ECE 208/408 Impressionist Painting Classifier

Xiaoshan Liu, Danielle Williams

Introduction

- People learn how to identify the creator of a painting through recognition of patterns in the brushwork, color palette, composition, and subject matter.
- ML algorithms does it's 'identification' based on pixel patterns and vectors.
  - It then assign labels to the elements that it detects based on its configuration.

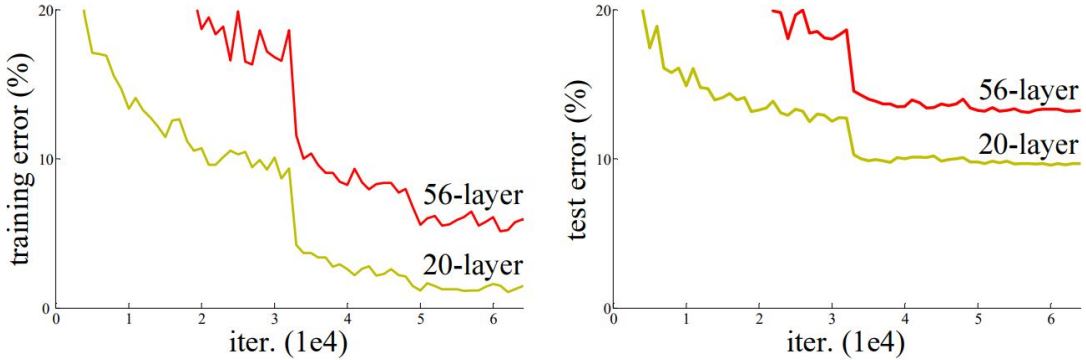
- Project Aim
  - To conduct a comparative analysis on the performance of pre-trained CNN networks on classifying paintings and their respective artists.
    - ResNet18
    - ResNet34
    - VGG16
    - VGG19

Method

- Data: Impressionist_Classifier_Data from Kaggle and Github
  - with 400 training images, 100 validation images and 70 test images respectively for each of the 10 artists.
- Preprocessing
  - transform
  - create dataloader
- Model configuration
  - loss: CrossEntropyLoss()
  - Optimizer: SGD
  - learning rate: 0.001
- Use Resnet and VGG

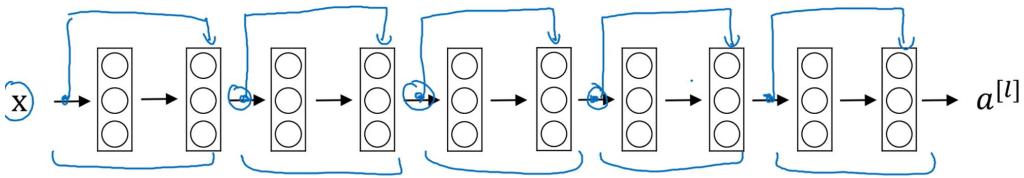
Resnet

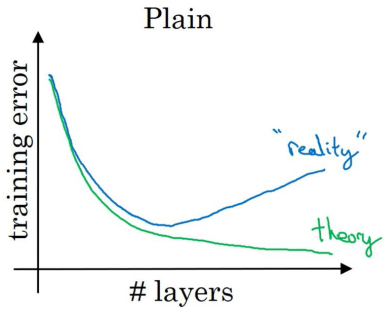
- Before:
  - The deeper the better
  - But vanishing gradient problem occurs when we stack too many layers
  - Not due to overfitting

