学习的方法，针对图像的不同分类层次的兴趣区域分别训练CNN模型，最后进行集成得到具有判别性的特征
- 该方法采用科、属、种3个层次进行训练（分类标签采用了外部信息）



研究现状

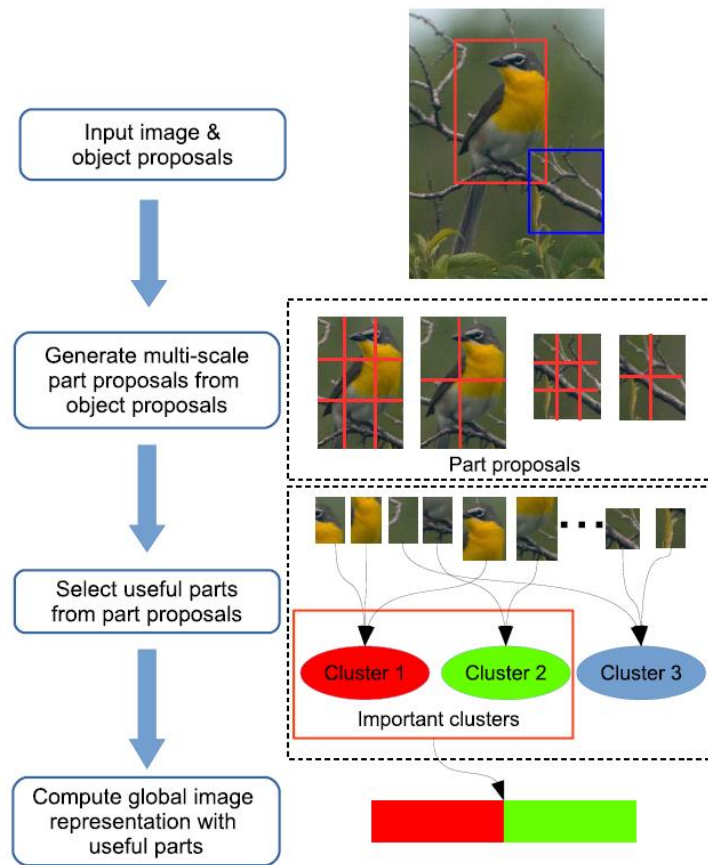
2.2 基于部件检测的方法

- 基于部件检测的方法
 - 训练物体的部件检测器，根据部件特征训练分类器
- [Zhang, et al. ICML 2014]首次利用R-CNN检测具有判别性的部件，然后通过几何约束筛选部件，最后利用部件的特征训练一个分类器



研究现状

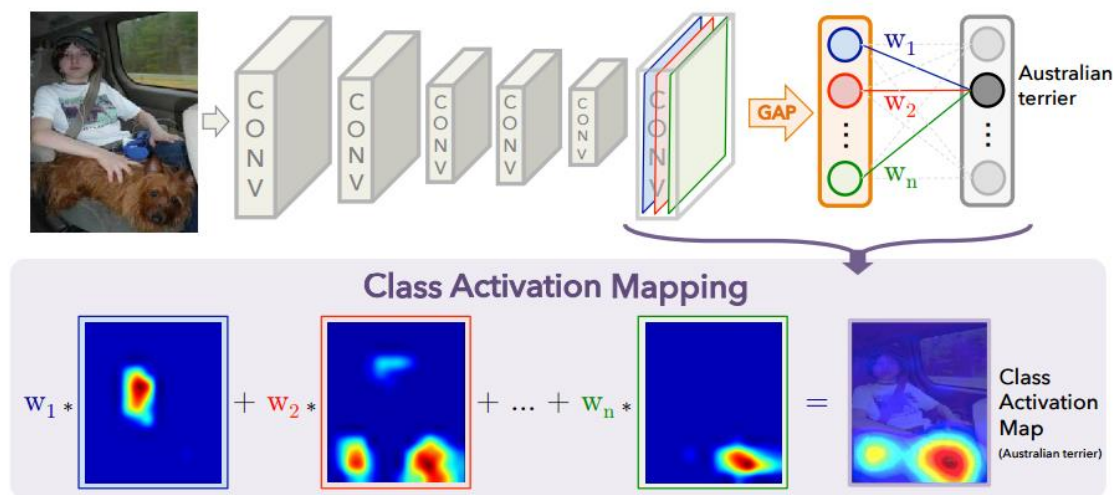
- [Zhang, et al. TIP 2016]在物体区域生成不同尺度的部件防止图像背景信息带来的影响，利用聚类方法去除无效部件
 - 生成物体候选块
 - 生成不同尺度的部件
 - 利用部件聚类筛选具有判别性的部件
 - 融合不同尺度的部件特征用于分类



