

# 表情视频分类（水一篇一区）

一张张图像，对于图像进行分类

以前的图像分类是对一张图像进行分类

现在的视频分类是对多张图像进行分类

多张图像之间有关系的（时间、空间）

第一步：我们依然先对图像进行分类

第二步：让我们的神经网络能够接纳多张图片作为输入

，只输出一个概率分布

? 图像的输入形式===【批次，通道数，高，宽】

视频的输入形式===【批次，时间维度，通道数，高，宽】

厚的图像输入===【批次\*时间维度，通道数，高，宽】

神经网络

输出概率分布

精度：序列模型：Mamba（2023, 2024 爆发，2025）

VideoMamba

损失函数：时间连续一致性：上一帧和下一帧的特征

差距应该足够小，空间一致性，语义分割 空间背景要一致，

速度：轻量化（用哪个都可以，但要说的动听）

不良条件下微表情数据集：晚上笑

---

开发拓展已有的数据集：用生成模型：controlnet 扩散模型

创新点：

1. 我们提出一张新颖的基于时空注意力机制的微表情分类模型  
一起个名字，
2. 我们拓展了已有的微表情分类数据集，弥补了在什么情况 gap
3. 在公开数据集上，我们产生了与 sota 方法有竞争力的结果，  
此外，在不良条件下，我们 outstanding。

快乐读博

深度学习比不过 sota，如何不造假提高自己的指标