Performance Evaluation of Simple BaselineSimple Baseline

Abstract. While comparing the Stacked hourglass model and the SimpleBaseline model, the performance of the SimpleBaseline model was too poor, so additional experiments were conducted.

1. Introduction (+ Background, Related works)

Previously, there was a stacked hourglass model that could detect human poses relatively accurately. However, we compared the performance of the SimpleBaseline model, which can detect human poses more efficiently and with fewer operations than the stacked hourglass. origin paper stacked hourglass, original paper simplebaseline my experiment my experiment notebook

2. Method

2-1. Benchmark dataset

The data for the experiment used the MPII dataset.

2-2. Stacked hourglass Model

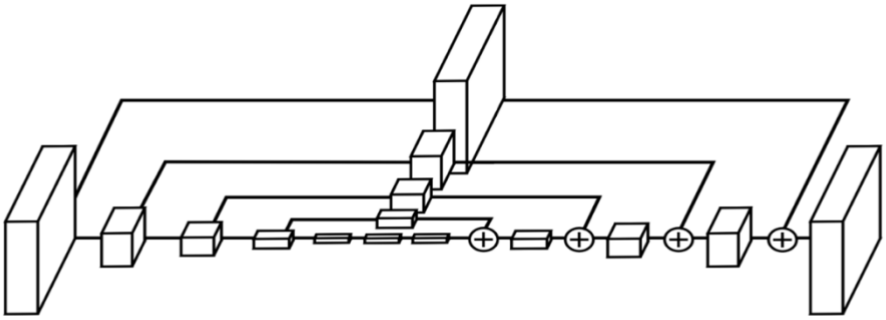
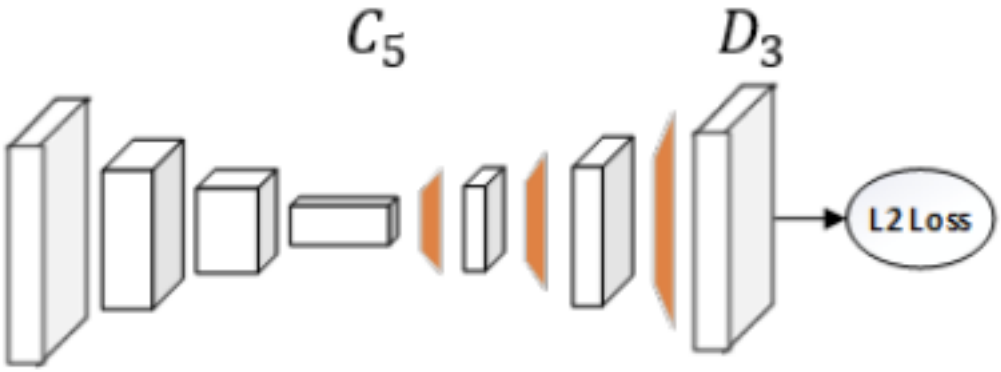


Fig. 3. An illustration of a single "hourglass" module. Each box in the figure corresponds to a residual module as seen in Figure 4. The number of features is consistent across the whole hourglass.

