Supplementary Materials for Learning to Terminate in Object Navigation

Yuhang Song
Department of Computer Science
University of Liverpool, UK

Anh Nguyen
Department of Computer Science
University of Liverpool, UK

Chun-Yi Lee
Department of Computer Science National Tsing Hua University,

1. Judge Model Prediction Analyse

Figure 1: Judge Model Prediction. An example of judge model inference after 2000 updates during training, with ground truth and predictions. Where the ground truth 0 stands for time steps that should terminate.

2. Judge Model Prediction Analyse

Given that our RL branch aligns with the baseline model and our contribution lies in the introduction of the judge model, we drew direct comparisons with the baseline. An ablation study on the essential components is detailed in Table 1.

* The author is affiliated with the Department of Computer Science, National Tsing Hua University as well in a duel Ph.D. program.
<table>
<thead>
<tr>
<th>Method</th>
<th>w/ Transformer</th>
<th>w/o Focal Loss</th>
<th>w/o Judge Model</th>
<th>DITA (Ours)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SR(%)</td>
<td>SPL(%)</td>
<td>SR(%)</td>
<td>SPL(%)</td>
</tr>
<tr>
<td>All</td>
<td>55.6</td>
<td>15.8</td>
<td>58.5</td>
<td>20.1</td>
</tr>
<tr>
<td>L&gt;=5</td>
<td>36.8</td>
<td>15.5</td>
<td>38.4</td>
<td>17.6</td>
</tr>
</tbody>
</table>

Table 1: An ablation study on the judge model. In this table, ‘w/ Transformer’ denotes a variant that employs a Transformer as the feature encoder. On the other hand, in the variant ‘w/o Focal Loss,’ we replaced the focal loss in the judge model with cross-entropy loss for comparison.