# 第2章 数据采集系统设计与数据标记

2.1. 数据采集基地介绍

为获取鱼类在工业生产环境下的鱼类摄食行为数据我们与国内领先的工厂化摄食养殖公司（烟山东东方海洋科技股份有限公司开发区分公司，以下检测实验基地）合作。实验基地位于烟台经济技术开发区，场区占地156亩，主要从事名贵鱼的育苗、品种改良和养成。该实验基地长期与中科院海洋研究所、中国海洋大学等知名院校联合建立了东方海洋鱼类研究中心（图2.1），致力于深海冷水名贵经济鱼类的健康苗种高效培育、工厂化养殖设施研制、技术体系构建等。本论文所有研究数据均从该实验基地日常生产循环养殖池获取。

图2.1 养殖基地鱼类研究中心

该实验基地2010年从挪威引进大西洋鲑鱼卵及全套大西洋鲑循环水育苗装备和管理技术，充分利用场区内罕见的恒低温12-15摄氏度的地下海水资源，与中科院海洋研究所共同研究开发大西洋鲑封闭循环清洁生产项目，现已建成养殖车间16个，养殖水体，实现国内乃至国际上大西洋鲑商品鱼的大规模工业化循环水养殖，养成的东方海洋三文鱼（大西洋鲑）在第十届、第十一届中国国际农产品交易会上获得金奖。

通过综合考察养殖基地各类养殖鱼的实际生产状况和研究需要我们选择了第16号车间作为实验采集车间。16号车间共有6个循环养殖池（RAS），养殖鱼类为大西洋鲑鱼，其中一个RAS现场如图2.2a所示。每个养殖池中大约有1500尾大西洋鲑鱼，17#车间总计养殖数量9000尾，平均养殖密度15尾/。RAS实际尺寸如图2.2b所示，养殖池深度为2m，底面积为约为100.5，容量为201.0，养殖水来源于近黄海地下海水，养殖温度12-15摄氏度。



（a）

**Tank**

**Water Treatment**

**8m**

**2m**

(b)

图2.2 16#养殖间1#养殖池(a)RAS现场实拍图（b）尺寸示意图

养殖所用饵料为膨化型颗粒配合饲料，饲料直径为8 mm。饵料投喂装置为轨道式自动投喂装置，每天投喂三次投喂时间分别为8:00/12:40/20:00，饵料投喂量由鱼类养殖专家根据鱼类的生长情况设定。投喂装置如图2.3所示，轨道位于养殖池正上方，投饵机通过电机驱动沿着轨道运动。当投饵机运动到RAS正上方时，饵料箱内部的饵料将根据设定值放出饵料，饵料经过饵料盘的旋转均匀撒在RAS中。



