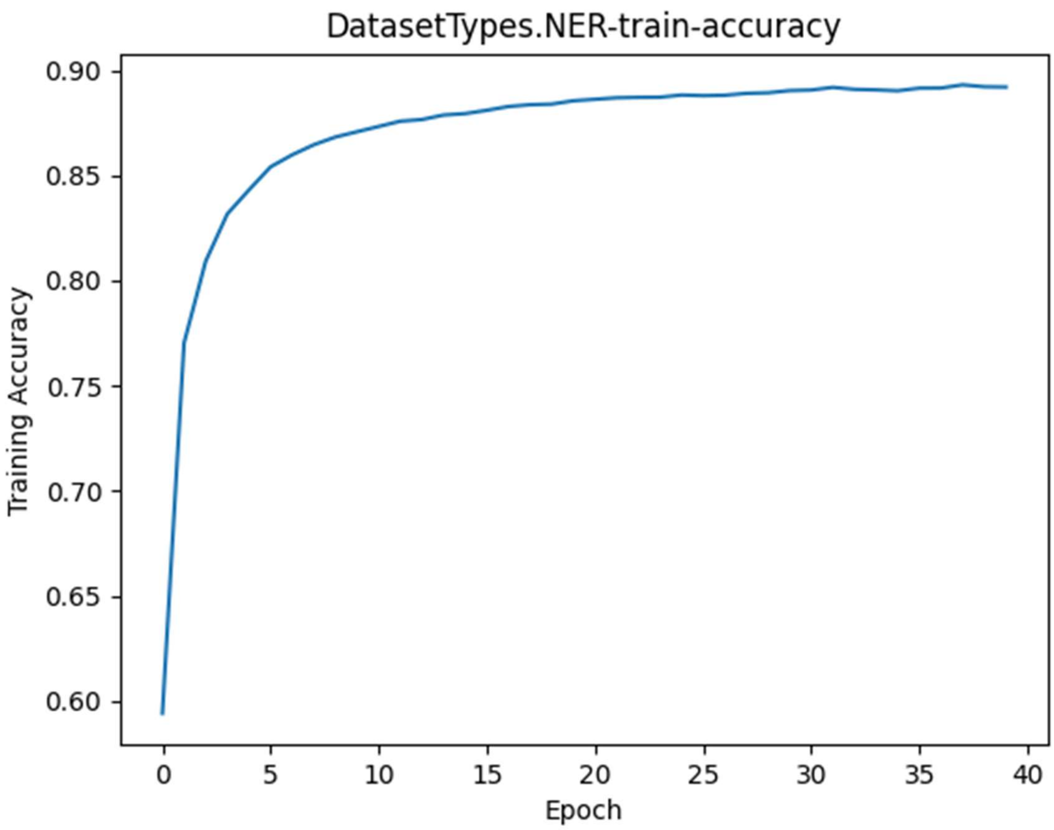
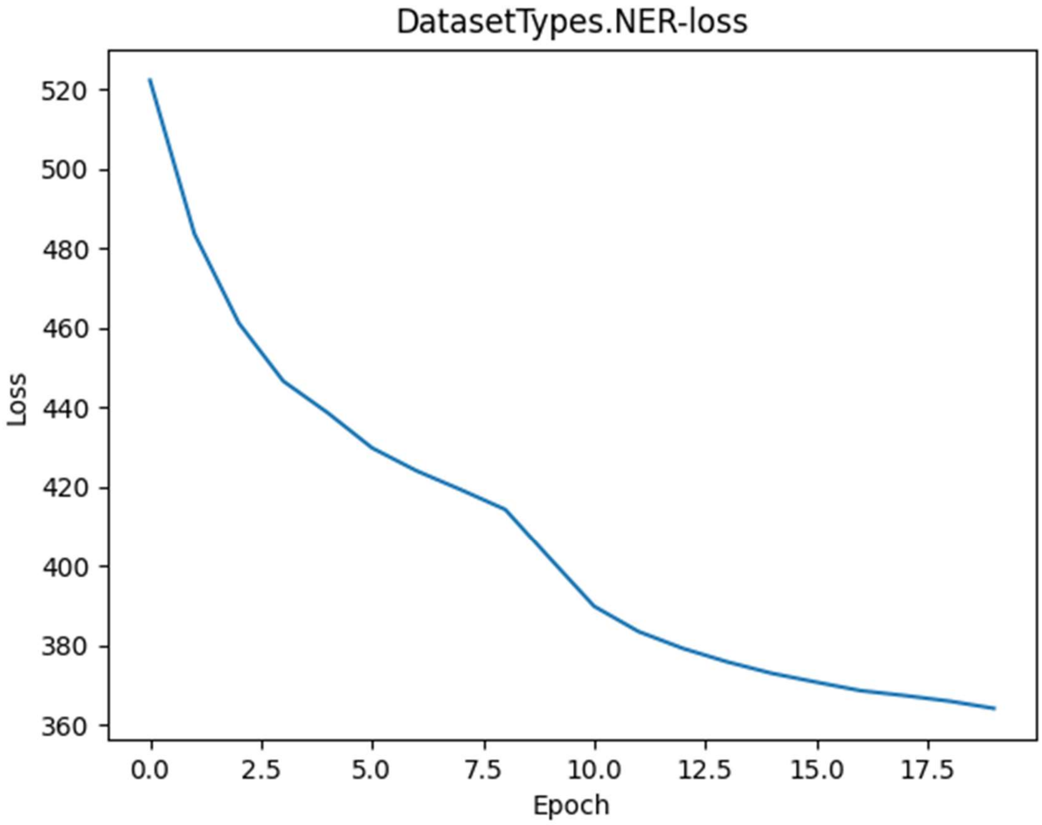
In our approach, we observed a notable phenomenon. The plots clearly illustrate a significant enhancement in performance, with accuracy surpassing 85% and exceeding the benchmark set by the previous method. Adjusting the window size and filter count was pivotal; reducing either led to a substantial drop 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.

