Supplementary information

Table 1 Comparison of CNN architectures

Architecture	Learning rate	Optimizing function	Tile-level accuracy	Time per epoch
ResNet18	4e-3	Adam	74.5 %	5 min
ViT	1e-4	Adam	75.7 %	5 min
ConvNeXt	**4e-3**	**Adam**	**81.0** %	**20 min**
MobilNetV3	**1e-3**	**Adam**	**79.3** %	**3 min**

Four architectures for comparison were chosen based on the literature. Training data included 10.000 image patches, and testing data 5.000 image patches. The two best-performing architectures at the tile level were ConvNeXt and MobileNetV3, the latter of which was chosen based on the lighter computational load.

Table 2 Hyperparameters of the individual models

Model	Learning rate	Optimizing function
TUM5x	1e-3	Adam
TUM20x	5e-2	SGD
OTHER5x	1e-2	Adam

The hyperparameters were chosen with 5-fold cross-validation on the validation data.

Table 3 Hyperparameters of the XGBoost layer

Model	Estimators	Depth	Learning rate	Objective	Scale	Subsample
XGBoost-5x	120	3	2e-2	'binary:logistic'	4.5	0.1
XGBoost-20x	65	3	2e-2	'binary:logistic'	4.5	0.1

Hyperparameters of the XGBoost models of the two branches (5x and 20x). Depth = maximum depth of the tree, objective = the objective function, scale = scaling factor in the case of imbalanced data, subsample = subsample ratio of the training instances. All other parameters were set to their default values.