饵料盘

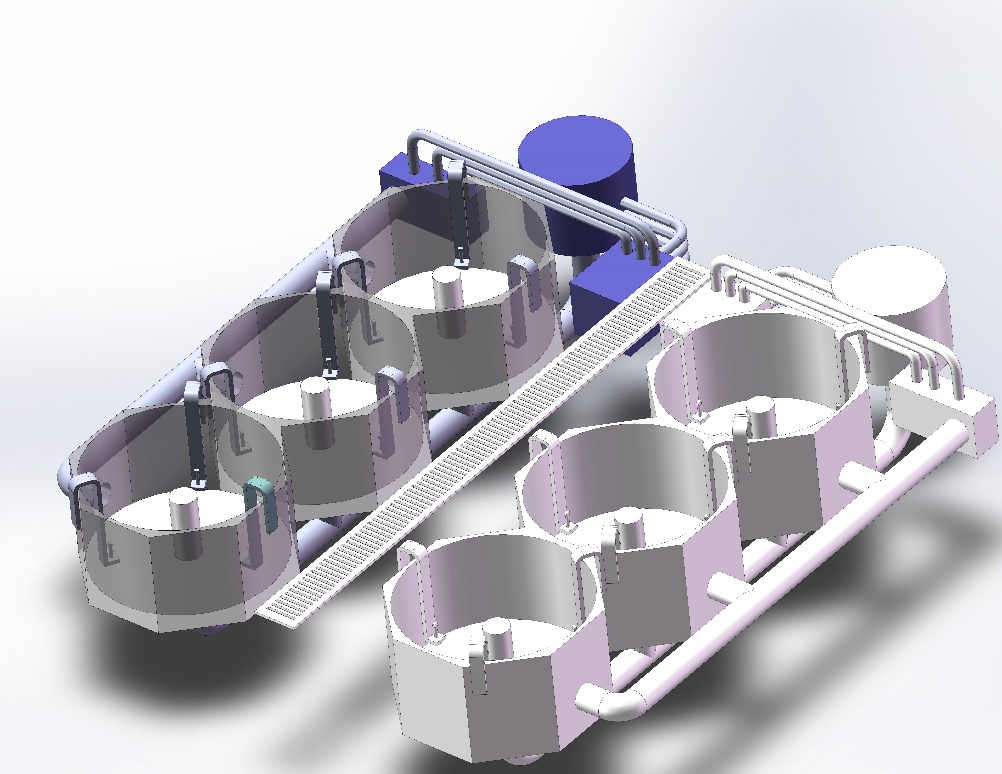
**投饵机运动方向**

**饵料箱**

图2.3 轨道投喂设备

2.2. 数据采集方案设计

现阶段针对鱼类行为的机器视觉研究所采用的方案大多数为将摄相机（普通摄像机和红外摄像机）安装在RAS正上方。但是在本论文的研究场景（基于真实生产环境）下这种方案有很多的局限性，除了在前一章所提到的由于饵料和鱼类共同造成的反光情况之外，更为突出的问题是在生产情况下会存在许多的干扰因素，比如投饵机的来回运动会使得视频内容干扰等。所以综合利弊，本论文选择了将摄像机布置在RAS底部，根据17#车间具体结构我们设计的三维仿真如图2.4所示。在每个RAS中我们布置了三个相机，获取不同视角的数据。



过道

水处理系统

摄像平台

图2.4 数据采集系统仿真设计

水下视频数据采集系统的具体设计如下。

第一步，摄像机选择。由于本论文所采取方案为水下方案，所以需要选择水下摄像机。最终所选择摄像机型号为GoPro HERO7 Black，其主要参数如下：传感器尺寸为0.393\*0.393英尺，光学防抖，感光元器件为CMOS，数码像素600万及以上。

第二步，摄像平台设计。为了将摄像机稳定的固定在水下，我们必须设计一个摄像平台的支架。支架设计采用Z字型设计，选用材料为2525型铝。安装相机后实物如图2.5。注意到在图2.5c中有两个摄像机，摄像机1可以通过数据线实时看到水中画面，但是由于其拍摄画质低感光度差，所以只是被用来测试和调整数据采集相机（摄像机2）的位置与角度。



（a）

（b）

（c）

摄像机2

摄像机1

图2.5 Z字型水下摄像平台，（a）实物侧视图（b）水下场景测试图（c）局部放大图，摄像机1为视频采集摄像机，摄像机2仅作为测试阶段角度位置调整之用。

第三步，相机视角调整。为了保证数据的完备性，我们在数据的采集过程中必须注意到在不同视角下的数据代表着不同的分布采样，所以数据集中应该包含在不同角度下的视频。根据养殖池的深度以及水下摄像平台离养殖池底部的高度，我们设计了三个不同的相机视角 则其对应的仰角为，其中为相机平视为相机中度仰视为相机最大仰角，分析设计见图2.6。

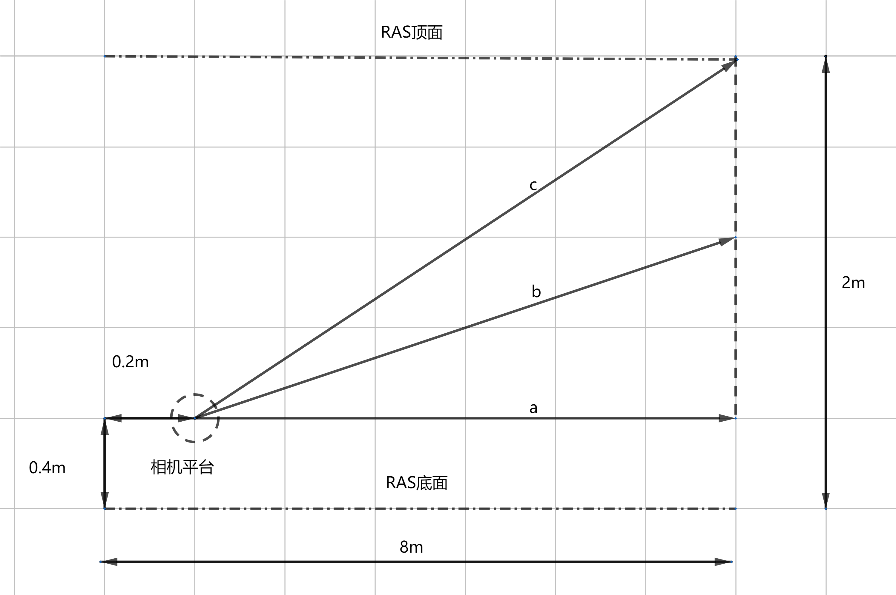


图2.6 相机视角设计计算图

2.3. 大西洋鲑鱼摄食行为水下视频数据库（UVDASSB）介绍

通过我实验平台我们采集到的未经处理的有效视频总长度为20269分钟，视频分辨率为2704x1520, 帧速率为30fps。UVDASSB将原始视频标记为两个类别，摄食状态和非摄食状态。每个样本长度为5秒即150张图片，样本分辨率为540x304，图2.7分别展示了摄食（eating）和非摄食样本（noneating），每张图中包含了从样本中均匀采样的15帧画面。为方便数据库的使用，采用了两种样本存储方式，第一种是将150帧图像保存为一个数组，然后将数组序列化为永久pickle文件[37]，每个样本对应一个pickle文件；第二种方式为直接储存图像，每个样本保存150帧图像。数据库标记完成的样本分为训练集,验证集,测试集三个子集存放，每个类别下的数量如表2.1所示，第一列为当视频样本的数量，第二列为图片样本的数量。