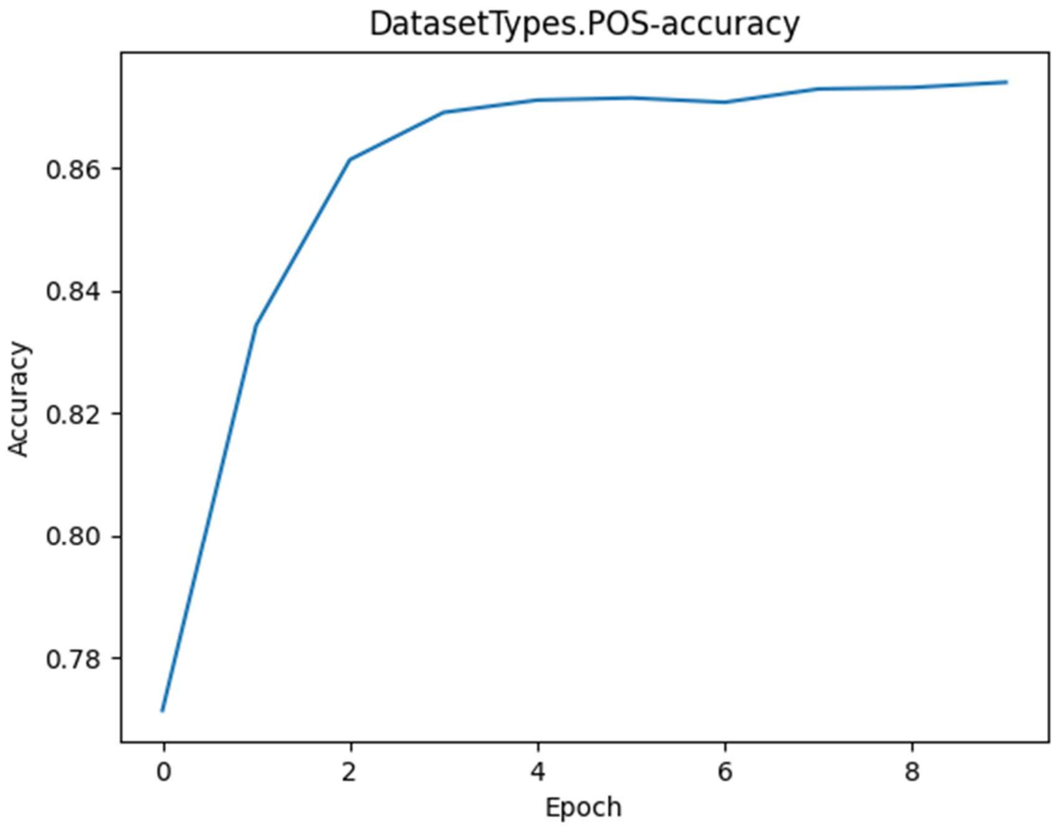
Important to note: Due to the overwhelming data of the O label. We had to ignore it during training to prevent collapse of the network to a single-label prediction. This obviously affected the prediction but we didn't find a way to get around it.

