

卒研第 2 稿に向けて

実験結果

精度：CER:0.6723

原因→黒板、もしあれやったら、JSON出力見せる

NA→Poolingを追加しmax, meanで比較

結果(CER)：max: 0.1944, mean: 0.1248

懸念：もとのTrOCRデコーダ使用してた場合：CER0.0433なので劣ってるのでこれがベースラインで本当にいいのか、、、

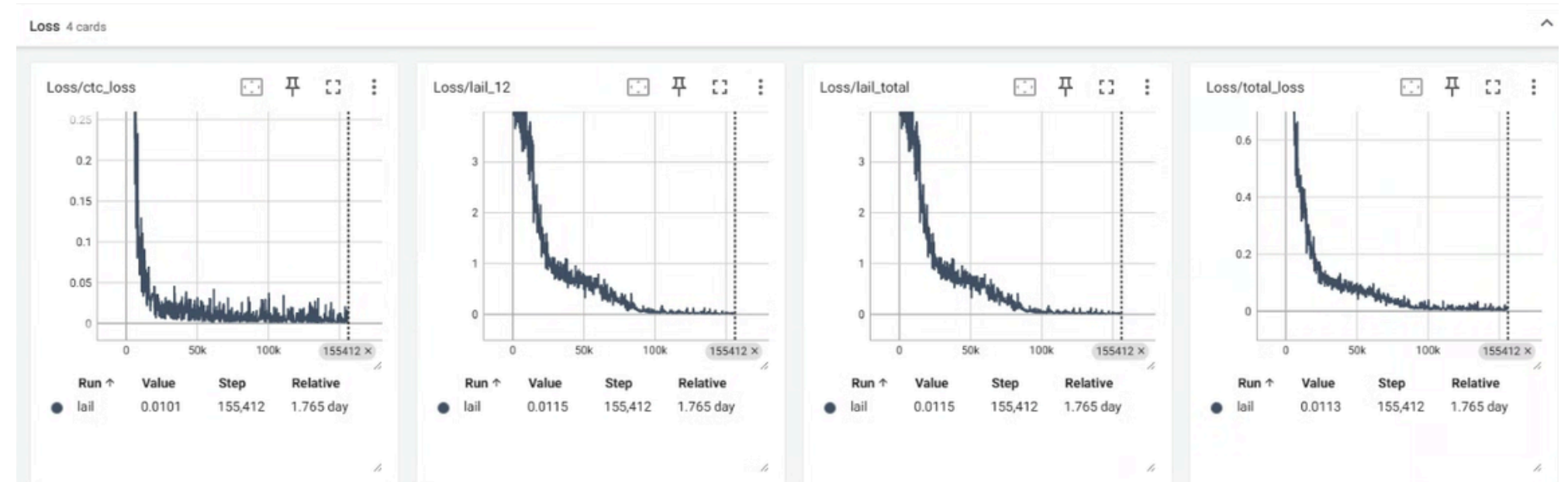
現状：+LLMの学習（共有PC）、デコーダの強化（個人PC）

ローカルだとLLM回せるGPUメモリないので共有マシン使用させていただいてます

現状

+LLMの学習

- encoderの最終層からLLMへ
- LLMモデル：Llama-3.2-1B
- 精度：0.2092(CER)
- 学習ログ（きれい！）



考察

- ハイパーパラメータがの比率ミスった

今後の展望

お悩み

- 卒論実験セクション
 - LLM無しとのアブレーション
 - LLMに渡す層のアブレーション
 - (2,4,6,8,10,12),(3,6,9,12),(4,8,12),(3,12),(6,12),(9,12),(12)
 - LLMの規模差によるアブレーション
 - LLM Size 1B,3B,8B
- ↑ どう組み合わせで実験すべきだろう
 - 推論速度のアブレーション
 - 学習時のみにLLMを使用する利点
 - TrOCRデコーダそのまま使ったやつとの比較はすべき??