图2.7. UVDASSB 视频数据实例。a) 鱼类被标记为摄食的视频样本；b) 鱼类被标记为未摄食的视频样本

a)

b)



**表2.1**

UVDASSB 视频样本数量详细表

|  |  |  |  |
| --- | --- | --- | --- |
|  | | Pickle 文件 | 图像 |
| 训练集 | 摄食 | 399 | 59850 |
| 未摄食 | 1882 | 282300 |
| 验证集 | 摄食 | 130 | 19500 |
| 未摄食 | 625 | 93750 |
| 测试集 | 摄食 | 130 | 19500 |
| 未摄食 | 625 | 93750 |

2.3.1.视频样本标记准则

UVDASSB将大西洋鲑鱼的摄食行为标记为摄食和未摄食，视频标记准则将在本节介绍。

1. 异常数据排除。由于原始数据采集地点为实际生产环境下的大西洋鲑鱼养殖基地所以在数据的采集过程中会存在部分异常数据，引起异常数据的原因主要有以下几类：a). 由于工人路过RAS对鱼群造成的惊吓； b). 由于养殖基地不规律的临时增氧活动； c).由于布置实验设备对鱼类造成的干扰。以上几类数据均被标记为异常数据并排除在数据库之外。
2. 当鱼群处于投喂期间时，除鱼类对周围食物没有反应之外的其他行为均被标记为摄食状态[38][39]。
3. 在非投喂阶段的鱼类正常巡游状态，鱼类的运动行为被标记为未摄食。

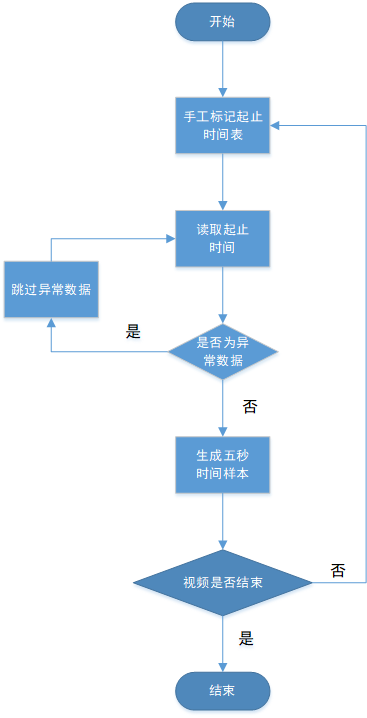
2.3.2.视频标注系统

为快速高效标记视频行为类别，本论文设计了如下的视频标注流程。如图2.8所示，根据大西洋鲑鱼养殖专家的指导和标记准则要求，对原始视频进行时间点标记，具体操作通过填写起止时间表完成（表2.2）

完成了所有的原始视频起止时间标注以后，通过设计python脚本文件自动从起止时间表中读取各个类的起止时间，如果出现异常数据类别则跳过异常数据，对其他类别的数据按照时间顺序剪切为5秒长的视频片段，直到视频结束。对所有原始视频都进行上述操作。

**表2.2**起止时间表 单位：秒（s） （**注：表格中#代表所标记视频不存在该类别**）

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| 视频  名称 | 数据  类别 | 开始1 | 结束1 | 开始2 | 结束2 |
| G071  5001 | 异常 | ## | ## | ## | ## |
| 摄食 | 25 | 55 | 100 | 130 |
| 未摄食 | 0 | 25 | 55 | 100 |

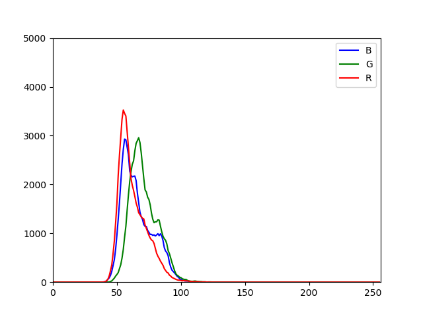
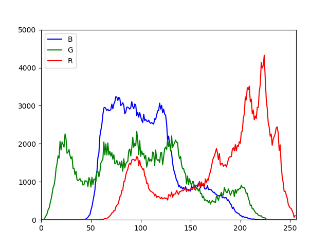
图2.3 视频标注系统流程图

