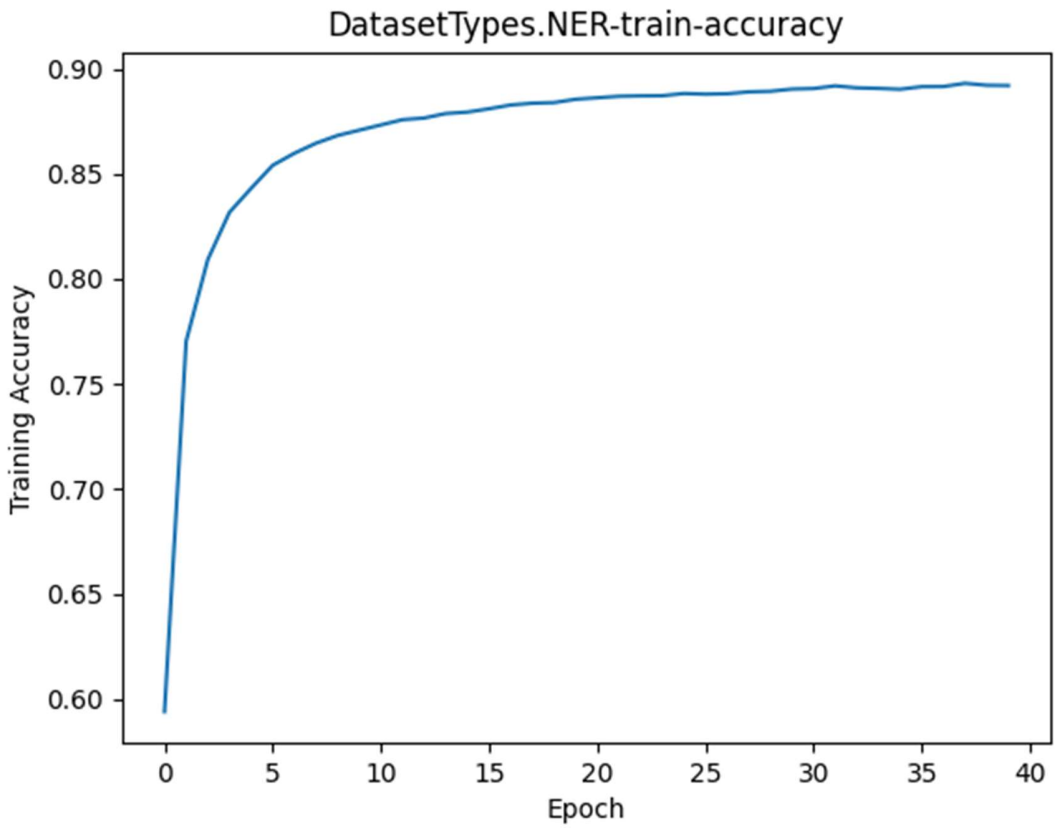
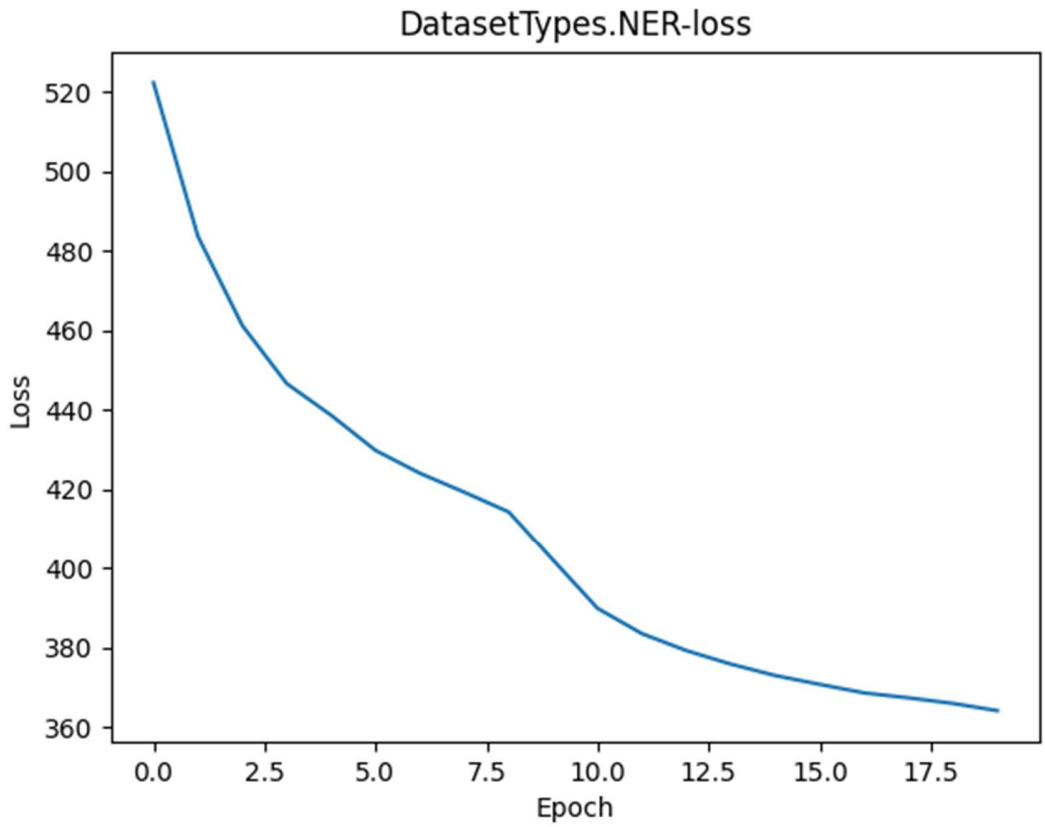
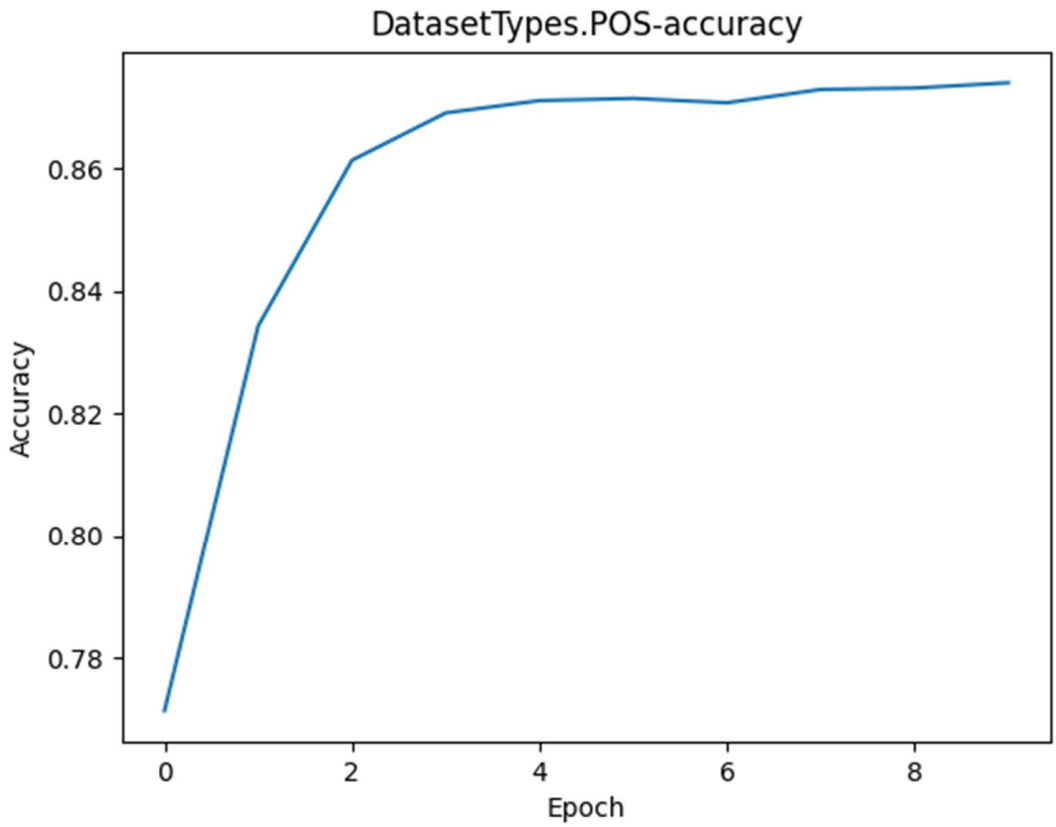
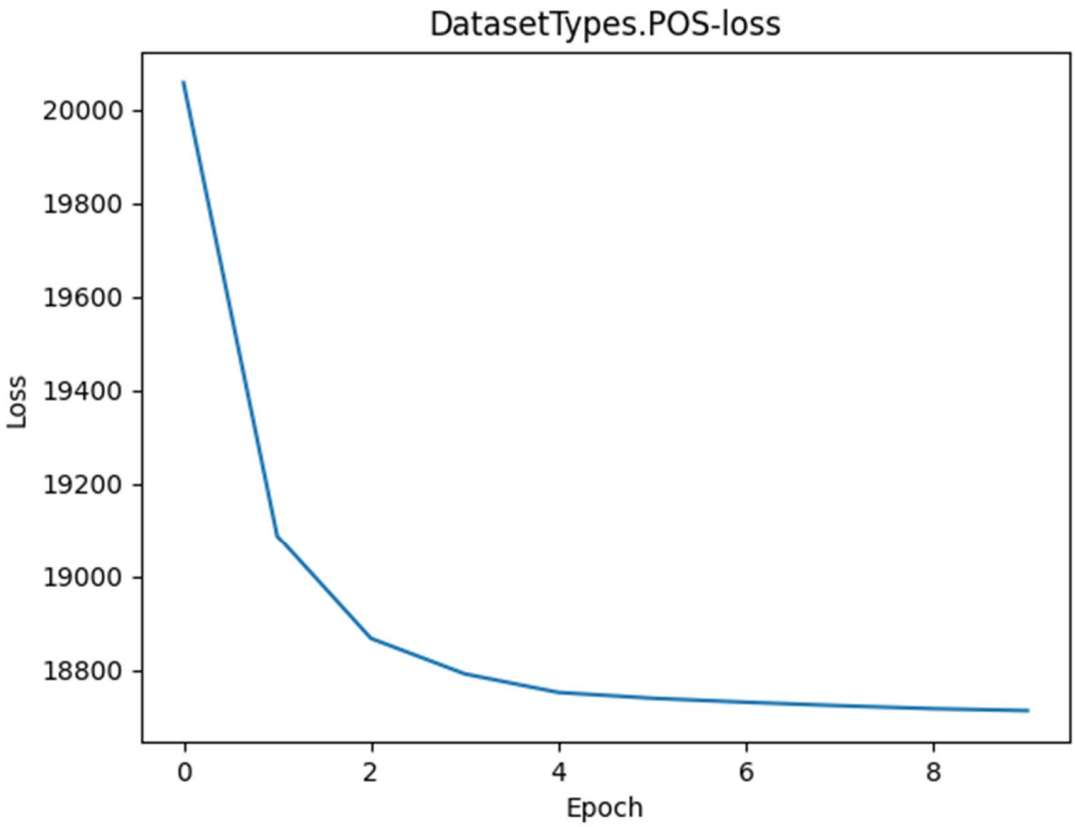
In our approach, we observed a notable phenomenon. The plots clearly illustrate a significant enhancement in performance, with accuracy surpassing 85% and exceeding the benchmark set by the previous method. Adjusting the window size and filter count was pivotal; reducing either led to a substantial drop in performance. While increasing the number of filters can initially boost accuracy, it eventually leads to a decline, although the optimal filter count remains undetermined. Additionally, we experimented with adding extra convolutional layers with ReLU activation but found that it negatively impacted training, leading us to discard this modification.

It is important to note that the POS dataset require far larger window size and more filters. It also took far more training to achieve convergence due to the complexity of the task




DatasetTypes.NER-loss
Loss
520
500
480
460
440
420
400
380
360
0.0
2.5
5.0
7.5
10.0
12.5
15.0
17.5
Epoch


DatasetTypes.POS-accuracy
Accuracy
0.86
0.84
0.82
0.80
0.78
0
2
4
6
8
Epoch


DatasetTypes.POS-loss
20000
19800
19600
19400
19200
19000
18800
Loss
0
2
4
6
8
Epoch