Audio-Visual Floorplan Reconstruction

Overview
Goal: given a short video, infer a full floorplan

AV-Map Model
Input: RGB+Audio Sequence
Output: 2D Binary Interior Map, N-d Room Map

Generating Realistic Audio-Visual Data

Interior Map Predictions

Room Map Predictions

References: