Unsupervised Learning of Object Landmarks through Conditional Image Generation

1. OVERVIEW

"Unsupervised discovery of semantically stable landmarks for visual objects"

CONTRIBUTIONS

- Object landmark discovery without manual annotations.
  Outperform state-of-the-art facial landmark detection methods using a simple method.
- Learn from synthetically warped images/videos directly.
  Applicable to a variety of datasets without modification — faces, humans, 3D objects, digits.
- Method factorizes object appearance and geometry: transfer style / pose.

2. METHOD

DISTILLING GEOMETRY

"SUBTRACT" pairs of images which share appearance, but differ in object pose / geometry.

Videos
Frames from a video of an object.

Synthetically Warped Images
Thin-plate spline warped versions of a single image.

3. RESULTS

Human Faces

Human Pose

3D Objects

4. DISENTANGLING STYLE & GEOMETRY

MAFL facial landmark detection

Unsupervised landmarks

Regressed landmarks

Sample efficiency for supervised regression

Replace keypoint bottleneck with FC-layer

Freeze parameters

Different style "source"

Geometric "target"