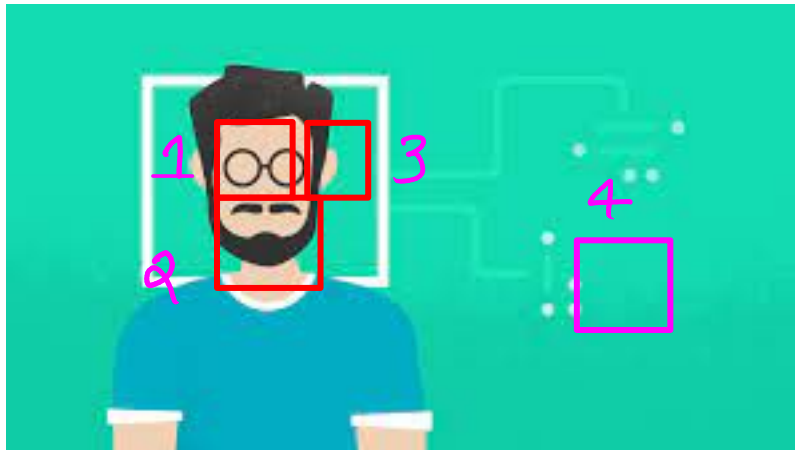


The convolution operation in CNN's is computed using pixels values in a neighborhood implying the presence of local features. Do you think this is a valid assumption? Explain your answer. Be precise. (10 points)



Yes, It is a valid assumption.

The pixels which are nearby when considered together helps in identifying features like ears, nose, vertical lines, horizontal lines etc. But pixels which are far from each other do not help in identifying any features as they do not have any relation with each other. Hence, it is valid to assume the presence of local features in images.

When we look at the above image in the marked rectangles, we can see that the pixels which are in regions are correlated. The pixels in region 1 helps in identifying feature, that it is eyes, pixels in region 2 helps in identifying ears etc. But if we look at region 1 and region 4, we cannot see any specific correlation between the pixels as they do not help in identifying a feature.