

视频分析的关键技术

视频动作识别

传统方法：特征分析

双流卷积网络：分别建模空间（外观）和时间（运动）信息

视频生成与预测

生成对抗网络（GAN）：通过分离前景和背景动态，实现了逼真的视频生成

变分自编码器（VAE）：一种基于自编码器结构的深度生成模型，结合了概率建模和变分推断的思想，主要用于数据生成、特征学习以及潜在空间的连续采样。

多尺度视频预训练（MVP）：通过自监督学习捕捉多尺度时间依赖关系

多模态学习：旨在融合视频、音频和文本等多种模态信息。

自监督学习：在于利用数据本身的结构或特性自动生成监督信号，而非依赖外部人工标注的标签