2.3.3.UVDASSB 特点分析

为进一步描述与分析UVDASSB特点，本节将从直方图，图像熵，灰度均值和方差均值比四个方面进行评估。

1. 直方图

图像亮度直方图给出图像中亮度值z的出现频率，一副有L个灰阶的图像亮度直方图由具有L个元素的一维数组表示，当图像为彩色图像时，则在不同的色彩空间中计算其对应的亮度直方图也叫灰度直方图。UVDASSB中的视频样本被随机选取并截取其中的一帧与Lena图[40]进行对比。在图2.3中分别展示了他们的原图与灰度直方图，通过对比容易发现，UVDASSB视频数据灰度等级较少而且比较集中，相比之下Lena图灰度等级更加丰富几乎涵盖了从0到255的灰度等级。造成这种情况的原因主要有两个：a). 鱼群与水体背景灰度比较接近所以得到的图像整体一致，灰度等级比较集中; b).摄像机在水下得到的视频片段与正常拍摄图像相比，成像过程受到水体的影响，由于水体对光线的吸收导致灰度等级较低比但是通过线性变换可以在一定程度上提高亮度，改善灰度直方图。



a

b

c

d

图2.3 直方图对比图，a)Lena 图， b)Lena 灰度直方图， c)水下视频帧， d)水下视频帧灰度直方图

1. 图像熵

熵的概率起源于热力学和统计力学，直到很长一段时间后才被香农[41]通过信息熵的形式引入到信息理论中。

如果将图像灰度直方图看作随机变量X，一副图像的灰度等级可能为, 设是出现的概率，则图像熵的定义如下：

熵度量随机变量实现的不确定性，对于图像熵而言，熵越高则图像包含信息越丰富。但是由于同一直方图可能对应不同图像，比如当图像背景为常数时物体位置的改变不会影响直方图，所以相同的熵可能对应不同的图像。经过计算图2.3c的水下视频帧的图像熵为0.27，为了进一步的讨论数据库视频的特点，我们随机选取了一个样本，并计算样本的所有帧图像熵。如图2.4所示，对于摄食和非摄食的视频样本而言，他们的图像熵都低于0.05并且在一个样本中熵的波动范围相近约为0.005.另一个方面值得注意的是，单纯依靠图像熵是很难区分摄食和非摄食样本，出现这样的情况原因在于虽然在画面中是以鱼类为拍摄主体，但是更加在直方图的分析中可以看出鱼类表面纹理和水体的灰度等级十分接近，所以得到的图像熵也是必然相似的。

