Humans can use local cues to distinguish image edges caused by a depth change from other types of edges (Vilankar et al., 2014). But which local cues? Here we use the SYNS database (Adams et al., 2016) to automatically label edges in images of natural scenes as depth or non-depth. We use this ground truth to identify the cues used by human observers and convolutional neural networks (CNNs) for edge classification. Eight observers viewed square image patches, each centered on an image edge, ranging in width from 0.6 to 2.4 degrees (8 to 32 pixels). Human judgments (depth/non-depth) were compared to responses of a CNN trained on the same task. Human performance improved with patch size (65%-74% correct) but remained well below CNN accuracy (82-86% correct). Agreement between humans and the CNN was above chance but lower than human-human agreement. Decision Variable Correlation (Sebastian & Geisler, in press) was used to evaluate the relationships between depth responses and local edge cues. Humans seem to rely primarily on contrast cues, specifically luminance contrast and red-green contrast across the edge. The CNN also relies on luminance contrast, but unlike humans it seems to make use of mean luminance and red-green intensity as well. These local luminance and color features provide valid cues for depth edge discrimination in natural scenes.