|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| 1 | first | 1,3,5 | √ | 0.945 | 0.987 | 1.000 | 0.055375 | 0.995291 |
| 2 | last | 1,3,5 | √ | 0.948 | 0.980 | 1.000 | 0.052117 | 0.9966432 |
| 3 | last | 1,3,5 | × | 0.928 | 0.987 | 1.000 | 0.071661 | 0.995544 |
| 4 | last | × | √ | 0.919 | 0.987 | 0.997 | 0.08143322 | 0.99476835 |
| 5 | last | 1 | √ | 0.932 | 0.977 | 0.997 | 0.0684039 | 0.99339431 |
| 6 | last | 3 | √ | 0.938 | 0.990 | 1.000 | 0.06188925 | 0.99723379 |
| 7 | last | 5 | √ | 0.925 | 0.977 | 0.993 | 0.0749186 | 0.9944704 |
| 8 | last | 7 | √ | 0.935 | 0.984 | 0990 | 0.0651466 | 0.9925593 |
| 9 | last | 3,5,7 | √ | 0.938 | 0.984 | 0.997 | 0.0618893 | 0.996415 |

Table 3. Ablation study. The top-k accuracy for k=1,3,5, test error, and AUC are shown for different experiment：1). The key features are obtained from the first M or the last M token features of the morphological features from the feature extractor, as shown in test 1 and 2;2) the location information under the neighborhood of convolution kernel size of 1, 3, 5 or its combination, and without the location information, as shown in test 2,4,5,6,7,8 and 9; 3). where the GAP layer is added: the class feature map or the WSIs feature map, as shown in test 2 and 3.