Supplementary Material for *Atlas of Digital Pathology: A Generalized Hierarchical Histological Tissue Type-Annotated Database for Deep Learning*

A Supplementary Materials

This document compiles the supplementary materials for the CVPR submission paper ID-6981 under the title of *Atlas of Digital Pathology: A Generalized Hierarchical Histological Tissue Type-Annotated Database for Deep Learning*. In the main paper, we studied the quality of the patch annotations by: (1) training a predictive convolutional neural network, and (2) collecting feedback on annotations from an expert pathologist. While the convolutional neural network learns to associate labels with observed visual patterns and makes consistent predictions while lacking high-level knowledge of label correctness, the pathologist is affected by human inconsistency but can draw on high-level histological knowledge. A label which was erroneously omitted by the ground-truth labeler would be detected by both the neural network and the pathologist, but a label consistently mislabeled as another type would only be detected by the pathologist. In the following three sections, we cover (a) the training error of the neural network, (b) the statistical analysis of the neural network and pathologist validation, and (c) association rule learning of individual WSIs.

A.1 Training Error Analysis

In Figure 1, we display the plots of the training and validation set accuracy and loss over the 80-epoch training process. All three network architectures (i.e. VGG16, ResNet18, and Inception-V3) are still improving in training accuracy and loss at 80 epochs but validation accuracy and loss have already converged. Also note that VGG16 starts the training slower than both ResNet18 and Inception-V3 but converges to a superior validation accuracy around the 50th epoch.

A.2 Statistical Analysis of Neural Network and Pathologist Validation

A.2.1 Overview

In this section, we consider the discordances of both the neural network (VGG16-layer-3+HBR with optimal thresholds on the test set of 1767 patches) and the pathologist (with their suggested label additions and subtractions on 1000 random patches) with the ground-truth labels, which enables us to find possible ground truth labeling errors. Ideally, perfect ground-truth labels would result in perfect confusion matrix metrics.

We also analyze possible reasons for the discordances by examining the residuals in the other classes using a novel metric we call the *mean prediction residuals (MPR)*. In cases where the model predicts a False Positive or a False Negative, a large residual error in another class going in the same direction (i.e. positive for FP, negative for FN) could indicate strong mutual inter-class support and possible label omission. Likewise, a large residual error in the opposite direction (i.e. negative for FP, positive for FN) could indicate strong mutual inter-class opposition and possible label swapping.

A.2.2 Confusion Matrix Metrics

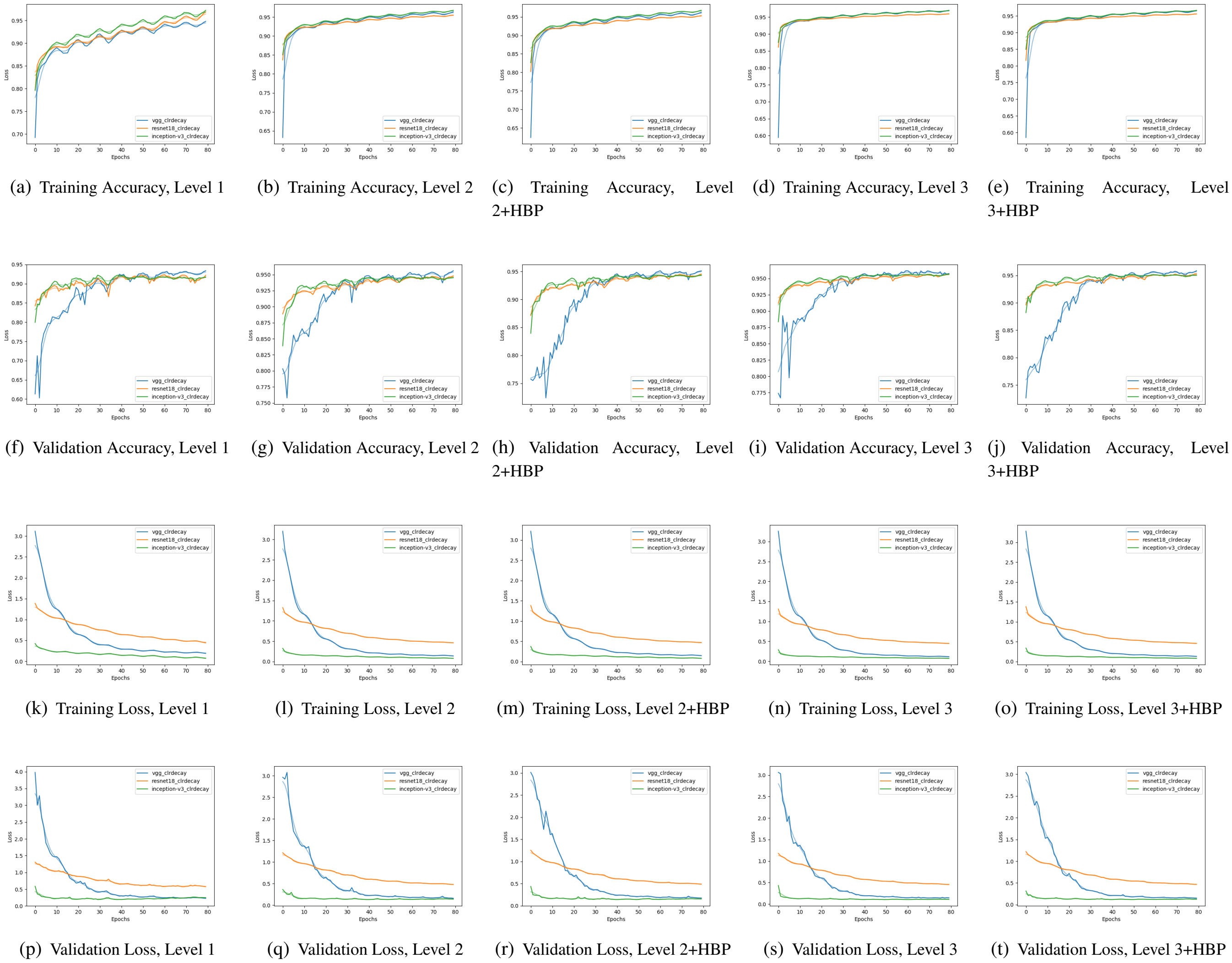
In Tables 1, 2, 3, 4, 5, 6, we analyze the confusion matrix metrics of the neural network and the pathologist for those classes with at least one ground-truth exemplar. The neural network does not give predictions for “Undifferentiated” tissue types (i.e. with codes ending in “X”). Overall, the confusion matrix metrical performance for both models is very good and the worst discordances exist for classes with either known consistent mislabeling errors (according to the pathologist) or few training examples (which disadvantages the neural network but not the pathologist).

