Learning Local Image Descriptors with Deep Siamese and Triplet Convolutional Networks by Minimizing Global Loss Functions

Supplemental Material

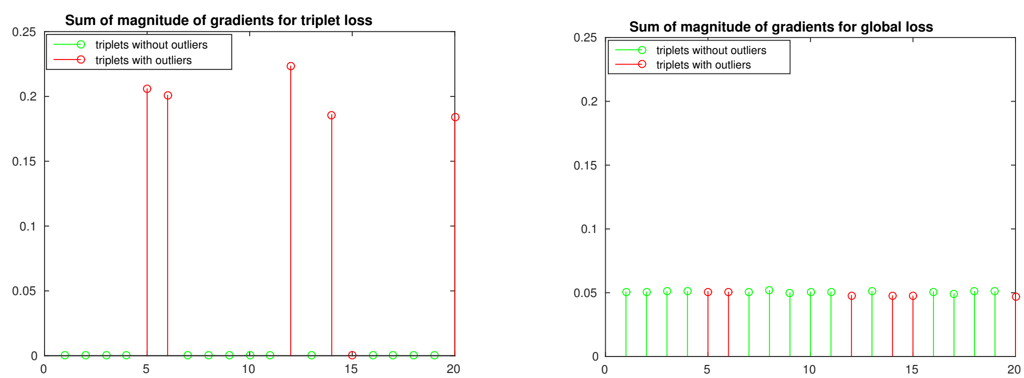
Vijay Kumar B G, Gustavo Carneiro, Ian Reid

The University of Adelaide and Australian Centre for Robotic Vision

{vijay.kumar,gustavo.carneiro,ian.reid}@adelaide.edu.au

1. Effect of Outliers on the Gradient Magnitude

In this section, we use the toy problem presented in Section 4 to demonstrate how the gradient magnitude (used for weight update during training) is affected by outliers for: (a) the triplet loss $J_1^t(.)$ in (5), (b) the global loss $J_1^g(.)$ in (6). In this study, we form a mini batch of 20 triplets, where six of them contain an outlier and the remaining 14 triplets do not contain any outlier. We plot the gradient magnitudes of the loss produced by the 20 triplets after 3 training epochs of triplet loss $J_1^t(.)$ of (5) (*i.e.*, $\|\partial J_1^t/\partial f(\mathbf{x}_i)\|$ + $\|\partial J_1^t/\partial f(\mathbf{x}_i^+)\|$ + $\|\partial J_1^t/\partial f(\mathbf{x}_i^-)\|$) in Fig. 1-(a); and the global loss $J_1^g(.)$ of (6) (*i.e.*, $\|\partial J_1^g/\partial f(\mathbf{x}_i)\|$ + $\|\partial J_1^g/\partial f(\mathbf{x}_i^+)\|$ + $\|\partial J_1^g/\partial f(\mathbf{x}_i^-)\|$) in Fig. 1-(b). Please note that the plots show the normalised magnitude of the gradients (*i.e.*, each triplet gradient magnitude is normalised by the sum of the 20 triplets). The red and green stems indicate the gradient magnitudes for triplets with and without outliers, respectively.



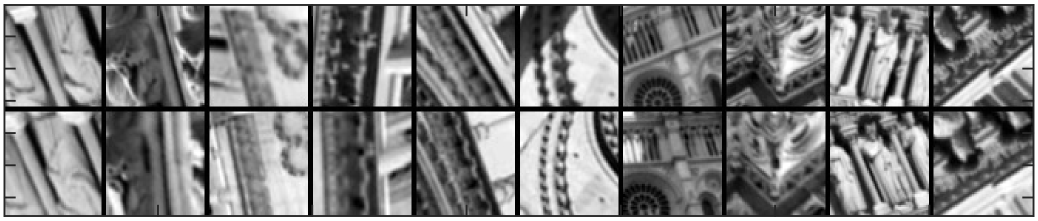
a) Triplet loss b) Global loss

Figure 1. Gradient magnitude after 3 training epochs for a) the triplet loss, b) the global loss.

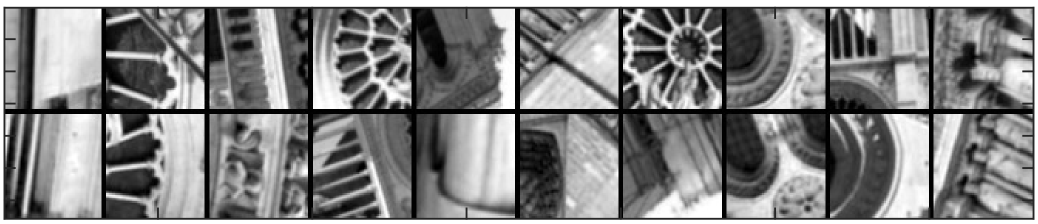
As discussed in Section 4, the gradients $\partial J_1^t/\partial f(\mathbf{x}_i), \partial J_1^t/\partial f(\mathbf{x}_i^+), \partial J_1^t/\partial f(\mathbf{x}_i^-)$ of the triplet loss in (5) depends only on the i^{th} triplet of the training set. After just a few training epochs, most of the triplets without outliers satisfy the condition in (5) and results in zero magnitude gradient, as indicated by the green stems in Fig. 1-(a); whereas the triplets containing outliers produce high magnitude gradients, as shown by the red stems in Fig. 1-(a). Since all non-zero magnitude gradients in Fig. 1-(a) are generated by triplets with outliers (spurious gradients), the weights of the network are affected only by these outliers after a small number of training epochs, as shown in (Fig 3-(b) of the paper. In the case of the global loss, the gradient $\partial J_1^g/\partial f(\mathbf{x}_i), \partial J_1^g/\partial f(\mathbf{x}_i^+), \partial J_1^g/\partial f(\mathbf{x}_i^-)$ is parameterised by μ^+ and μ^-, which means that it depends on the global statistics of the training set. This makes the global loss less sensitive to outliers as shown in Fig. 1-(b), where only $\approx 30\%$ (as opposed to 100% in the case of triplet loss) of the gradient magnitude is generated by triplets containing outliers, indicated by red stems in Fig. 1-(b). Thus the global loss function is more robust to outliers than the triplet loss.

2. UBC Patch Dataset

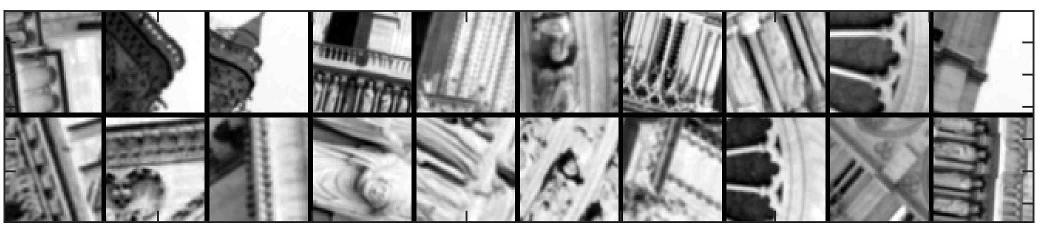
In this section, we show a few top ranked matching and non-matching pairs from different datasets where the matching score was obtained using CS SNet GLoss network.



