

Project proposal

Definition.

We need to develop a software that, given an image, it retrieves from our image dataset the most relevant images. We will develop this using convolutional neural networks, more specifically, ResNet50 along with a neural network that hashes ResNet's outputs. We will do this with our dataset that has INRIA Holiday Dataset (1) and the Oxford Buildings Dataset as a basis.

Preliminary plan of work.

We will train a siamese network that binarizes the input, taking into account the similarity with respect the other photos. For that we will use the TensorFlow library, as well as numpy. Our loss function will be Triplet Loss (2). In order to do that, we need to modify the dataset into triplets. Each sample will consist of three images: a reference one, a positive (similar) one, and a negative (different) one). Gonzalo Gandul will implement the part of the ResNet50, Sergio Escalante will implement the part of the siamese network and Corentin Anaclet will develop the function that allow us do the backpropagation. All of us will modify the dataset, expand it; and we will all test and analyze the results of the app too.

Evaluation Plan.

We are going to use INRIA Holidays dataset and the Oxford Buildings Dataset, because is use free and also specifically created to evaluate image search and retrieval algorithms. As evaluation metrics, we'll use the mean Average Precision (mAP) as well as a human-based metric, in which a person will give a score from 1 to 10 regarding how similar are the retrieved images. One reference project could be Supervised Learning of Semantics-Preserving Hash via Deep Convolutional Neural Networks (3), which has results in the Oxford Buildings Dataset.

References.

- (1) <https://thoth.inrialpes.fr/~jegou/data.php.html>
- (2) <https://qdrant.tech/articles/triplet-loss/>
- (3) Yang, H. F., Lin, K., & Chen, C. S. (2017). Supervised learning of semantics-preserving hash via deep convolutional neural networks. *IEEE transactions on pattern analysis and machine intelligence*, 40(2), 437-451.