Pixel-wise losses, i.e., cross-entropy or L2, have been widely used in structured prediction tasks as a spatial extension of generic image classification or regress