

鱼类研究综述

鱼类分析的难点分析及其主要挑战：鱼类分类的难点主要存在于两个方面

1. 首先各种鱼的大小不一致，有的鱼类很小，只占整个图片的很小部分，有的鱼类很大，占据了整个图片的绝大部分区域，另外鱼的外形多变，与人脸检测分类问题不同，人脸具有对称性，且每张照片中的人脸大小差异不大，这降低了分类方法去检测和分类的难度，所以相较于人脸检测分类，鱼类检测分类有更多的难点，首先我们需要模型去定位到鱼区域，保证兴趣区域是鱼区域。这个难点需要让我们的模型去自动区分辨别前景和背景。
2. Low-resolution Underwater Images
图片质量过低，大多数鱼类照片是在外界环境比较复杂的情况拍摄的，鱼的背景环境复杂，噪声较多，有多种因素的干扰。一些照片是在水下拍摄的，水下环境黑暗，鱼类照片不清晰，一些照片像素值很低，不能提取出有效信息，一些照片拍摄时鱼类正在移动，具有照片模糊。鱼类的相似度也对于分类产生了一定的影响。照片的扭曲，噪声，重叠，遮挡，物体之间的混淆性也造成了分类的困难。

以上两点造成了鱼类分类的照片分类的难点。

分类方法发展：

传统方法：传统方法一般基于图片本身的特征，在像素值特征上下功夫，利用不同的特征提取方式去提取不同的特征，HARR，HOG，小波滤波器都被应用于特征提取中。原始的分类方法从一开始的未经过任何特征处理方式的分类算法开始，KNN、K-means 等一些简单的聚类算法被应用于鱼分类，在这之上，有的人对其做出了一些改进，根据鱼的外形特征进行分类。从鱼的体长，以及各种部位占总体的一个比例来进行分类。

轮廓匹配和直方图统计都是解决分类问题的一种有效方案。获取鱼的一部分。但是这种方案对数据集要求比较高，数据集不能太大，且数据集内部的差异必须要足够大，若数据集内部的数据差异不大，则使用这种方案的效果会很差。

从同一张照片中提取不同类型的特征，如一张图可以分为颜色特征和骨架特征，纹理特征，然后就可以得到不同的特征和训练集，这样就可以尽可能得获得更多的特征。

一些传统方法侧重于局部特征的选取和学习，选取 local intersection，然后从中学习到鱼类的特征，一些论文中选取 ISV（Inter-Session Variability）去描述局部特征。在进行分类和拟合特征的时候，加

上了许多参数进行描述和拟合局部特征。在进行实验的时候，加上了 epoch learn 的思想，对整个数据集的照片进行编号，然后在学习到一定 epoch 的特征的时候再更新参数。

LDA（线性分类器）：线性分类方法

GMM 方法：高斯混合模型，高斯混合模型的使用，将图片的特征映射到更高维的空间中从而存储图片特征。然后再加上一个 SVM 作为分类器，最后用一个 BGOT 作为一个评价反馈函数，在这种方法中，它还使用了一个独特的通道去回传 error，将分类错误的信息回传到分类器中，作为 rejection function.

前景检测方法：利用一个参数矩阵去判别和提取图像特征，在参数矩阵中，调整参数的值，使图片经过右乘参数矩阵的输出能够正确的区分前景和背景。去掉背景因素的干扰。前景检测方法还可以应用到主成分分析中。在主成分分析方法之前，对数据集做一个前景检测，能提高主成分分析的效果。

传统方法一般基于鱼的外形特征，对照片的整个特征进行提取，然后进行根据颜色，体以及多层向量机进行分类型，在传统方法中分类一般使用的方法就是 SVM，改进一点的方法就是 C-R classification + rejection.在分类的时候将不属于原来种类的也添加到新的类别中去。不过这种方法比较依赖于参数的选择，参数选择得好，正确率才比较高，反之亦然。(2007 IEEE)。

在传统方法中，使用的分类方法以 SVM 为主，不管前面做的工作是什么，使用的特征提取方法有何不同，在模型的后面都是增加一个 SVM 网络进行分类。在改进的方法中，基于这种 SVM 分类方法，使用最多的就是 SIFT+BoW+SVM，首先去提取 sift 特征，sift 特征具有尺度不变性，即使改变旋转角度，图像亮度或拍摄视角，仍然能够得到好的检测效果，可以一定程度上解决鱼类大小变化过大的问题，BoW 是对特征进行编码，将提取到的特征进行编码皈依，然后将特征向量送到 SVM 进行分类

操作。这种方法在一定程度上依赖于特征的提取，分类效果主要跟特征提取效果有关。许多后续实验工作都是基于 SVM 展开的。

基于 SVM 的一些实验工作：

1.two multi-class SVMs。一个 SVM 是一一对一（1-vs-1），另外一个是一对多（1-vs-R）。两个 SVM 相结合

2.HOG+SVM，HOG 提取特征，SVM 进行分类。

3.MSVM，用多个 SVM 去分类提取到的不同的特征，最后几个 SVM 分类器的分类结果用不同的评价参数进行串联到一起作为最后的评价输出。

神经网络方法发展：

原始的神经网络方法从 MLP 开始，最先使用到鱼类分类的是 MLP 和置信神经网络(PNN)，接下来人工神经网络（ANN）登上舞台，这些网络在小数据样本集的表现并不好，甚至比不上传统的提取特征然后再进行分类的方法。这些神经网络表现不好的原因主要有两点，一是网络设计简单，参数少，不能有效的拟合数据的特征，二是训练数据量太小，模型容易对一些特征过拟合。导致了在测试的时候测试结果很差。这些方法在分类算法中取得的效果并不好。

