

The background of the entire image is a dark blue field filled with a pattern of red dots. These dots are arranged in a way that they form a large, faint, stylized 'H' shape, which is the primary visual element behind the text.

HUST

ĐẠI HỌC BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY OF SCIENCE AND TECHNOLOGY

ONE LOVE. ONE FUTURE.



ĐẠI HỌC
BÁCH KHOA HÀ NỘI
HANOI UNIVERSITY
OF SCIENCE AND TECHNOLOGY

Computer Vision

Nguyễn Hà Phú Thịnh

ONE LOVE. ONE FUTURE.

● Structure

- Project 1.2 for Classical Computer Vision
 - Picture 1
 - Picture 2
 - Picture 3
 - Picture 4
- Content-Based Image Retrieal
 - Dataset
 - Architecture
 - Evaluation
 - Results

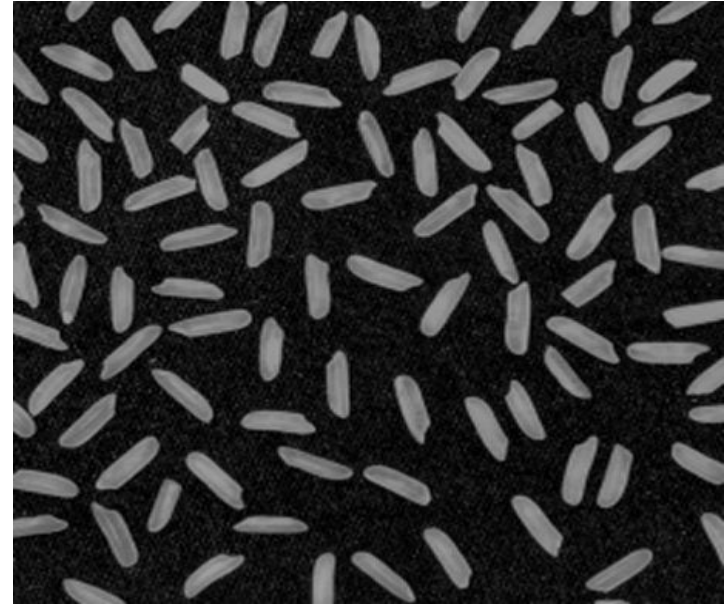
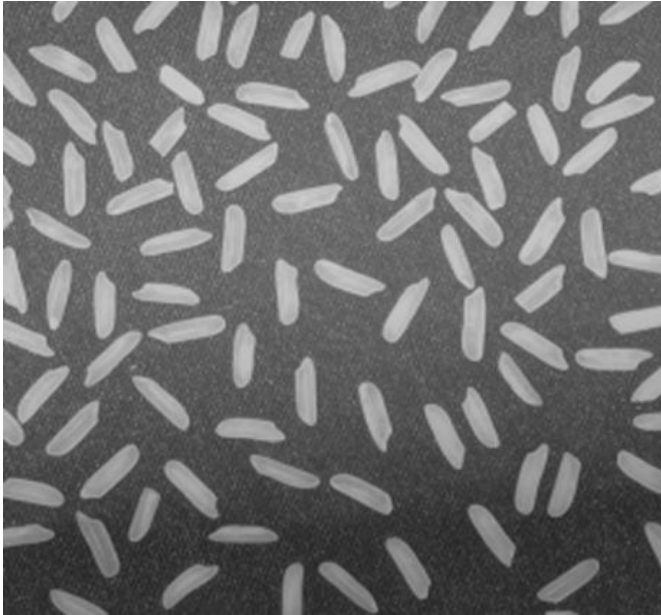
● Project 1.2

Pic 1 + 2:

- Top Hat transform: fix uneven light
- Median Blur: remove noise
- Otsu's Thresholding: verify rice vs. background
- Morphological Opening: clean up
- Watershed Algorithm: separate touching grains