A.2.3 Mean Prediction Residual

Overview

The Mean Prediction Residual (MPR) is a metric that we devised to measure the prediction residuals in the other (consequent) classes whenever a discordance exists for a given (antecedent) class. For each discordance (either a false positive or a false negative), the prediction residual is the difference between a target label and its predicted score, where their mean is the MPR. There are two types of MPR:

1. FP-MPR (for false positives)
2. FN-MPR (for false negatives)



(a) Training Accuracy, Level 1 (b) Training Accuracy, Level 2 (c) Training Accuracy, Level 2+HBP (d) Training Accuracy, Level 3 (e) Training Accuracy, Level 3+HBP

(f) Validation Accuracy, Level 1 (g) Validation Accuracy, Level 2 (h) Validation Accuracy, Level 2+HBP (i) Validation Accuracy, Level 3 (j) Validation Accuracy, Level 3+HBP

(k) Training Loss, Level 1 (l) Training Loss, Level 2 (m) Training Loss, Level 2+HBP (n) Training Loss, Level 3 (o) Training Loss, Level 3+HBP

(p) Validation Loss, Level 1 (q) Validation Loss, Level 2 (r) Validation Loss, Level 2+HBP (s) Validation Loss, Level 3 (t) Validation Loss, Level 3+HBP

Figure 1. *Training progress plots for all three network architectures across all five training configurations and four metrics: the rows of the figure correspond to different metrics and the columns correspond to different training configurations. Original values are shown as solid lines, smoothed values are shown as translucent lines.*

Table 1. *True Positive Rate (TPR)*

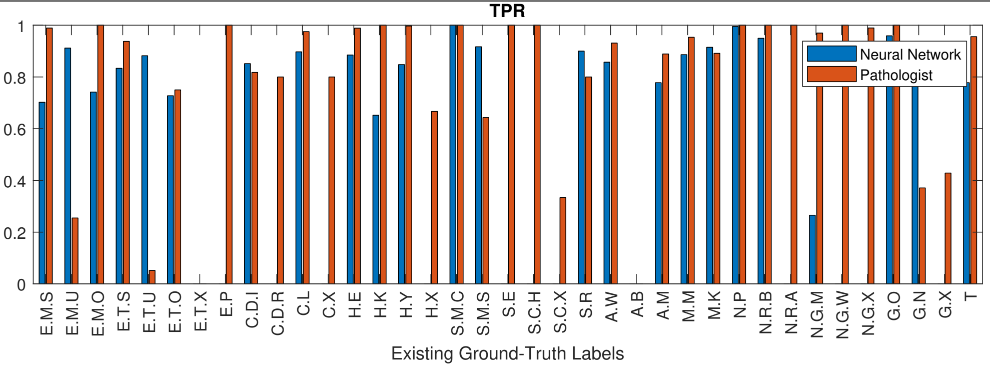
Confusion Matrix Metric	
Neural Network	Agrees well with ground-truth positive labels except for N.G.M (which has few training examples)
Pathologist	Agrees with ground-truth positive labels except for E.M.U, E.T.U, E.T.O, H.X, S.M.S, S.C.X, G.N, and G.X (which are known to have systematic mislabeling errors)

Table 2. *False Positive Rate (FPR)*

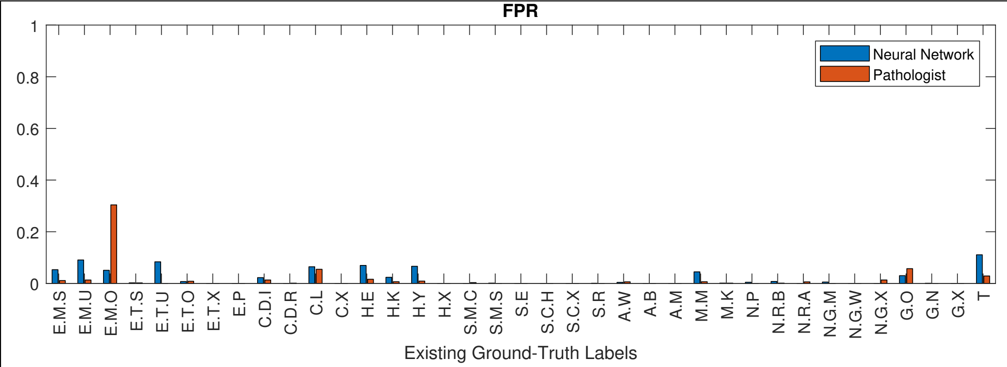
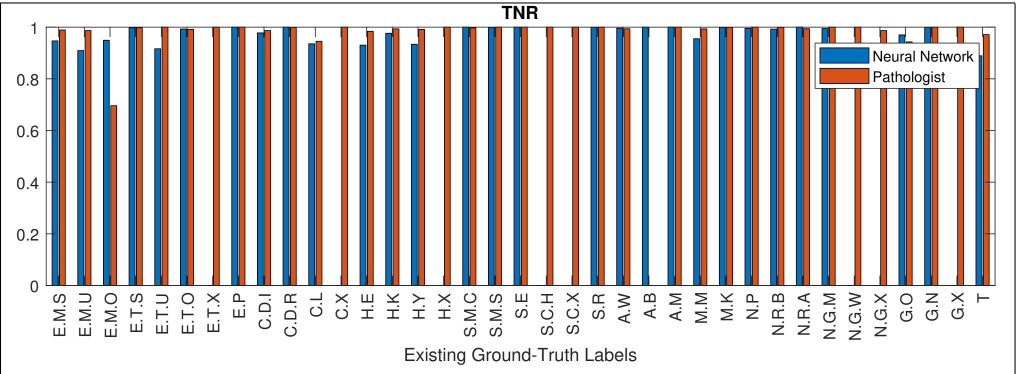
Confusion Matrix Metric	
Neural Network	Predicts more false positives in general (perhaps it is overly sensitive to small regions of tissue than humans)
Pathologist	Agrees well with most ground-truth positive labels except for E.M.O (which are known to be consistently mislabeled as E.M.U in the ground truth)

Table 3. *True Negative Rate (TNR)*

Confusion Matrix Metric	
Neural Network	Highly agrees with ground-truth negative labels
Pathologist	Highly agrees with ground-truth negative labels, except for E.M.O (related to consistent mislabeling with E.M.U mentioned above)

The FP-MPR between an antecedent class A and a consequent class B is defined as follows:

$$\begin{aligned} \text{FP-MPR}(A,B) &= \mathbb{E}[\text{Pred}_B(i) - \text{Targ}_B(i)]\forall i \\ &\text{s.t. Targ}_A(i) = 0, \text{Pred}_A(i) = 1 \end{aligned} \quad (1)$$

The FN-MPR between an antecedent class A and a consequent class B is defined as follows:

$$\begin{aligned} \text{FN-MPR}(A,B) &= \mathbb{E}[\text{Pred}_B(i) - \text{Targ}_B(i)]\forall i \\ &\text{s.t. Targ}_A(i) = 1, \text{Pred}_A(i) = 0 \end{aligned} \quad (2)$$

These pairwise relationships between antecedent and consequent classes can be displayed in a matrix, with each row corresponding with an antecedent for which a discordance exists and the columns corresponding to the consequent classes for which the prediction residual is calculated - this is known as the MPR matrix. Antecedent classes without any discordances can be shown as black rows but still re-appear as consequent classes in the columns. Another way to understand the MPR matrix is to show the consequent classes with the highest and lowest MPR value for each row of the MPR matrix. This is called the Max/Min MPR Table.

