Master in Computer Vision Barcelona

M5 Project: Cross-modal Retrieval

Week 4

Image retrieval

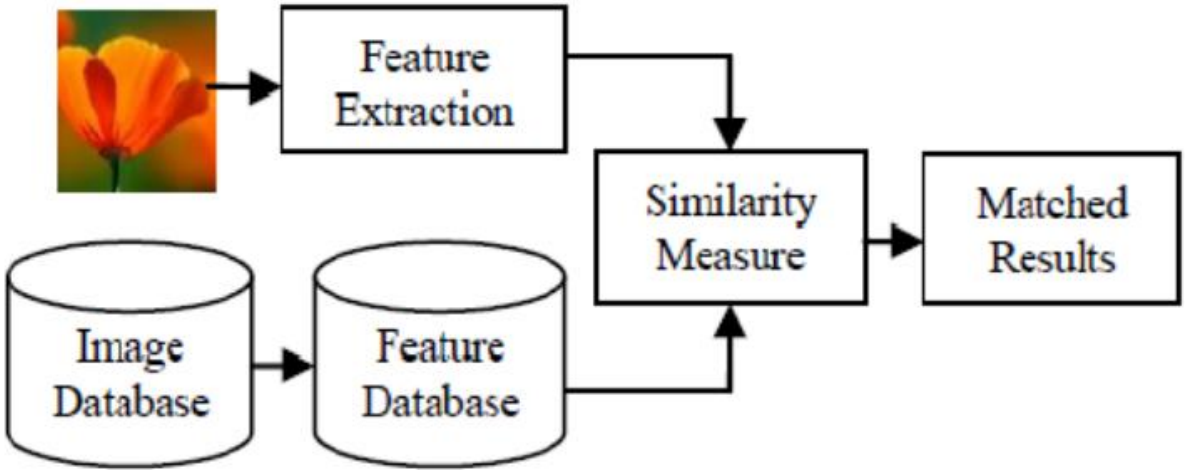
Rubèn Pérez Tito
rperez@cvc.uab.es

Ernest Valveny
ernest@cvc.uab.es

M5 – Image Retrieval

Application approach

- Extract features from database images (train set).
- Extract features of the query image (val/test set).
- Retrieve the most similar images from the database.

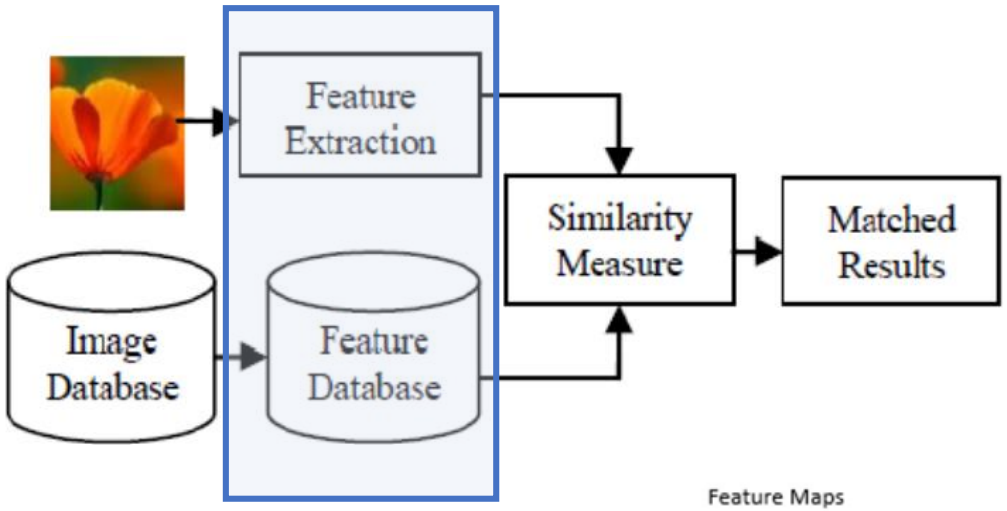


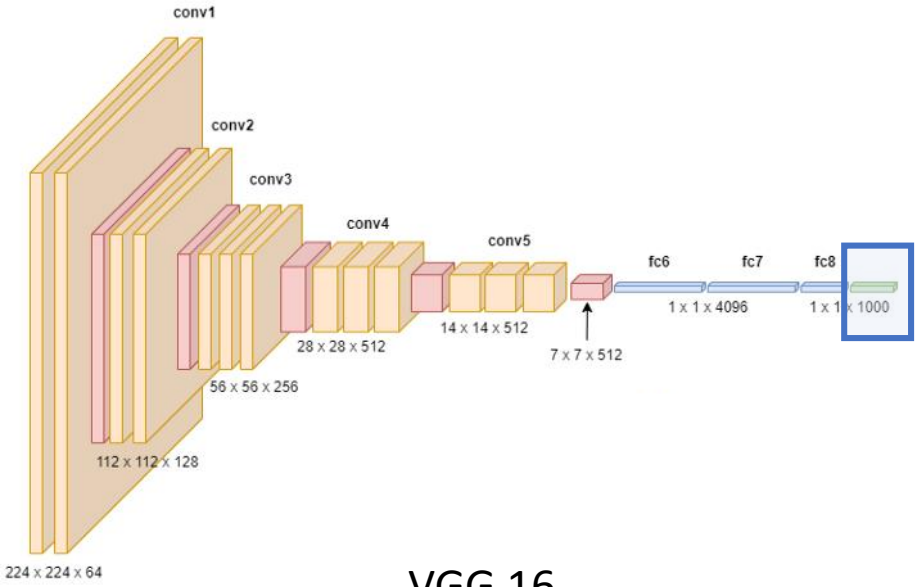
Notice that image retrieval is not a training methodology, but an application!