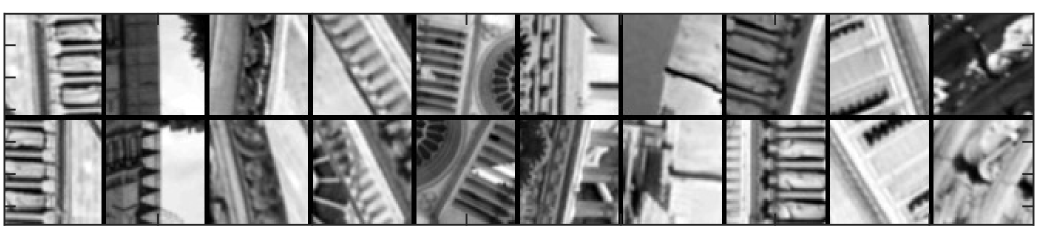
True positives



False positives

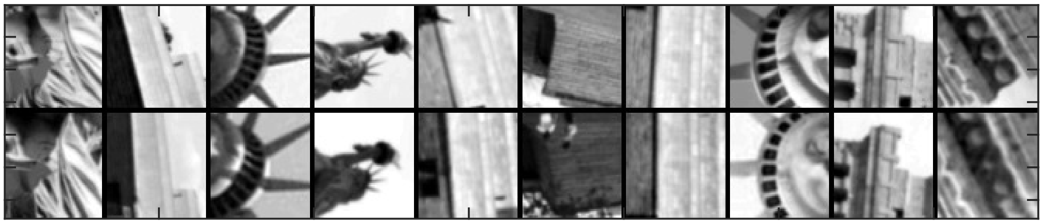


True Negatives



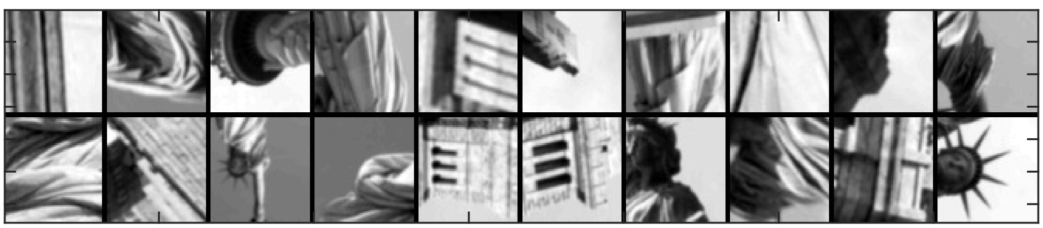
False Negatives

Figure 2. Sample pairs from Notredame dataset with Yosemite dataset used for training



True positives

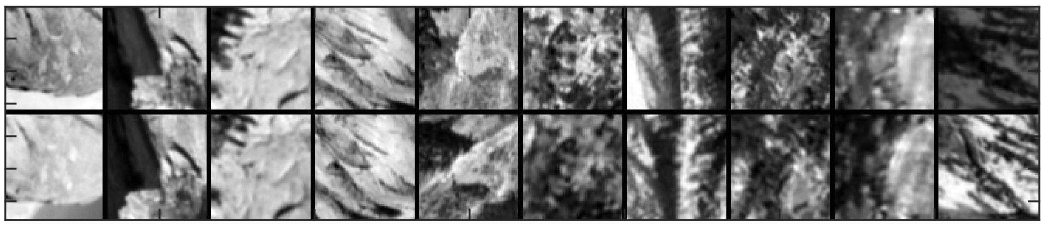
False positives



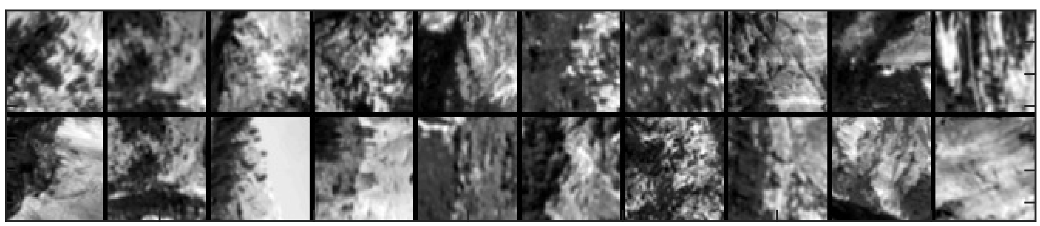
True Negatives

False Negatives

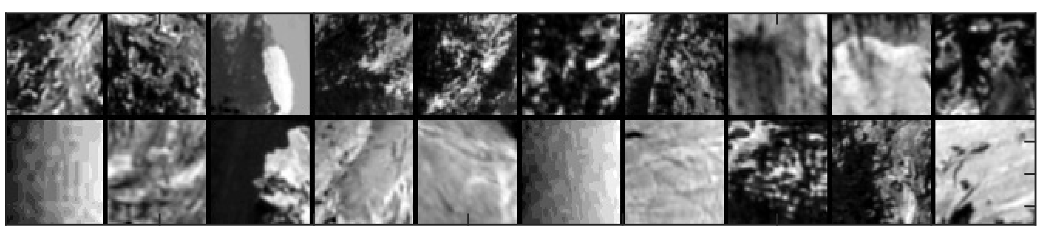
Figure 3. Sample pairs from Liberty dataset with Yosemite dataset used for training



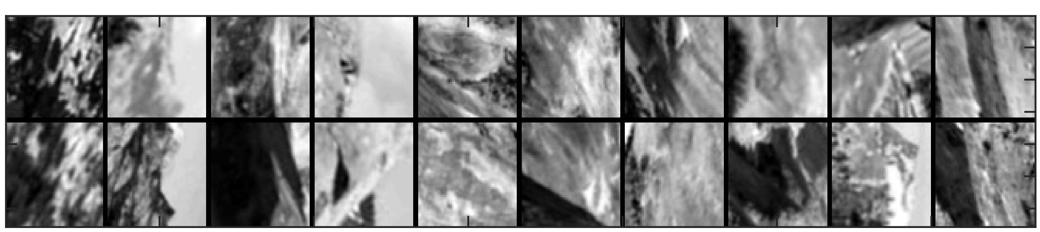
True positives



False positives



True Negatives



False Negatives

Figure 4. Sample pairs from Yosemite dataset with Liberty dataset used for training