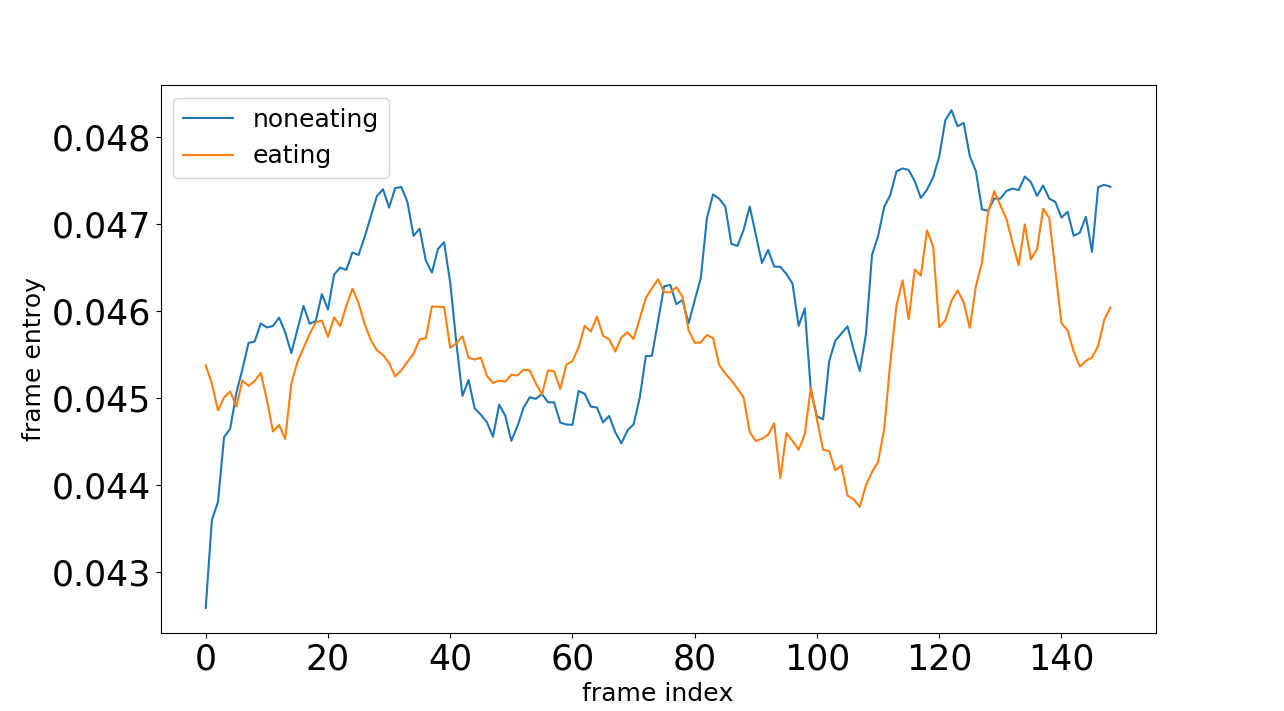


图2.4 视频帧图像熵分析对比图

1. 灰度均值

为评估本论文所创建的UVDASSB样本的灰度分布状态，下面将从两个方面介绍样本的灰度均值，第1是从单个样本角度计算视频帧的灰度平均值，单帧视频的灰度平均值定义如下

