|  |  |
| --- | --- |
|  |  |

视频分析实验方案

|  |  |
| --- | --- |
| 姓名： |  |
|  |  |
|  |  |

2020年 6 月 20 日

1. 实验背景及目的

利用眼动仪可以采集到反映驾驶员心理活动的眼动数据，如瞳孔变化、眨眼频率、关注点位置等，同时还能同步获取视野内的视频。对视频数据进入场景分类可以建立心理变化与场景的关联，如上坡、下坡、左转、右转、直行等。对场景图片进行语义分割，可以得到视野图片中不同语义对象如道路、建筑物、行人、车辆、后视镜、仪表盘等对象的像素区域，依据这些区域再结合眼动仪捕捉到的视觉关注点，可以获知驾驶员实时关注的对象。

要对视频中的场景进行分类和语义对象分割，如果全部由人工进行标注则代价太高，不仅费时，而且人力成本极高。

场景分类的目的是根据检测到的帧可以定位特定场景在视频中的位置，从而得到特定场景下驾驶员在该场景下操作的视频片断，从而为后续分析驾驶员前后的心理变化，以及在特定场景下的心理状态。

图像语义分割的目的是可以分析在持续的时间片断内，驾驶员关注的对象变化、频率及其他的心理状态，为更深入的分析心理状态提供依据。

利用深度学习模型对视频帧进行分类和**语义分割**，可以帮助我们达成上述目的。这样只需要人工标注少量数据，通过对现有的图片分类和对象分割的预训练模型进行微调，则可以使模型适应本课题中的应用场景。这也正是深度学习模型在图像和视频分析领域中最常见的应用。

实验的目标就是利用人工标记的标本进行训练，然后利用训练的模型对视频帧自动分类和对象分割，人工再对自动分析的结果进行校对，将校对后的结果作为新的样本重新加入训练集进行再训练。整个过程不断迭代，最后得到准确相当于人工水平的自动分析工具。

意义：行车安全

目的：对从来没有去过的路的眼动视频能自动识别场景

训练出来的模型具有通用性。跟自动驾驶技术重叠的技术去查询资料。



1. 实验准备

眼动仪同步采集的视频数据。

图片标注工具labelme。

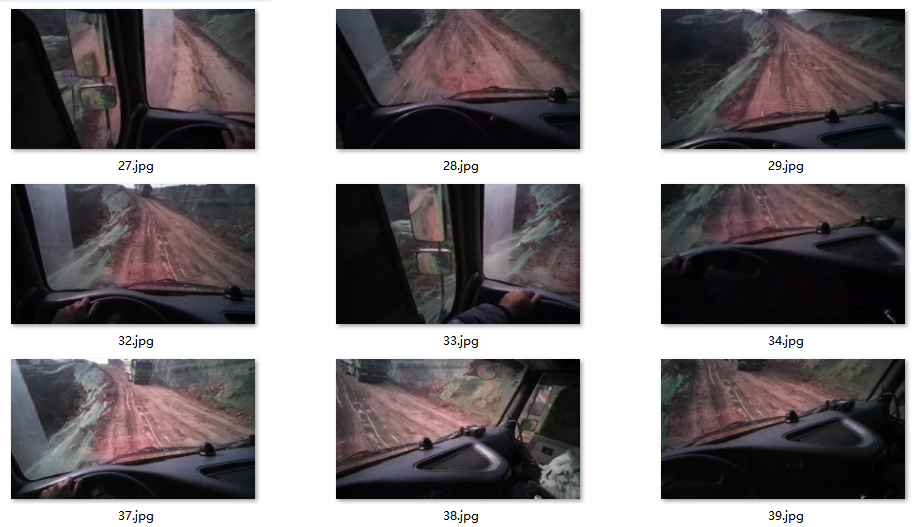
深度学习工具paddle，预训练模型VGG，fast-RCNN。

1. 实验步骤
   1. 任务1：场景分类

此任务属于图像分类，利用VGG模型可以实现。

第一步：帧采样

对视频按固定帧间隔进行采样，得到图片集合，图片名称按帧顺序进行编号。



第二步：人工标注图片类别

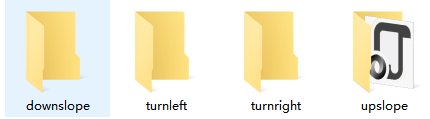
根据预先建立的目录如上下坡、左右转弯以及各种场景的组合等，将采样帧拖入到不同目录，有些帧可能会有多个标记，如上坡同时左转，对多标记的图片则复制到多个目录。

叠加分类可能更合适。

视频分析，利用透视关系来分析，几何空间，利用数学和物理关系的结合研究，辅助视频分析效果。

障碍物的问题，道路的宽窄、过桥、过隧道、对方来车情况、交叉路口、交通指挥设施（永久或者临时）：

车辆本身的问题，驾驶员的位置、座位的高度等是否影响效果。





如上图就可以复制到upslope目录。

每类场景提供500帧，其中400帧用于训练，100帧用于验证。

图片分类比较简单，直接利用系统的目录结构就可以完成，不需要借助其他辅助工具。

第二步：微调预训练模型。

第三步：模型训练后，利用它进行自动分类。

第四步：人工校验。并由人工对分类结果进行校验，检验只需要挑选出分类错误的图片。此步可以减轻提高人工标注的效率。这些错误分类的图片对提升模型泛代能力要强于相似度高的训练样本，可迭代训练中可以增加采样的概率或损失权重。

重复二、四步，直到自动分类可以达到或接近人工分类的准确率。

* 1. 任务二：图像语义分割

此任务采用fast-rcnn模型。

第一步：人工标注

利用辅助工具labelme，对图片进行多边形标记，标记出我们关注的语义对象，如道路、车辆等。从每类场景挑选图像差异大的图片作为样本，标注500帧，其中400帧作为训练样本，100帧作为测试样本。







识别左右后视镜

第二步，微调图像分割模型

第三步，用模型进行自动分割。

第四步，人工微调自动分割的结果。

重复二到四步，得到接近人工分割效果的模型。

* 1. 任务三：扫视分析

由于视频的视角是驾驶员，因此由于头部的和车辆颠簸会带同样的效果的图像晃动。这对分析人员的头部活动带来困扰。

同时如何根据视频获取头部的转动还是一个挑战。

此任务是通过分类光流变化来分析头部转动。***此任务目标和模型选择有等讨论。***

第一步：根据前两个任务得到的视频片断，人工标记视频中人员的头部转动。

第二步：对视频进行检测。

第三步：微调模型。

第四步：用模型视频进行检测，检测出头部转动的视频起始帧位置。

第五步：人工校验。并重复二到四步。

1. 实验目标

利用接近人工标注准确率的图像场景分类和语义对象分割模型。

1. 创新点

一是应用创新，将现有的图像领域的成果应用于新的应用场景。

二是为提升分类和分割效果，可以对图片进行预处理。这个预处理方法需要根据模型验证结果，并结合处理对象的特征进行研究。

三是提供施工车辆驾驶员心理状态和行为分析的数据集。