origin paper stacked hourglass

Train a Stacked hourglass model to compare with the Simple Baseline model

2-3. SimpleBaseline Model



original paper simplebaseline

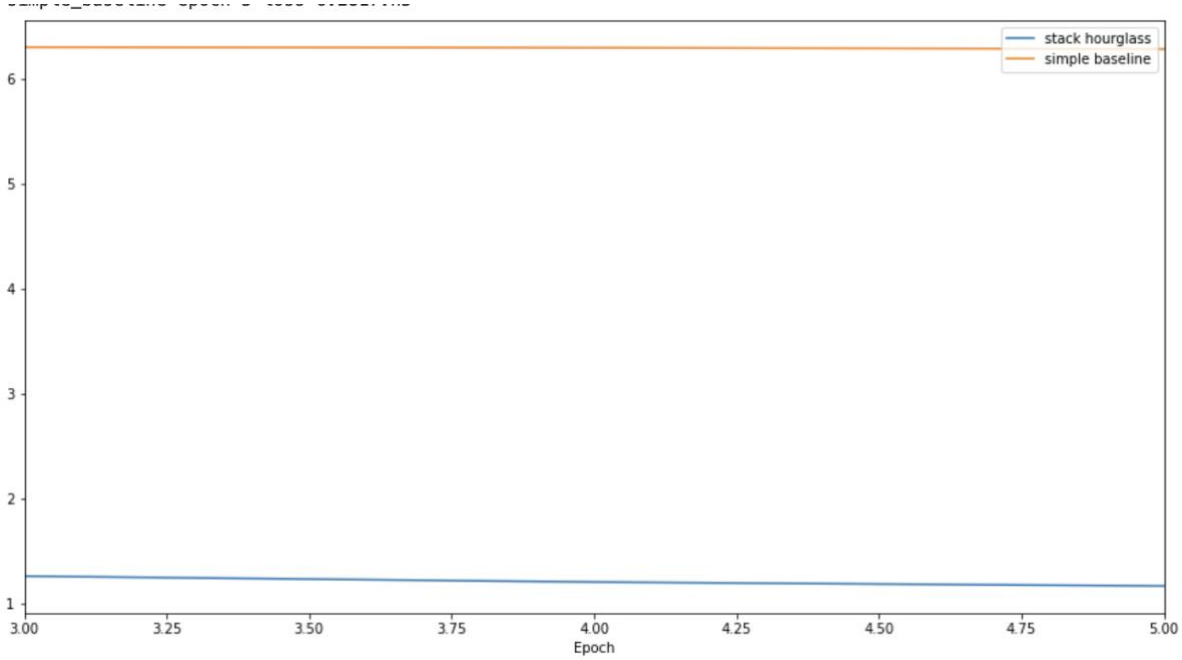
Training a SimpleBaseline model

The performance of the SimpleBaseline model was so poor that only SimpleBaseline was additionally trained.

3. Result

3-1. Comparison of Stacked hourglass and SimpleBaseline model validation loss

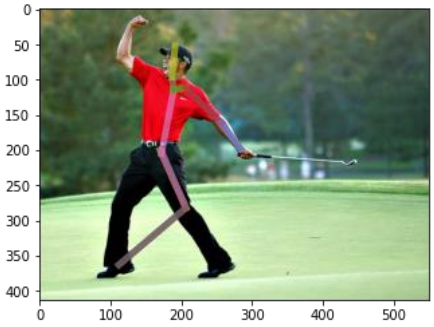
Stakced hourglass SimpleBaseline validation loss comparison based on Epoch 5



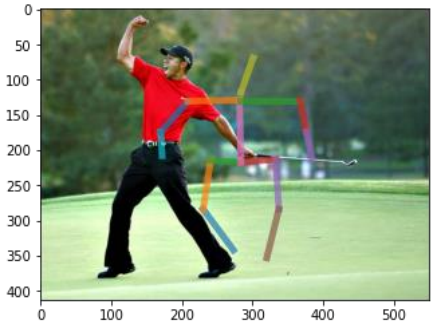
- The validation loss of stacked hourglass is 1.2~1.1.
- SimpleBaseline's validation loss is 6.3~6.2

3-2. Visualize and compare Stacked hourglass and SimpleBaseline models

Stakced hourglass(epoch 5)

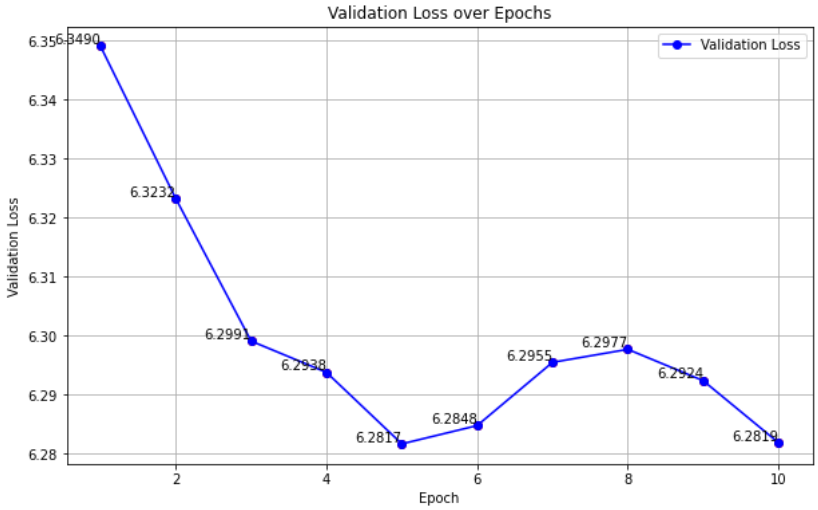


Simple Baseline(epoch 5)



3-3. The performance of the SimpleBaseline model is not good, so only SimpleBaseline is additionally trained.

Simple Baseline Additional Learning (Total Epoch 10)

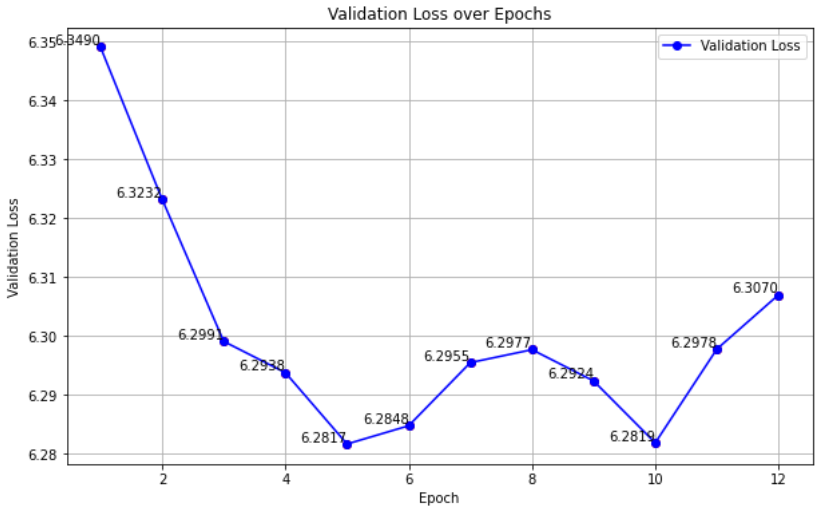


Additional SimpleBaseline learning was performed, and the loss actually increased by 0.0002 compared to the previous epoch 5.