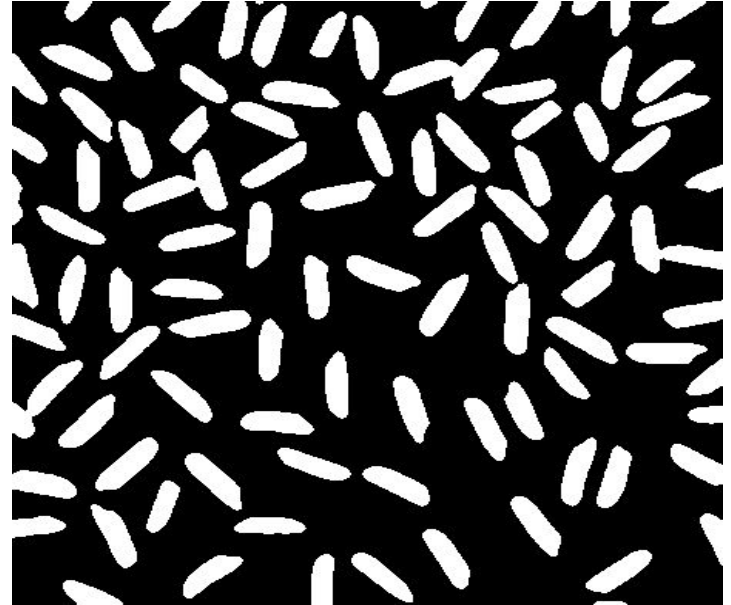
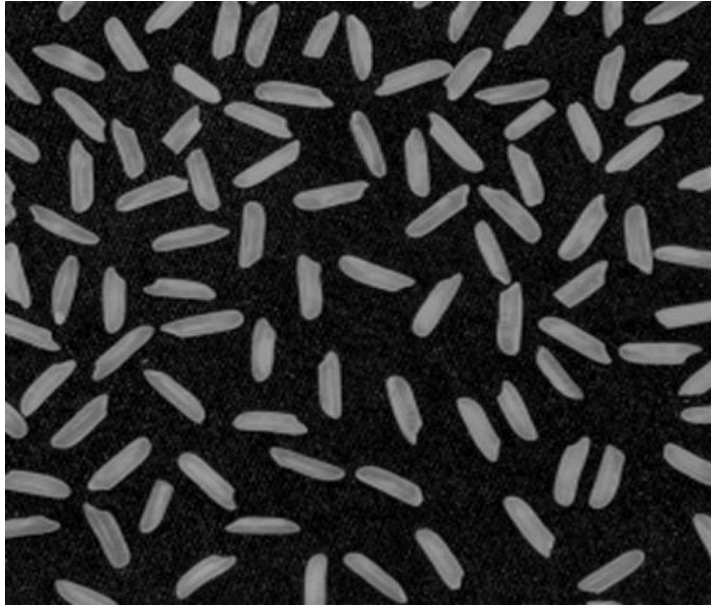
Pic 1

Top-Hat transform: kernel (40x40)



Pic 1

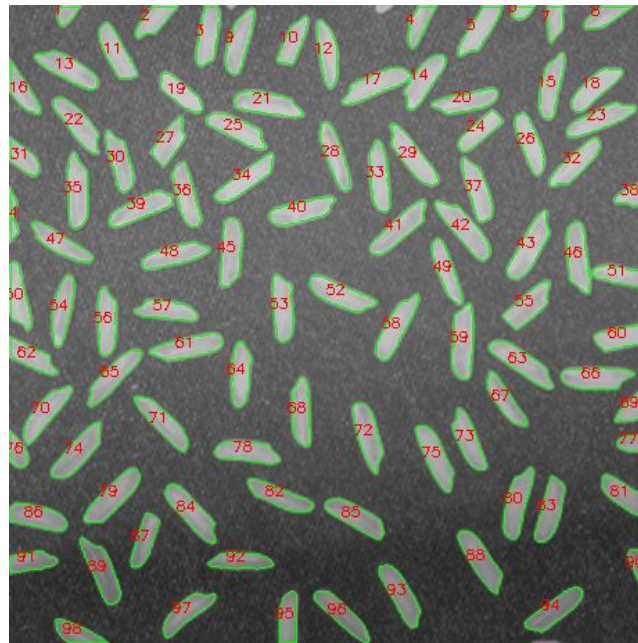
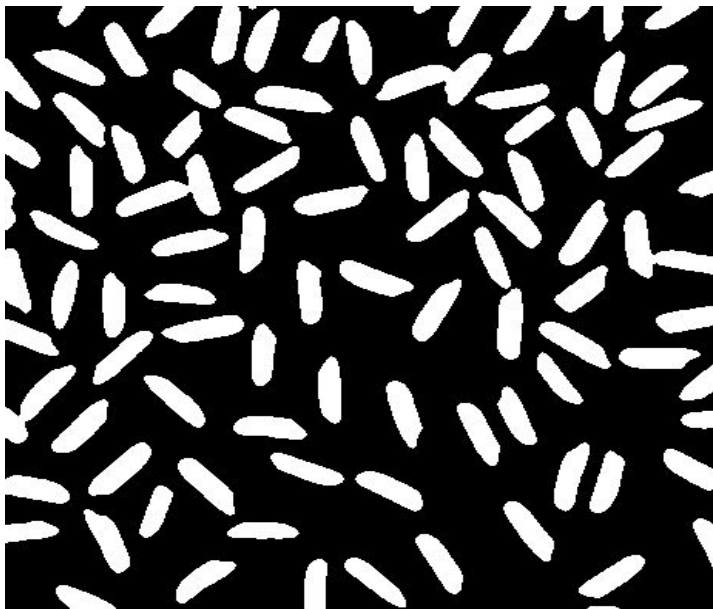
Median Blur (kernel: 5) + Otsu thresh hold (50%)