研究现状

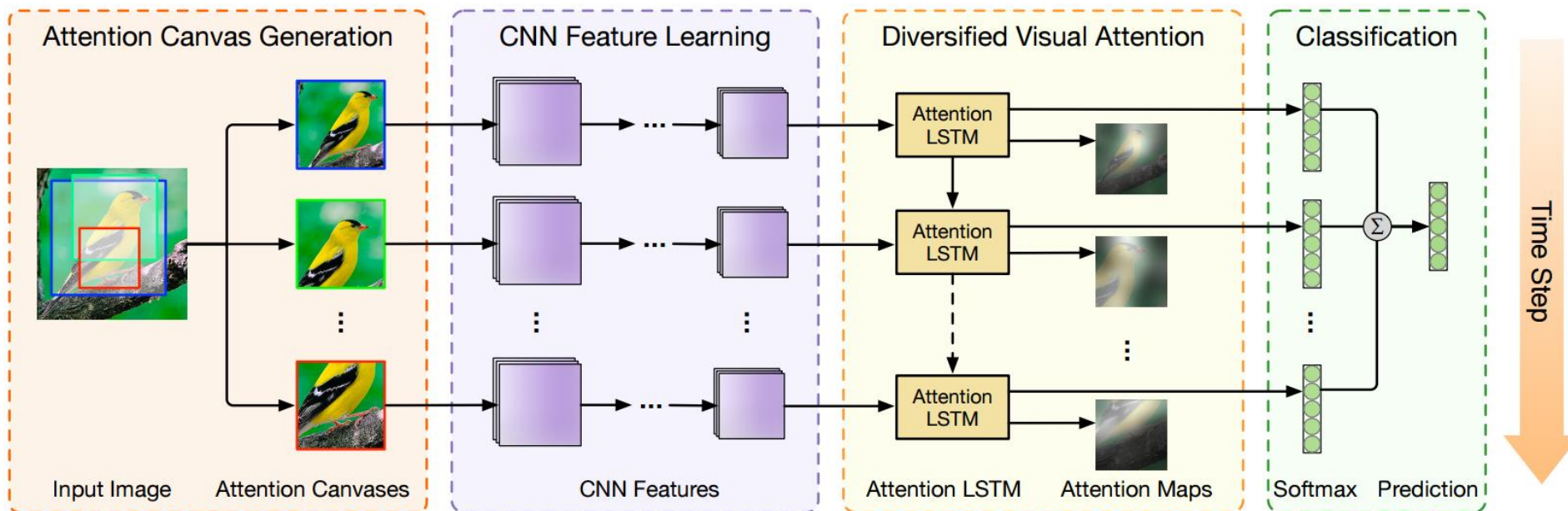
2.3 基于视觉注意力模型的方法

- 基于视觉注意力模型的方法
 - 利用视觉注意力模型检测图像具有判别性的区域
- [Zhou, et al. CVPR 2016]利用全局均值池化（global average pooling, GAP）生成和物体相关的显著图，从而获得具有判别性的区域



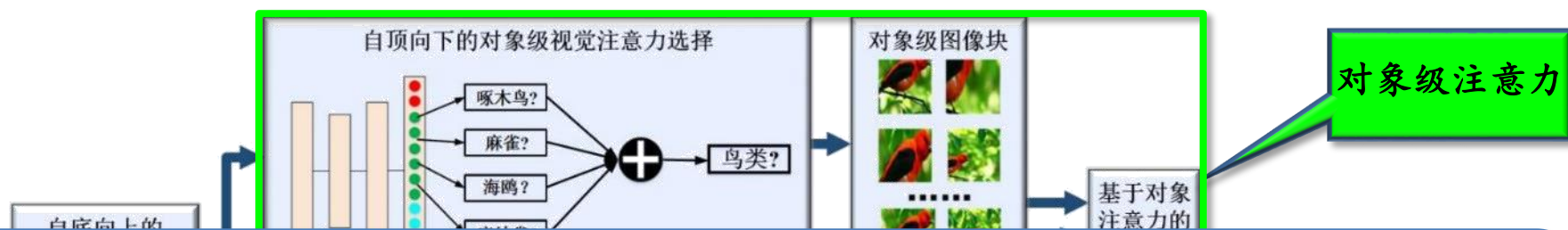
研究现状

- [Zhao, et al. TMM 2016] 提出一个多样化的视觉注意力网络（diversified visual attention network, DVAN），同时得到物体和判别性部件，该方法端到端的训练自底向上和自顶向下的注意力模型，优化了候选块的生成；同时，LSTM的使用使得分类逐渐由粗粒度趋向于细粒度



研究现状

- 针对细粒度特征学习问题，**提出了一种物体—局部两级视觉注意力的深度模型**，在不需要物体与局部位置标注的情况下，利用深度卷积网络的特征检测能力实现了物体级别与局部级别的视觉注意力，提升细粒度图像识别效果

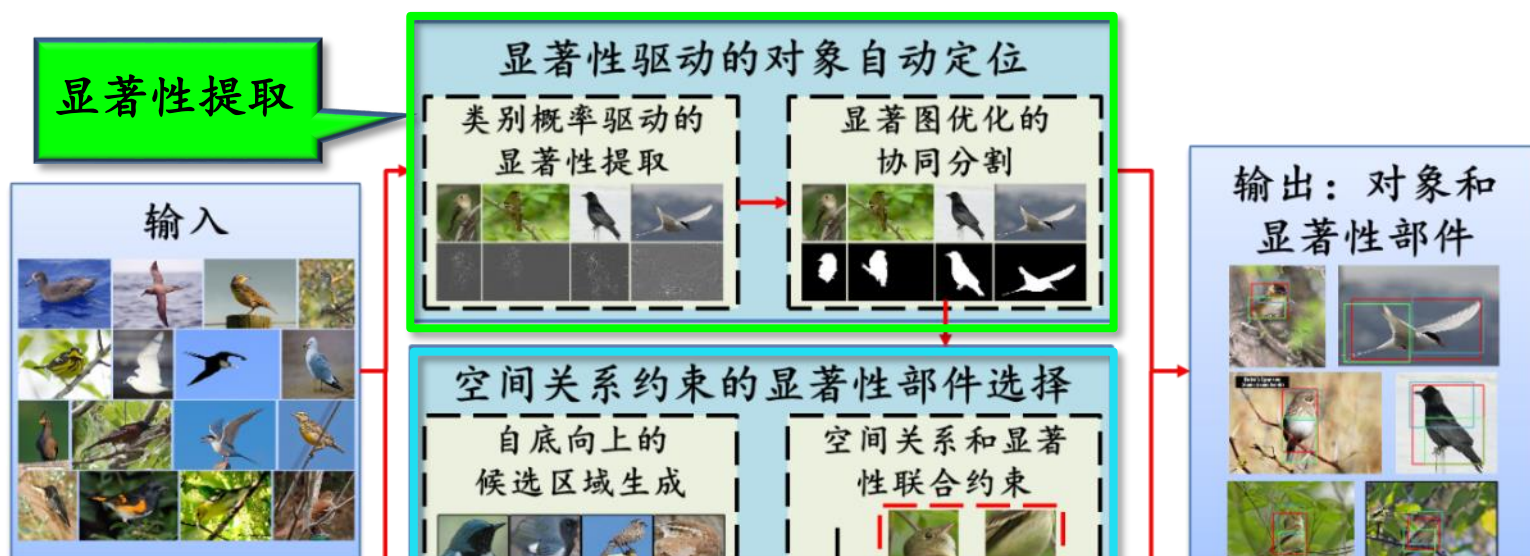


Yuxin Peng, Xiangteng He, and Junjie Zhao, “Object-Part Attention Model for Fine-grained Image Classification”, *IEEE TIP*, Vol. 27, No. 3, pp. 1487-1500, 2018

