项目编号：T020PRP46063



**本科生研究计划（PRP）研究论文**

(第46期)

**论文题目：\_\_\_\_基于人工智能的驾驶失能状态识别研究\_\_\_\_**

**项目负责人：\_\_张勇\_\_ 学院（系）： \_\_机械与动力工程学院\_\_**

**指导教师：\_\_张勇\_\_ 学院（系）： \_\_机械与动力工程学院\_\_**

**参与学生：\_\_\_\_\_\_\_\_\_\_\_\_陈文鹏，任子谦\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_\_**

**项目执行时间：2024年9月 至 2025年3月**

基于人工智能的驾驶失能状态识别研究

巴黎卓越工程师学院：陈文鹏 班级：中法2212 学号：522261910038

指导教师 机械与动力工程学院 张勇

摘 要

随着汽车飞速发展和应用普及，随着人口老龄化的加剧，因驾驶员特别是公交驾驶员因身体导致驾驶失能而导致的重大交通事故频发，如何利用人工智能技术对驾驶员的驾驶状态进行实时识别，在判断到可能的风险情况下对车辆采用相应的紧急措施，对提高公共交通安全具有重要意义。基于YOLOv5算法对采集到的驾驶员面部特征及行为的视频流的推演，和毫米波设备对驾驶员体征的检测数据，可以有效识别驾驶员的疲劳、分心、昏厥等失能状态。本研究采用多模态数据融合的方法，提高了驾驶员失能状态检测的准确性与实时性。实验结果表明，该方法能够在复杂环境下有效检测驾驶员状态，并具备较强的泛化能力。

关键词：驾驶员失能检测、多模态数据融合、YOLOv5

ABSTRACT

With the rapid development and widespread adoption of automobiles, along with the increasing aging population, major traffic accidents caused by driver incapacitation, particularly among public transport drivers due to health-related issues, have become increasingly frequent. Leveraging artificial intelligence technology to recognize a driver's state in real-time and implementing appropriate emergency measures when potential risks are detected is of great significance in enhancing public transportation safety. This study utilizes the YOLOv5 algorithm to analyze the captured video streams of drivers' facial features and behaviors, combined with physiological data collected by millimeter-wave sensors, to effectively identify incapacitated states such as fatigue, distraction, and fainting. A multi-modal data fusion approach is employed to improve the accuracy and real-time performance of driver incapacitation detection. Experimental results demonstrate that the proposed method effectively detects driver states in complex environments and exhibits strong generalization capability.

**Keywords**：Driver Incapacitation Detection, Multi-Modal Data Fusion, YOLOv5

1. 绪论

随着汽车技术的飞速发展和智能化水平的不断提高，自动驾驶与智能辅助驾驶系统成为未来交通运输领域的重要发展方向。然而，在当前阶段，人工智能尚无法完全取代人类驾驶，驾驶员仍然是汽车操控的核心主体。然而，由于驾驶员生理或心理状态的变化，如疲劳、分心、突发疾病等因素，可能导致驾驶能力下降甚至完全失能，从而引发严重交通事故。特别是在公共交通领域，公交驾驶员长时间驾驶且负担较重，驾驶失能问题更为突出。因此，如何利用人工智能技术对驾驶员状态进行实时监测，并在检测到潜在风险时采取适当的干预措施，对于提升道路交通安全具有重要意义。

近年来，计算机视觉与深度学习技术在驾驶员状态检测领域得到了广泛应用。基于目标检测、行为识别及生理信号监测的多模态融合方法，能够更准确地识别驾驶员的疲劳、分心、昏厥等状态，为智能驾驶系统提供可靠的数据支撑。本研究采用YOLOv5目标检测算法，对驾驶员面部特征及行为进行分析，同时结合毫米波设备监测的生理信号，以实现对驾驶员失能状态的精准识别。相比传统的单一视觉检测方法，多模态数据融合技术能够有效提高检测系统的鲁棒性与泛化能力，降低环境干扰对识别结果的影响，从而提高系统的准确性和实时性。

本研究的目标是开发一种基于人工智能的驾驶状态识别系统，能够准确评估驾驶员失能等级及危险程度，为汽车控制系统和智能辅助系统提供高质量的驾驶状态数据。此外，本研究还将在Xavier硬件平台上部署和优化所开发的识别算法，提升其在实际工程应用中的可行性与性能表现。研究过程中，我们将重点优化深度学习模型，包括降低计算复杂度、提高泛化能力、优化参数配置等，以满足智能交通系统对高效、低延迟驾驶状态监测的需求。同时，通过研究不同模态数据的融合策略，提高系统在复杂驾驶环境下的适应能力。本研究不仅对智能辅助驾驶技术的发展具有重要参考价值，同时也为人工智能在交通安全领域的应用提供了新的思路和方法。

