MarioNette: Self-Supervised Sprite Learning

Dmitriy Smirnov\textsuperscript{1} Michäel Gharbi\textsuperscript{2} Matthew Fisher\textsuperscript{2} Vitor Guizilini\textsuperscript{3} Alexei A. Efros\textsuperscript{4} Justin Solomon\textsuperscript{1}

\textsuperscript{1}Massachusetts Institute of Technology \textsuperscript{2}Adobe Research \textsuperscript{3}Toyota Research Institute \textsuperscript{4}UC Berkeley

**Summary**

We propose a self-supervised deep learning approach that decomposes sprite-based video animations into a disentangled representation of recurring graphic elements. By jointly learning a dictionary of patches and a model that places them onto a canvas, we deconstruct sprite-based content into a consistent and explicit representation that can be easily used in downstream tasks, like editing or analysis.

**Pipeline Overview**

![Pipeline Diagram]

**Comparisons**

![Comparison Images]

**Results: Sprite-based Game Deconstruction**

![Deconstruction Results]

**Results: Unsupervised Natural Video Segmentation**

![Segmentation Results]
