Online Supplemental Data

Description of Ensemble Learning Methods

Majority Voting

In Majority Voting each class prediction of the 70 CNNs is taken into account and the class receiving the most votes is determined as the final class.

Decision Tree

In Decision Tree a tree like model is created where each branch represents a decision. At each node, some attribute is tested and the outcome determines which branch to follow. For instance, in the case of this research, at a certain node it might be tested if the prediction score for 'No ICH' of CNN number 65 is bigger or smaller than 0.73. If this is the case, we continue along the left side of the branch to the next node with the next test, otherwise we continue along the right. This is repeated until no new node is encountered, which indicates that we encountered an outcome; in the case of this research, a class label for either 'ICH' or 'No ICH'.

Support Vector Machine

The Support Vector Machine (SVM) tries to separate the 'ICH' and 'No ICH' class using boundaries where the margin of separation between the classes is largest. In the case of a two-dimensional problem, it might fit a straight boundary or a polynomial to best separate the cases of the several classes. This boundary is known as the hyperplane, but instead of two dimensional, the space is mostly multi-dimensional.

Multi Input – Deep Neural Network

The last ensemble learning method we tested was a Multi Input Deep Neural Network. Unlike the ensemble learning methods described earlier, Multi Input Deep Neural Network requires multiple inputs, namely, a CT image and the probability scores for the 'No ICH' class from each of the 70 CNNs. The method can be visualized as a multi-branch neural network at the input, one branch with the image as input and the other with the probability score as input. At some point before the final prediction is given, the branches merge into a single unit. The Multi Input Deep Neural Network provides a probability score for 'No ICH' and 'ICH' classification. The Multi Input Deep Neural Network was trained for 25 epochs, with a batch size of 32, and using stochastic gradient descent with momentum as optimizer and a learning rate of 0.001.

Due to the random weight initialization when training a Multi Input Deep Neural Network, the network was trained 5 times and the average accuracy was computed to provide a more robust estimate of the accuracy.

Supplemental Figures

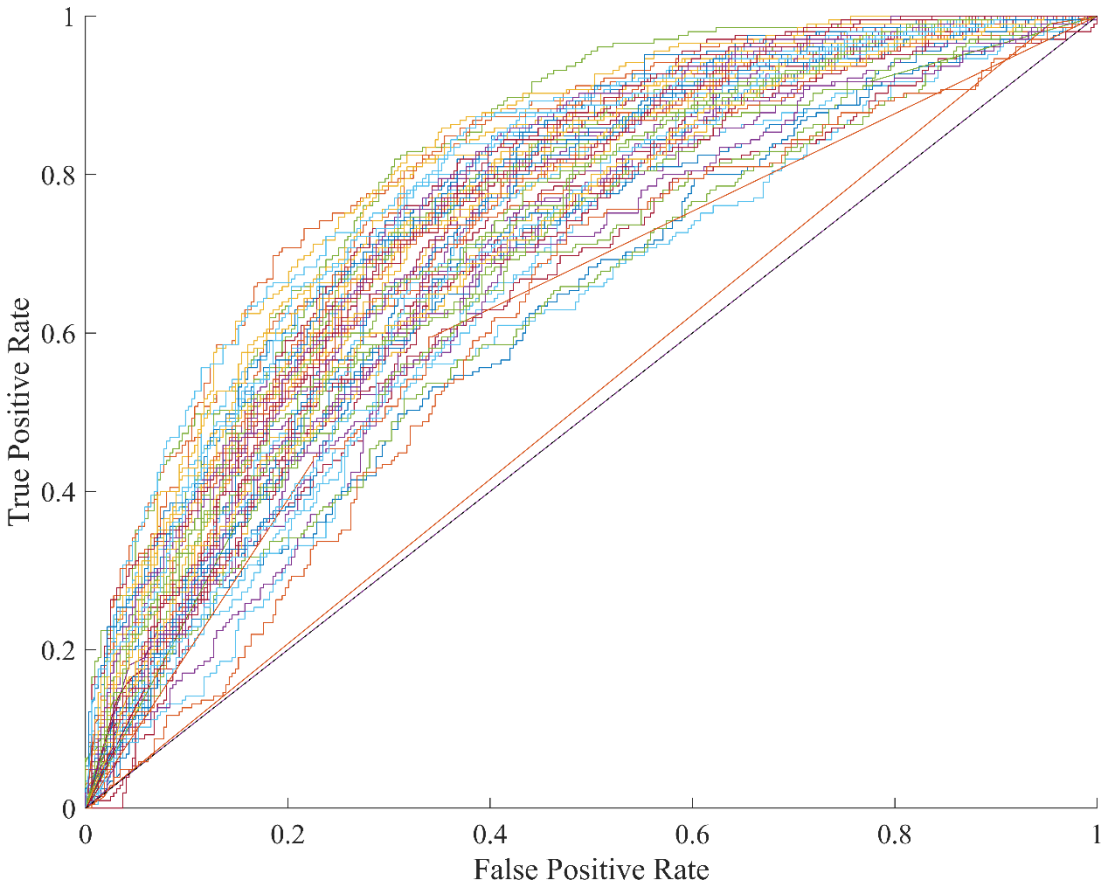

Figure S1: Receiver Operating Characteristic curves of all 70 individual CNNs.