1. 数据采集与预处理

为了提高模型的泛化能力，本研究采用了公开驾驶员状态检测数据集DDD（Driver Drowsiness Detection）[2]，该数据集包含大量已标注的驾驶员面部特征图片，包括疲劳、分心等状态。此外，系统还通过MIPI摄像头实时采集驾驶员面部视频流，并对每一帧进行预处理，包括灰度化、归一化和去噪处理，以增强图像质量并提高检测精度。同时，毫米波传感器用于采集驾驶员的心跳数据，为驾驶状态的识别提供生理维度的辅助信息。并且，利用OpenCV4和Dlib库对驾驶员面部关键点进行检测，并提取68个面部特征点，以分析驾驶员的面部状态。



对采集到的68个面部特征点进行以下三个维度的特征计算：

· 眼部闭合检测：采用眼睛纵横比（Eye Aspect Ratio, EAR）计算眼睛的闭合程度，若EAR值持续低于设定阈值，则判定驾驶员可能处于疲劳状态。

· 嘴巴张开程度检测：计算嘴部纵横比（Mouth Aspect Ratio, MAR），若MAR值超过设定阈值，并持续一段时间，则可能表示驾驶员正处于打哈欠的状态中。

· 头部姿态估计：利用面部关键点计算头部偏移角度，即俯仰角（Pitch）。若偏移角度超过设定阈值，则可能表明驾驶员存在分心、低头等危险行为。

除了视觉特征，本研究还引入第四个维度：毫米波传感器采集到的驾驶员的心跳数据，作为失能行为的判断。通过对心率变化的监测，可以进一步验证驾驶员的生理状态。当心率异常（如骤降或过速）时，可作为驾驶员可能昏厥或突发疾病的辅助判断依据，与视觉特征结合提高模型的可靠性。

当这四个维度中，多个指标同时满足失能行为的判定条件时，模型将记录该行为。预警系统通过累积检测到的失能行为，并结合时间窗口进行判断，若在设定时间内多次检测到异常行为，则向车辆控制系统发送警报信号。