False Positive-Mean Prediction Residual (FP-MPR)

In Table 7, we show the FP-MPR matrices for both the neural network and the pathologist in the second column, and the max/min FP-MPR tables in the third column. For the neural network, we observe that the most mutually-supporting classes already co-occur frequently in the data while while those in strong opposition are similar in appearance. A similar pattern is observed for the pathologist False Positives. However, since the neural network picks up inconsistencies in the labeling and the pathologist picks up both inconsistencies and systematic mislabeling (as mentioned earlier), then classes with much more negative FP-MPR values for the pathologist than the neural network are likely to have systematic mislabeling, such as the FP-MPR from G.O to G.N, which drops from just -0.16129 for the neural network to -0.75758 for the pathologist.

False Negative-Mean Prediction Residual (FN-MPR)

Table 4. *False Negative Rate (FNR)*

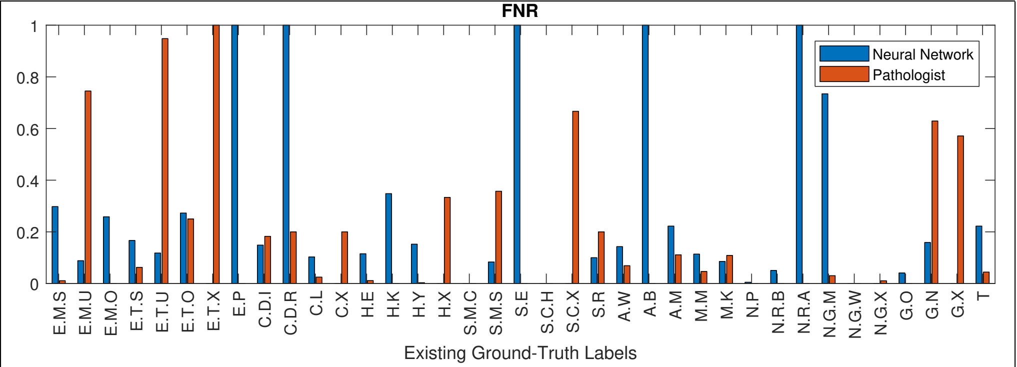
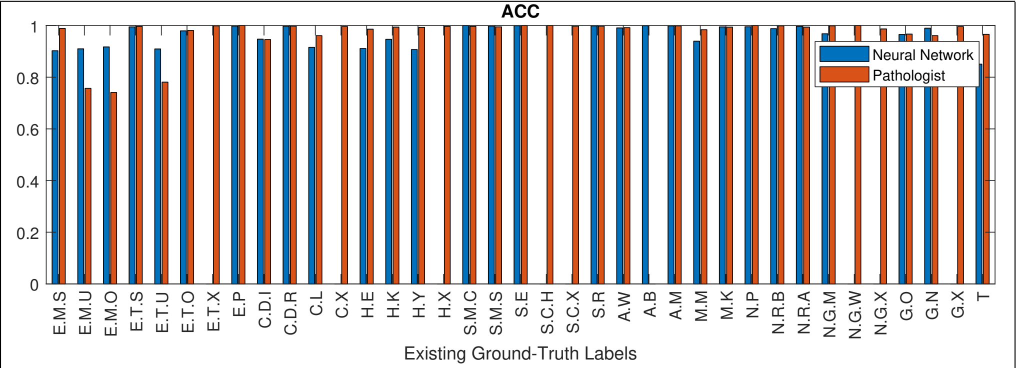
Confusion Matrix Metric	
Neural Network	Agrees well with ground-truth negative labels except for N.G.M (which has few training examples)
Pathologist	Generally agrees with the ground-truth positive labels, except for E.M.U, E.T.U, E.T.O, H.X, S.M.S, S.C.X, G.N, and G.X (which are known to have systematic mislabeling errors)

Table 5. *Accuracy (ACC)*

Confusion Matrix Metric	
Neural Network	Highly accurate across all classes
Pathologist	Highly accurate across all classes

In Table 8, we show the FN-MPR matrices for both the neural network and the pathologist in the second column, and the max/min FN-MPR tables in the third column. For the neural network and the pathologist, we observe (similarly to FP) that classes with strong mutual support tend to co-occur frequently in the data while those in strong opposition are similar in appearance. Again, as for FP-MPR, those classes with much more positive FP-MPR values for the pathologist than the neural network are likely to have systematic mislabeling. For example, the FN-MPR from E.M.U to E.M.O rises from just 0.21277 for the neural network to 0.79487 for the pathologist.

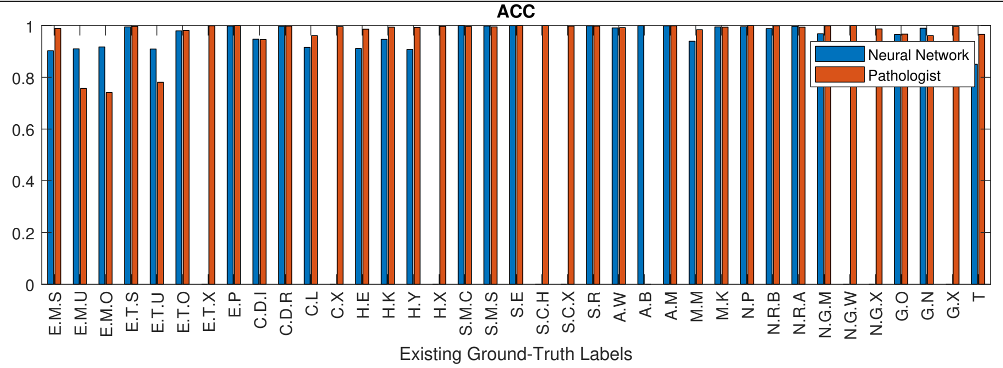
B Association Rule Learning of WSI Scoring

In this section we study (1) co-occurence network and (2) associate rule learning (ARL) (introduced in the submitted paper draft) using six different WSIs shown in in Table 9 (first row). Note that slide numbers 1, 2, and 3 here are the same slides used to demonstrates the heatmap representation in Figure 5 of the submitted paper draft. The circular co-occurrence of each slide is demonstrated in Table 9 using different number of image patches per slide, where the labels of each patch (extracted from individual WSI) are predicted by different levels of VGG16 trained network. The nodes of co-occurrence network here share the similar connections of the ADP co-occurrence shown in Figure 2 of the submitted paper draft.

