

LM in Artificial Intelligence
Image Processing and Computer Vision
05/02/2024

- 1) (A) Describe how SIFT descriptors are computed and (B) how they can be matched between a target view and a reference (or model) view. In this respect, (C) highlight how potentially wrong matches may be filtered out.
- 2) Describe grouped convolutions with G groups, discuss how many kernels of which shape they use to realize a $C_{out} \times C_{in} \times 3 \times 3$ convolution and compare them to standard convolutions in terms of number of parameters and FLOPS, motivating your answer. Then, discuss how grouped convolutions are used to realize depthwise separable convolutions. Finally, draw the inverted residual block and the squeeze and excitation resnet block, highlighting if, where and why they use grouped convolutions.