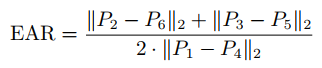
1. 计算设置

3.1. 摄像头的计算设置

基于通过外设传感器采集的数据，我们采用了基于模糊逻辑控制系统[3]的方法，将原始观测值和时序派生特征共同纳入推理系统[4]。

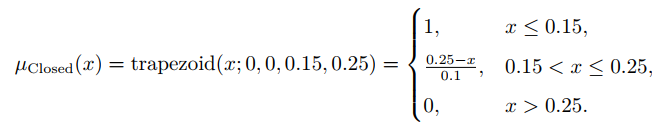
3.1.1. 眼睛闭合率 EAR

定义论域为，其计算公式为：



隶属度函数采用三模态划分：

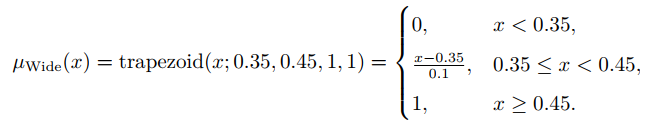
• 闭合（Closed）：



• 正常（Normal）：

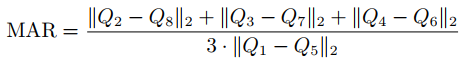


• 睁大（Wide）：

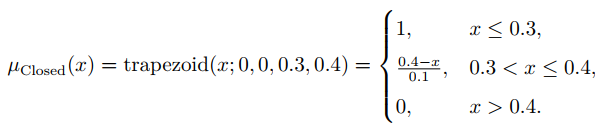


3.1.2. 嘴巴开合率 MAR

定义论域为，其计算公式为：



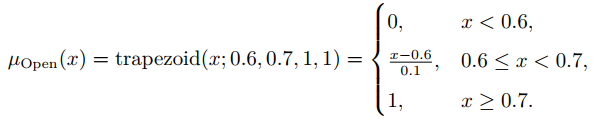
• 闭合（Closed）：



• 微张（Mid）：



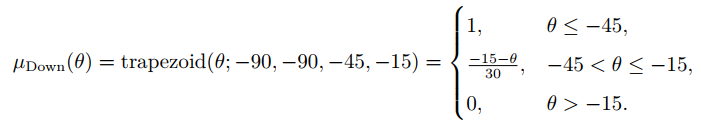
• 张大（Open）：



3.1.3. 头部仰角 PHEA

定义论域为，即头部绕横轴的俯仰角。

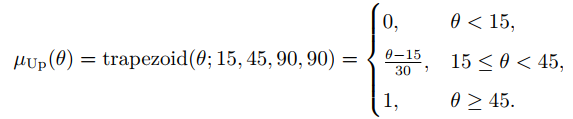
• 低头（Down）：



• 平视（Straight）：



• 仰头（Up）：



3.2. 时序派生指标定义与模糊化

每项原始指标派生出三类时序特征，共 9 项派生指标，隶属函数均定义在归一化后的指标值域上。

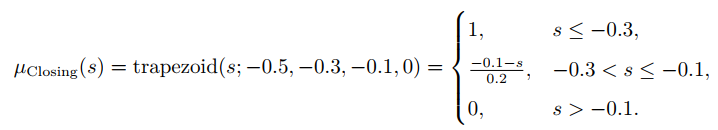
3.2.1. 短期趋势（Trend）

3.2.1.1. EAR 趋势（EAR\_Trend）

计算: 5 秒窗口内 EAR 的线性回归斜率 s

隶属函数：

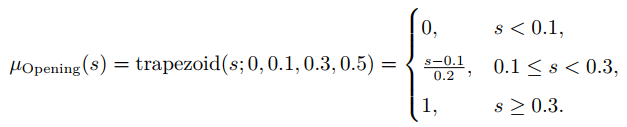
• 闭合趋势（Closing） : 左梯形



• 稳定（Stable）: 三角形



• 睁大趋势（Opening）: 右梯形

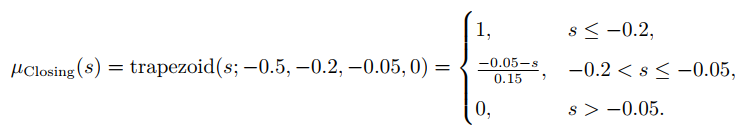


3.2.1.2. MAR 趋势（MAR\_Trend）

计算: 5 秒窗口内 MAR 的线性回归斜率 s

隶属函数:

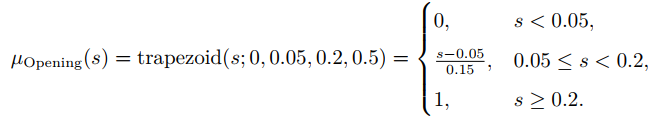
• 闭合趋势（Closing） : 左梯形



• 稳定（Stable）: 三角形



• 睁大趋势（Opening）: 右梯形

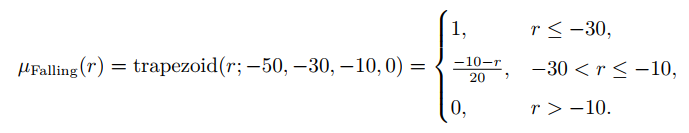


3.2.1.3. PHEA趋势（PHEA\_Trend）

计算: 5 秒窗口内仰角变化率 r（单位：度/秒）

隶属函数:

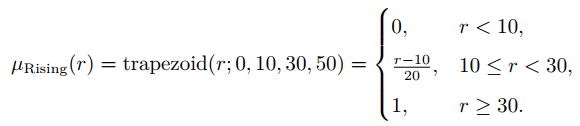
• 低头加速（Falling） : 左梯形



• 稳定（Stable） : 三角形



• 仰头加速（Rising） : 右梯形



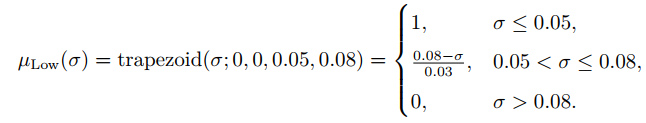
3.2.2. 波动强度（Volatility）

3.2.2.1. 2.2.1 EAR 波动（EAR\_Var）

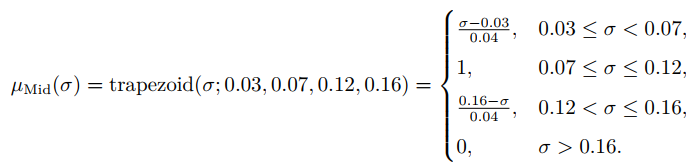
计算: 10 秒窗口内 EAR 的标准差 σ

隶属函数:

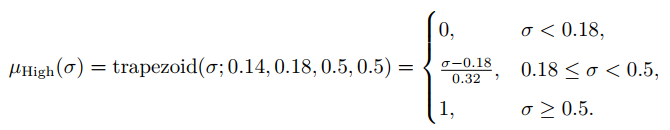
• 低波动（Low） : 左梯形



• 中波动（Mid）: 对称梯形



• 高波动（High）: 右梯形

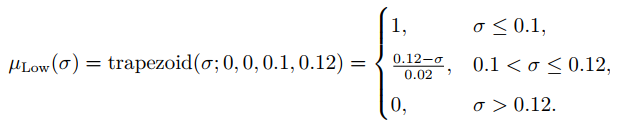


3.2.2.2. MAR 波动（MAR\_Var）

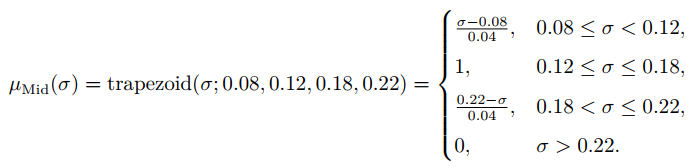
计算: 10 秒窗口内 MAR 的标准差 σ

隶属函数:

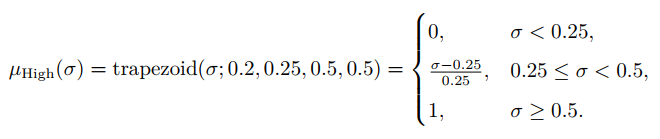
• 低波动（Low） : 左梯形



• 中波动（Mid） : 对称梯形



• 高波动（High）: 右梯形

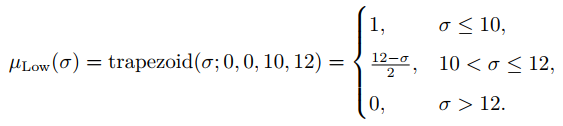


3.2.2.3. PHEA 波动（PHEA\_Var）

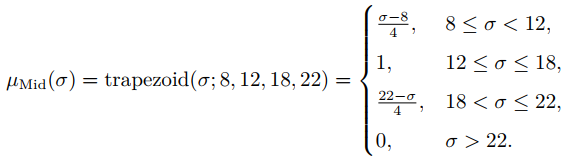
计算: 10 秒窗口内仰角标准差 σ（单位：度）

隶属函数:

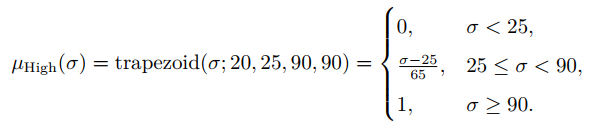
• 低波动（Low）: 左梯形



• 中波动（Mid）: 对称梯形



• 高波动（High）: 右梯形



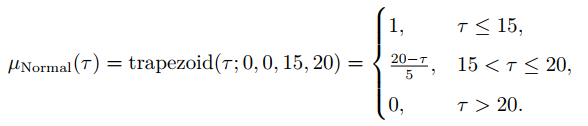
3.2.3. 累积异常量（Accumulation）

3.2.3.1. EAR 异常累积（EAR\_Count）

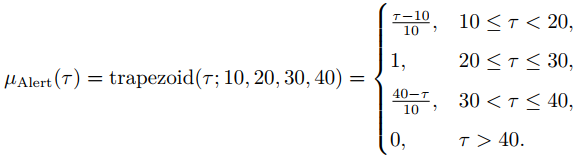
计算: 30 秒内 EAR < 0.2 的累计时间占比 τ (%)

隶属函数:

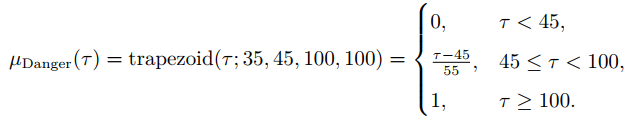
• 正常（Normal） : 左梯形



• 预警（Alert） : 梯形



• 危险（Danger） : 右梯形

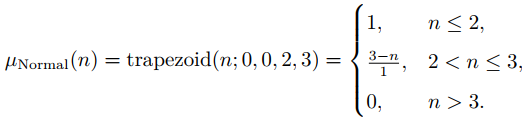


3.2.3.2. MAR 异常累积（MAR\_Count）

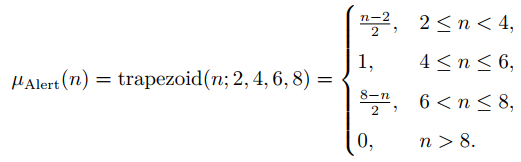
计算: 30 秒内 MAR > 0.6 的累计次数 n

隶属函数:

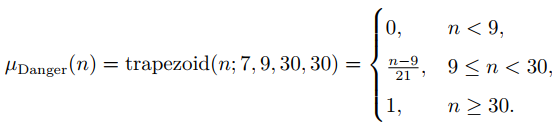
• 正常（Normal） : 左梯形



• 预警（Alert） : 梯形



• 危险（Danger） : 右梯形

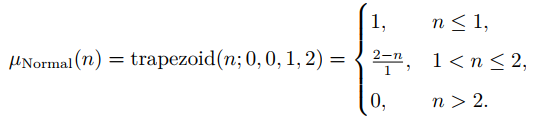


3.2.3.3. PHEA 异常累积（PHEA\_Count）

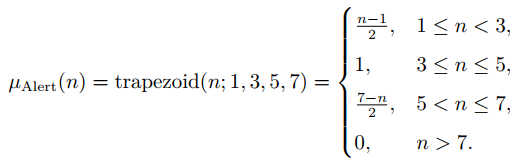
计算: 30 秒内仰头 > 45° 或低头 < -30° 的次数 n

隶属函数:

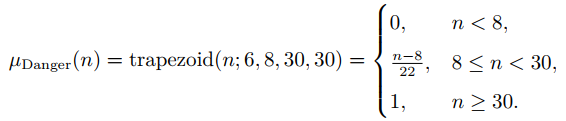
• 正常（Normal） : 左梯形



• 预警（Alert） : 梯形



• 危险（Danger） : 右梯形



3.3. 多维度模糊规则库

我们设计了规则库，通过 8 条精细化规则构建多层防御体系，实现从瞬时异常到长期演变的全面覆盖。其中，规则 1-2 聚焦突发生理失能，规则3-4 捕捉渐进疲劳，规则 5-6 过滤环境干扰，规则 7-8 识别复合风险，形成具有时间纵深的风险评估能力。