M5 – Image Retrieval

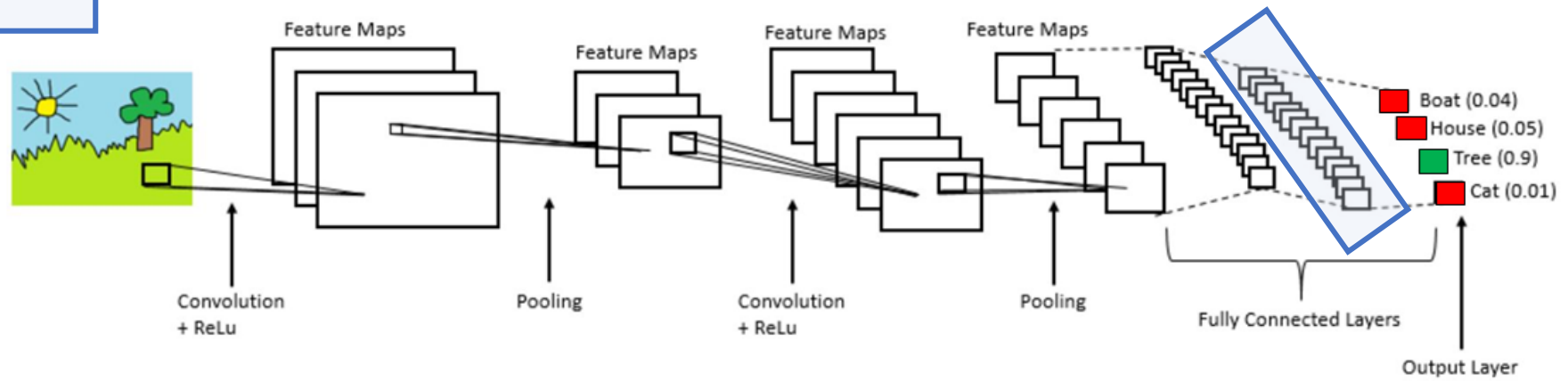
Training strategies for image retrieval

- Classification





VGG 16

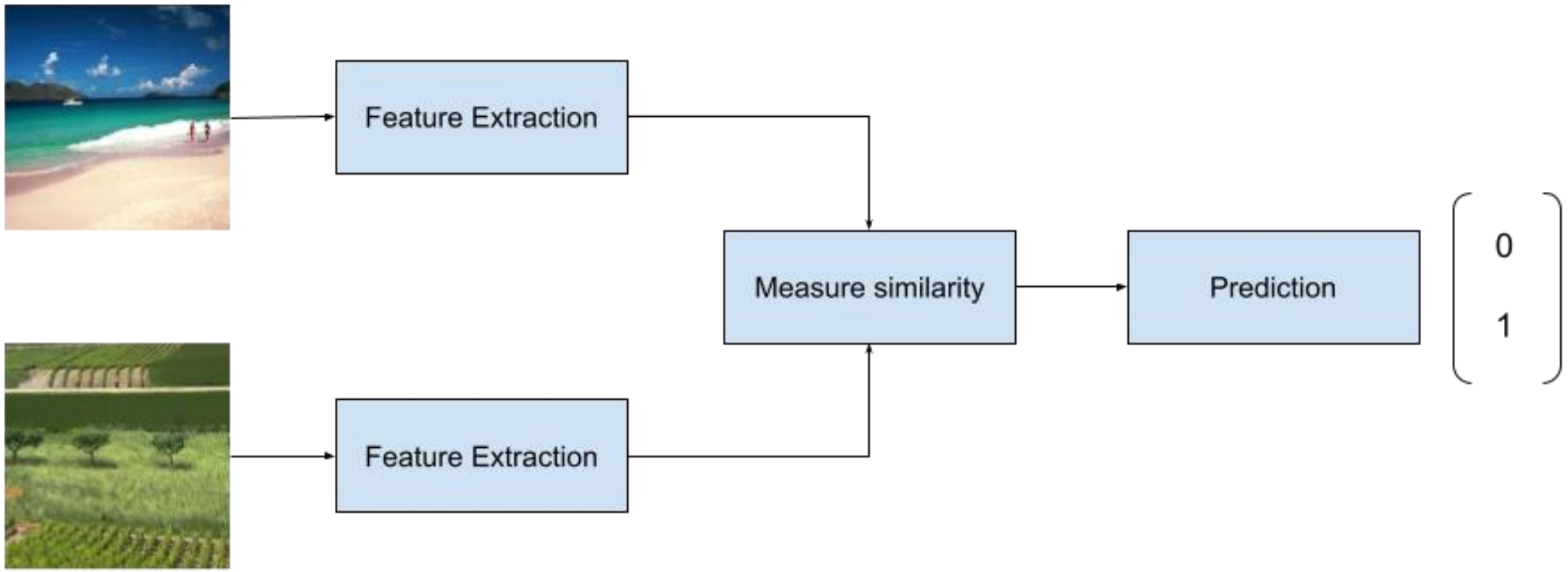


We can train your network to classify. It will implicitly learn an image representation that is representative to perform retrieval.

