Are you Van Gogh's?

Team 27: Wing Hong Leung, Francisco Martinez, Jiawei Sun, Ling Zhou

Building a model to determine whether a painting is by Van Gogh or not

[https://github.com/francis2martinez/BootCampProject2020]

The Problem and Goals

Vincent van Gogh is a legendary Dutch painter from the mid 19th century. His huge legacy includes more than 860 oil paintings characterised by bold colors and dramatic, impulsive, and expressive brushwork [Wikipedia 2020].

Problem and Goals: Using digitalizations of Van Gogh's paintings from the "***Web Gallery of Art"*** *(https://www.wga.hu/)*, create a model than can recognize if a painting is in Van Gogh's style or not. This as a first step to eventually being able to recognize not only his style, but also if a work is a counterfeit or not.

Our Approach and Teamwork

Data cleaning [Francisco]

Feature Extraction:

- **Count Brush strokes with persistence homology** [Ling, Francisco]
- **Face counting** [Francisco]
- **K-NN of color histogram** [Joseph, Jiawei]

Combine Features [All]: Create a Decision Tree that takes these 3 features.

Data Gathering

1. Database: https://www.wga.hu
2. Form sample database:
   a. Techniques used by van Gogh for more than 10 times
   b. 1000 non-van Gogh paintings with the same techniques

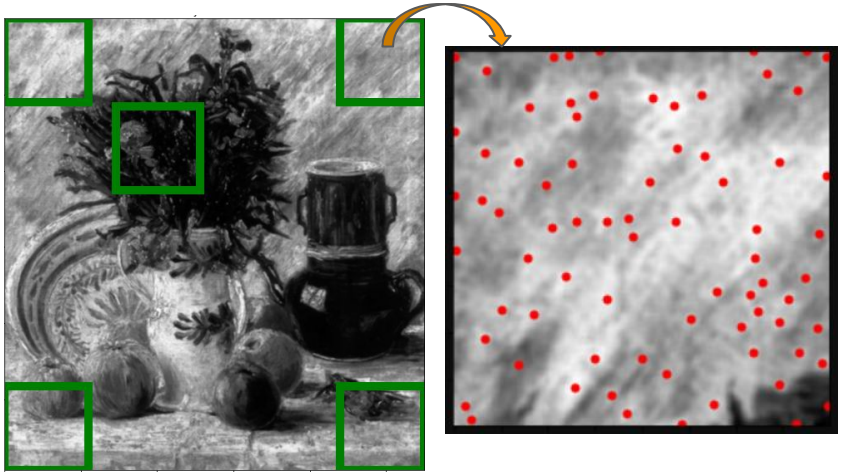
	TECHNIQUE	Count
13	Oil on canvas	293
15	Oil on canvas on panel	18
28	Pencil	14
0	Black chalk	10
5	Black pencil	10
12	Lithograph	9

AUTHOR	URL	Vangogh
SUSTRIS, Lambert	https://www.wga.hu/art/s/sustris/lambert/venus...	0
FRANGIPANE, Niccolò	https://www.wga.hu/art/f/frangipa/pieta.jpg	0
GRECO, El	https://www.wga.hu/art/g/greco_el/07/0703grec.jpg	0
GOGH, Vincent van	https://www.wga.hu/art/g/gogh_van/06/paris20.jpg	1

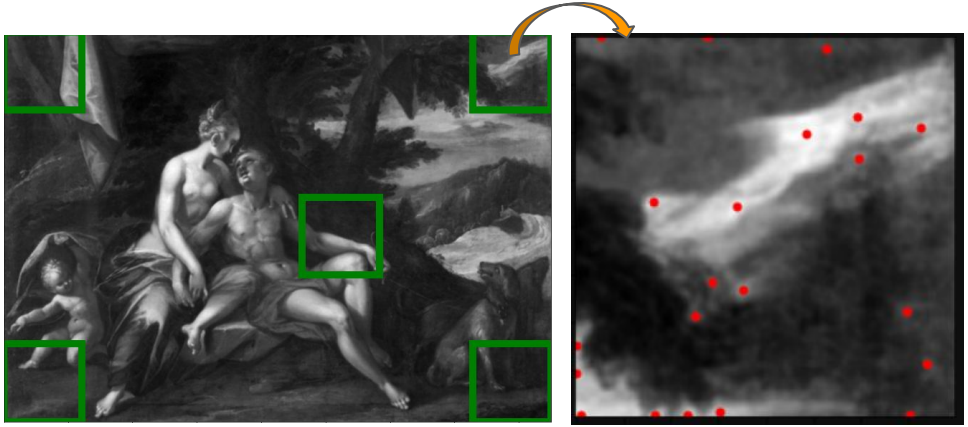
	Sample	Train	Test
Van Gogh	325	260	65
Other	1000	800	200
Total	1325	1060	265

Count Brush Strokes

Grayscale → Cut patches on corners and center → Count local minima, using lower star persistence → Compute mean