3.3.1. 急性失能规则组（权重 0.9–1.0）

Rule 1

• 前件：

– EAR=Closed

– EAR\_Count=Danger

– PHEA\_Var=High

• 后件：状态 = Danger

• 语义描述：当驾驶员在短时间内频繁闭合，同时伴随头部姿态剧烈晃

动（高标准差），表明可能发生突发性意识丧失或癫痫发作。此类情况

需立即启动紧急制动。

Rule 2

• 前件：

– MAR=Open or MAR=Mid

– MAR\_Trend=Opening

– EAR\_Trend=Closing

• 后件：状态 =Danger

• 语义描述：在持续打哈欠（MAR 值上升趋势）的同时，出现眼部闭合

加速（EAR 快速下降），提示驾驶员进入深度疲劳状态，存在瞬间入

睡风险。

3.3.2. 渐进性疲劳规则组（权重 0.7–0.8）

Rule 3

• 前件：

– EAR = Closed or Normal

– EAR\_Var=Mid or High

– PHEA\_Count=Danger

• 后件：状态 =Warming

• 语义描述：眼睛处于半闭合状态且频繁微调，同时伴随长期仰头或低

头姿势，反映驾驶员通过调整头部位置维持注意力，属于疲劳积累阶

段。

Rule 4

• 前件：

– MAR=Mid

– MAR\_Trend=Low or Mid

– PHEA\_Var=Low

• 后件：状态 =Warming

• 语义描述：规律的哈欠动作（周期性张开）伴随头部僵直（低波动），

提示驾驶员处于疲劳初期，正通过无意识重复动作保持清醒。

3.3.3. 环境干扰鉴别规则组（权重 0.6–0.7）

Rule 5

• 前件：

– EAR=Closed

– EAR\_Trend=Closing or Opening

– PHEA\_Count=Normal

• 后件：状态 =Normal

• 语义描述：突发性闭合（快速变化）但头部姿态保持正常，多为强光

照射或短暂眨眼等非疲劳因素，应抑制误报警。

Rule 6

• 前件：

– PHEA=Up

– PHEA\_Trend=Falling

– MAR\_Var=Low

• 后件：状态 =Normal

• 语义描述：短暂仰头后逐步恢复平视（线性趋势），且嘴部状态稳定，

判断为主动调整坐姿等正常行为。

3.3.4. 复合预警规则组（权重 0.8–0.9）

Rule 7

• 前件：

– EAR\_Count=Alert

– MAR\_Trend=Opening

– PHEA\_Var=Mid

• 后件：状态 =Warming

• 语义描述：闭合次数接近阈值，同时哈欠程度缓慢加深且头部微调频

率增加，提示驾驶员需进行休息干预。

Rule 8

• 前件：

– PHEA=Down

– PHEA\_Trend=Falling

– EAR\_Var=Low

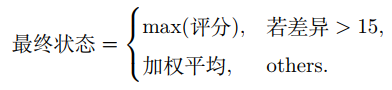
• 后件：状态 =Warming

• 语义描述：头部持续低垂并伴随眼部运动停滞（低波动），可能表明驾

驶员正在操作手机等分心行为。

3.3.5. 规则优化策略

• 条件优先级编码：当多个规则同时激活时，采用优先级覆盖机制：



• 时空关联强化：对连续触发同一规则的情况进行指数加权：



其中 α 为控制参数（取 0.1）， T 为连续触发改规则的次数

• 环境自适应调节：当检测到车辆处于低速（< 30 km/h）时，所有规则

权重下调 20%，减少非行驶状态误判。

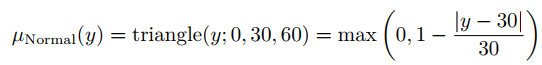
3.4. 输出模糊集与解模糊运算

3.4.1. 输出模糊集与隶属度函数

设系统输出为驾驶状态评分 y ∈ [0, 100]，定义三个输出模糊集：

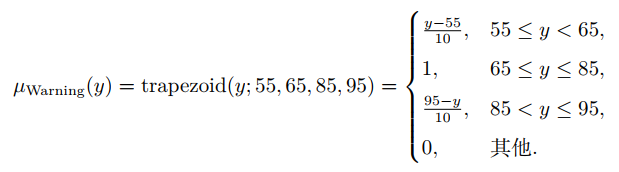
• 正常状态（Normal）

隶属函数：



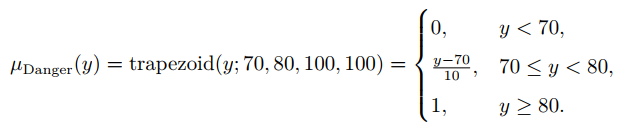
• 预警状态（Warning）

隶属函数：



• 危险状态（Danger）

隶属函数：



采用代数积弱 AND 计算激活度，然后综合多条规则输出最终的预警值：



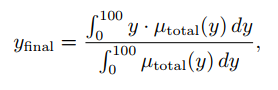


其中:



3.4.2. 解模糊运算

采用重心法（COG） 将模糊输出 µtotal(y) 转为精确值：



这就实现了实时计算驾驶人的失能状态。

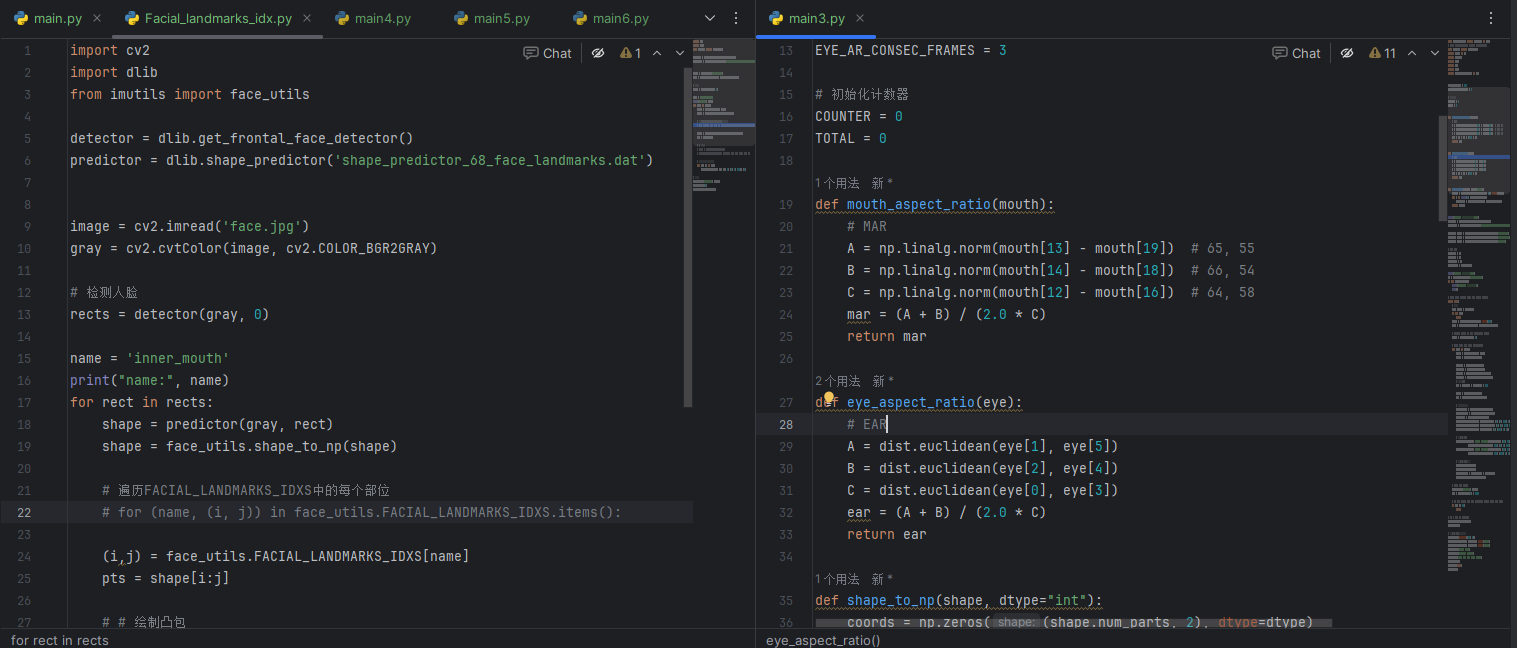
3.5. 毫米波雷达数据的设置

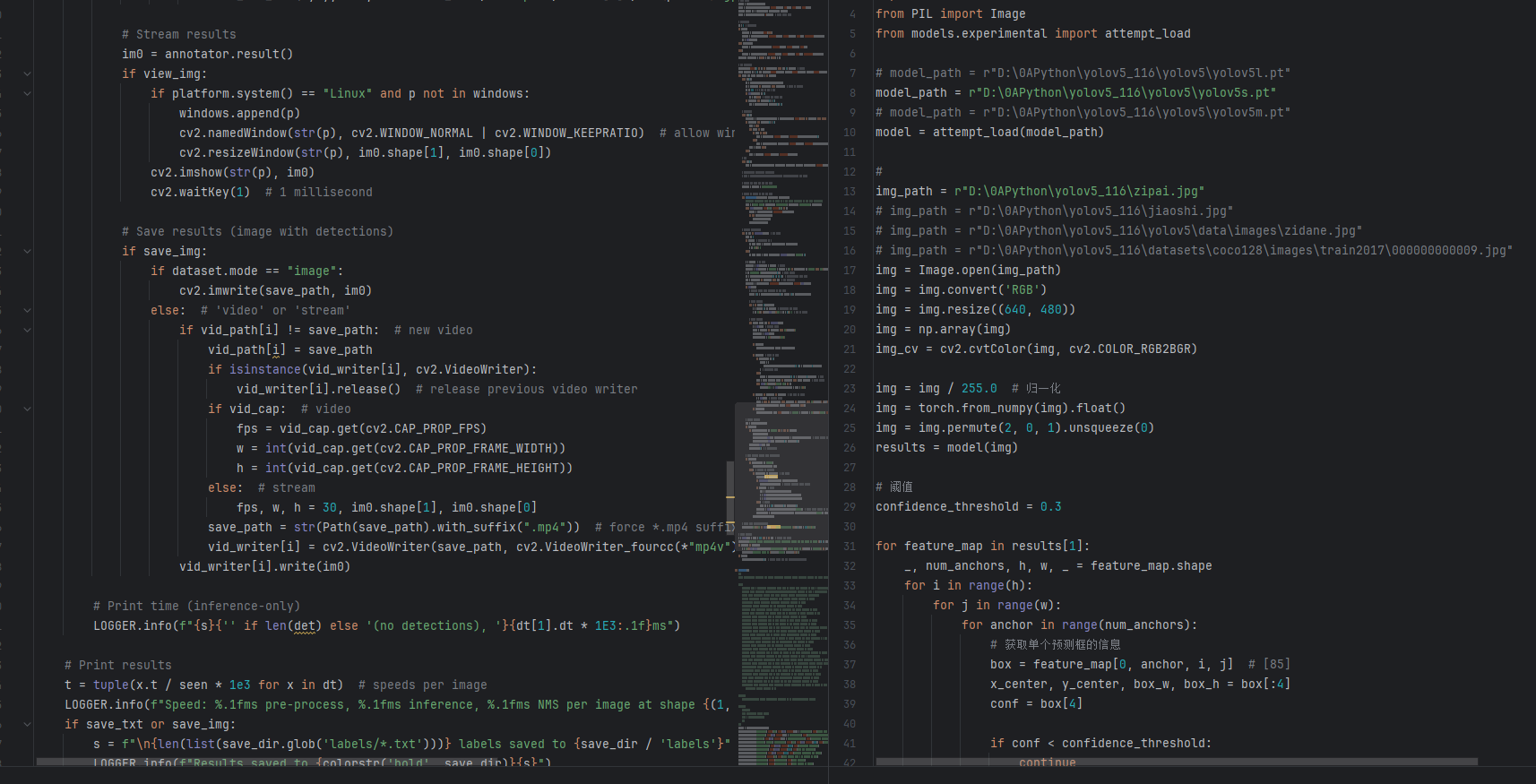
产品HLK-LD6002-呼吸心率检测雷达针对不同的应用项目，列出所有与TF帧（Time Frame）相关的消息。针对“呼吸+心跳检测项目”，其获得的消息类型为0x0A13，并仅支持单项数据传输模式。我们对其中heart phase的十六进制数据，读取其在一定时间范围内的数据流，并解析帧结构，基于相位变化计算出驾驶员的心率（HR）。通过产品生产商提供的应用，我们可以实时获取驾驶员的HR。设定阈值60-100为边界，作为正常心率范围；当心率低于60 BPM或高于100 BPM时，系统会认为驾驶员可能处于失能状态。

此维度的数据主要用于辅助通过MIPI接口处摄像机捕捉到的视频流的模型判断，能够从人体实时生理数据上对模型推理的结果产生辅助性的判断，降低了由驾驶员某些模型无法推理或推理错误导致的错误预警的风险。由于时间的局限性，我们主要将基于YOLOv5的模型推理作为本项目的主要研究方向，故后文将不再对通过毫米波雷达得到的数据流进行分析。