Pic 1

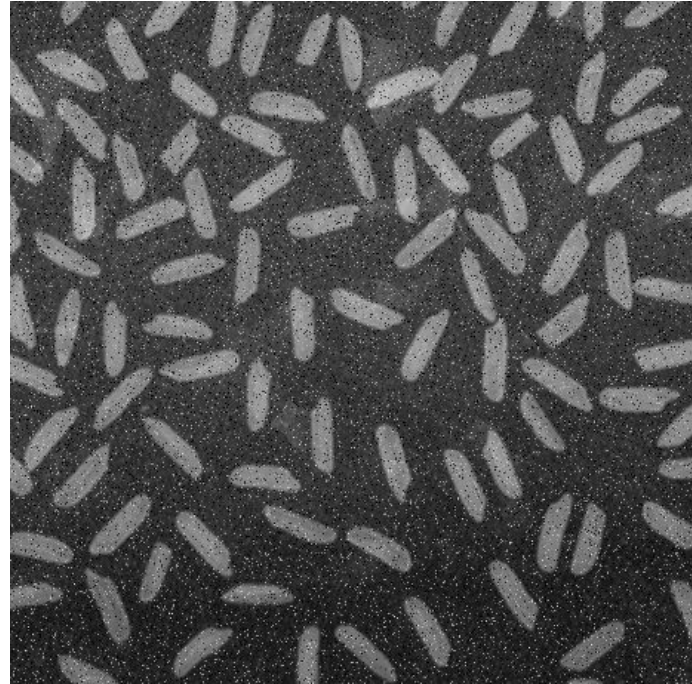
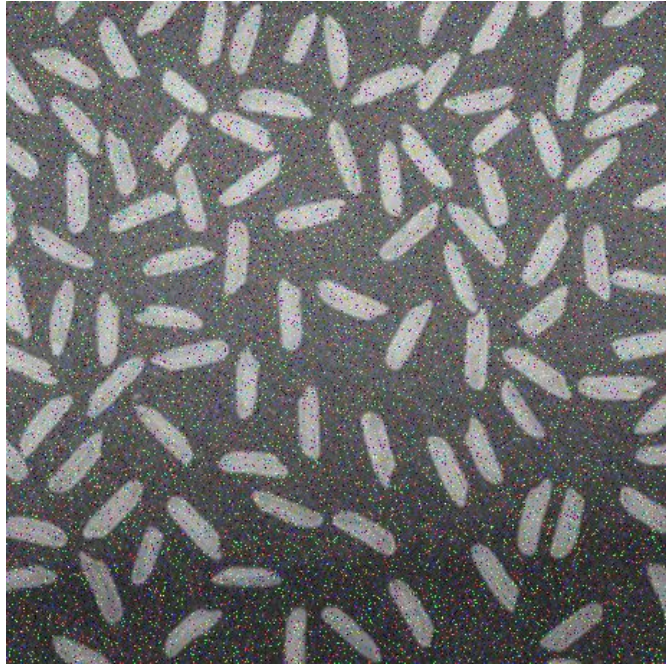
Clean: Morphological Opening + Watershed Algorithm

=> 97



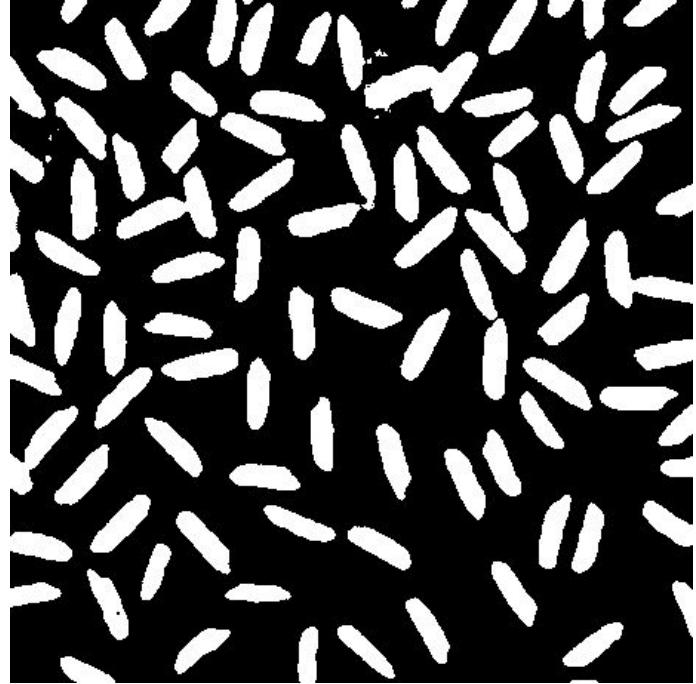
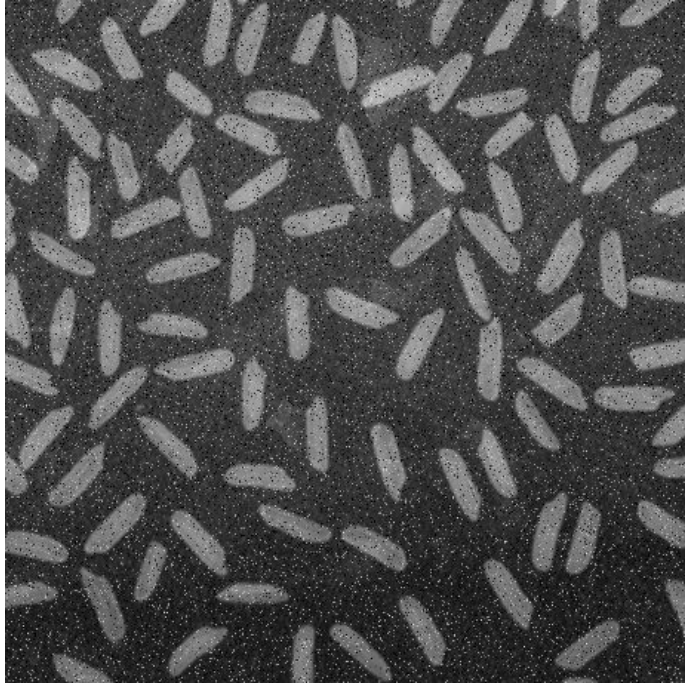
Pic 2

Top-Hat transform: kernel (40x40)



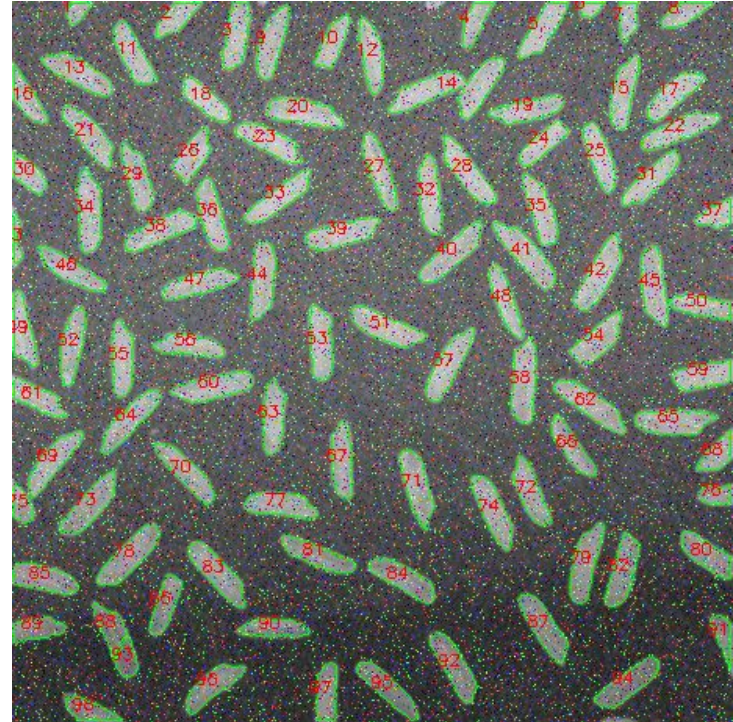
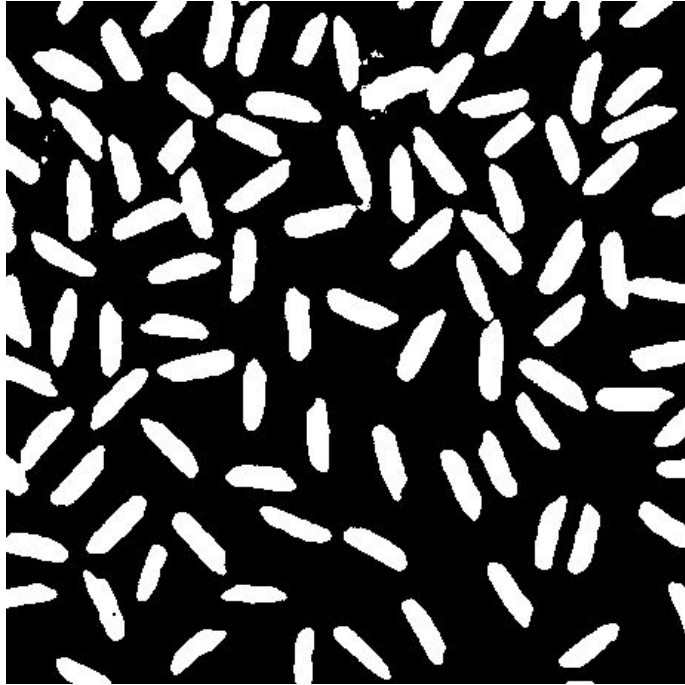
Pic 2

Median Blur (kernel: 5) + Otsu thresh hold



Pic 2

=> 98



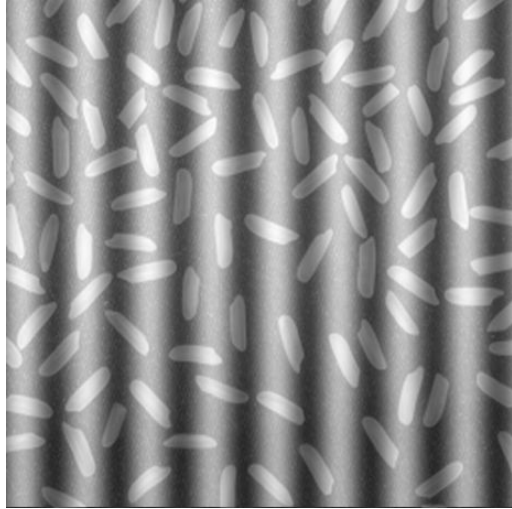
Project 1.2

Pic 3:

- Vertical Kernel: To isolate the waves
- Image subtraction: To clean the image
- Otsu's Thresholding: verify rice vs. background
- Morphological Opening: clean up
- Watershed Algorithm: separate touching grains

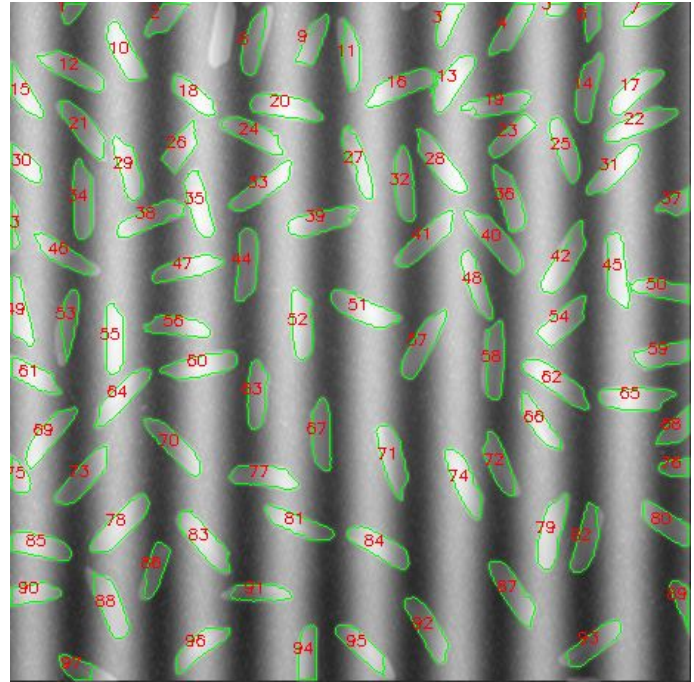
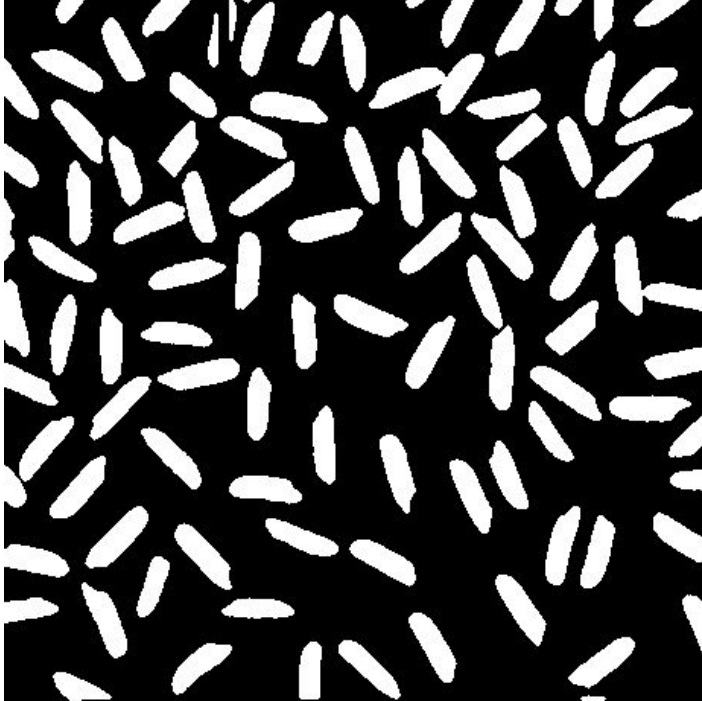
Pic 3

Vertical kernel (1, 80)