The results of applying Apriori ARL algorithm to the predicted labels (driven from VGG16-level-3+HBP) of each WSI are also shown in Tables 10 to Tables 15. The selected consequent labels here are mainly similar to the ARL results of ADP Atlas demonstrated in Table 2 of the submitted paper draft. However, we notice dissimilarities in the antecedent itemsets between the selected WSIs here and the ones used to populate the ADP database. This is mainly because (a) they are related to different tissue cases; (b) different levels of confidence are selected as the best candidates. In fact, the majority of WSIs here are selected from GI tracts. In conclusion, we observe the following

1. there are no G.N for E.M.U (because no endocrine glands in GI)

Table 6. *F1 Score (F1)*

Confusion Matrix Metric	
Neural Network	Has low F1 score for E.M.S, E.M.O, H.K, and N.G.M (which were accidentally omitted in the ground truth)
Pathologist	Has low F1 score for E.M.U, E.M.O, E.T.U, S.C.X, N.R.A, G.N, and G.X (which are known to have systematic mislabeling errors)

2. there are no M.K, T for C.D.I and instead has M.M, A.W, and T (because no skeletal muscle in GI)

3. there are no N.P and N.R.B (because generally no nervous tissue in GI)

4. there are no H.Y for G.O and T (because they have less lymphocytes near exocrine gland and transport vessels)

In particular, between the ARL of Slide #4 shown in Table 13 and the other slides, we observe

1. Slide #4 ARL has E.M.S, C.L, H.K instead of E.M.U, E.T.U, H.E, H.K for H.Y (because Slide #4 lacks the exocrine glands with those epithelia - H.Y occurs with C.L instead)

2. Slide #4 ARL has H.K, H.Y, M.M instead of H.K, H.E, E.M.S, E.M.O for C.L (because Slide #4 again lacks the exocrine glands with those epithelia - C.L occurs with M.M instead)