Tianjun Xiao, Yichong Xu, Kuiyuan Yang, Jiaxing Zhang, Yuxin Peng, and Zheng Zhang, “The Application of Two-level Attention Models in Deep Convolutional Neural Network for Fine-grained Image Classification”, *28th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 842-850, Boston, MA, USA, June 7-12, 2015.

研究现状

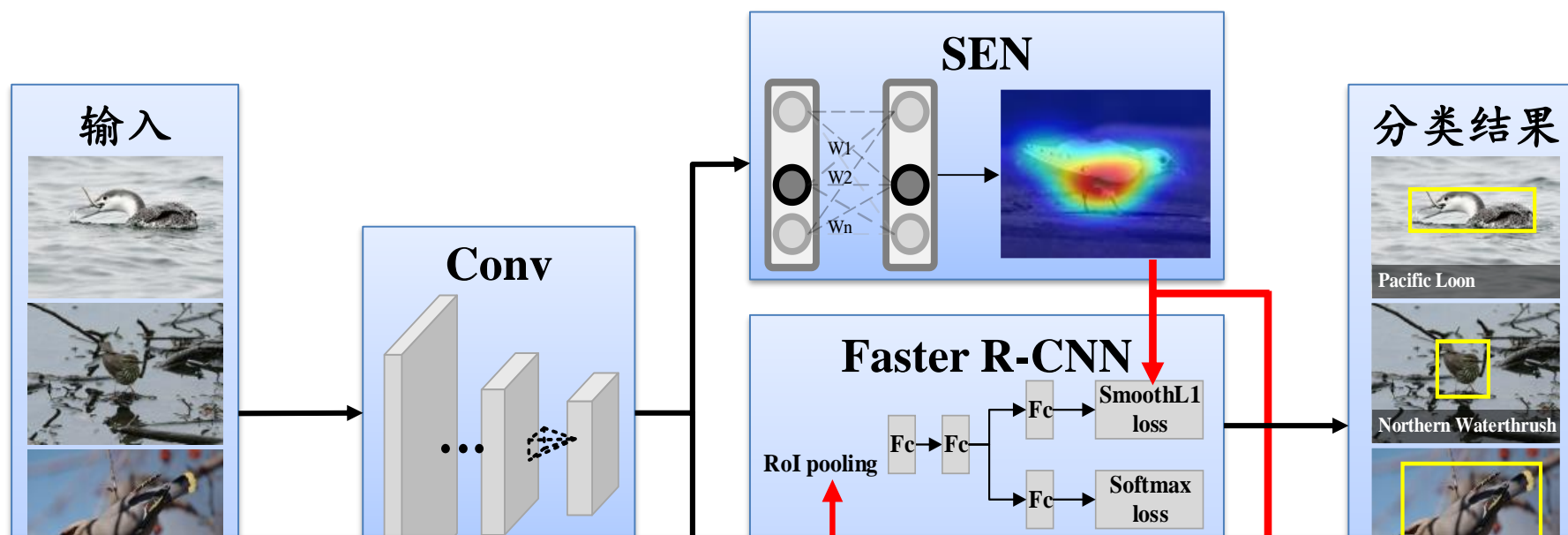
- 在显著性部件选择上，进一步提出了**空间约束的显著性部件选择模型**，联合建模空间关系和显著图，实现了显著性部件的有效选择，形成了**辨识力更强**的细粒度表示，提升了细粒度图像识别效果



Xiangteng He and Yuxin Peng, “Weakly Supervised Learning of Part Selection Model with Spatial Constraints for Fine-grained Image Classification”, *31th AAAI Conference on Artificial Intelligence (AAAI)*, pp. 4075-4081, San Francisco, California, USA, Feb. 4–9, 2017.