Pic 3

=> 97



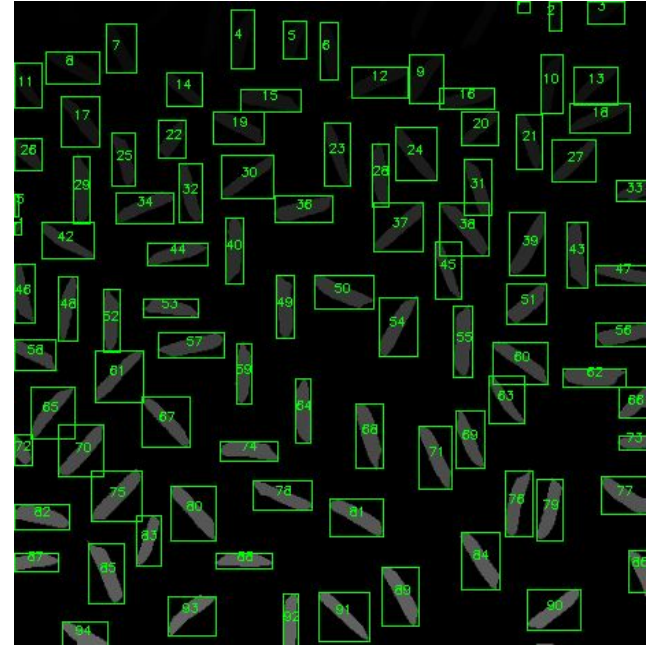
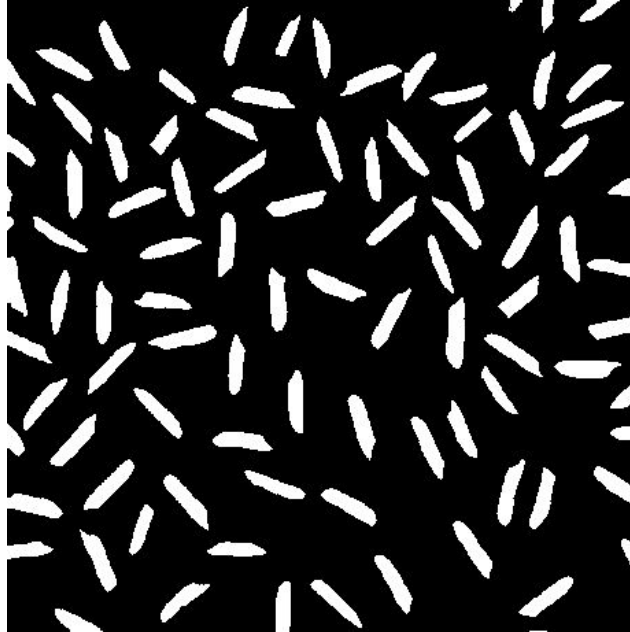
Project 1.2

Pic 4:

- Otsu's Thresholding: Key
- Morphological Opening: clean up
- Watershed Algorithm: separate touching grains

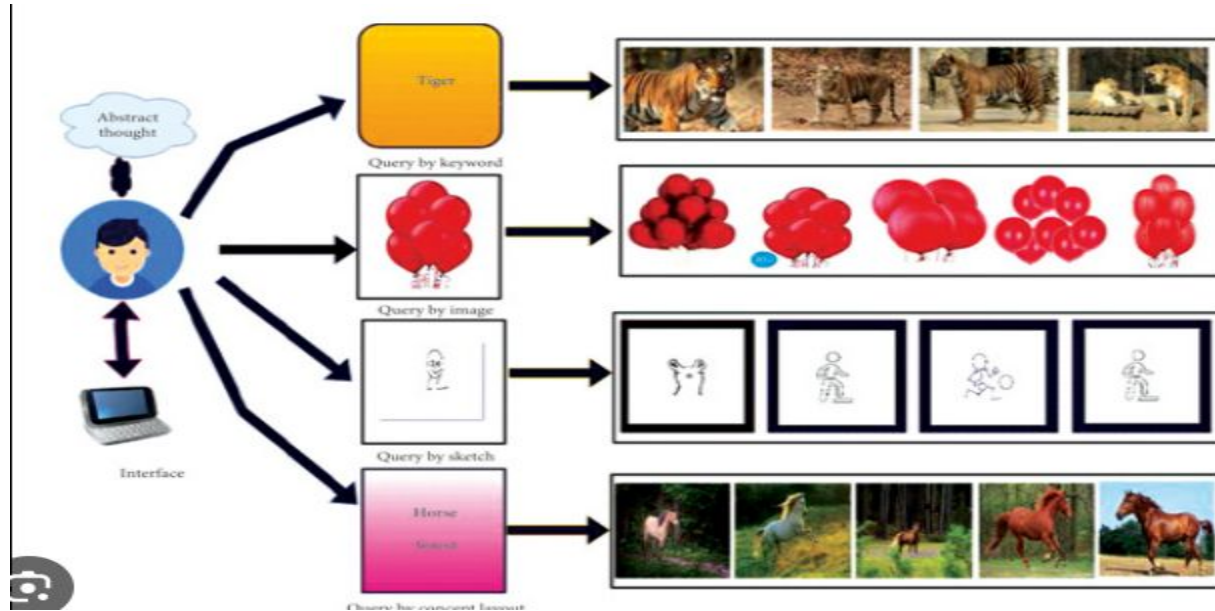
Pic 4

=> 94



● CBIR

- **Goal:** Build a system to retrieve images from a large database that are visually similar to a query image.
- **Challenge:** Semantic gap—pixels are just numbers; the model must understand "architecture," "shape," and "object identity."
- **Application:** Digital asset management, visual search engines (like Google Lens), e-commerce.