其中，m, n为视频帧的高度与宽度I为视频帧。

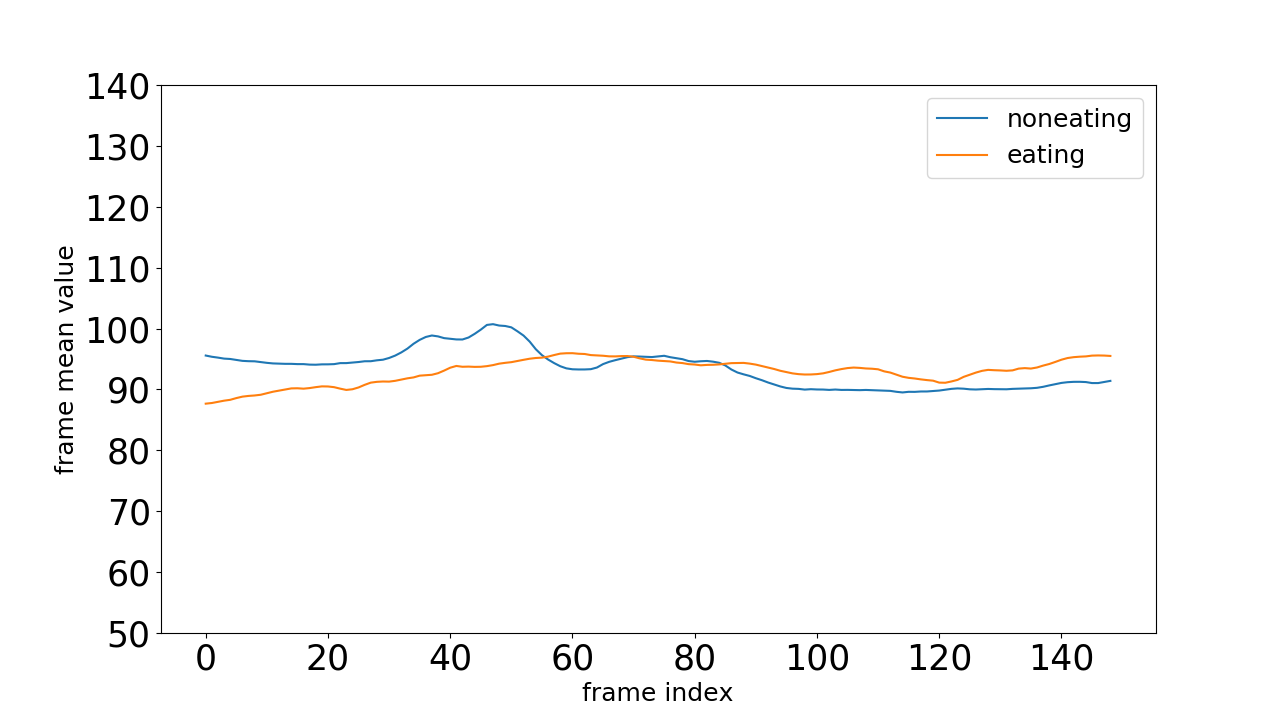


图2.5 单个样本的灰度平均值比较

第2从视频样本的宏观角度分析，UVDASSB的灰度平均值分布。视频样本的灰度平均值定义如下

式中t为视频帧序列时间，为第t帧的视频帧灰度平均值，为视频样本的长度。

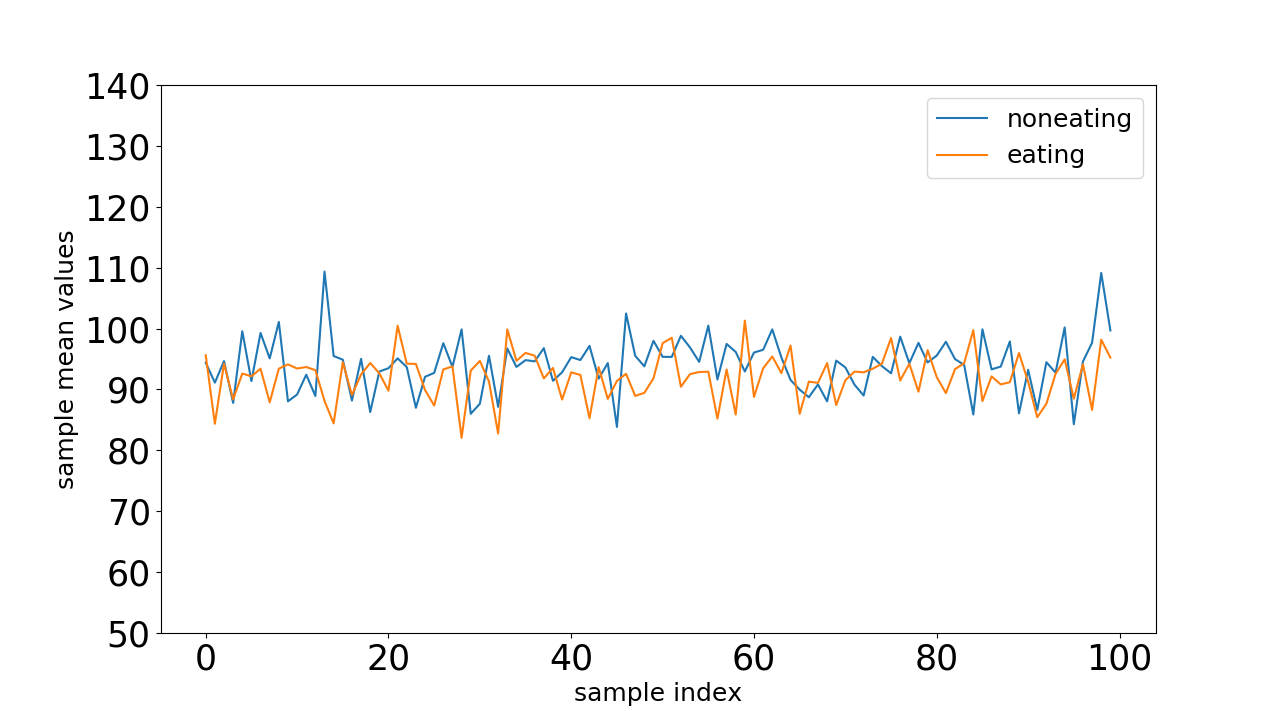


图2.5 摄食和未摄食样本灰度平均值比较

本论文中UVDASSB中分别随机选择了100个摄食样本和非摄食样本计算其样本灰度平均值，计算结果如图2.6所示。

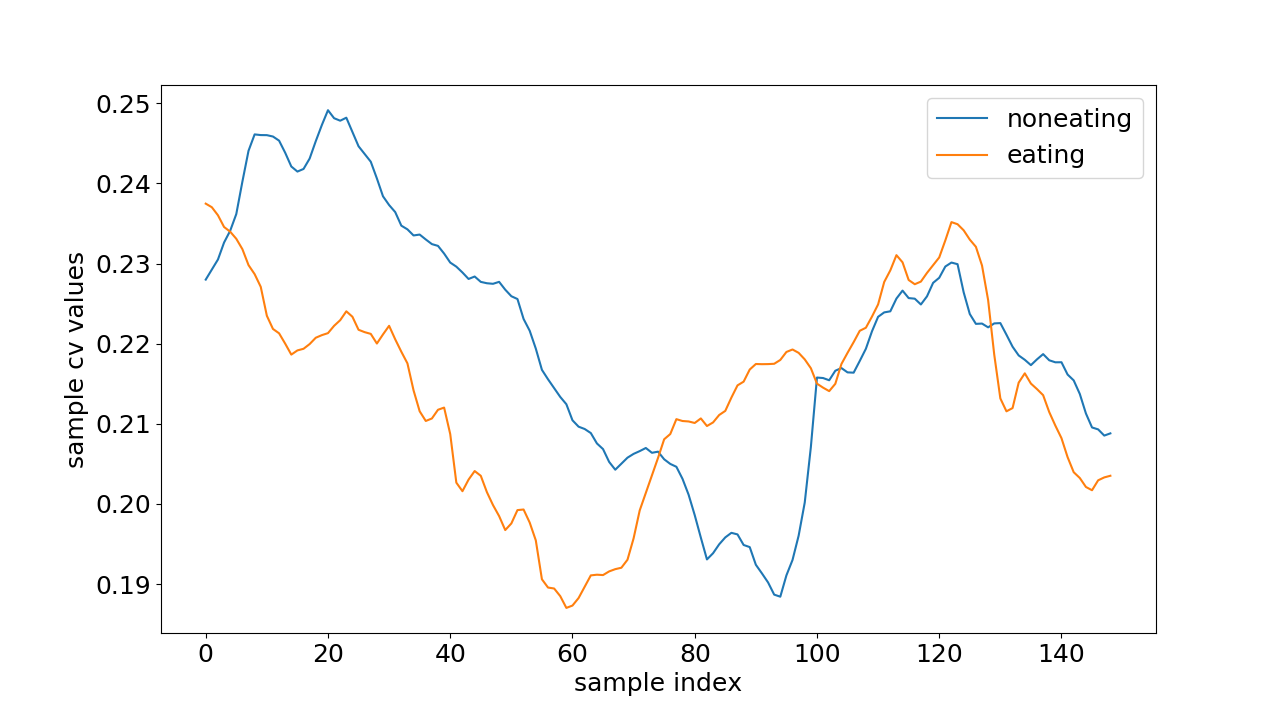
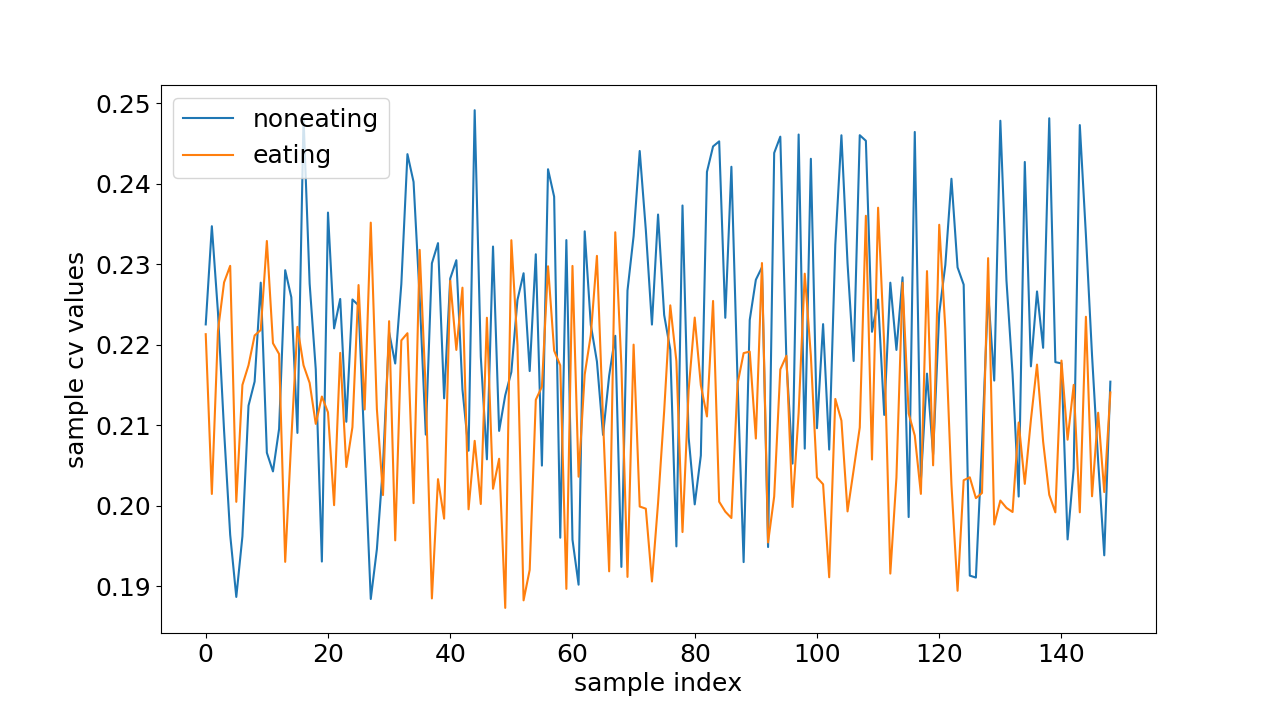
基于单个视频帧的灰度均值和样本的灰度均值特征我们可以总结出UVDASS样本的以下两个特点：a). 从微观看视频帧间的灰度平均值变化比较平而从宏观角度样本的平均灰度值变化更大，引起这种现象可以解释为视频帧间为连续的采样而样本之间为非连续，虽然所有的视频属于同一个类别但是样本间的相关关系与视频帧间的相关关系更弱所以平均灰度值变化更加剧烈；b). 但是从另一个角度可以发现无论是宏观或者微观灰度平均值的变化范围都是比较小的，均小于30，这意味着所有的物体在视频数据中的灰度等价是较近的而且画面中不会存在突然的画面切换等现象。

1. 方差均值比(信噪比)

图像的方差均值比又被称为图像的信噪比，源于统计学的变异系数（Coeffficient of Variation）。在统计学中当两组数据的测量尺寸相差太大或者数据量纲不一致时，变异系数被用于比较两组数据的离散程度的大小。在这里我们对单帧图像计算图像的变异系数，其计算公式如下

式中为视频帧的方差值，fm为视频的均值，为图像的变异系数。

计算视频帧的灰度均值比的意义主要为衡量视频帧之间的相互关系，当视频帧具有相对连续的画面时则灰度均值比为较光滑曲线反之则曲线变化快。

a) 

b)