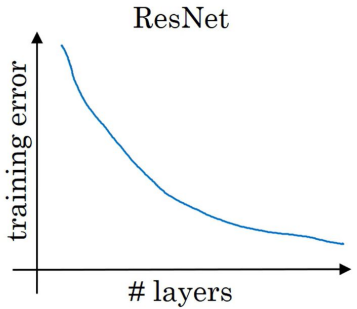


Properties of Resnet

- Proposed in 2015 by a Microsoft research team
- Skip connections



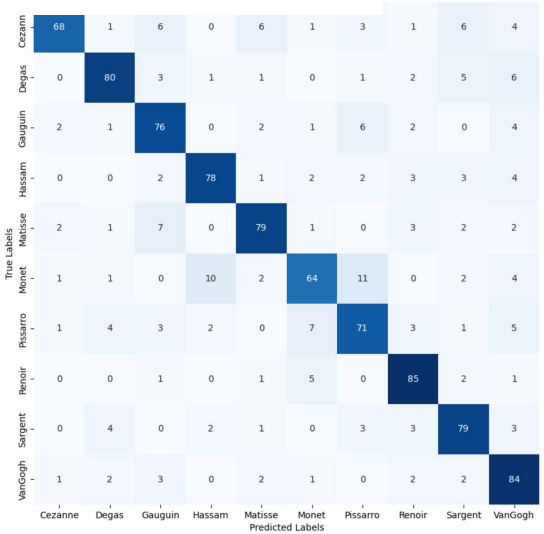




Preliminary Results

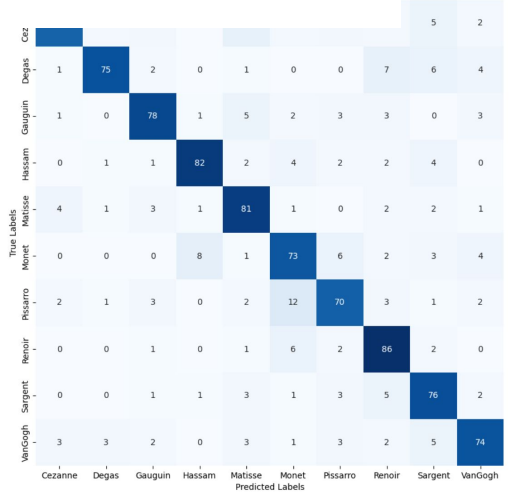
Confusion Matrix (training/validation) and accuracy (test) for pre-trained Resnet18

```
Accuracy of the network on the 714 test images: 79.97%
Confusion Matrix
----------------
[[49  1  4  0  3  1  3  1  6  4]
 [ 0 60  2  1  0  0  1  1  4  3]
 [ 2  1 59  0  2  1  3  2  0  2]
 [ 0  0  2 58  1  2  2  2  2  3]
 [ 1  0  6  0 58  1  0  3  1  2]
 [ 0  1  0  8  1 50  8  0  2  2]
 [ 1  2  2  2  0  5 53  3  1  3]
 [ 0  0  1  0  1  2  0 65  2  1]
 [ 0  3  0  2  2  0  2  0 60  3]
 [ 0  2  2  0  2  0  0  0  1 59]]
```



Confusion Matrix (training/validation) and accuracy (test) for pre-trained Resnet34

```
Accuracy of the network on the 714 test images: 80.39%
Confusion Matrix
----------------
[[53  2  2  0  3  2  2  2  4  2]
 [ 1 59  0  0  1  0  0  5  5  1]
 [ 1  0 58  1  5  1  2  3  0  1]
 [ 0  1  1 61  2  2  1  2  2  0]
 [ 4  1  3  1 58  1  0  2  1  1]
 [ 0  0  0  6  1 54  4  2  3  2]
 [ 1  0  3  0  2  9 54  3  0  0]
 [ 0  0  1  0  1  3  0 65  2  0]
 [ 0  0  1  1  3  1  2  1 61  2]
 [ 1  3  2  0  3  0  3  0  3 51]]
```



VGG

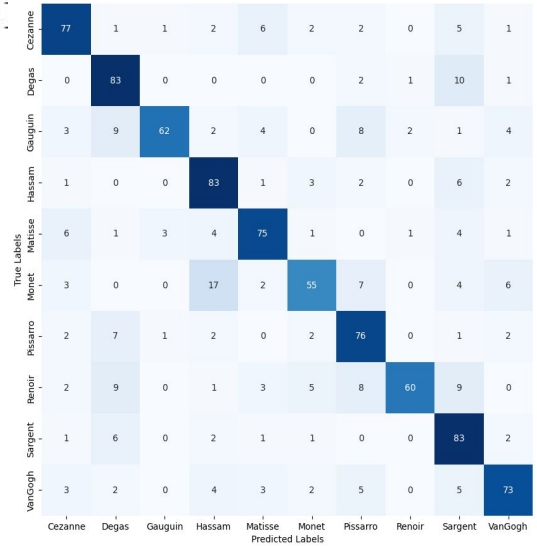
- VGG came as a result of the exploration of deeper CNNs
- It contains almost exclusively 3x3 convolutions
  - Smaller filters
    - Induced more non-linearity
      - More degrees of freedom for the network.
    - Limits the number of parameters.
  - Stacking layers enabled the network to see more:
    - With two, receptive field = 5x5
    - With three, receptive field = 7x7
  - Top-5 error on ImageNet database = 7.3%