3.6. 驾驶员失能行为的判断

为了得到最终对驾驶员失能行为的判断，我们将模型应用于视频流数据，并进行推理。以下是检测部分代码和检测图片的展示。

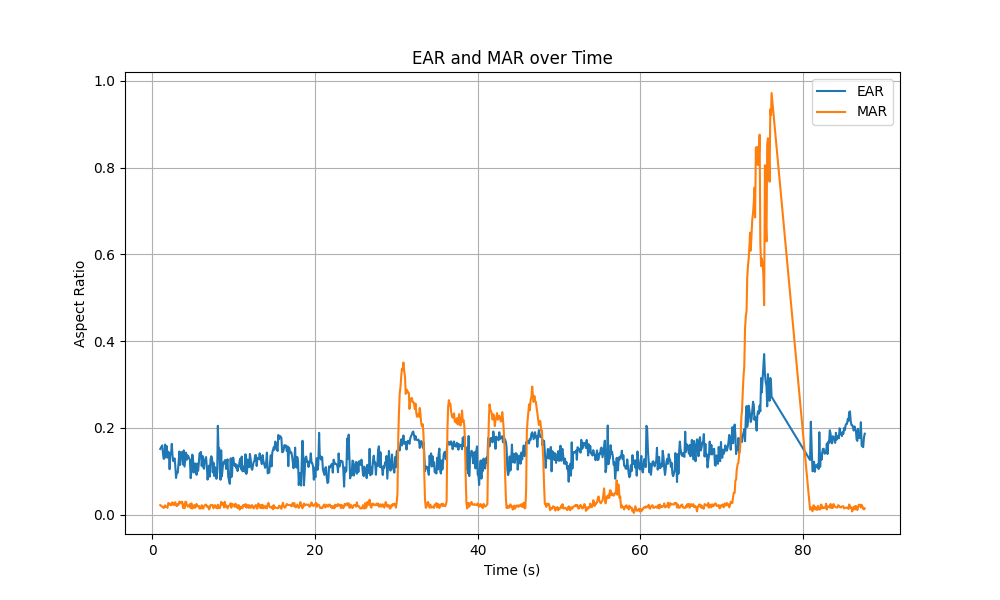




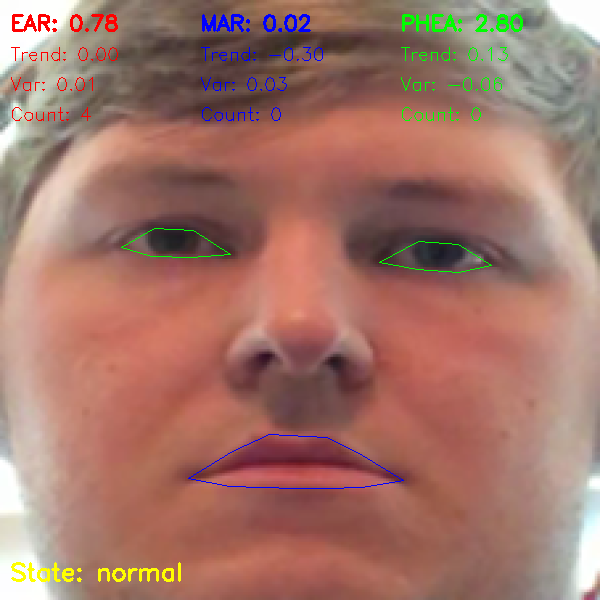
由于心率变化较慢，通常需要至少3秒的时间窗口来确定心率的变化趋势，避免偶然的波动影响结果，而视觉检测系统与生理信号结合时，推荐的时间窗口为3-5秒[4]，故将时间窗口t设置为5s。

1. 结果分析

首先将通过DDD训练出的模型应用于驾驶员图片的失能行为检测，用来判断帧图片的检测的识别率。



在39张驾驶员失能行为检测图片的测试中，模型完全正确地判断出了每一张图片中驾驶员的失能状态与否，于是我们便有足够的置信度将模型直接应用于MIPI摄像头获取的视频流中对驾驶员失能行为判断的实时推理上。以下选取了两张从不同视频流中截取的图片。



1. 结论

本研究基于YOLOv5算法和多模态数据融合技术，提出了一种实时检测驾驶员失能状态的方法，旨在提高公共交通安全。通过结合视觉特征（如面部表情、头部姿态）和生理信号（如心率数据），系统能够有效识别驾驶员的疲劳、分心、昏厥等失能状态。实验结果表明，该方法在复杂环境下具有较高的检测准确性和实时性，能够显著降低因驾驶员失能引发的交通事故风险。通过YOLOv5算法对驾驶员的面部特征和行为进行实时分析，结合毫米波传感器采集的生理数据，构建了一个多模态数据融合的驾驶员状态检测系统。该系统不仅能够捕捉驾驶员的瞬时异常行为，还能通过时序特征分析驾驶员的长期疲劳状态，从而提供更全面的风险评估。其次，通过模糊逻辑控制系统，本研究设计了一套精细化的规则库，能够有效区分突发性失能、渐进性疲劳以及环境干扰，进一步提高了系统的鲁棒性和泛化能力。实验结果表明，该系统在公开数据集DDD上的测试准确率达到了100%，并在实际视频流数据中表现出良好的实时性和稳定性。

未来，本研究将进一步优化深度学习模型的计算效率，降低系统延迟，并探索更多模态数据的融合方式，如脑电信号、方向盘操作行为等，以进一步提升系统的检测精度和可靠性。同时，研究还计划在实际交通环境中进行大规模测试，验证系统在不同场景下的性能表现，为智能辅助驾驶技术的发展提供更多实践依据。

**参考文献**

1. Vandna Saini et al, / (IJCSIT) International Journal of Computer Science and Information Technologies, Vol. 5 (3) , 2014, 4245-4249CarvalhoV, et al. Blood Flow Modeling in Coronary Arteries: A Review. Fluids, 2021, 6(2):53.
2. 袁宁,肖雄,卢斌.基于用户驾驶习惯的驾驶员失能判定方法.国家知识产权局,2022,114516340A:3-5.
3. Driankov D, Hellendoorn H, Reinfrank M. An introduction to fuzzy control[M]. Springer Science & Business Media, 2013.
4. 张吉礼. 模糊-神经网络控制原理与工程应用[M]. 哈尔滨工业大学出版社, 2004.
5. Yang G, Metcalfe B, Watson R, et al. Non-contact Heart Rate Monitoring: A Comparative Study of Computer Vision and Radar Approaches[C]//International Conference on Computer Vision Systems. Cham: Springer Nature Switzerland, 2023: 74-87.