M5 – Image Retrieval

Training strategies for image retrieval

- Classification
- Metric learning:
  - Siamese networks

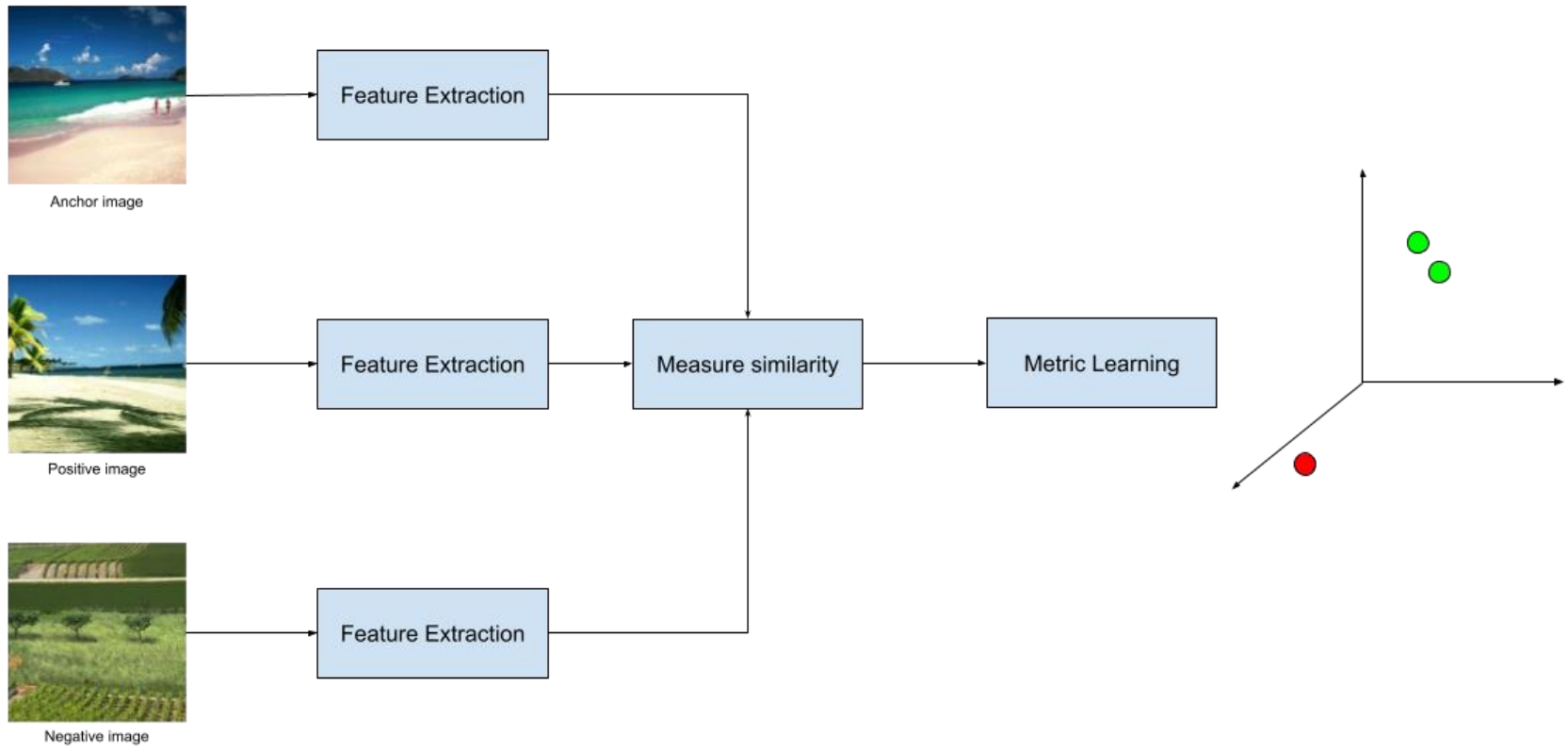


We can explicitly learn a representation that facilitates the retrieval of the images.

M5 – Image Retrieval

Training strategies for image retrieval

- Classification
- Metric learning:
  - Siamese networks
  - Triplet Networks

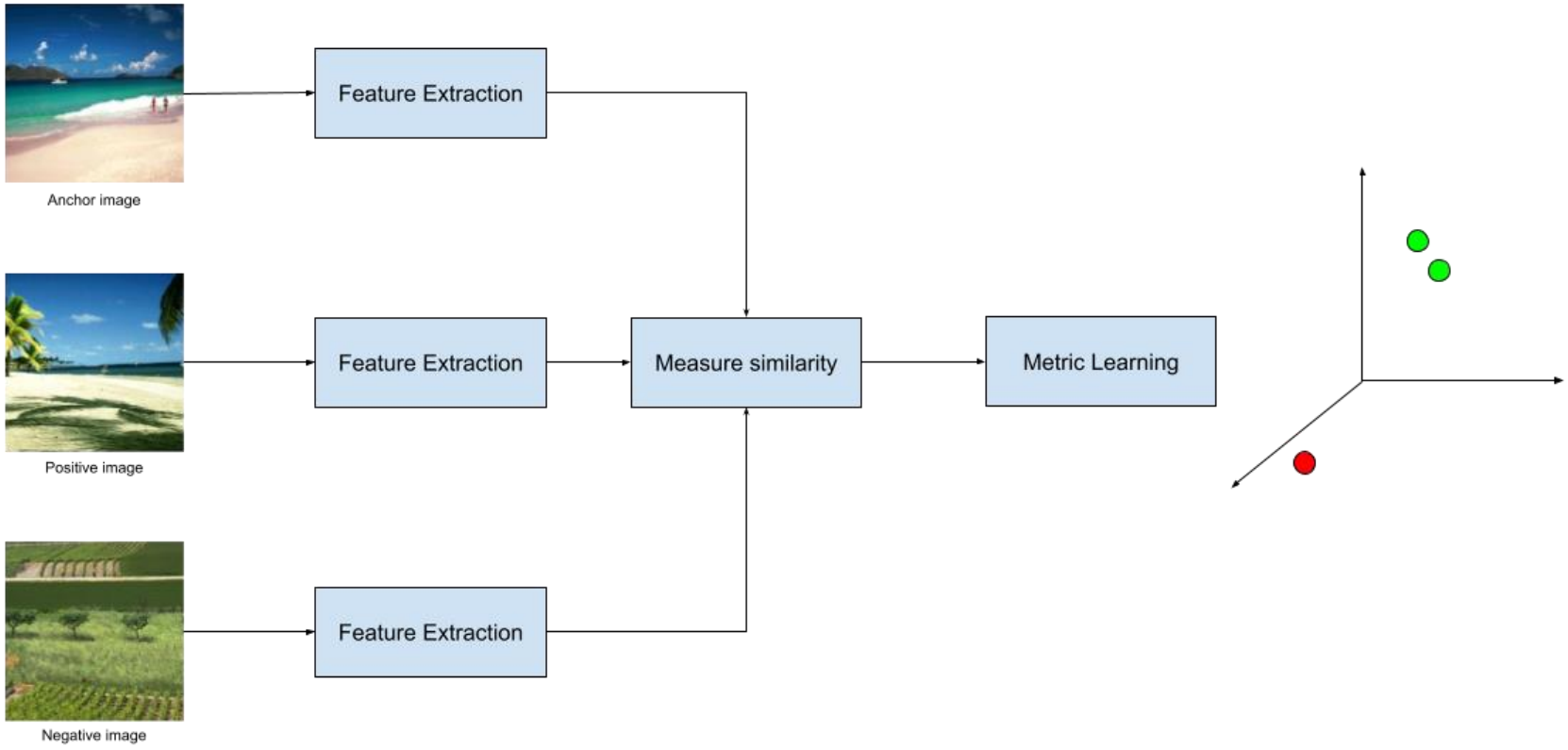


We can explicitly learn a representation that facilitates the retrieval of the images.

M5 – Image Retrieval

Training strategies for image retrieval

- Classification
- Metric learning:
  - Siamese networks
  - Triplet Networks
  - Quadruplet Networks
  - Etc.

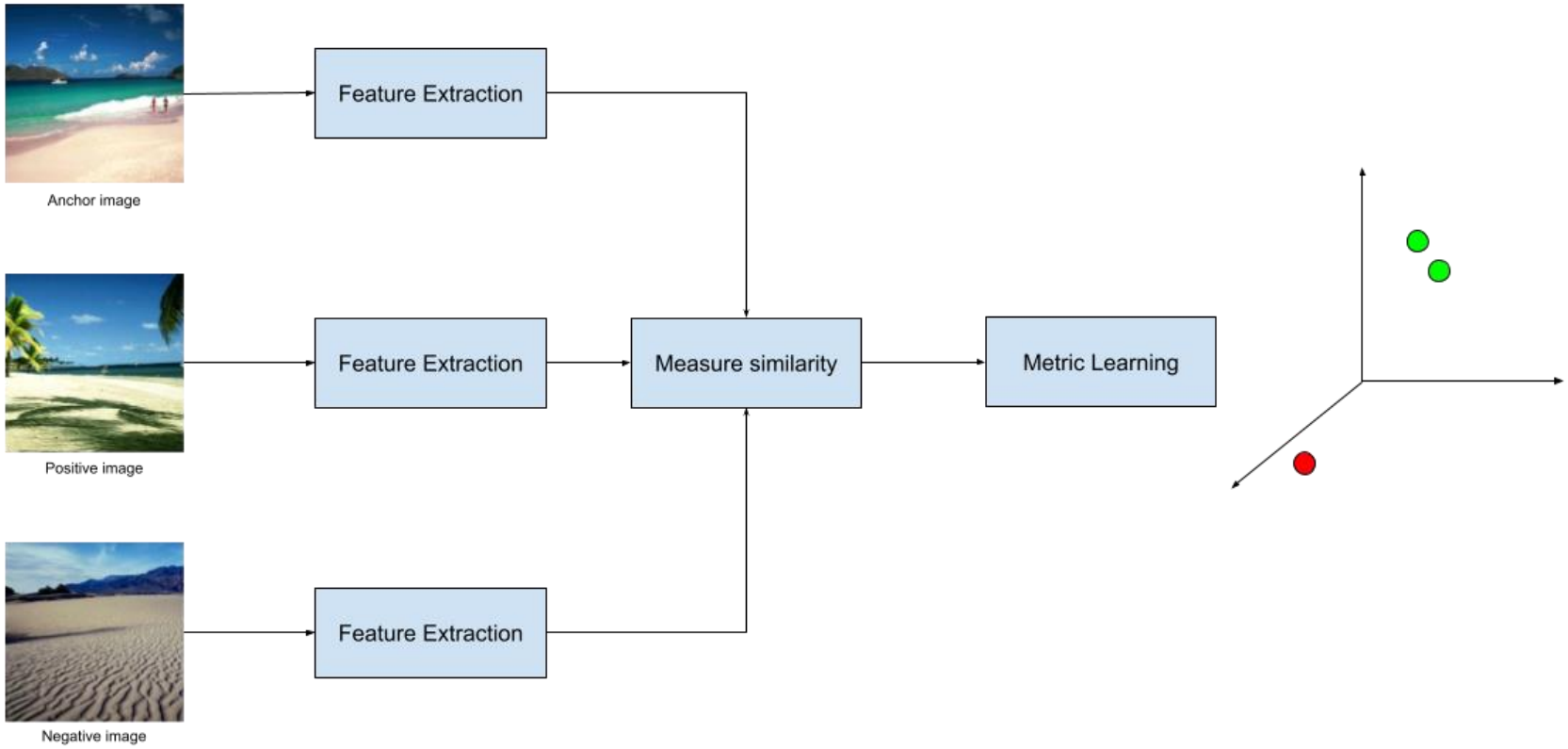


We can explicitly learn a representation that facilitates the retrieval of the images.

M5 – Image Retrieval

Training strategies for image retrieval

- Classification
- Metric learning:
  - Siamese networks
  - Triplet Networks



We can explicitly learn a representation that **facilitates** the retrieval of the images → **Hard negatives!**

M5 – Image Retrieval

Training strategies for image retrieval

- Classification
- Metric learning:
  - Siamese networks
  - Triplet Networks

Note: When is said that the models share parameters, you can use the same model.

```
1. img1_emb = model(img1)
2. img2_emb = model(img2)
3. loss = criterion(img1_emb, img2_emb)
```

M5 – Image Retrieval

Retrieval procedure

- Extract features from database images (train set).
- Extract features of the query image (val/test set).
- Retrieve the most similar images from the database.
  - NN, KNN…
  - (Facebook AI Similarity Search (FAIS), getting started documentation.

M5 – Image Retrieval

Retrieval procedure

- Extract features from database images (train set) → with torch.no_grad()
- Extract features of the query image (val/test set) → with torch.no_grad()
- Retrieve the most similar images from the database.
  - NN, KNN…
  - (Facebook AI Similarity Search (FAIS), getting started documentation.

