

《计算机视觉》大作业 2022-2023

内容：共一题，涉及所学 CV 的理解、应用及拓展

考核：（大作业：汇报：平时）=80：10：10

作业形式：电子版 & 相关非保密代码

时间：最晚第十六周交给各班负责人

（如提前完成，发送至邮箱 kqian@jiangnan.edu.cn）

题目：

通过查阅文献调研等手段，针对公共数据集进行监督学习或无监督学习等计算机视觉任务的仿真实现及理解分析，完成小论文。

任务：

（1）一切计算机视觉的任务，可从课本、知网、三大会议网站以及各种期刊官网上寻找；

（2）也可以是某一个领域的方法综述

[1] <https://github.com/amusi/CVPR2022-Papers-with-Code>

[2] <https://zhuanlan.zhihu.com/p/31727402>

[3] <https://github.com/amusi/CVPR2023-Papers-with-Code#VT>

数据：

所能查阅到的一切图像数据集或者自己搜集到的数据集（RGB 图像、红外图像、高光谱图像数据等）。

一般确定好任务后，都能在对应的文献里找到相应的数据集，下载好、复现代码并分析结果。

[1] <https://github.com/xingchenzhang/VIFB> (可见光与红外图像融合)

[2] <https://www.cnblogs.com/isLinXu/p/15893506.html> (目标检测数据集)

[3] <https://anti-uav.github.io/Evaluate/> (反无人机挑战 CVPR)

要求：

（1）切忌直接简单使用公共源码生成的原始结果，以及大量复制粘贴文字!! 杜绝抄袭他人!

（2）注意格式排版以及内容逻辑性，鼓励以小论文形式呈现（Word 和 latex 模板见附件），实验报告形式也可。

（3）最终提交小论文（字数不限，若使用 word 编辑，请提交 pdf；若使用 latex，请提交文件包；），Word 或 pdf 命名格式为“学号_姓名_题目”，放至文件夹“学号_姓名”中，非保密代码请尽量标注清晰。

（4）可两人为小组合作完成，各自贡献或合作方式请在最后参考文献后说明。

若是小论文形式，主要内容和格式如下：

（1）标题

20 字以内，例：发现了并熟悉了 RGB 领域最新 CV 方法 A，将其应用在多源数据信息融合，发现通过对 A 进行 B 改进可提升效果，则取标题“基于 A 和 B 的多源信息融合”。

核心：老方法解决老问题；老方法解决新问题；新方法解决老问题；新方法解决新问题；

(2) 摘要（可不用英文摘要）&关键词

主要包括以下几部分：（按下述 abcd 写，逻辑性较强）

- a. 存在问题或研究目的；
- b. 方法描述；
- c. 结果说明；
- d. 结论表述

关键词 3-5 个：领域；方法；**；...

(3) 引言

建议内容：

- a. 所选领域背景综述；
- b. 其他学者已有研究成果的优缺点描述；
- c. 陈述为什么需要进行进一步的研究，即所描述方法的目的；
- d. 引出本文开展的研究工作。

引言切忌与摘要、结论重复；文字描述客观，不出现“首次”“我们”等主观性强的词。

(4) 方法描述和问题改进

- a. 基础方法的原理或应用描述；
- b. 存在的问题，可是算法本身缺陷或算法在其他领域的不足；

c. 针对存在问题的改进（加分项，若时间紧凑暂无创新点，则描述清楚问题即可，但仍需给出实验结果图表及分析等形式）

(5) 实验结果&分析

- a. 数据描述&参数设置；
- b. 定量及定向分析，可用图表加以描述说明

(6) 结论&参考文献

a. 描述清楚所采用的方法达到了何种效果，具有何种优势即可；切忌与摘要引言一样，以及重复描述所提方法步骤等！

- b. 参考文献适量即可（注意书写形式一致）

(7) 格式（一切以美观为核心，若是 latex 则按模板来就行，无需调整格式）

- a. 除标题黑体之外，其余均可用（宋体+Times New Roman），以美观简洁为目的；
- b. 小标题段前段后可有 0.5 行，图表（小五号）均需比正文（5 号）小一号字体；表标题在表前而图标题在图后；所给图表、公式均需给出序号，在之前给出引用话语（例：流程图如图 1 所示），并在之后给出进一步描述语句（例：图 1 中**表明***）。