Since we confirmed that the loss decreases from epoch 8, we expect the loss to decrease as we proceed with additional learning, and proceed with additional learning.

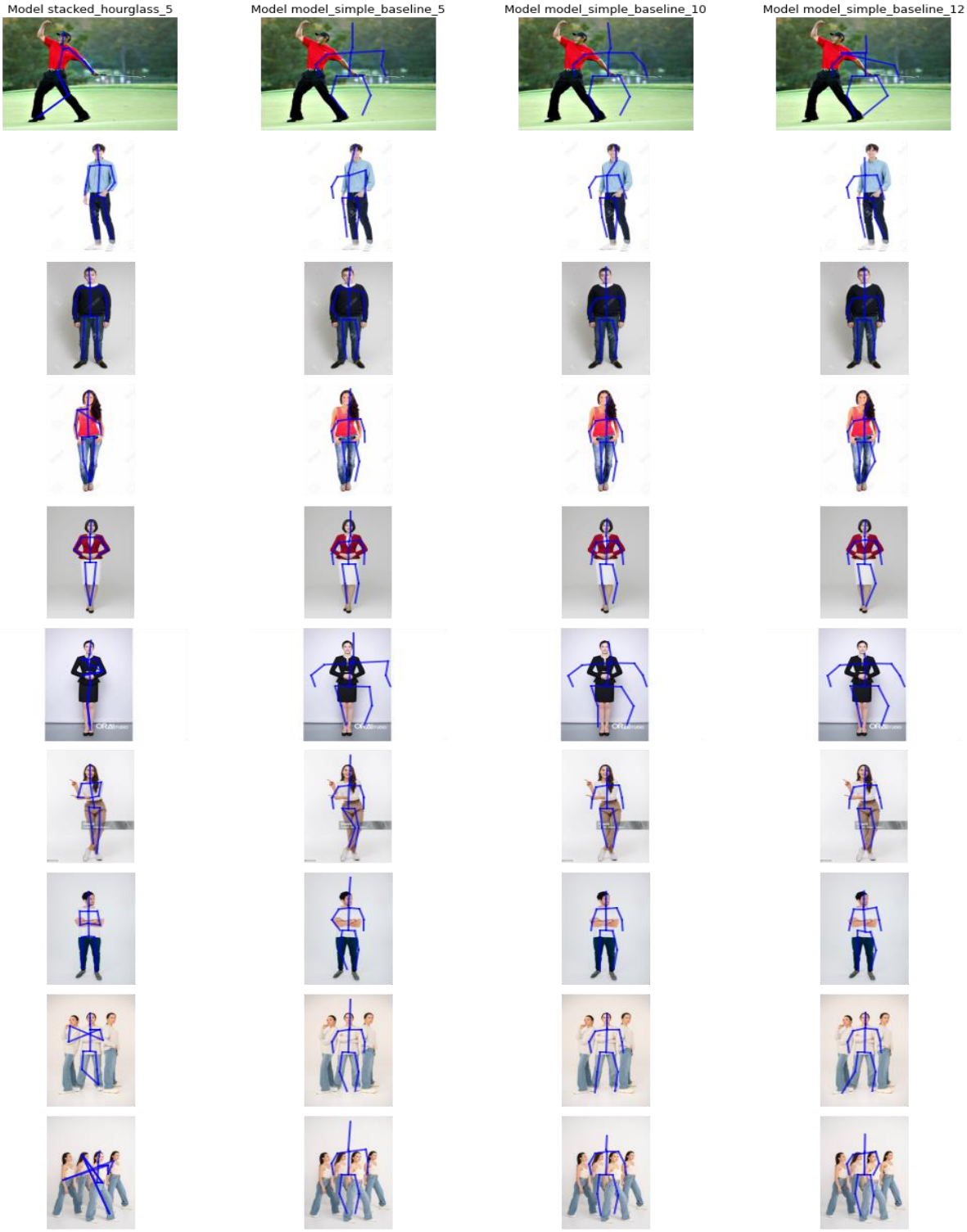
3-4. SimpleBaseline Additional Learning 2 times

Simple Baseline Additional Learning (Total Epoch 12)



Contrary to expectations, it increased again as of epoch 10. I decided that further study was meaningless.

3-5. Summary of qualitative comparison between Stacked hourglass and Simple Baseline



Stacked hourglass has higher accuracy on images with one person.

In images with multiple people, both models cannot predict pose.

4. Conclusion (+ Discussion, Future works)

For both models, the pose cannot be predicted for photos with multiple people.

SimpleBaseline's training is not progressing properly for some reason (loss is not reduced)