CS280 - Final Project Milestone

Xiangyi Zhang
Shanghaitech University
zhangxy9@shanghaitech.edu.cn

1. Introduction

We propose to work on Amodal segmentation, which plays an indispensable role in scene understanding, depth estimationetc. To be more specific, the problem requires a model which can predict the missing information and tackle the large-scale occlusion pattern. We plan to propose a novel semi-supervised model approach to infer multiple amodal masks from a single RGB image, which consists of three parts: A reference set constructed from existing amodal data, visual concept based network for similarity and a generative model to infer multiple solution.

2. Problem Statement

Problem statement consists of three parts. Firstly, it's hard to predict scope of an object due to missing information and lack of prior knowledge. Secondly, there exists large-scale variance on occlusion patterns. Thirdly, labeling such variant dataset is challenge and expensive.

3. Technical Approach

Based on these challenges, first for occlusion, we adopt part-based representation instead of enumerating all occlusion patterns. Second, we construct a refenrence set without any amodal annotations to relieve the labeling burden. Third, we utilize the visual concept based network based deep voting network to and generative model to infer multiple solution in one-to-many mapping.

we want to propose a novel semi-supervised model approach to infer multiple amodal masks from a single RGB image, which consists of three parts: A reference set constructed from existing amodal data, visual concept based network for similarity learning and a generative model to infer multiple solution.

4. Intermedia Result

Considered result on votenet for voting the center of proposal bounding box and groung truth bounding box. Predict scale which means that we use the ratio from scalenet to resize the bounding box and vote the center. And oracle scale use the ground truth scale ratio.

Version	Predict Scale	Oracle Scale
mindelta	0.044	0.058
rgood(0.1) at best epoch	0.821	0.772
loss L2 at best epoch	0.2102	0.3430
bestepoch	3	12

Version	Predict Scale	Oracle Scale
pos_acc	0.756	0.725
neg_acc	0.913	0.836
acc_all	0.901	0.805
bestepoch	12	95

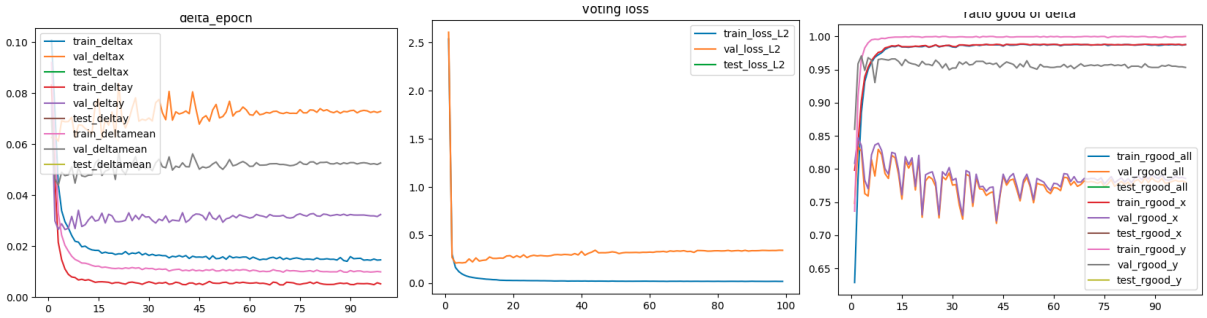

Figure 1: Votenet Branch Predict result

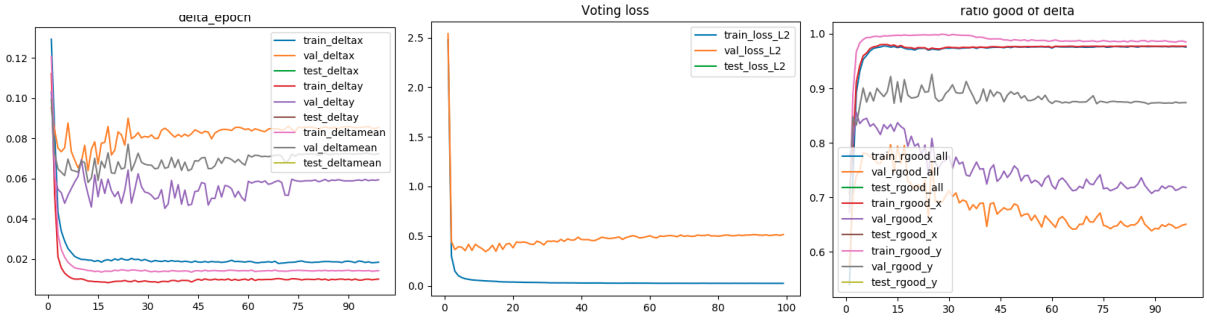

Figure 2: Votenet Branch Oracle result

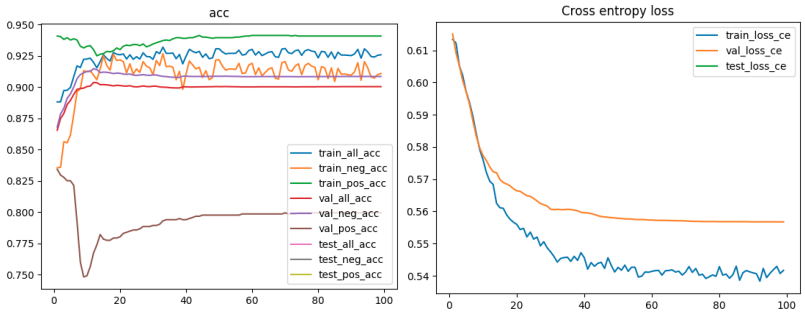

Figure 3: Scorenet Branch Predict result

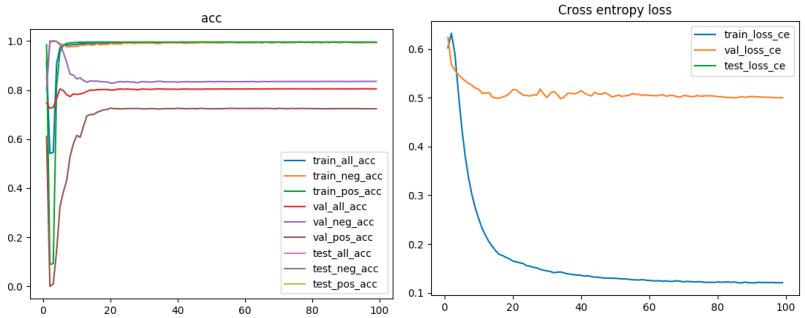

Figure 4: Scorenet Branch Oracle result