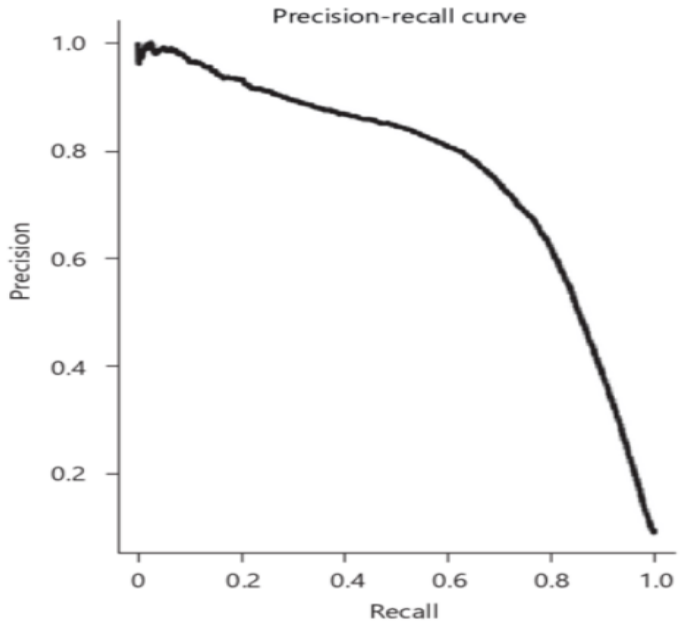
M5 – Image Retrieval

Retrieval procedure

- Extract features from database images (train set) → with torch.no_grad()
- Extract features of the query image (val/test set) → with torch.no_grad()
- Retrieve the most similar images from the database.
  - NN, KNN…
  - (Facebook AI Similarity Search (FAIS), getting started documentation.

Evaluation / Metrics

- Mean Average Precision (**MAP**)
- Precision@K
- Recall@K
- Difference between object detection and information retrieval metrics link.



Week 4. Image retrieval

Details on tasks, deliverables, and marks for this week

Week 1	**Introduction to Pytorch - Image Classification**
Week 2	**Object Detection, Recognition and Segmentation I**
Week 3	**Object Detection, Recognition and Segmentation II**
	Object Classification, Detection and Segmentation Report
Week 4	**Image Retrieval**
Week 5	**Cross-modal Retrieval**
Week 6	**Image and Cross-modal retrieval Report**
	Final Presentation

M5 Project: Goals per week

Week 4: Image retrieval

Goals

(a) Perform retrieval with pre-trained classification model.
(b) Train model on metric learning (Siamese)
(c) Train model on retrieval task. (Triplet)
(d) Visualize the learned embeddings.
(e) Extra features to analyze
(f) Start writing the retrieval report.

Marks

(C) Achieve (a)-(c), (f) goals
(B) Achieve (a)-(d), (f) goals
(A) Achieve (a)-(e), (f) goals

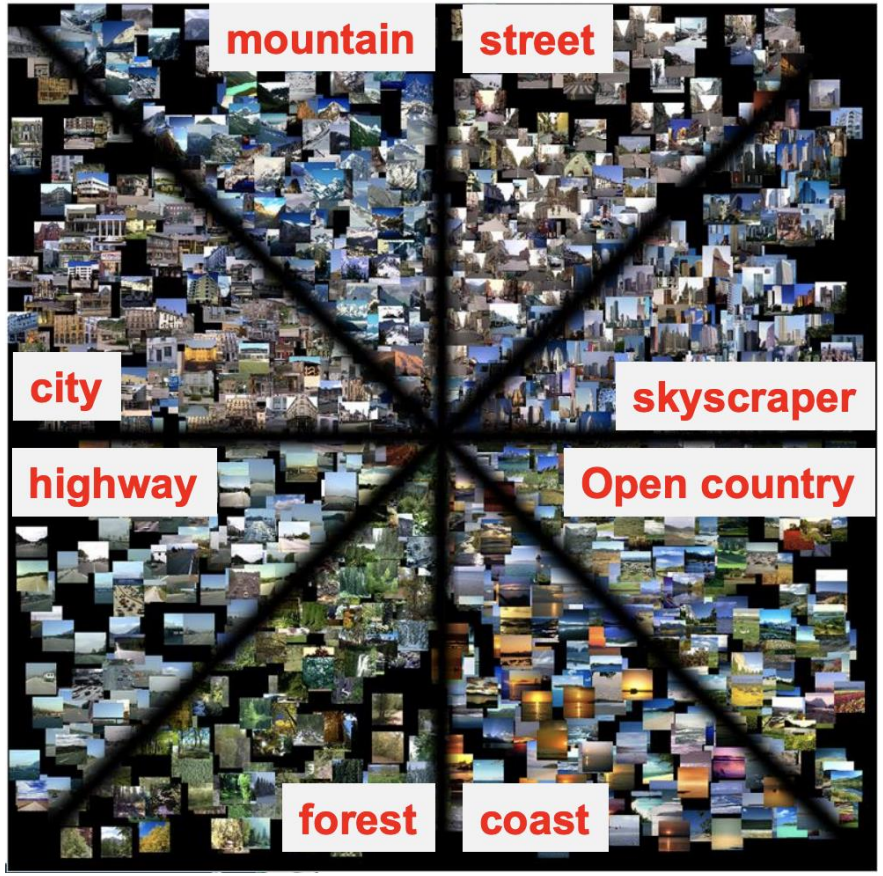
Deliverable (for next week)

- **Github** repository (code explanation & instructions)
- Presentation with information about models and results.
  - 1 minute slide presentation
- Report on overleaf about image retrieval.