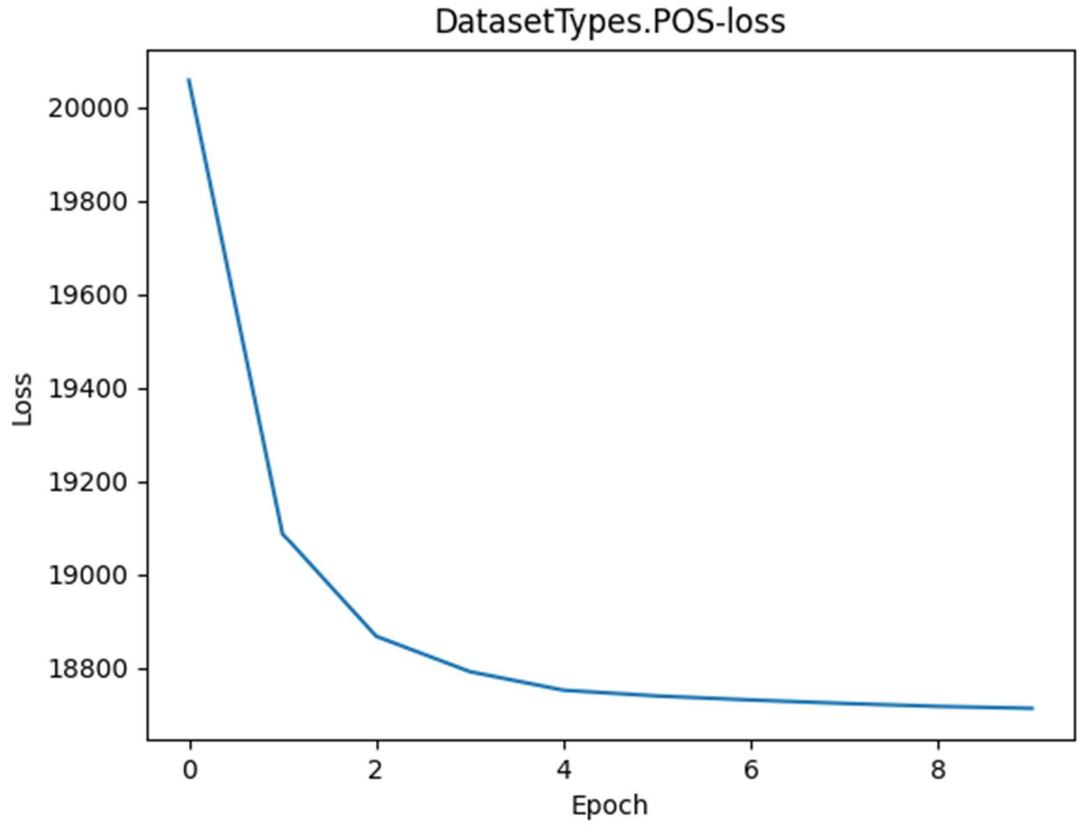
It is important to note that the POS dataset requires a far larger window size and more filters. It also took far more training to achieve convergence due to the complexity of the task

When we tried to find the meaning of the filters, we created a heat map of each filter, we included an example of a filter as an image below. As you can see there is a weight on the beginning and the end. This means the filter looks at both of these locations particularly to determine the classification of this word. I.E. the character of this location will trigger the filter. There are many types of filters like this, all dedicated to multiple locations within the word to trigger an indication of the classification of this word


DatasetTypes.NER-train-accuracy
Training Accuracy
Epoch
0.60
0.65
0.70
0.75
0.80
0.85
0.90
0
5
10
15
20
25
30
35
40


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
Loss
20000
19800
19600
19400
19200
19000
18800
0
2
4
6
8
Epoch