Week 1

Museum painting retrieval

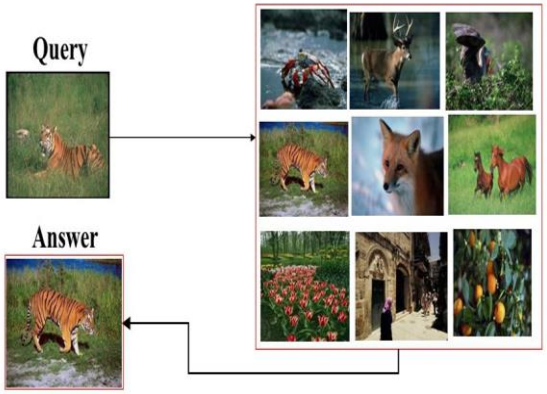
Group 8: José Manuel López Camuñas, Marcos V. Conde, Alex Martin Martinez

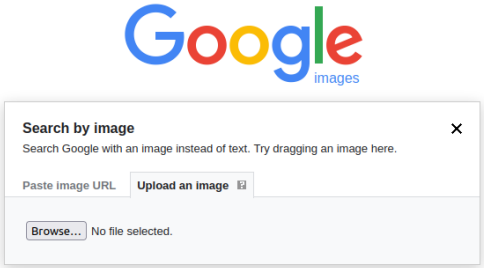
Index

1. Image retrieval
2. Method for QSD1
3. Methods for QSD2 - Segmentation
   - Otsu's threshold
   - Difference between pixels
4. Results
5. Discussion and conclusions

Image retrieval

The Image Retrieval problem consists on finding the most similar images (in a database) to a given query image. This search is based only on images' features.





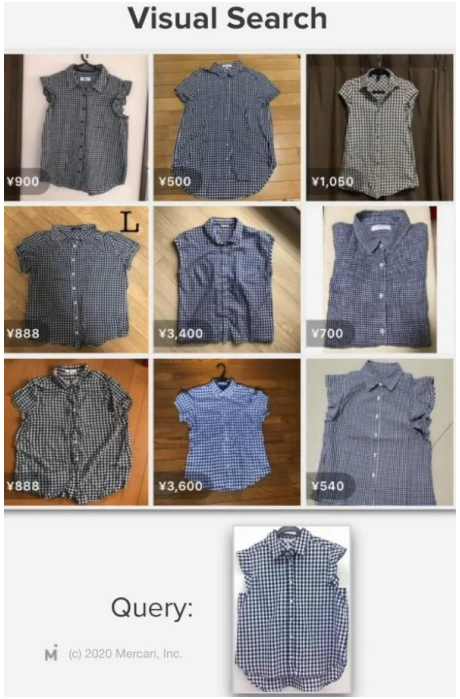
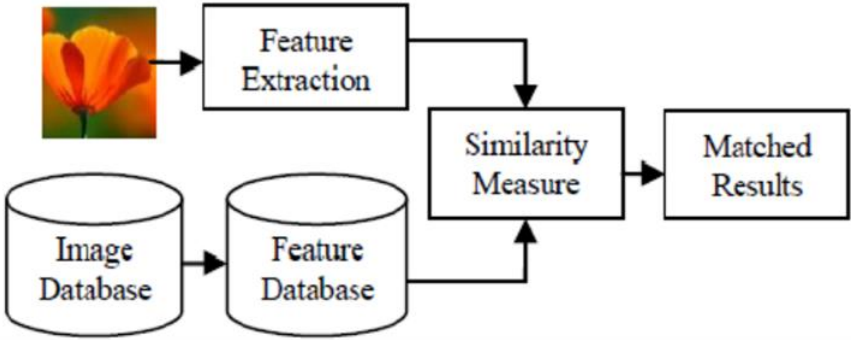
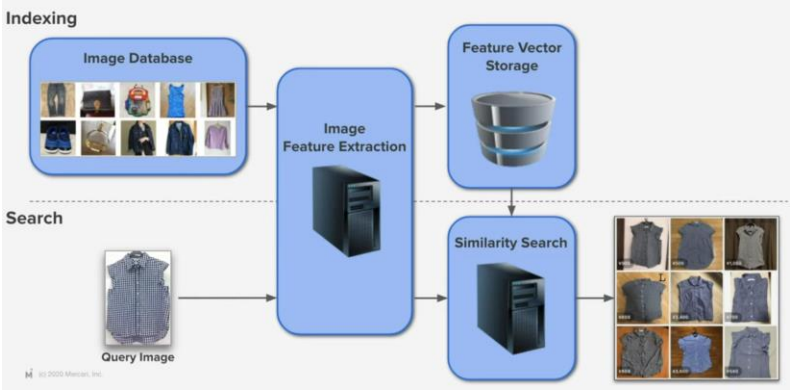


Image retrieval

One common procedure is to extract the features for each database image, then measure the similarity between the input image and the database images to extract the most similar one.



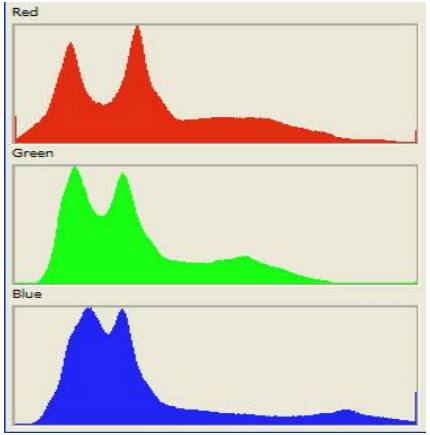


CVPR 2020, Scalable Image Retrieval Tutorial

Method for QSD1

For the curated dataset where the images are already cropped the strategy was:

- First, we obtain the color histogram for the R, G and B channels and the gray version of the image is extracted separately.



\+ Gray Image Histogram

Method for QSD1

Each of the histograms are then stacked one after another creating a **feature vector** for the image.

This process is applied to the entire database generating a feature matrix for all the images in the dataset.

Feature Matrix

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}$$

Method for QSD1

To search an image in the database then the cosine similarity is calculated with the feature vector of the input image and each vector from the feature matrix to compare which image of the dataset corresponds to the input image.

$$\cos(\theta) = \frac{\mathbf{A} \cdot \mathbf{B}}{\|\mathbf{A}\|\|\mathbf{B}\|}$$

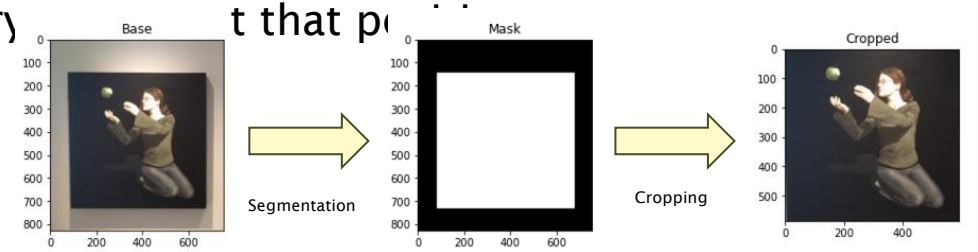
The cosine range of value varies from -1 to 1 having the most similarity when it is equal to 1 and the least when it is equal to -1. The image from the database with the cosines closer to 1 is will be the match for the query.

Methods for QSD2

In this dataset, the images are more realistic and contain backgrounds, in order to improve the retrieval performance we aim to remove the background (via Segmentation) before the feature extraction and query are performed.

In the first attempt an automatic thresholding technique named **Otsu Threshold** was used to extract the foreground from the image to obtain the painting.

In the second attempt the boundaries where detected calculating the **difference between pixels** from each of the boundaries to the center and if the difference surpassed a threshold the boundary t that p



Otsu's threshold

Otsu thresholding [1] starts by calculating the intensity distribution, ω_i, for the image from 0 to 255 and from 255 to 0. Previously the histograms were calculated using 256 bins in both.

$$\omega_0(t) = \sum_{i=0}^{t-1} p(i)$$

After that the mean, μ_i, is calculated for each of the histograms.

$$\mu_0(t) = \frac{\sum_{i=0}^{t-1} ip(i)}{\omega_0(t)}$$

[1] N. Otsu, "A threshold selection Method for Gray-level Histograms".

Otsu's threshold

Using the intensity distribution and the mean for each histogram the inter-class variance can be calculated for each difference of intensities:

$$\sigma = \omega_0(t)\omega_1(t)[\mu_0(t) - \mu_1(t)]^2$$

This function is intended to be maximized so the intensity with the greatest inter-class variance will be selected as the threshold value.

Difference between pixels

For this method the difference between neighbouring pixels are being calculated for each of the directions.

90	91	92	93	94	95	96	97	98	99
80	81	82	83	84	85	86	87	88	89
70	71	72	73	74	75	76	77	78	79
60	61	62	63	64	65	66	67	68	69
50	51	52	53	54	55	56	57	58	59
40	41	42	43	44	45	46	47	48	49
30	31	32	33	34	35	36	37	38	39
20	21	22	23	24	25	26	27	28	29
10	11	12	13	14	15	16	17	18	19
0	1	2	3	4	5	6	7	8	9

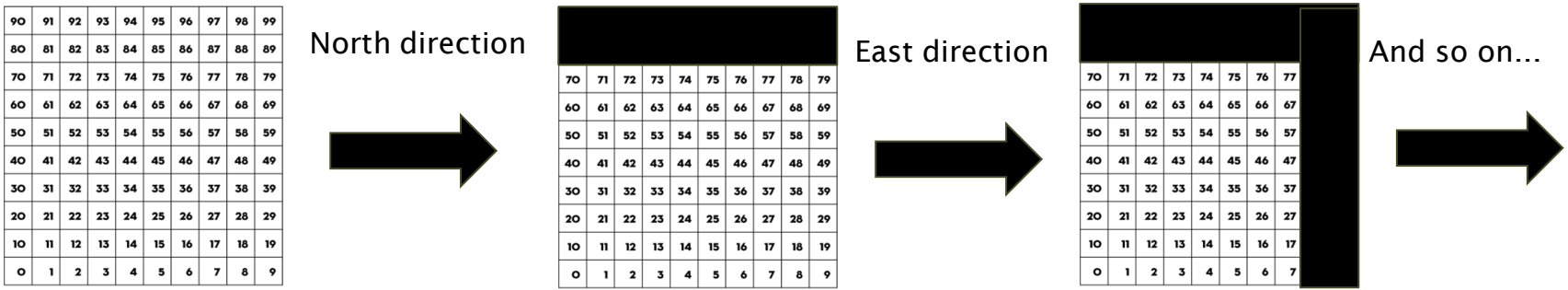
North direction:
$D(i,j) = p(i,j) - p(i+1,j)$

being $p(i,j)$ the pixel value

If D(i,j) surpasse a certain threshold value that in this case was 0.15 the border of the painting in that direction established at that pixel.

Difference between pixels

The process will produce 4 borders delimiting the zone where the painting is located.



Which will produce a mask without 0 values in surrounded by 1s which is how the ground truth masks are.

Results

	Dataset	MAP@1	MAP@5	Comment
Method 1 - Similarity	QSD1	0.366	0.428	Baseline Retrieval Method using the curated QSD1 dataset.
Method 1 - Similarity	QSD2	0.1	0.117	Performance decay due to “wild” images with background.
Method 2 - Otsu	QSD2	0.433	0.461	After cropping the images the performance improves notably.
Method 3 - Pixel difference	QSD2	0.30	0.386	Overall worst performance than the Otsu method.

Table 1: *Comparison of Quantitative Retrieval.*

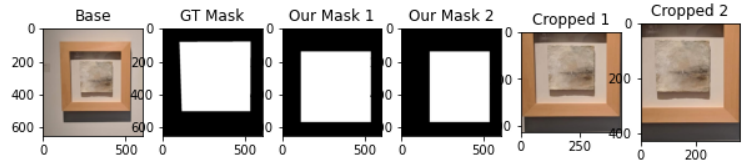
Results

	Precission	Recall	F1-measure	IOU
Otsu Method	0.9022	0.9595	**0.9266**	**0.8716**
Pixel Difference Method	0.9538	0.8608	0.8825	0.8264

Table 2: We report: *Precision, Recall, F1- score and Intersection over Union (IOU) metrics for the segmentation methods performed to the QSD2.*

Results

```
Threshold found: 0.4695622870710784 (119.73838320312498 if not normalized)
Method 1 IOU 0.7335376743737028 - F1 0.8462898559602603 Prec 0.8393740886394768 Recall 0.853320530745863
Method 2 IOU 0.6392869798023544 - F1 0.7799573688792819 Prec 0.8380462061535087 Recall 0.7293993452711847
```




```
Threshold found: 0.4138076953125 (105.5209623046875 if not normalized)
Method 1 IOU 0.9891233909898285 - F1 0.994531958620848 Prec 0.9970791198798267 Recall 0.9919977782740396
Method 2 IOU 0.7416129501699239 - F1 0.8516392233964118 Prec 0.9955069865784152 Recall 0.7441035507532682
```

Figure 1 : Examples of the Segmentation methods. We show (base) original image, (GT mask) the ground-truth segmentation mask, our masks using Otsu Method (1) and Pixel difference (2), and the crops obtained by applying each mask. We also provide the quantitative segmentation metrics (IOU, F1, Precision and Recall) for each sample.

Results

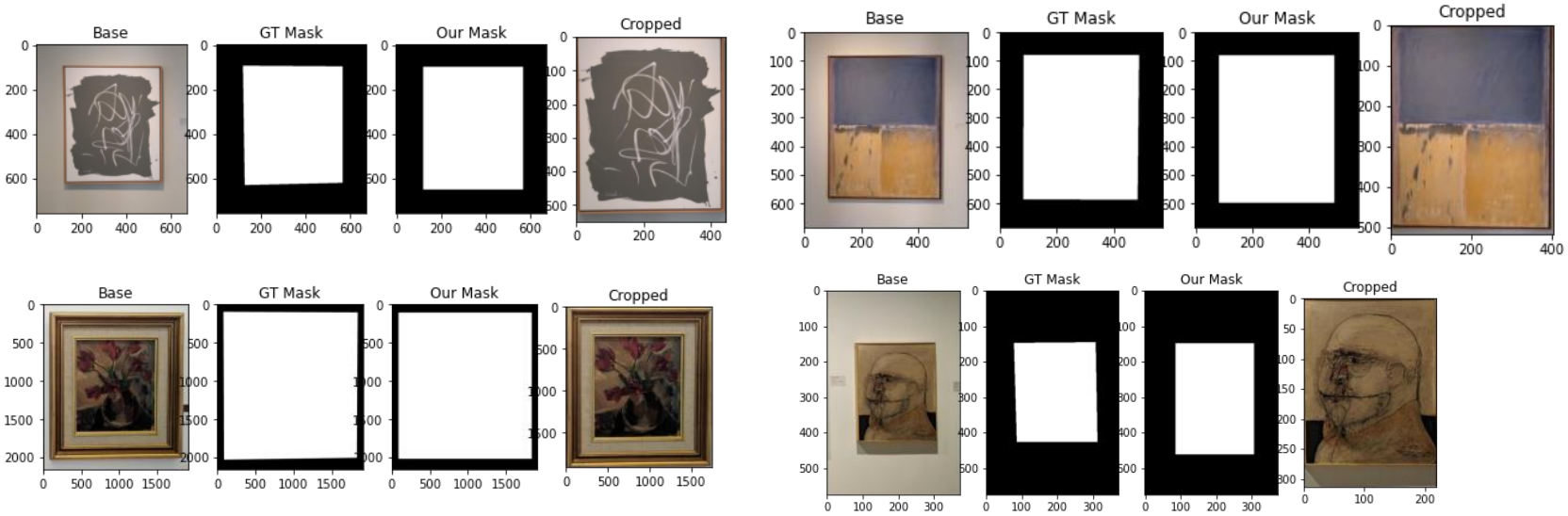


Figure 2 : Qualitative Samples of the segmentation masks using the Otsu Method.

Results

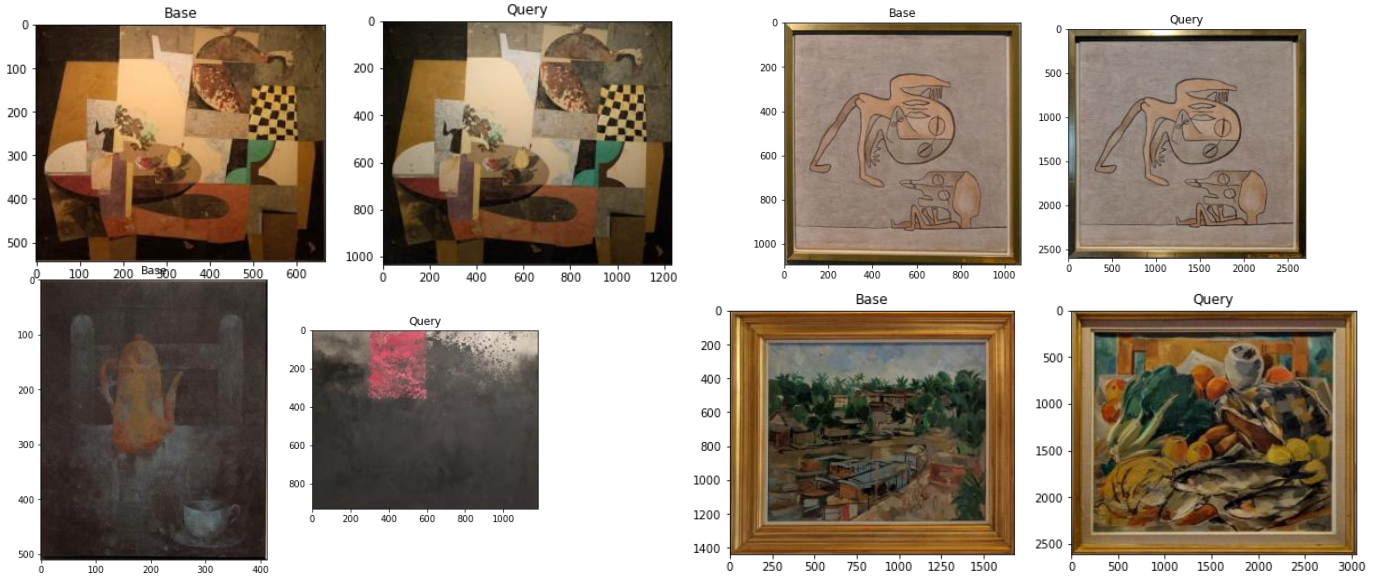


Figure 3 : Qualitative Image Retrieval samples. We show the Query image and the most similar (top-1) image in our database.

Discussion and conclusions

- For the **Image Retrieval**, the usage of only the histograms of each RGB channel, and the grayscale, is not enough to achieve a precise retrieval system.
- We noticed minimum performance changes related to the chosen distance/ similarity.
- Adding more features such as textures, edges, and other properties of the image, might improve performance.
- **Image Segmentation** (background segmentation) improves considerably the retrieval performance. The proposed strategies resulted to be successful with the majority of the images.