AIP Assignment 1

Badrinath Singhal

January 2024

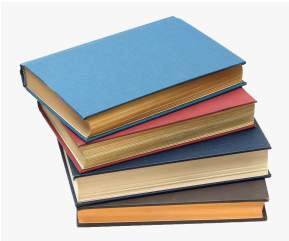
Question 1

Since first three steps of SIFT and PCA-SIFT are same, implmentation of SIFT is done as usual. We calculated Gaussian pyramids across different scales, then we calculated difference of gaussians (DoG) in respective pyramids. For each scale we calculate the extremum points and mark that as a keypoints while ignoring the process of thresholding using some metric. We use that keypoints and calculated gradients on 41×41 patch and saved the horizontal and vertical gradients as a vector. Having all the vector across all the non overlapping patches we get series of vectors where we apply PCA to reduce dimension appropriately. And those reduced dimensions representation represent our features of the image (for those respective keypoints).

Different manipulation of images are tried to see if they keypoints detected are similar in any way, images were manipulated by rotating, blurring and scaling. Two parent images are used which can be seen in fig 1

Result of SIFT algorithm implementation is shown in fig 2 and in fig 3

(a) Image 1



(b) Image 2

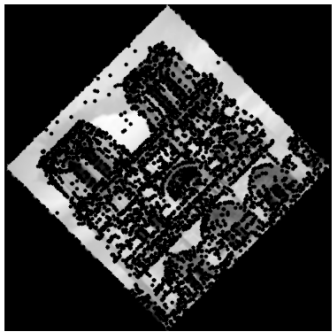
Figure 1: Parent images used in PCA-SIFT

We can see that number of keypoints are changed when we manipulate images, but among the key points detected on same image and its manipulated version, most of the keypoints are in the same location as it was in the original one.

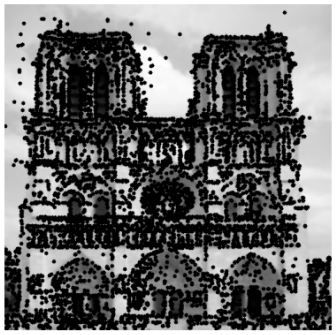
When we manipulate the image on the basis of scale and blur, the number of keypoints doesn't change much, one of the reason for that is because the keypoints detected in the original image is invariant to scale and gaussian blur (as we already process image through those operation across different scales).

When we manipulate the image by rotating it, the keypoint reduces drastically compared to the original one. Although the keypoints in rotated image detected are similar to the keypoints detected in original one (which is surprising).

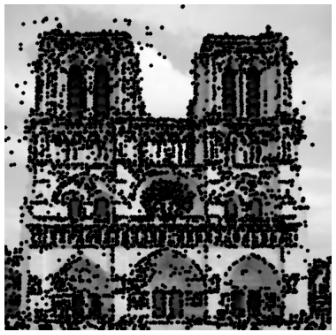
(a) No. of keypoint: 9384



(b) No. of keypoint: 5250



(c) No. of keypoint: 8590



(d) No. of keypoint: 8231

Figure 2: Key point on image and its multiple manipulation for 1st image

One interesting observation is that for image 2, keypoint detected in gaussian blur image is much less than what was expected. One reason that I can guess is that the parameter used for gaussian blur (sigma 5 with kernel size 7) was different than that of the sigma used in sift implementation

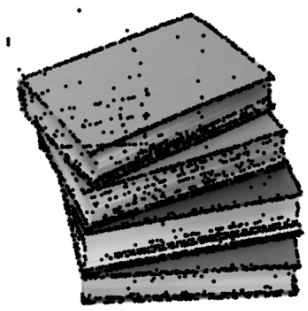
Question 2

A custom CNN was built for classification purpose on CIFAR-10 dataset. The purpose is to understand the effect of different activation functions on the network and its learning. The network contains 2 conv layers and 2 fully connected layers. We were suppose to compare the performance with sigmoid and relu activation functions.

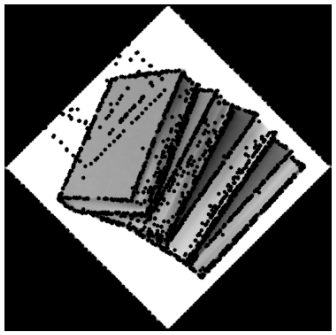
Image dimensions were 32×32, cross entropy loss was used with Adam optimizer for training. Models were trained for 5 epochs each to get the indication about the effect of activation rather then getting best numbers by training and fine tuning. The loss graph for training is shown in fig 4 whereas the performances on validation and test set can be seen in table 1.

Activation Function	validation set	test set
Sigmoid	42.66%	42.31%
ReLU	61.39%	57.45%

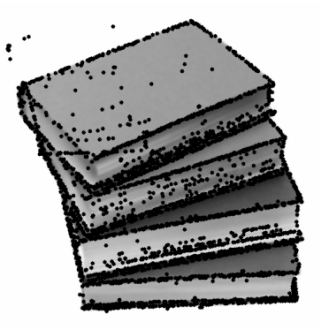
Table 1: Performance on validation and test set



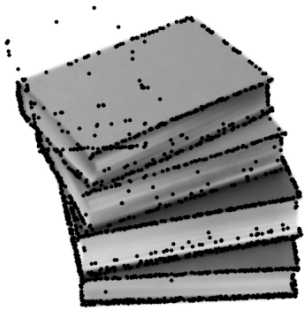
(a) No. of keypoint: 3515



(b) No. of keypoint: 2559



(c) No. of keypoint: 3160

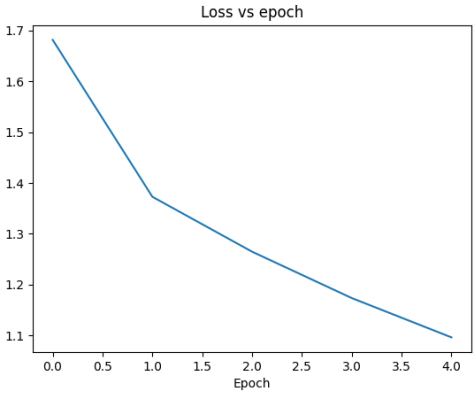


(d) No. of keypoint: 1676

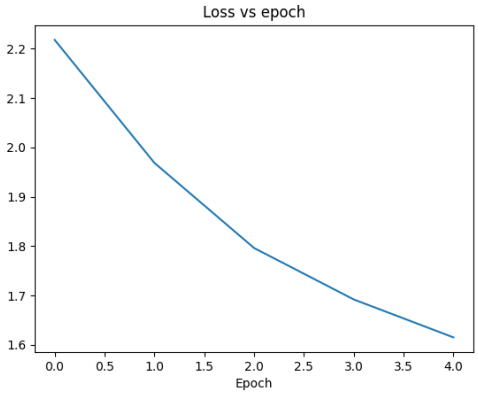
Figure 3: Key point on image and its multiple manipulation for 2nd image

We can see significant uptick in performance when we switched activation from sigmoid to relu in table 1. The loss graph can be refered to see that lowest loss value in sigmoid during training for 5th epoch is the starting loss value when we used relu and decresing when we trained it for 5 epochs. Even if we have trained the model for 1 epoch in relu activation we would have received better performance metric than sigmoid so using 5 epochs outperforms in every way.

Since the dataset provided were in pickle format, it took a while to understand the format and how to extract information from it. Online resources were used such as stack overflow, LLMs etc to get indication about resolving the issues.



(a) Loss with relu activation



(b) Loss with sigmoid activation

Figure 4: Loss for different activation