Table 7. *FP-MPR Matrices and Max/Min Tables for Neural Network and Pathologist*

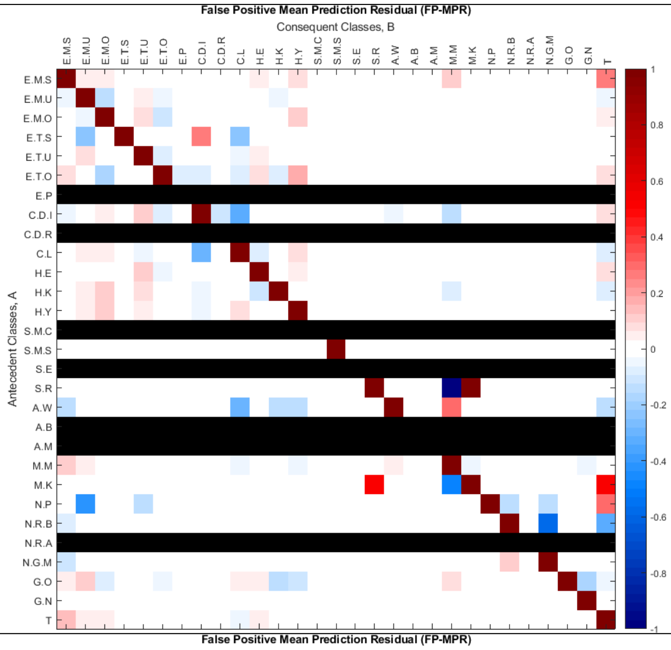
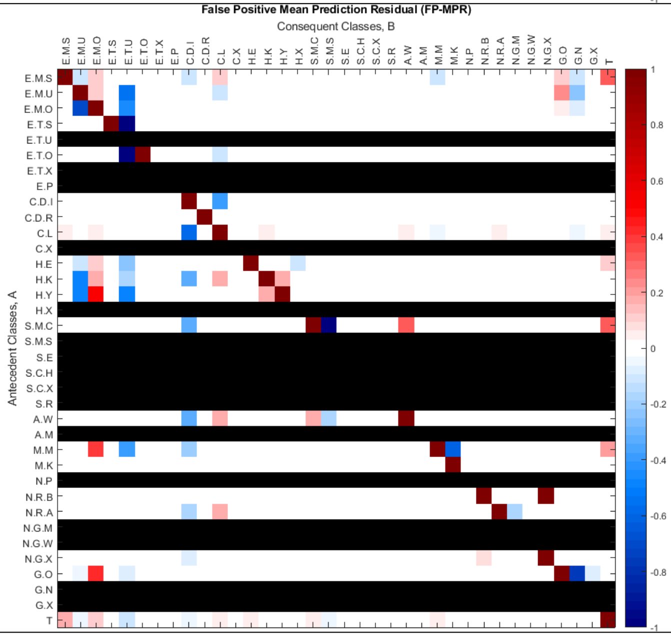
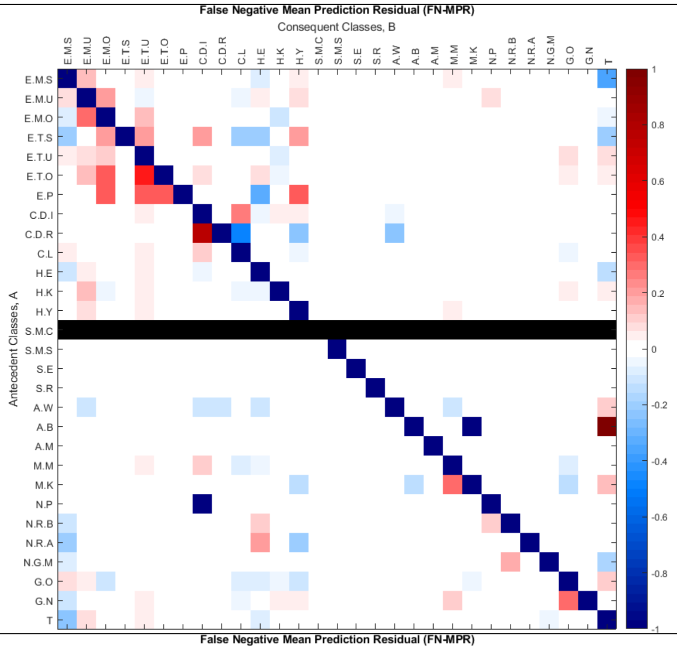
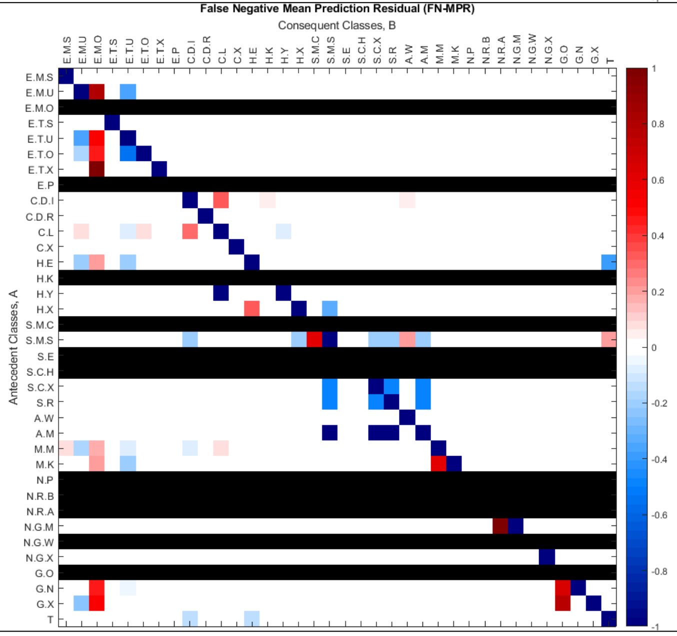
<table>
<tr><th>Model</th><th>FP-MPR Matrix</th><th>Max/Min FP-MPR Table</th></tr>
<tr><td>Neural Network</td><td>
</td><td>
<table>
<tr><th>Antecedent
Class</th><th>Max FP-MPR
Class</th><th>Max FP-MPR</th><th>Min FP-MPR
Class</th><th>Min FP-MPR</th></tr>
<tr><td>E.M.S</td><td>T</td><td>0.25974</td><td>C.D.I</td><td>−0.02597</td></tr>
<tr><td>E.M.U</td><td>E.T.U</td><td>0.053571</td><td>E.M.O</td><td>−0.13393</td></tr>
<tr><td>E.M.O</td><td>H.Y</td><td>0.118421</td><td>E.T.O</td><td>−0.10526</td></tr>
<tr><td>E.T.S</td><td>C.D.I</td><td>0.25</td><td>E.M.U</td><td>−0.25</td></tr>
<tr><td>E.T.U</td><td>E.M.U</td><td>0.09322</td><td>E.T.O</td><td>−0.08475</td></tr>
<tr><td>E.T.O</td><td>H.Y</td><td>0.166667</td><td>E.M.O</td><td>−0.16667</td></tr>
<tr><td>C.D.I</td><td>E.T.U</td><td>0.1</td><td>C.L</td><td>−0.33333</td></tr>
<tr><td>C.L</td><td>H.Y</td><td>0.072727</td><td>C.D.I</td><td>−0.29091</td></tr>
<tr><td>H.E</td><td>E.T.U</td><td>0.112676</td><td>E.T.O</td><td>−0.04225</td></tr>
<tr><td>H.K</td><td>E.M.O</td><td>0.105263</td><td>H.E</td><td>−0.10526</td></tr>
<tr><td>H.Y</td><td>E.M.O</td><td>0.098765</td><td>C.D.I</td><td>−0.03704</td></tr>
<tr><td>S.M.S</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>S.R</td><td>M.K</td><td>1</td><td>M.M</td><td>−1</td></tr>
<tr><td>A.W</td><td>M.M</td><td>0.285714</td><td>C.L</td><td>−0.28571</td></tr>
<tr><td>M.M</td><td>E.M.S</td><td>0.098361</td><td>C.L</td><td>−0.03279</td></tr>
<tr><td>M.K</td><td>S.R</td><td>0.5</td><td>M.M</td><td>−0.5</td></tr>
<tr><td>N.P</td><td>T</td><td>0.285714</td><td>E.M.U</td><td>−0.42857</td></tr>
<tr><td>N.R.B</td><td>E.M.U</td><td>0</td><td>N.G.M</td><td>−0.58333</td></tr>
<tr><td>N.G.M</td><td>N.R.B</td><td>0.111111</td><td>E.M.S</td><td>−0.11111</td></tr>
<tr><td>G.O</td><td>E.M.U</td><td>0.096774</td><td>G.N</td><td>−0.16129</td></tr>
<tr><td>G.N</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>T</td><td>E.M.S</td><td>0.148438</td><td>C.L</td><td>−0.04688</td></tr>
</table></td></tr>
<tr><td>Pathologist</td><td>
</td><td>
<table>
<tr><th>Antecedent
Class</th><th>Max FP-MPR
Class</th><th>Max FP-MPR</th><th>Min FP-MPR
Class</th><th>Min FP-MPR</th></tr>
<tr><td>E.M.S</td><td>T</td><td>0.33333</td><td>E.M.U</td><td>−0.11111</td></tr>
<tr><td>E.M.U</td><td>G.O</td><td>0.22222</td><td>E.T.U</td><td>−0.55556</td></tr>
<tr><td>E.M.O</td><td>G.O</td><td>0.054054</td><td>E.M.U</td><td>−0.71429</td></tr>
<tr><td>E.T.S</td><td>E.M.S</td><td>0</td><td>E.T.U</td><td>−1</td></tr>
<tr><td>E.T.O</td><td>E.M.S</td><td>0</td><td>E.T.U</td><td>−1</td></tr>
<tr><td>C.D.I</td><td>E.M.S</td><td>0</td><td>C.L</td><td>−0.4</td></tr>
<tr><td>C.D.R</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>C.L</td><td>E.M.S</td><td>0.038462</td><td>C.D.I</td><td>−0.57692</td></tr>
<tr><td>H.E</td><td>E.M.O</td><td>0.11111</td><td>E.T.U</td><td>−0.22222</td></tr>
<tr><td>H.K</td><td>E.M.O</td><td>0.16667</td><td>E.M.U</td><td>−0.5</td></tr>
<tr><td>H.Y</td><td>E.M.O</td><td>0.5</td><td>E.M.U</td><td>−0.5</td></tr>
<tr><td>S.M.C</td><td>A.W</td><td>0.33333</td><td>S.M.S</td><td>−1</td></tr>
<tr><td>A.W</td><td>C.L</td><td>0.16667</td><td>C.D.I</td><td>−0.33333</td></tr>
<tr><td>M.M</td><td>E.M.O</td><td>0.4</td><td>M.K</td><td>−0.6</td></tr>
<tr><td>M.K</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>N.R.B</td><td>N.G.X</td><td>1</td><td>E.M.S</td><td>0</td></tr>
<tr><td>N.R.A</td><td>C.L</td><td>0.16667</td><td>C.D.I</td><td>−0.16667</td></tr>
<tr><td>N.G.X</td><td>N.R.B</td><td>0.083333</td><td>C.D.I</td><td>−0.083333</td></tr>
<tr><td>G.O</td><td>E.M.O</td><td>0.42424</td><td>G.N</td><td>−0.75758</td></tr>
<tr><td>T</td><td>E.M.S</td><td>0.15789</td><td>E.T.U</td><td>−0.10526</td></tr>
</table></td></tr>
</table>

Table 8. *FN-MPR Matrices and Max/Min Tables for Neural Network and Pathologist*

<table>
<tr><th>Model</th><th>FN-MPR Matrix</th><th>Max/Min FN-MPR Table</th></tr>
<tr><td>Neural Network</td><td>