Still-Life, by Vincent van Gogh

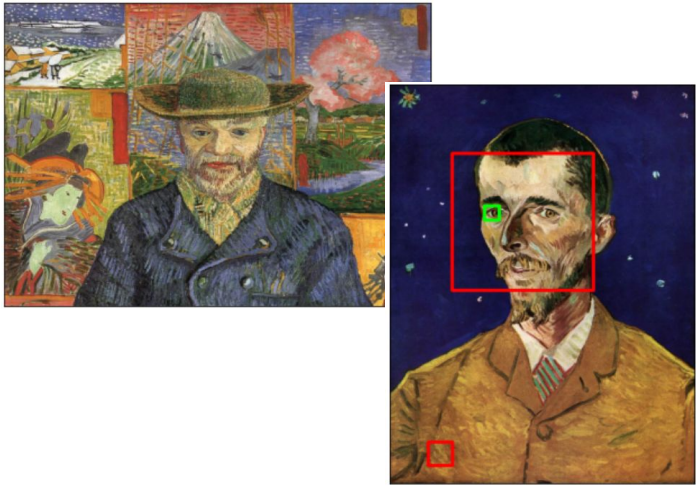


Venus and Adonis, by Hans von Aachen

Face Counting

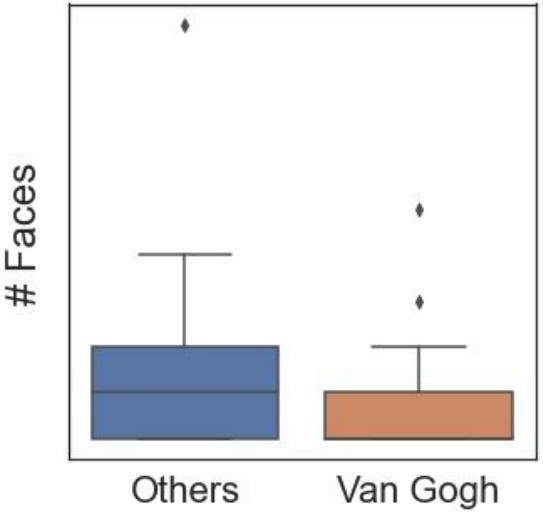
Detect and count faces, using Haar Cascade

Less detectable faces in Van Gogh's paintings



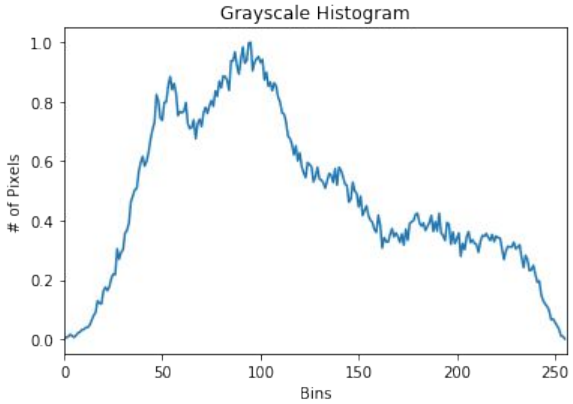
Van Gogh

Others

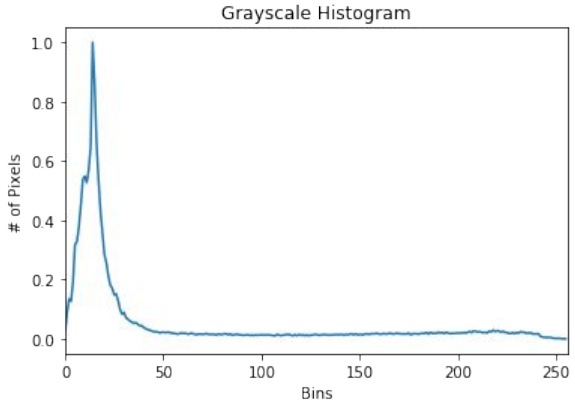


K-NN of Color Histograms

Grayscale → Compute normalized histogram → Predict using 3-NN, with χ^2 distance.



Reaper with Sickle, by Vincent van Gogh



David with the Head of Goliath, by Bernardo Strozzi

Methods/Packages Used

1. OpenCv (computer vision), scikit-learn (machine learning), ripser (persistence diagrams) , pandas, numpy, matplotlib, url, pickle.
2. Haar Cascades with pre-trained weight to recognize frontal and profile faces from OpenCv.
3. K-NN Neighbors from scikit-learn.
4. DecisionTree from scikit-learn.

Results and Challenges

Accuracy:

On test data: **80.4%** (> 75.5% trivial)

- **P**[Van Gogh | Predict Van Gogh] =59.4%
- **P**[Predict Van Gogh | Van Gogh] =63.1%

On train data: **88.5%**

Confusion Matrix:

	Pred. Other	Pred. Van Gog
True Other	172	28
True Van Gogh	24	41

Challenges:

- Overfitting.
- Quality of data set.
- Image processing is time consuming.

Next Steps

1. Get higher resolution images to extract shapes of brushstrokes.
2. Find a way to characterize brushstroke, probably using geometric parameters.
3. Improve features with different algorithms.
4. Expand the program to detect other famous painters' artwork

THANK YOU!

TEAM 27: Van Gogh's Ear and Eyes

[https://github.com/francis2martinez/BootCampProject2020]

Wing Hong Leung
leung.179@osu.edu
OSU

Francisco Martinez
martinezfigueroa.2@osu.edu
OSU

Jiawei Sun
sun.2261@osu.edu
OSU

Ling Zhou
zhou.2568@osu.edu
OSU