We interpret the image-based asset pricing factors produced with convolutional neural network (CNN) by proposing a methodology through which the aggregate pattern learned by vision models can be visualized in a heat map. We show and quantify the CNN captured price and volume paths, including the level, the slope, the curvature, and, most importantly, the last day change in the difference between close price and high-low average. Four simple geometric measures can explain 31% (24%) of variation in predicted probability of positive weekly (monthly) returns, and 73% (48%) of variation in weekly (monthly) CNN returns. We further hypothesize and confirm that the machine can identify time series persistence and reversal patterns. Separating them enables the formation of long-short portfolios with better performances than does the existing CNN factor.