接下来就是卷积神经网络的出现，对于图片分类问题，我们先选用一些卷积层去提取图像特征，再将卷积得到的特征进行下采样，并将采样得到的特征送到全连接层进行分类操作。与 ANN 和 MLP 方法相比，CNN 多了卷积层，利用这些卷积层进行提取特征。CNN 另一个优势就是更深的网络层，更多的参数，更多的参数保证了能够去拟合更多的特征，能够在弱标记的数据集中去提取特征。传统的分类方法都需要手工标记特征，这种强标记的数据集是传统方法取得良好结果的保证，如在上文提到的根据鱼的外形和体长进行分类的方法，作者在实验中对鱼的一些重要部位进行了标记，从而让模型能够记住鱼的体长信息和各个重要区域的特征。但是这种方法太过于依赖人的手工标记，在拥有大量数据需要处理的条件下不适合应用于分类问题。因此我们的卷积神经网络能够自动去提取特征这

一点显得十分重要。基于卷积神经网络，又有许多提升准确率的改进和技巧。

1.一些使用神经网络使用多个提取特征的网络层来进行提高正确率，比如在 (Accurate underwater live fish recognition with a deep architecture) 一文中，使用两个不同的卷积层来提取特征，一个用 $k1*k2$ 的 patch 去卷积，去减去这个 patch 的平均值，卷积的时候使用 PCA filter 进行卷积，有两个卷积阶段。接下来是一个阶梯函数，然后把得到的区块的特征连接成一个向量。还使用特征融合方法的论文是 (Fish Detection and Identification Using Neural Networks-Some Laboratory Results)，在这里，它一个在原始的数据中去训练提取特征，另外一个就是在处理之后的数据上提取特征，检测使用 a peak detector。检测一张图片最具有表现力的区域，因为鱼的数据集照片背景大多是海水，海水的背景颜色大致均匀，鱼的区域是相较于背景区域是有变化和显著特征的，这种方法可以一定程度上帮助提取兴趣区域。

2.BEOTR (balance-enforced optimized tree with reject option)能解决数据集不平衡的问题，在该论文中，作者主要使用对比的方法也是 SVM,PCA 方法，使用不同的评价指标 F1-score, AP, AR 来评价各种不同方法的优劣，之所以作者选用的方法行之有效，作者在论文中提到他们将分类器分类多层处理，他们在分类器上加上了一个基于高斯

混合模型的 rejection system。并将贝叶斯原理运用到模型中来，一些后验信息也作为分类器的输入特征进行分类。

3.DCNN (deep convolutional neural network)

加深网络深度，换用 GoogleNet, VGGNet 等更深的网络，这些网络相较于传统的 CNN 网络更深，有更多的参数，能过更好的提取卷积特征，进行调参优化各种网络，这也是一种提升分类准确率的方案。

4.region proposal

类似于 faster-rcnn 的原理，在第一阶段使用一个卷积层去提取特征，去进行分类的迭代，在第二阶段生成分类的候选区域，与 faster-rcnn 有所不同的是在第一阶段可以使用不同的方案去获取更加鲁棒性的特征，有些论文是人为标定特征图和轮廓图，人工加标签，有的是通过传统的特征检测方法来获取边缘信息，语义信息。有的是通过用不同的网络去提取不同的特征值，然后混合特征去训练网络，比如鸟分类，我们可以用一个网络专门去训练提取鸟腿和羽翼的特征。最后选取更具有特征代表的 region 进行分类会取得更好的分类结果，而且避免了学习到没有价值的区域的信息。

5.对网络层进行分层处理

将 cnn 进行分层，然后给予一个稀疏惩罚来促进训练，对于分类效果不好的节点，给予一定的惩罚，降低权值，强化分类正确的节点的连接。

对 CNN 进行分支训练。branching component deals with a subset of categories confusing to each other。

Conclusion:

通过讨论传统方法和神经网络方法，我们可以发现传统方法主要是基于 SVM，在 SVM 的基础上进行改进，SVM 适用于一些较小的数据集，在更小的数据集中相较于神经网络能取得更好的结果，所以可以将神经网络的最后一层换成 SVM 能够提升分类效果。传统方法侧重于从照片中提取特征，通过特征检测子来实现这一效果，所以在传统方法中，决定正确率高低的往往是特征选取的好坏以及数据集的难易程度。而神经网络方法不同，其就像是一个黑匣子，有大量的参数和网络节点去拟合提取出数据特征，其中怎么提取特征和拟合特征的过程我们不得而知，做的工作也是调参，更改网络模型。尝试更改网络模型中的一些参数。以达到更好的分类效果。

发展趋势大致是加深网络，增加网络参数，让模型有更强的泛化能力，另外需要做的工作就是将传统方法中一些有助于提升准确率的方法与神经网络相结合。将不同的模型融合，将局部信息和全局信息相融合。以获得更好的分类结果。另外需要解决的难点就是定位问题，从一张照片中定位到一个检测物体的分类，然后就是根据检测区域进行一个细粒度分类，根据局部区域的特征去进行分类，这要求我们的模型有极强的检测和分辨

能力。所以鱼类分类问题就是一个检测定位问题+细粒度识别问题。未来我们可以从这两个方面着手来提高模型和方法的准确度。