研究现状

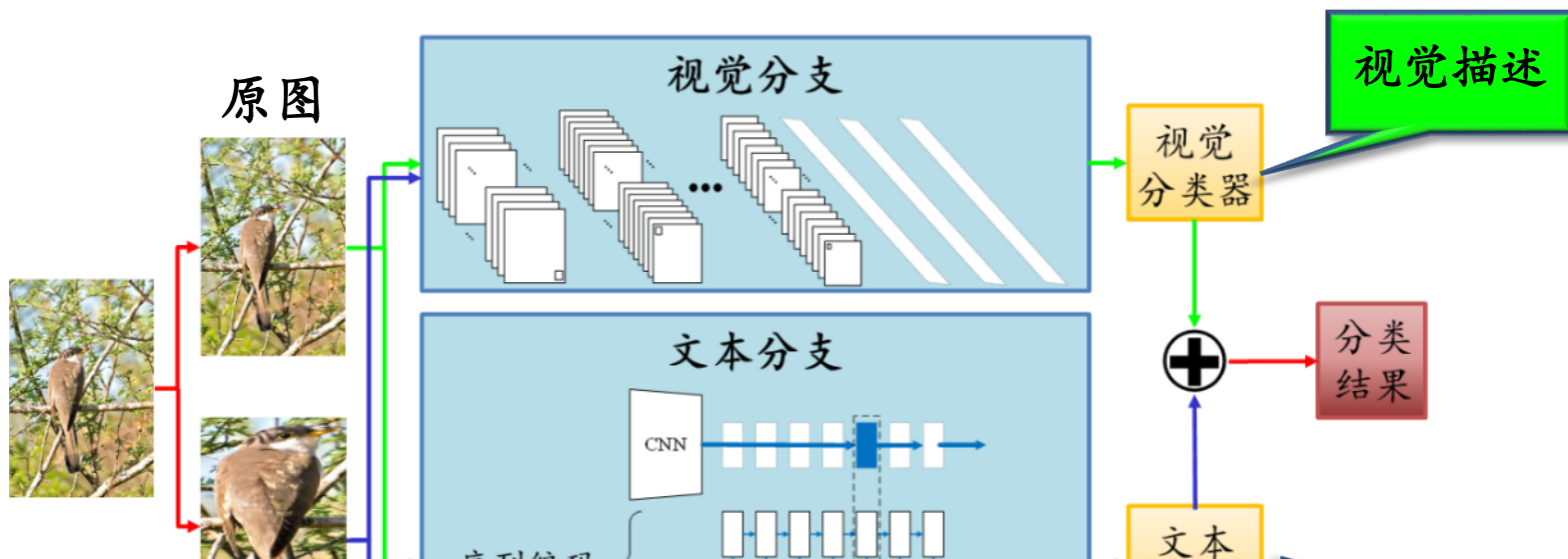
- 针对图像细粒度分类实时性问题，提出了**显著性引导的细粒度辨识性定位方法**，实现了细粒度分类与辨识性区域定位的协同促进，在提升分类准确率的同时实现**实时分类**



Xiangteng He, Yuxin Peng and Junjie Zhao, “Fine-grained Discriminative Localization via Saliency-guided Faster R-CNN”, *25th ACM Multimedia Conference (ACM MM)*, Mountain View, CA, USA, Oct. 23-27, 2017.

研究现状

- 引入文本描述信息，提出了**视觉-文本联合建模的图像细粒度表示方法**，通过卷积-循环网络共同建模视觉信息和文本信息，提供视觉对象的细粒度文本描述，实现了**多粒度、多角度、多层次**的图像描述



Xiangteng He and Yuxin Peng, “Fine-grained Image Classification via Combining Vision and Language”, *30th IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 5994-6002, Honolulu, Hawaii, USA, Jul. 21-26, 2017.