图2.6 灰度均值比对比图, a) 连续视频帧灰度均值比; b) 随机扰乱视频帧灰度均值比

如图2.6所示，我们随机选取了摄食与未摄食样本计算样本的均值方差比，图2.6a为原始视频数据，2.6b为将视频帧顺序随机打乱后得到的灰度均值方差比。通过对比两图，我们可以得出无论是摄食或非摄食样本其视频样本的变化过程或者所离散程度都是在缓慢连续变化的视频的前后帧存在一定相关关系的结论。同时我们可以看到，虽然所有的视频都是在同一个场景下采集，他们的灰度均值以及直方图分布都比较接近但是其灰度均值比还是存在不同的相关关系。

2.3.4．基于卷积神经网络（CNN）对UVDASSB分类

为研究鱼类的摄食行为，已有许多的研究者基于仿真环境下的鱼类摄食行为进行了研究，其中周超等人[42]所提出的基于卷积神经网络的鱼类摄食行为分类方法在仿真环境下的鱼类摄食行为数据达到了90%的准确率，为评估UVDASSB的数据在一般模型下的表现，我们基于UVDASSB的数据按照文献[42]的方法设置CNN的结构以及超参数，通过实验说明UVDASSB在CNN模型下的表现。为清楚说明问题下面将首先将简要介绍CNN的结构与数据集设置，然后将介绍评估方法，最后展示实验结果。

1. 卷积神经网络结构与数据集设置

由于深度学习的飞速发展，卷积神经网络近几年在图像分类[43]，目标检测[44]和图像分割[45]等任务取得惊人的成果。卷积神经网络通过数据集自动学习特征提取方法，与传统的手工特征比如LBP、HOG[46]和SIFT[47]等方法，CNN 的特征提取不仅限于某一个层次的特征，基于数据CNN结构可以同时提取图像的浅层和深层特征。文献[45]的CNN结构如图2.6所示。鱼类摄食行为图像首先将会通过5x5的卷积网络，第一层卷积网络输出为五个通道特征图，之后经过极大值池化层减小图像尺寸，之后再重复前述卷积和池化炒作，将得到的特征图再送入到两个全连接层，两个全连接层的输出通道分别为120和84，最后一层的全连接网络特征将被输入到softmax层最后输出鱼类摄食行为类别。

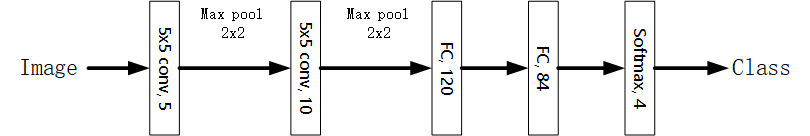


图2.7 卷积神经网络结构图

由于UVDASSB是视频数据集，所以训练卷积神经网络本轮文从视频样本中降采样创建了一个鱼类摄食行为图像数据库，其中训练集包含摄食和未摄食图像各2800张，验证集包含摄食和未摄食图像各600张，测试集包含摄食和未摄食图像各600张。将训练集数据用于训练CNN模型，验证集用于决定模型训练何时终止，测试用于评估模型在该数据集下的表现。

1. 评估方法

为科学合理的评估CNN 模型在真实的生产环境下的鱼类摄食行为分类问题的表现。在本文采用根据混淆矩阵计算的Accuracy， Precision， Recall， Specificity评估CNN在UVDASSB下的表现（在后续我们所提出的算法也将从这四个维度进行评价），混淆矩阵定义如下，

|  |  |  |  |
| --- | --- | --- | --- |
|  |  | Predicted class | |
|  |  | 摄食 | 未摄食 |
| Ground truth | 摄食 |  |  |
| 未摄食 |  |  |

Accuracy， Precision， Recall， Specificity计算公式如下所述。

1. 实验结果

根据如前所述的评估方法，通过训练集得到的CNN模型在测试集的表现如图2.8，其accuracy，recall和specificity达到了75%而precision达到了85%。与在文献[42]中的90%相比，CNN的模型准确率有所下降，其原因主要在于在真实的生产环境下鱼类的养殖密度更高，现场光环境更为特殊，所以最后模型表现不如在仿真环境下的实验结果。

图2.8 CNN baseline表现

2.4.本章小结

本章我们首先介绍了大西洋鲑鱼摄食行为水下视频数据集，数据采集系统的水产养殖基地以及数据采集实验系统的设计与实现；第二介绍了如何将原始数据制作为一个标记为摄食和未摄食的数据标记流程；第三基于灰度平均值、图像熵和灰度均值方差比等方面分析所建数据集的特点，总的来说UVDASSB的特点有灰度等级较少，图像亮度较低但是视频帧间具有连续的帧间关系，帧间特征变化是缓慢的；第四基于CNN的图像分类模型我们对UVDASSB的数据进行了数据集Baseline的实验，在CNN模型下的准确率为75%。