</td><td>
<table>
<tr><th>Antecedent Class</th><th>Max FN-MPR Class</th><th>Max FN-MPR</th><th>Min FN-MPR Class</th><th>Min FN-MPR</th></tr>
<tr><td>E.M.S</td><td>E.M.U</td><td>0.12632</td><td>T</td><td>−0.34737</td></tr>
<tr><td>E.M.U</td><td>E.M.O</td><td>0.21277</td><td>E.T.U</td><td>−0.042553</td></tr>
<tr><td>E.M.O</td><td>E.M.U</td><td>0.3</td><td>H.K</td><td>−0.1</td></tr>
<tr><td>E.T.S</td><td>E.M.O</td><td>0.2</td><td>E.M.S</td><td>−0.2</td></tr>
<tr><td>E.T.U</td><td>E.M.O</td><td>0.11905</td><td>H.K</td><td>−0.071429</td></tr>
<tr><td>E.T.O</td><td>E.T.U</td><td>0.45833</td><td>H.K</td><td>−0.041667</td></tr>
<tr><td>E.P</td><td>E.M.O</td><td>0.33333</td><td>H.E</td><td>−0.33333</td></tr>
<tr><td>C.D.I</td><td>C.L</td><td>0.25397</td><td>H.E</td><td>−0.031746</td></tr>
<tr><td>C.D.R</td><td>C.D.I</td><td>0.75</td><td>C.L</td><td>−0.5</td></tr>
<tr><td>C.L</td><td>C.D.I</td><td>0.10638</td><td>H.Y</td><td>−0.042553</td></tr>
<tr><td>H.E</td><td>E.T.U</td><td>0.05814</td><td>T</td><td>−0.12791</td></tr>
<tr><td>H.K</td><td>E.M.U</td><td>0.125</td><td>E.M.O</td><td>−0.053571</td></tr>
<tr><td>H.Y</td><td>E.M.U</td><td>0.072289</td><td>E.M.O</td><td>−0.024096</td></tr>
<tr><td>S.M.S</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>S.E</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>S.R</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>A.W</td><td>T</td><td>0.11111</td><td>E.M.U</td><td>−0.11111</td></tr>
<tr><td>A.B</td><td>T</td><td>1</td><td>M.K</td><td>−1</td></tr>
<tr><td>A.M</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>M.M</td><td>C.D.I</td><td>0.1087</td><td>C.L</td><td>−0.086957</td></tr>
<tr><td>M.K</td><td>M.M</td><td>0.28571</td><td>H.Y</td><td>−0.14286</td></tr>
<tr><td>N.P</td><td>E.M.S</td><td>0</td><td>C.D.I</td><td>−1</td></tr>
<tr><td>N.R.B</td><td>H.E</td><td>0.11111</td><td>E.M.S</td><td>−0.11111</td></tr>
<tr><td>N.R.A</td><td>H.E</td><td>0.2</td><td>E.M.S</td><td>−0.2</td></tr>
<tr><td>N.G.M</td><td>N.R.B</td><td>0.17021</td><td>T</td><td>−0.17021</td></tr>
<tr><td>G.O</td><td>T</td><td>0.1</td><td>E.M.O</td><td>−0.1</td></tr>
<tr><td>G.N</td><td>G.O</td><td>0.29412</td><td>E.M.S</td><td>−0.11765</td></tr>
<tr><td>T</td><td>E.M.U</td><td>0.088235</td><td>E.M.S</td><td>−0.25</td></tr>
</table>
</td></tr>
<tr><td>Pathologist</td><td>


</td><td>
<table>
<tr><th>Antecedent Class</th><th>Max FN-MPR Class</th><th>Max FN-MPR</th><th>Min FN-MPR Class</th><th>Min FN-MPR</th></tr>
<tr><td>E.M.S</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>E.M.U</td><td>E.M.O</td><td>0.79487</td><td>E.T.U</td><td>−0.36325</td></tr>
<tr><td>E.T.S</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>E.T.U</td><td>E.M.O</td><td>0.52055</td><td>E.M.U</td><td>−0.3653</td></tr>
<tr><td>E.T.O</td><td>E.M.O</td><td>0.45455</td><td>E.T.U</td><td>−0.54545</td></tr>
<tr><td>E.T.X</td><td>E.M.O</td><td>1</td><td>E.M.S</td><td>0</td></tr>
<tr><td>C.D.I</td><td>C.L</td><td>0.34091</td><td>E.M.U</td><td>−0.022727</td></tr>
<tr><td>C.D.R</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>C.L</td><td>C.D.I</td><td>0.30769</td><td>E.T.U</td><td>−0.076923</td></tr>
<tr><td>C.X</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>H.E</td><td>E.M.O</td><td>0.2</td><td>T</td><td>−0.4</td></tr>
<tr><td>H.Y</td><td>E.M.S</td><td>0</td><td>C.L</td><td>−1</td></tr>
<tr><td>H.X</td><td>H.E</td><td>0.33333</td><td>S.M.S</td><td>−0.33333</td></tr>
<tr><td>S.M.S</td><td>S.M.C</td><td>0.6</td><td>C.D.I</td><td>−0.2</td></tr>
<tr><td>S.C.X</td><td>E.M.S</td><td>0</td><td>S.M.S</td><td>−0.5</td></tr>
<tr><td>S.R</td><td>E.M.S</td><td>0</td><td>S.M.S</td><td>−0.5</td></tr>
<tr><td>A.W</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>A.M</td><td>E.M.S</td><td>0</td><td>S.M.S</td><td>−1</td></tr>
<tr><td>M.M</td><td>E.M.O</td><td>0.18182</td><td>E.M.U</td><td>−0.18182</td></tr>
<tr><td>M.K</td><td>M.M</td><td>0.6</td><td>E.T.U</td><td>−0.2</td></tr>
<tr><td>N.G.M</td><td>N.R.A</td><td>1</td><td>E.M.S</td><td>0</td></tr>
<tr><td>N.G.X</td><td>E.M.S</td><td>0</td><td>E.M.S</td><td>0</td></tr>
<tr><td>G.N</td><td>G.O</td><td>0.64103</td><td>E.T.U</td><td>−0.051282</td></tr>
<tr><td>G.X</td><td>G.O</td><td>0.75</td><td>E.M.U</td><td>−0.25</td></tr>
<tr><td>T</td><td>E.M.S</td><td>0</td><td>C.D.I</td><td>−0.13333</td></tr>
</table>
</td></tr>
</table>

Table 9. *Circular co-occurrence networks of six different WSIs predicted by three different levels of network training (including augmented HBP layers). Within each WSI, different number of image patches are extracted for label prediction.*