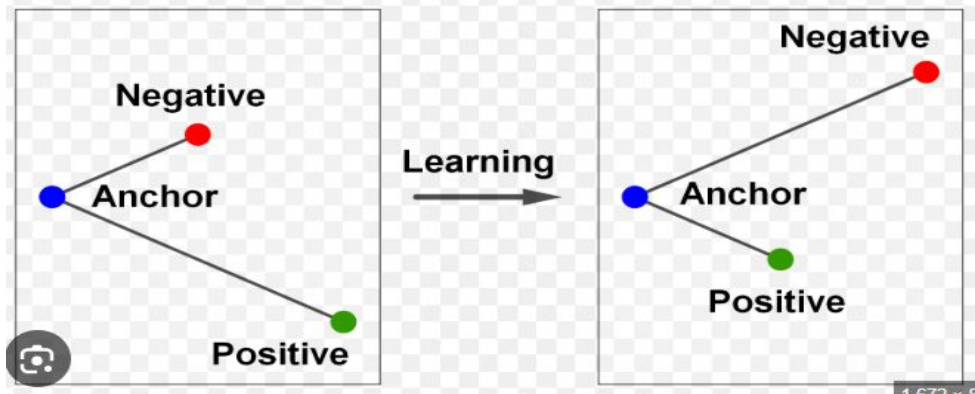
- Dataset:
- Oxford5k:
 - The Oxford Buildings Dataset consists of 5063 images collected from Flickr by searching for particular Oxford landmarks.
 - The dataset contains 11 landmarks. For each landmark there are *ok*, *good*, *junk* and *query* images.
- Paris 6k:
 - buildings dataset with ~6,000 images.

Methodology

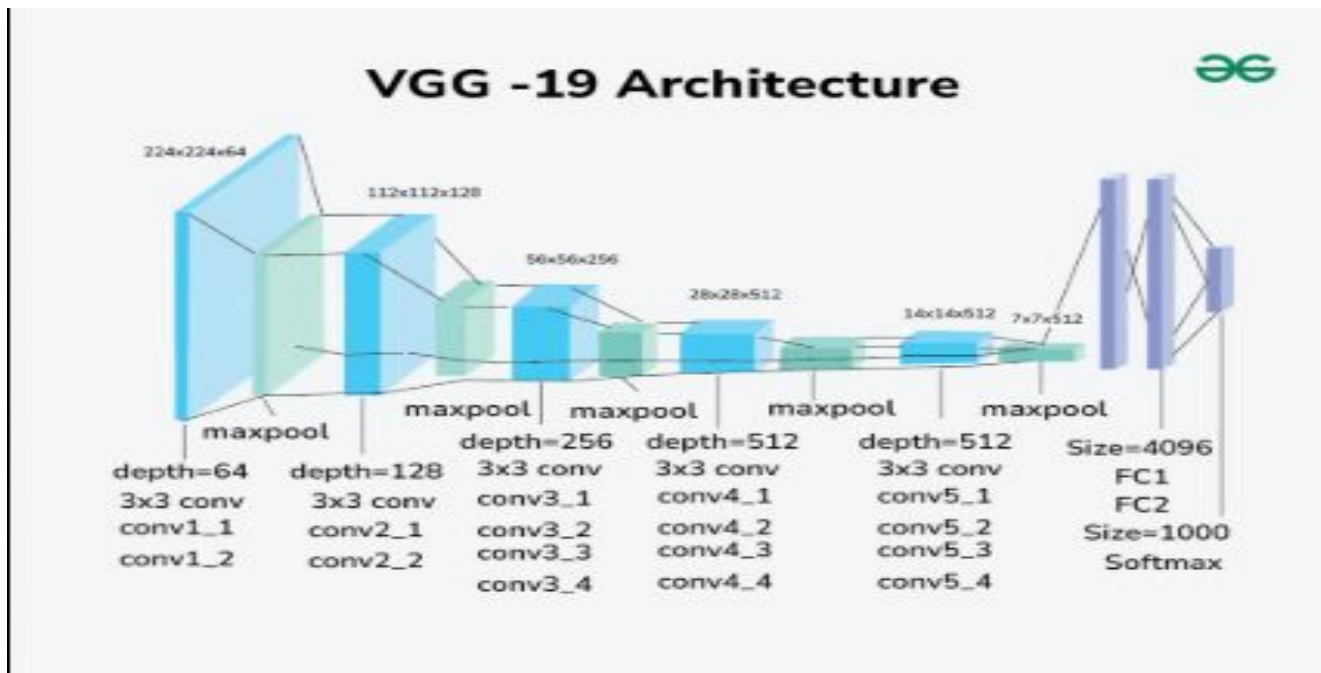
Visual Suggestion: (Three images entering the network: Anchor, Positive, Negative).