(8) Latex（省去调整格式的时间）

a. TeXstudio2.12.22 和 overleaf 在线编辑均可用，所给模板可以上传到 overleaf 中打开，其余 latex 方式可试试；

- b. 图片用 visio 软件或者在线 draw.io 画，并保存成 eps、png 和 jpg 格式；
- c. 推荐学习一下线上 overleaf

例：（仅供参考）

（1）选用一种 CV 方法对所给多源信息数据（多光谱或 RGB）进行处理：参考回归、聚类、分类、降维等方法，以及**物体分割**（如基于颜色的）、检测跟踪识别、信息融合增强、超分辨等衍生手段

数据：

（1）高光谱序列以及对应的 RGB 序列，具体读写操作见示例代码。若需更多高光谱数据，可自行网站下载。

（2）多光谱相机集数据速度为 30 帧/秒，每帧图像包含 460-630nm 区间段的 16 个波段信息。RGB 和多光谱数据已作匹配，这也导致了图像清晰度的降低；

（3）每组数据中，包含多光谱源数据、RGB 图以及多光谱伪彩色图，均给出目标真实状态 $[x,y,w,h]$ ，如图 1 所示。



图 1 正样本状态

要求：

可选线性回归、 k 均值、Bayes 分类、PCA 以及 CNN 等算法对该数据进行相关处理

（2）红外与可见光图像融合技术研究

红外图像主要依靠物体自身的热辐射进行成像，其不受光照条件、天气的影响，但对比度较低，纹理细节不丰富。可见光图像通过反射可见光进行成像，纹理细节和对比度更适合人类的视觉感知，但可见光图像在烟雾、夜间等条件下的成像效果差。基于此，两者融合后能够获得一幅既有可见光图像边缘、细节信息又有红外热辐射目标信息的互补融合图像。军事领域，红外与可见光融合图像可实现恶劣环境下隐藏目标的识别与跟踪；智能交通领域，红外与可见光融合图像应用于行人检测、车辆识别与车距检测、道路障碍物分类；农业生产领域，红外与可见光融合图像可应用于水果的成熟度检测、植物的病态检测等。

近几十年来，大量红外与可见光图像的融合方法相继被提出，并在实际应用中得到推广。基于此，可对传统的融合算法及流行的深度学习算法进行探索，完成有效地多源图像信息融合。或者，对已有的方法进行总结比较，并撰写相关综述。

任务具体包括：熟悉一些传统的融合方法；熟悉一种深度学习图像融合的框架；进行仿真实验对比；撰写相关报告或小论文

[1] Zhang X, Ye P, Xiao G. VIFB: A visible and infrared image fusion benchmark[C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops. 2020: 104-105.

[2] Yu Liu, Xun Chen, Hu Peng, Zengfu Wang, Multi-focus image fusion with a deep convolutional neural network, Information Fusion, vol. 36, pp. 191-207, 2017.

[3] 综述网站: <https://github.com/xingchenzhang/VIFB>

[4] 周航. 基于深度学习的红外与可见光图像融合[D]. 沈阳理工大学, 2021.

[5] Zhang S, Li X, Zhang X, et al. Infrared and visible image fusion based on saliency detection and two-scale transform decomposition[J]. Infrared Physics & Technology, 2021, 114: 103626.

(3) Kaggle 网站寻找相关方法进行学习

[1] <https://www.kaggle.com/>

考察 CV 领域的基本算法的理解及拓展;

考察阅读广度、算法细节的把握以及小论文(或实验报告)的撰写能力, 请尽量保证内容的完整性以及格式的正确性。

钱琨

kqian@jiangnan.edu.cn

20230309