input (224 × 224 RGB image)					
conv3-64	conv3-64 **LRN**	conv3-64 **conv3-64**	conv3-64 conv3-64	conv3-64 conv3-64	conv3-64 conv3-64
maxpool					
conv3-128	conv3-128	conv3-128 **conv3-128**	conv3-128 conv3-128	conv3-128 conv3-128	conv3-128 conv3-128
maxpool					
conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256	conv3-256 conv3-256 **conv1-256**	conv3-256 conv3-256 **conv3-256**	conv3-256 conv3-256 conv3-256 **conv3-256**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512	conv3-512 conv3-512 **conv1-512**	conv3-512 conv3-512 **conv3-512**	conv3-512 conv3-512 conv3-512 **conv3-512**
maxpool					
FC-4096					
FC-4096					
FC-1000					
soft-max					

Source: https://stanford.edu/~shervine/blog/evolution-image-classification-explained

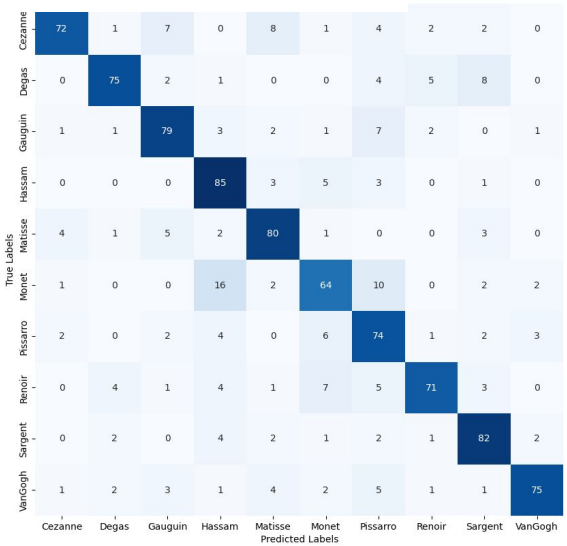
Confusion Matrix (training/validation) and accuracy (test) for pre-trained VGG16

```
Accuracy of the network on the 714 test images: 76.33%
Confusion Matrix
----------------
[[56  1  1  2  4  1  1  0  5  1]
 [ 0 62  0  0  0  0  1  0  9  0]
 [ 3  5 46  2  4  0  7  2  1  2]
 [ 1  0  0 62  1  2  1  0  4  1]
 [ 4  1  3  1 58  1  0  1  2  1]
 [ 1  0  1 14  0 42  7  0  4  3]
 [ 1  4  1  2  0  1 62  0  1  0]
 [ 1  6  0  1  3  2  6 44  9  0]
 [ 0  5  0  2  1  1  0  0 61  2]
 [ 0  2  0  4  3  1  2  0  2 52]]
```



Confusion Matrix (training/validation) and accuracy (test) for pre-trained VGG19

```
Accuracy of the network on the 714 test images: 78.57%
Confusion Matrix
----------------
[[55  1  3  0  5  1  3  2  2  0]
 [ 0 56  1  1  0  0  4  3  7  0]
 [ 1  0 58  3  2  0  6  2  0  0]
 [ 0  0  0 64  2  4  2  0  0  0]
 [ 4  1  5  0 60  1  0  0  1  0]
 [ 1  0  0 16  0 44  8  0  2  1]
 [ 1  0  3  4  0  4 57  1  1  1]
 [ 0  2  2  2  1  4  4 54  3  0]
 [ 0  2  0  4  1  1  2  1 59  2]
 [ 0  2  2  1  3  0  3  0  1 54]]
```