Bullet Points:

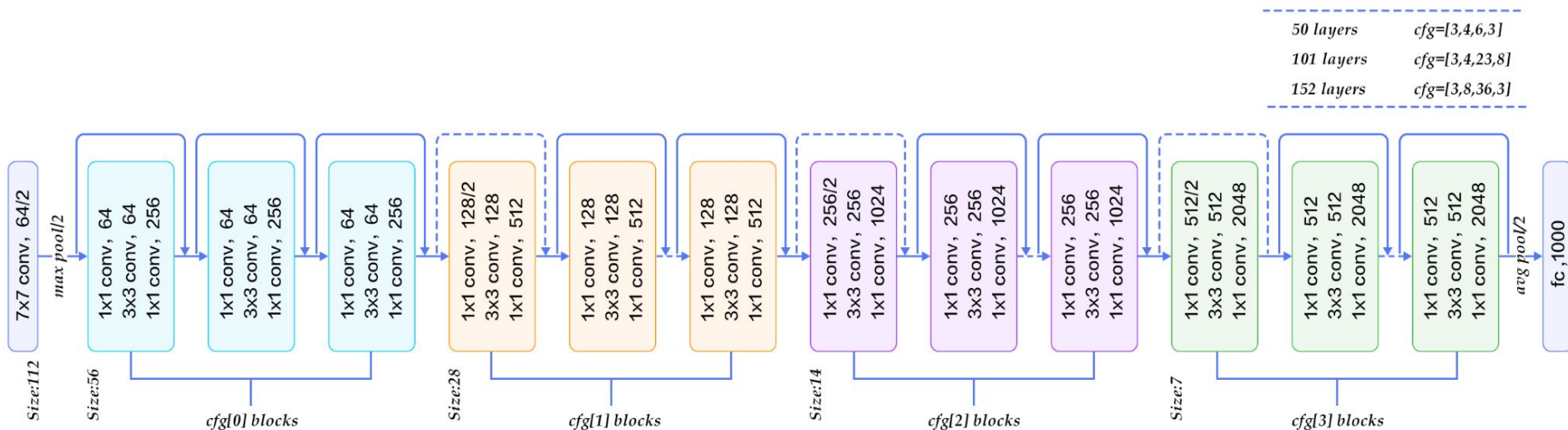
- **Approach:** Deep Metric Learning.
- **Architecture:** Siamese Network with Triplet Architecture.
- **Objective:** Learn an embedding space where similar images are close, and dissimilar ones are far apart.
- **Loss Function:** Triplet Margin Loss.



Backbone: VGG-19



Backbone: Resnet-50



Configuration

```
# Configuration
CONFIG = {
    'oxford_data_dir': '/kaggle/input/oxford5k/images/',
    'oxford_gt_path': '/kaggle/input/oxford5k/groundtruth.json',
    'paris_data_dir': '/kaggle/input/paris6k/Paris6k/',
    'batch_size': 32, # Larger batch size for 2 GPUs
    'num_epochs': 80,
    'learning_rate': 1e-5,
    'num_workers': 4,
    'embedding_dim': 2048,
    'backbone': 'resnet50', # Options: 'resnet50', 'vgg19'
    'device': torch.device("cuda" if torch.cuda.is_available() else "cpu")
}
```

Evaluation: Quantitative Results

Metric	ResNet-50	VGG-19
<i>System Metrics</i>		
Index Size (MB)	54.90	54.90
Vector Dimension	2048	2048
Number of Images	6960	6960
<i>Quality Metrics</i>		
mAP	0.8640	0.8002
MRR	0.9864	0.9301
NDCG@10	0.9879	0.9277
NDCG@20	0.9851	0.9314

Evaluation: qualitative assessment

Query
(Model: VGG-19)



#1 | 1.00



#2 | 1.00



#3 | 1.00



#4 | 1.00



#5 | 1.00



#6 | 1.00



Query
(Model: ResNet-50)



#1 | 1.00



#2 | 0.99



#3 | 0.99



#4 | 0.98



#5 | 0.98



#6 | 0.98



Evaluation: qualitative assessment

Query
(Model: VGG-19)



#1 | 1.00



#2 | 0.98



#3 | 0.98



#4 | 0.97



#5 | 0.97



#6 | 0.97



Query
(Model: ResNet-50)



#1 | 1.00



#2 | 0.96



#3 | 0.95



#4 | 0.95



#5 | 0.95



#6 | 0.95



Evaluation: qualitative assessment

Query
(Model: VGG-19)



#1 | 1.00



#2 | 0.96



#3 | 0.96



#4 | 0.95



#5 | 0.95



#6 | 0.95

Query
(Model: ResNet-50)



#1 | 1.00



#2 | 0.93



#3 | 0.92



#4 | 0.92



#5 | 0.92



#6 | 0.92

Evaluation: qualitative assessment

Query
(Model: VGG-19)



Query
(Model: ResNet-50)



A decorative graphic on the left side of the slide. It features a dark blue background with a large, stylized circular shape composed of many small red dots. The dots are arranged in a way that creates a sense of depth and movement, with some dots appearing larger and more concentrated than others. The word "HUST" is written in white, bold, sans-serif capital letters across the center of this circular graphic.

HUST

THANK YOU !