M5 – Image Retrieval

Dataset

- MIT_Split



M5 – Image Retrieval

Task (a): **Image retrieval with pre-trained image classification model**

- Use P1 or standard Image Classification method (ResNet) pre-trained for Image Classification on the MIT_Split dataset.
  - You might need to remove the last linear layer where you projected the hidden size into the output (num_classes) size.

- Show (and analyze) precision-recall curve.
- Show qualitative results in your presentation.
- Show quantitative results in your presentation.
  - At least MAP, Prec@1, Prec@5
    - For MAP use the average_precision_score() function from the Sklearn library
      - Sklearn: Metrics, Basic models (NN, KNN, K-Means, SVMs)…
      - You will have to turn your integer targets $[7, 3, 1, 3, \ldots]_{bs}$ to binary $[0, 1, 0, 1, \ldots]_{database\_size}$

- You can choose the retrieval method you prefer (NN, KNN, FAIS…)

M5 – Image Retrieval

Task (b): **Train the model on metric learning (Siamese network)**

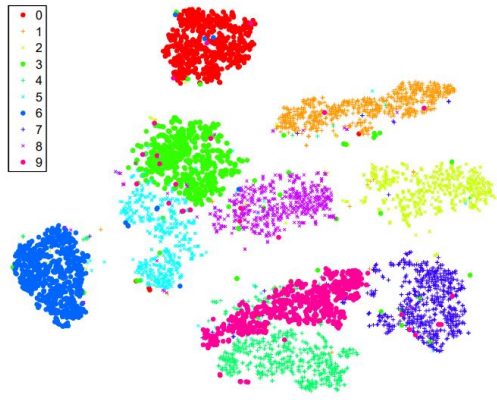
- Include precision-recall curve, quantitative and qualitative results in your presentation.

Task (c): **Train the model on metric learning (Triplet network)**

- Include precision-recall curve, quantitative and qualitative results in your presentation.

Task (d): **Visualize the learned image representation of each of the previous tasks a-c.**

- You can use PCA, TSNE or another you choose.
  - TSNE paper and implementation in sklearn



(a) Visualization by t-SNE.

M5 – Image Retrieval

Task (e): **Interesting features to analyze**.

- Implement 2 out of the 4 following options:

1. Compare your Project 1 method against a standard one (ResNet).

- Analyze how affects the use of different metric learning setups in a particular experiment:

2. Different losses, different distances (Euclidean, Mahalanobis), give different weights or margins.

3. Usage of hard negative and different hard-negative mining strategies.

2. Analyze how affects the performance the use of different retrieval methods with the same learned image representations.

4. Analyze different visualization models with the same learned image representations.

M5 – Image Retrieval

General information requirements for the presentation

- Describe your method.
  - Was it necessary to perform any change? (remove the last fully connected layer).
- Describe the training strategies (loss function).
  - Did you use any hard negative strategy? Which one?
- Describe the retrieval method.
- Describe the visualization method.

M5 – Image Retrieval

Extra material

- Siamese, Triplet examples (AdamBielski)
- Pytorch-metric-learning library (Kevin Musgrave)
  - Official Github repository
  - CIFAR10 example

M5 – Image Retrieval

Task (f): **Start the second report.**

- Use the CVPR format.
- Start writing the introduction, related work of training strategies for image retrieval and include the image retrieval experiments of this project.
- Remember to include only those experiments which are **relevant** to reach a conclusion or that have a particular behavior interesting to analyze.

M5 – Image Retrieval

- **Code** on Github project

- Report your results in your **presentation**.
  - Remember **1 minute slide** to present.

- **Overleaf** link on your Github

Due date: Monday 4th April before 10:00 AM