Compare

Article

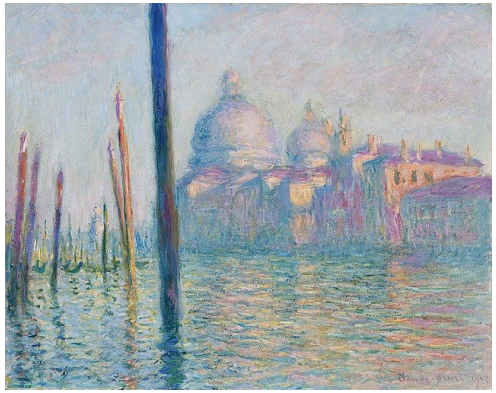
	VGG19	34 layer plain	34 Resnet
Properties	-Downsize the resolution -Increase filter -Lots of parameters, hence requires lots of computation -19.6billion floating points	-3.6 billion floating points -not much parameters per layer -less filters	-3.6 billion floating points -not much parameters per layer -less filters -add skip connections -output has to have the same size as input
Performance	middle	worse	best
Time	slowest	middle	fastest

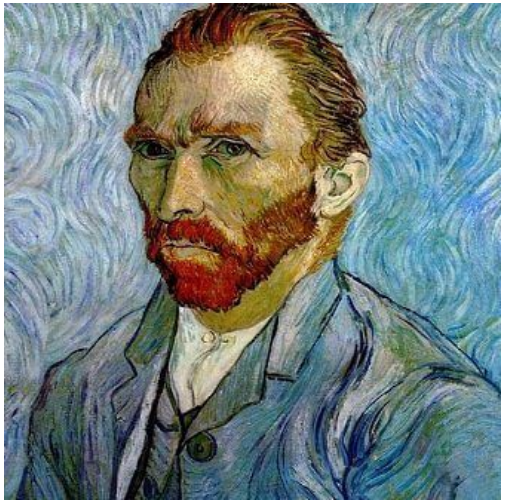
Our implementation

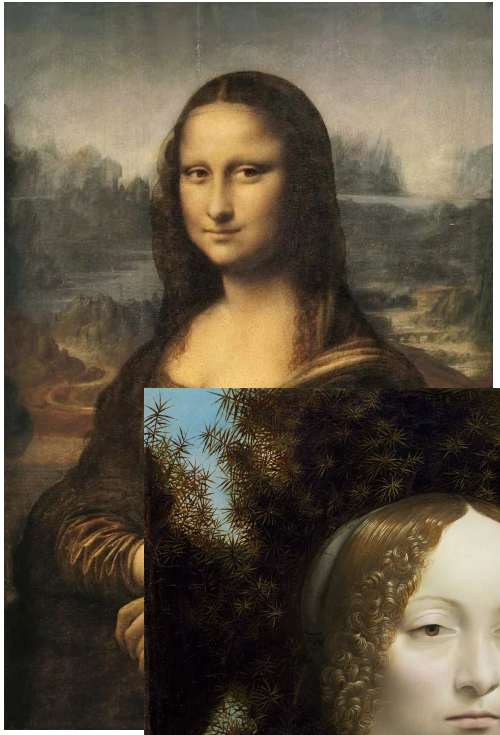
	VGG16	VGG19	Resnet18	Resnet 34
Accuracy On test data	76.33%	78.57%	79.97%	80.39%
Best Validation Accuracy	73.83%	76.56%	77.37%	77.68%
Training time	78m 57s	106m 1s	34m	38m
# of epochs	10	10	10	10

Final Steps & Results

Why do we choose Impressionist Artist?