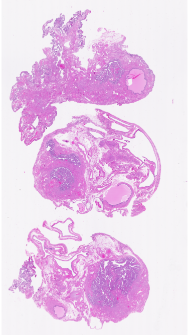
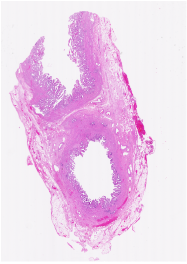
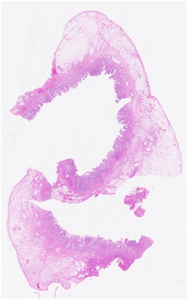
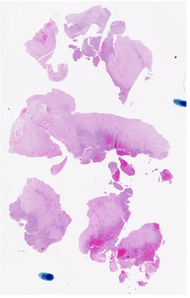
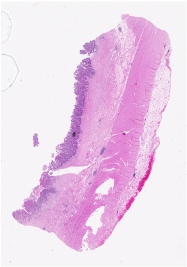
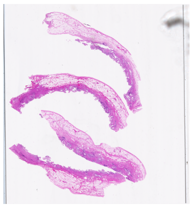
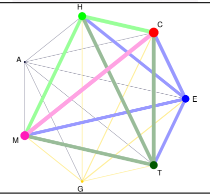
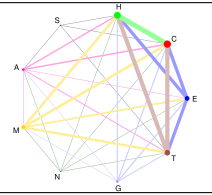
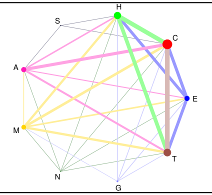
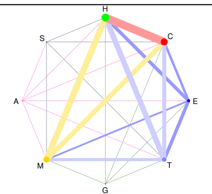
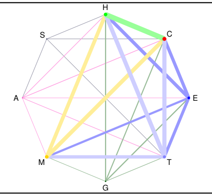
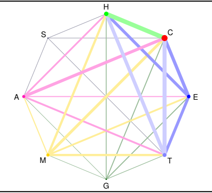
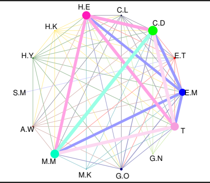
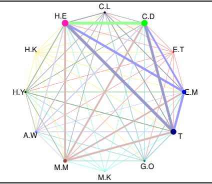
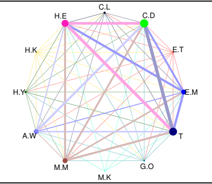
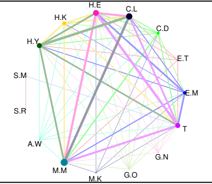
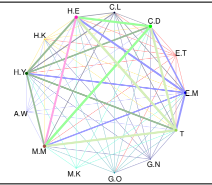
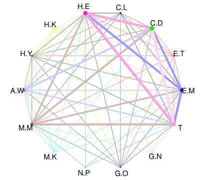
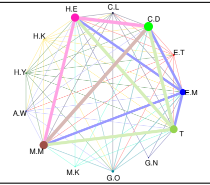
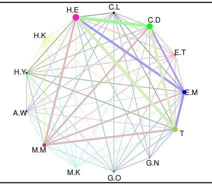
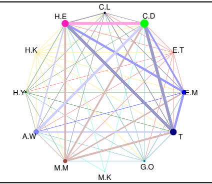
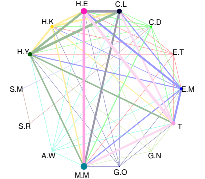
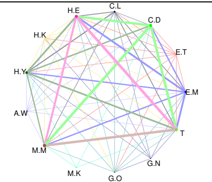
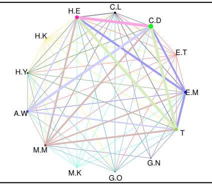
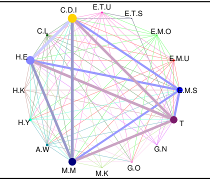
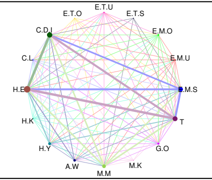
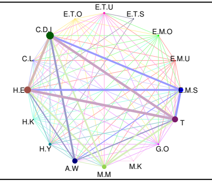
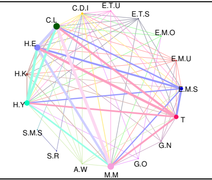
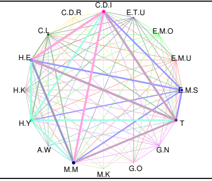
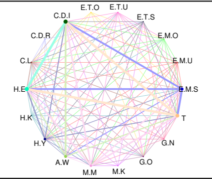
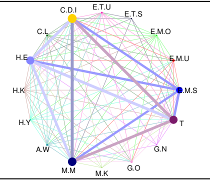
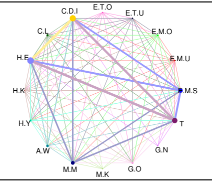
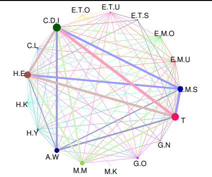
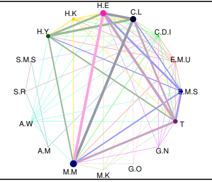
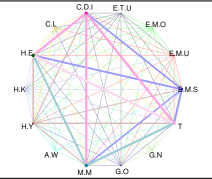
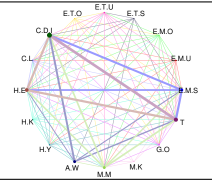
	Slide-1	Slide-2	Slide-3	Slide-4	Slide-5	Slide-6
# patch	4347	1734	4919	3236	2565	4587
WSI-Preview						
Level-1						
Level-2	 H.E C.L H.K C.D H.Y E.T S.M E.M A.W T M.M G.N M.K G.O	 H.E C.L C.D H.K E.T H.Y E.M A.W T M.M M.K G.O	 H.E C.L C.D H.K E.T H.Y E.M A.W T M.M M.K G.O	 H.E H.K C.L H.Y C.D E.T S.M E.M S.R T A.W G.N M.M M.K G.O	 H.E C.L H.K C.D E.T H.Y E.M A.W T M.M G.N M.K G.O	 H.E C.L H.K C.D H.Y E.T A.W E.M M.M T M.K N.P G.N G.O
Level-2+HBP	 H.E C.L C.D H.K E.T H.Y E.M A.W T M.M G.N M.K G.O	 H.E C.L C.D H.K E.T H.Y E.M A.W T M.M G.N M.K G.O	 H.E C.D H.K E.T H.Y E.M A.W T M.M G.O M.K	 H.E C.L H.K C.D H.Y E.T S.M E.M S.R T A.W G.N M.M G.O	 H.E C.L H.K C.D E.T H.Y E.M A.W T M.M G.N G.O	 H.E C.L H.K C.D E.T H.Y E.M A.W T M.M G.N M.K G.O
Level-3	 C.D.I E.T.U E.T.S C.L E.M.O H.E E.M.U H.K E.M.S H.Y T A.W G.N M.M M.K G.O	 E.T.O E.T.U E.T.S C.D.I E.M.O C.L E.M.U H.E E.M.S H.K T H.Y G.O A.W M.K M.M	 E.T.O E.T.U E.T.S C.D.I E.M.O C.L E.M.U H.E E.M.S H.K T H.Y G.O A.W M.K M.M	 C.D.I E.T.U C.L E.T.S E.M.O H.E E.M.U H.K E.M.S H.Y T S.M.S G.N S.R A.W G.O M.M	 C.D.I C.D.R E.T.U C.L E.M.O H.E E.M.U H.K E.M.S H.Y T A.W G.N M.M M.K G.O	 E.T.O E.T.U C.D.I E.T.S C.D.R E.M.O C.L E.M.U H.E E.M.S H.K T H.Y G.N A.W G.O M.M M.K
Level-3+HBP	 C.D.I E.T.U E.T.S C.L E.M.O H.E E.M.U H.K E.M.S H.Y T A.W G.N M.M M.K G.O	 C.D.I E.T.O E.T.U C.L E.M.O H.E E.M.U H.K E.M.S H.Y T A.W G.N M.M M.K G.O	 E.T.O E.T.U C.D.I E.T.S C.L E.M.O E.M.U H.E E.M.S H.K T H.Y G.N A.W G.O M.M M.K	 H.E H.K C.L H.Y C.D.I S.M.S E.M.U S.R E.M.S T A.W A.M G.N M.M M.K G.O	 C.D.I E.T.U C.L E.M.O H.E E.M.U H.K E.M.S H.Y T A.W G.N M.M G.O	 C.D.I E.T.O E.T.U E.T.S E.M.O C.L E.M.U H.E E.M.S H.K T H.Y G.O A.W M.M M.K

