Interpretable Deep Learning Uncovers Cellular Properties in Label-Free Live Cell Images that are Predictive of Highly Metastatic Melanoma

Deep learning has emerged as a powerful technique to identify hidden patterns in complex cell imaging data, but is criticized for the lack of insights it provides on the machine's prediction. Here, we demonstrate that a generative adversarial neural network captures subtle details of cell appearance that allow classification of melanoma metastatic efficiency of patient-derived xenograft models that reflect clinical outcome. We used the network to generate "in-silico" cell images that amplified the cellular features critical for the classification. These images unveiled pseudopodial extensions and increased light scattering as functional hallmarks of metastatic cells. We validated this interpretation using live cells spontaneously transitioning between states indicative of low and high metastatic efficiency. Together, this data demonstrates how the application of Artificial Intelligence can support the identification of processes that are essential for the execution of complex integrated cell functions but are too subtle to be identified by a human expert.