Supplementary Materials: Hierachical Multi-label Learning for Incremental Multilingual Text Recognition

Anonymous Authors

1 MORE ALGORITHM ANALYSIS

1.1 Analysis of Rehearsal Size

In replay-based incremental methods, the size of the rehearsal set determines the extent to which old knowledge is retained. We conduct analysis experiments to assess the effect of the rehearsal set size M on HAMMER. The results in Fig. 1(a) and 1(b) for MLT17 and MLT19 demonstrate that as the size increases, more old knowledge is retained, resulting in improved accuracy on both old and new tasks and significant mitigation of the catastrophic forgetting. However, larger rehearsal sets may encounter issues like limited access to old data or storage constraints. Therefore, we select $M=2000$ to balance performance gains and data constraints as the default rehearsal set size.

1.2 Analysis of Trade-off Parameter

We investigate the impact of different trade-off parameters on MLT17 and MLT19. As depicted in Fig. 1(c) and 1(d), larger values of λ lead to relatively better prediction results. Essentially, the parameter λ controls the strength of the word and character language evaluation. Therefore, a larger λ may optimize new and shared knowledge more accurately to guide the final predictions of the multiple specific text recognizers. In this paper, our default setting is $\lambda=15$.
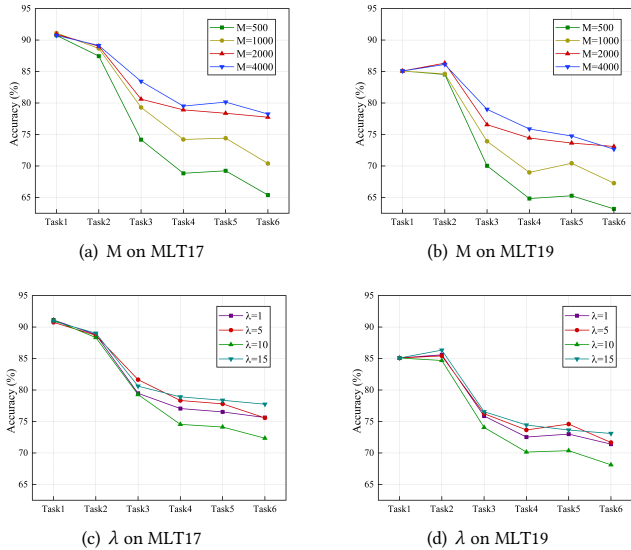

Figure 1: (a)(b): Results of different size of the rehearsal set M. (c)(d): Results of different trade-off parameter λ.

Figure 2: Visualization of Recognition Results. The character sequences located on the right side of the image from top to bottom are labels, baseline model, and HAMMER. Green and red indicate correctly and incorrectly recognized characters, respectively.

2 VISUALIZATION OF PREDICTION RESULTS

We randomly select two images for each incremental language to visualize the prediction results. On the left are the images for each language, and on the right are the corresponding results, from top to bottom: label, baseline model, and HAMMER. As can be seen from Fig. 2, the baseline model easily recognizes characters with similar glyphs no matter in which language, such as mistakenly recognizing 'n' as 'h' in English 'Enginering'. Our proposed HAMMER can accurately recognize the correct sequence of characters, especially in the earliest-learned Chinese, which illustrates the effectiveness of our proposed incremental MLTR framework.