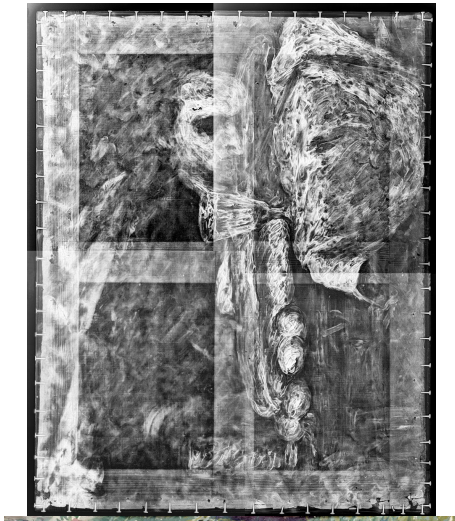


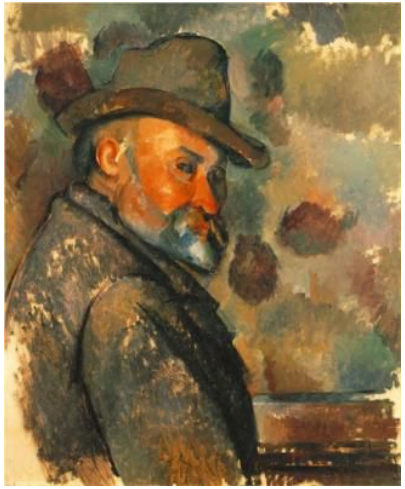
Train on another dataset

- Use Another dataset to train the model
  - 40 artist
  - 50 training images, 20 validation images for each artist

- Smaller dataset could still achieve good accuracy as long as the artist's style is well-distinguished from others
  - E.g. Monet 55.6% vs 69.4%
  - 5 artists from Renaissance achieved over 70% accuracy.

VGG - Potential Application





Conclusions
&
Future Applications

- Achieved relatively high accuracy in identifying paintings by different artists.
- Can be used for a variety of applications
  - Art authentication
  - Online Galleries
  - Historical research
  - Education
- Future research could focus on improving the model's performance, addressing the limitations and challenges, and expanding its applications.
  - For example, it could be interesting to apply similar techniques to other types of artwork, such as sculpture or architecture, or to study the evolution of artistic styles over time.

References (1/2)

[1]https://canvas.saatchiart.com/art/art-history-101/paul-gauguin-defined-by-his-paintings

[2]https://www.britannica.com/list/10-famous-artworks-by-leonardo-da-vinci

[3]https://www.tripimprover.com/blog/madonna-of-the-pinks-by-raphael

[4]https://stanford.edu/~shervine/blog/evolution-image-classification-explained

[5]https://arxiv.org/pdf/1512.03385.pdf

[6]https://www.youtube.com/watch?v=GWt6Fu05voI

[7]https://thedatafrog.com/en/articles/image-recognition-transfer-learning/

[8] https://claudemonetgallery.org/thumbnail/81000/81119/mini_icon/San-Giorgio-Maggiore-At-Dusk.jpg?ts=1

References (2/2)

[9]https://claudemonetgallery.org/thumbnail/81000/81120/mini_icon/Sandviken-Village-In-The-Snow.jpg?ts=1459229076
[10]https://www.vangoghstudio.com//Files/6/102000/102147/ProductPhotos/1920x1080/307257209.jpg
[11]https://uploads4.wikiart.org/images/paul-cezanne.jpg!Portrait.jpg
[12]https://th-thumbnailer.cdn-si-edu.com/oYqdpb4yivg--V7XZGFfM1g8-mk=/fit-in/1072x0/filters:focal(2117x2702:2118x2703)
[13]https://th-thumbnailer.cdn-si-edu.com/oYqdpb4yivg--V7XZGFfM1g8-mk=/fit-in/1072x0/filters:focal(2117x2702:2118x2703)/https://tf-cmsv2-smithsonianmag-media.s3.amazonaws.com/filer_public/6d/c2/6dc21502-b34d-407e-b305-103bfa1d208b/cezanne_xray_for_label.jpeg
[14]https://cdn.saam.media/8Dugj35lovqUxxWVZSyGskFc3Cc/2600/0/center/cover/webp/https%3A%2F%2Fd3ec1vt3scx7rr.cloudfront.net%2Ffiles%2Ffiles%2Fimages%2F1929%2FSAAM-1929.6.56_1.jpg