Table 10. *Results of applying to the Apriori Association Rule Learning algorithm to the predicted labels of all WSI #1, displaying only the most significant rule for each unique consequent label where such a rule exists.*

Antecedent Itemsets	⇒	Consequent Labels	Confidence
{A.W, M.M, E.M.S}	⇒	C.D.I	1
{H.E, E.M.O}	⇒	C.L	0.5540
{C.L, G.O}	⇒	E.M.O	0.5382
{C.D.I, H.E, A.W, M.M, T}	⇒	E.M.S	0.8811
{C.D.I, C.L, G.O, E.M.S}	⇒	E.M.U	0.8795
{C.D.I, E.M.O}	⇒	E.T.U	0.6776
{E.M.O}	⇒	G.O	1
{H.Y, E.M.S, C.L}	⇒	H.E	1
{E.M.S, C.D.I, H.E}	⇒	M.M	0.9421
{E.M.S}	⇒	T	1

Table 11. *Results of applying to the Apriori Association Rule Learning algorithm to the predicted labels of all WSI #2, displaying only the most significant rule for each unique consequent label where such a rule exists.*

Antecedent Itemsets	⇒	Consequent Labels	Confidence
{A.W, M.M}	⇒	C.D.I	1
{H.K, E.M.S, E.M.O}	⇒	C.L	0.9487
{E.T.U, C.L, H.K}	⇒	E.M.O	0.9429
{E.M.U, H.K,T}	⇒	E.M.S	1
{C.D.I, C.L, H.Y, M.M, G.O, E.M.S}	⇒	E.M.U	0.6061
{C.D.I, G.O}	⇒	E.T.U	0.5994
{E.M.O}	⇒	G.O	1
{H.K}	⇒	H.E	1
{E.M.U, H.K}	⇒	H.Y	1
{T, E.M.U, E.T.U, C.D.I, C.L}	⇒	M.M	0.8478
{E.M.S}	⇒	T	1

Table 12. *Results of applying to the Apriori Association Rule Learning algorithm to the predicted labels of all WSI #3, displaying only the most significant rule for each unique consequent label where such a rule exists.*

Antecedent Itemsets	⇒	Consequent Labels	Confidence
{C.D.I}	⇒	A.W	0.5402
{A.W, T}	⇒	C.D.I	0.9861
{H.K, G.O}	⇒	C.L	0.9485
{H.K, G.O, E.M.S}	⇒	E.M.O	0.8933
{H.E, H.Y, A.W}	⇒	E.M.S	0.9811
{C.D.I, M.M, G.O}	⇒	E.M.U	0.5000
{C.D.I, H.E, E.M.O}	⇒	E.T.U	0.6000
{E.M.U, E.M.O}	⇒	G.O	1
{H.K}	⇒	H.E	1
{H.Y, M.M, T, E.M.O}	⇒	H.K	0.5926
{E.M.S, C.D.I, H.K}	⇒	H.Y	0.9844
{E.M.S, C.D.I, H.K}	⇒	M.M	0.9219
{E.M.S}	⇒	T	1

Table 13. *Results of applying to the Apriori Association Rule Learning algorithm to the predicted labels of all WSI # 4, displaying only the most significant rule for each unique consequent label where such a rule exists.*

Antecedent Itemsets	⇒	Consequent Labels	Confidence
{H.K, H.Y, M.M}	⇒	C.L	0.9303
{C.D.I, H.E, H.K, M.M, T}	⇒	E.M.S	0.8596
{H.K, E.M.S}	⇒	H.E	0.9932
{H.Y, M.M, C.L, H.E}	⇒	H.K	0.7462
{E.M.S, C.L, H.K}	⇒	H.Y	0.9790
{E.M.S, C.D.I}	⇒	M.M	0.8780
{E.M.S}	⇒	T	1

Table 14. *Results of applying to the Apriori Association Rule Learning algorithm to the predicted labels of all WSI #5, displaying only the most significant rule for each unique consequent label where such a rule exists.*

Antecedent Itemsets	⇒	Consequent Labels	Confidence
{H.Y, A.W, T}	⇒	C.D.I	0.9796
{E.M.S, E.M.O}	⇒	C.L	1
{A.W, M.M}	⇒	E.M.S	0.9643
{C.D.I, C.L, G.O}	⇒	E.M.U	0.9231
{H.E, H.K, M.M, E.M.U}	⇒	E.T.U	0.6923
{E.M.O}	⇒	G.O	1
{E.M.S, E.M.U}	⇒	H.E	1
{M.M, E.M.U, E.T.U}	⇒	H.K	0.9231
{E.M.S, H.K}	⇒	H.Y	1
{E.M.S, C.D.I}	⇒	M.M	0.8223
{E.M.S}	⇒	T	1

Table 15. *Results of applying to the Apriori Association Rule Learning algorithm to the predicted labels of all WSI #6, displaying only the most significant rule for each unique consequent label where such a rule exists.*

Antecedent Itemsets	⇒	Consequent Labels	Confidence
{C.D.I }	⇒	A.W	0.5163
{H.Y,A.W,E.M.S }	⇒	C.D.I	1
{H.Y,E.M.O,E.T.U }	⇒	C.L	0.9655
{C.L,H.Y,G.O }	⇒	E.M.O	0.5437
{H.E,A.W,M.M,T }	⇒	E.M.S	0.9617
{H.Y,M.M,G.O,E.M.S }	⇒	E.M.U	0.8571
{C.D.I,G.O }	⇒	E.T.U	0.5288
{E.M.U,E.M.O }	⇒	G.O	1
{H.Y,E.M.S,E.M.U }	⇒	H.E	1
{E.M.O,E.T.U,C.L,H.E}	⇒	H.Y	0.5625
{T,E.M.U,E.T.U,C.D.I}	⇒	M.M	0.8727
{E.M.S,E.M.U }	⇒	T	1