



Machine Learning

Lab3

Fall 2023

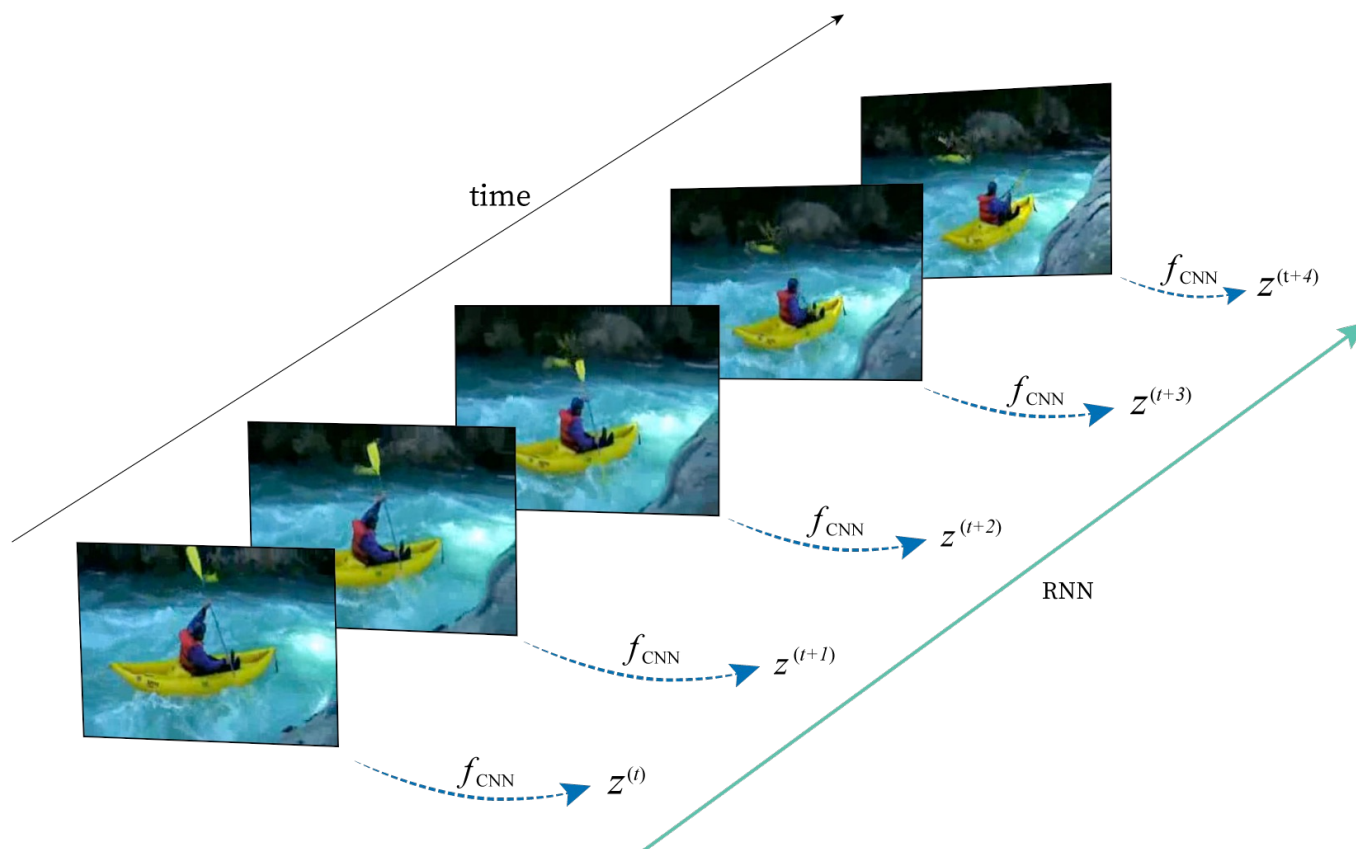
Instructor: Xiaodong Gu



基于CNN-LSTM的视频分类



基于CNN-LSTM的视频分类





基于CNN-LSTM的视频分类

- **任务:**

参考示范代码实现CNN-LSTM的视频分类系统并进行调参实验，其中CNN要求分别使用自行定义的普通CNN和调用预训练CNN进行尝试，最后选择样例视频进行分类结果展示。

- **数据:**

我们从 UCF-101 Dataset 中抽取了10类视频，见 canvas，请按照 8:2 进行随机 train-test split (已在提供的代码中实现)。

- **示范代码:**

<https://www.kaggle.com/code/nguyenmanhcuongg/pytorch-video-classification-with-conv2d-lstm>

<https://github.com/pranoyr/cnn-lstm>

有问题请联系助教石雨凌 yuling.shi@sjtu.edu.cn



基于CNN-LSTM的视频分类

- 提交：SID_NAME.zip

- 代码、环境及运行说明（数据集不用放入压缩包但请指明路径）
- 实验报告。包括但不限于系统设计、训练过程（如loss曲线）、调参实验及结果（不同模型结构/参数下的accuracy等指标）、样例展示等。**训练结果指标仅作为参考，不是主要的评价标准！**

- 系统设计

- 模型设计思路

- 训练方法

- 实验结果

- 训练过程（如loss曲线）

- 调参实验及结果（如不同模型结构/参数的精确度）

- 视频分类效果展示

- 截止日期：2023年12月24日



基于CNN-LSTM的视频分类

注意事项：

1. 要求使用 PyTorch, 建议使用 Kaggle/Colab 的 GPU 进行训练
2. 可以 resize 图像至更低像素来加快训练速度, 但可能损失一定精度, 需要权衡
3. DataLoader 中可以设置如 num_workers=4 提高计算效率
4. 在架构较新的GPU上 (18 年后发布的NIVDIA GPU), 可以使用混合精度训练提高速度
5. 为了减轻工作量, 可以参考示范代码等开源代码 (报告中注明来源), 但要有自己的发挥。



基于CNN-LSTM的视频分类

- 评分：综合评价功能、质量和工作量

功能：

代码无法运行



能完成功能、鼓励举一反三、尝试其他新结构



质量：

准确率极低、报告质量低



普通CNN准确率>65%, 预训练CNN准确率>75%、报告完整思路清晰



工作量：

直接提交示例代码或完全照搬开源代码



显示出对代码有理解、重构、或改进

