1.1. *Key Transformative features of all CNN versions employed to classify TCGA BRCA patients across different patient attributes*

CNN	Key Features
GoogLeNet (2014) with modified stem	Native *22*-layer model, with modified initial layers (stem) to accept tile size of *512* x *512* by gridding WSI. Modification made to additional dropout or terminal layer to prevent over fitting.
Inception (v3 2015) with no pre-training	Newer version of GoogLeNet, with no additional modifications. Tiles of size *299* x *299* extracted by gridding WSI. No pre-training to learn image features.
Inception (v3 2015) with pre-training	Network's weights pre-trained with *1.28* million images sourced from ImageNet[1], and trained on fully connected network . Input tiles are *299* x *299* in size.
Inception (v3 2015) with pre-training and boosted data	Along with pre-training, tiling is performed by a sliding window approach. We boost the input data by including 4 (rotated) versions of each tile.
Inception (v3 2015) with pre-training and only epithelial regions	Along with pre-training, only tiles from epithelial regions of the WSI are considered given tumors proliferate in the epithelium breast tissue.

References

1. Russakovsky O, Deng J, Su H, et al. ImageNet Large Scale Visual Recognition Challenge. *Int J Comput Vis*. 2015;115(3):211-252. doi:10.1007/s11263-015-0816-y.