Project 3: Image Geolocalization

**TA: Carlo Masone, Gabriele Berton, Gabriele Trivigno**

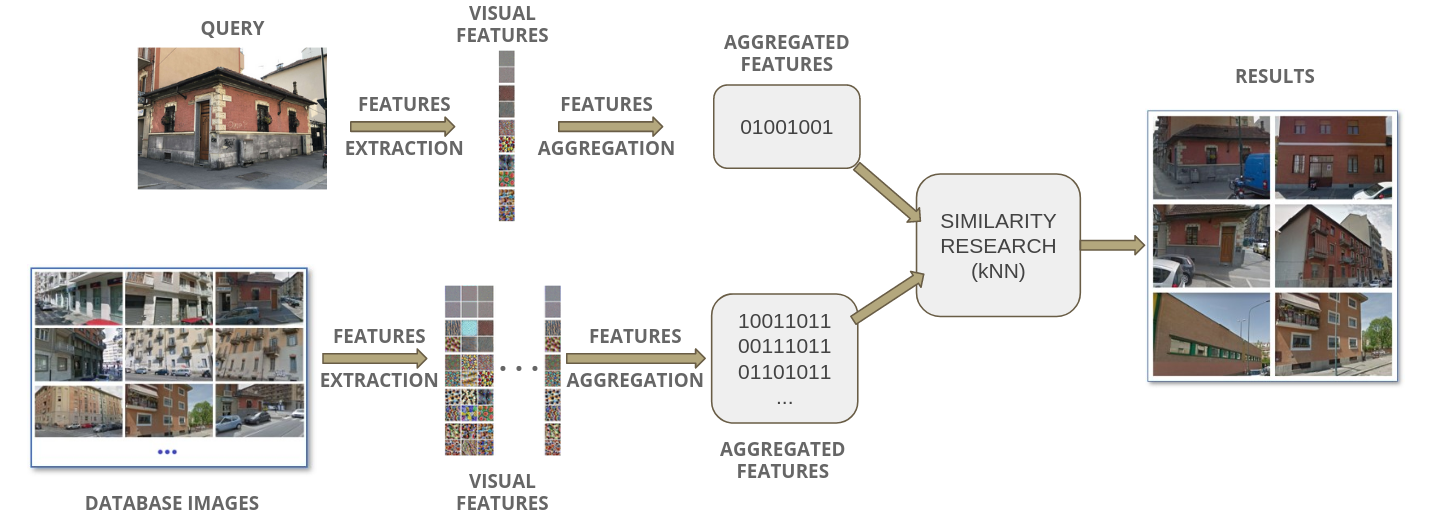
# **OVERVIEW**

Image geo-localization is a technique for finding the geographical location where a photo was taken. This is usually treated as an image retrieval problem. Given an unseen photo (the query), it is matched against a database of photos with known locations (geo-tagged images). The top matches, along with their geo-tags (usually GPS coordinates), provide a hypothesis for the location of the query.



The task is approached through image retrieval: the query is transformed into a set of features that describe its appearance, and a neural network is trained to recognize the location of photos with similar visual features. During inference, the photo is compared to the database in features space to find the most likely location.

The retrieval is done using a k-Nearest Neighbor search in a learned features space that represents the visual similarity of different locations. Each photo is processed by a network that extracts features and aggregates them into a global descriptor of the image. The training process is performed by dividing the geographical area into squares and classifying the location of the images based on their heading (spatial orientation).



The neural network to extract the features can be trained in two ways:

A) **Image geolocalization through retrieval**, by forcing the features of the same place to be near to each other, using a metric learning approach and losses like the triplet loss;

B) **Classification for retrieval in image geolocalization**: this builds on the idea that a neural network that is a good classifier for places produces informative features that can be used for image retrieval.

After choosing the approach you prefer, you will start running experiments with an existing codebase, to learn about the task of image geo-localization and how the system works. Then, you'll focus on improving one or more aspects of the system, such as robustness to changes in lighting, different perspectives, or occlusions. We'll provide the code for the baseline so you can focus on understanding the code and tackling real research problems. If you make significant progress, you may even have the opportunity to publish your findings.