KLE Technological University

Creating Value
Leveraging Knowledge

Neural Style Transfer

Team

- Shreyas Dhotrad - 01fe18bec170
- Shridatha Hegde - 01fe18bec171
- Vishwajitt Kamat - 01fe18bec212
- Vishal S Teli - 01fe18bec209

Introduction

KLE Technological University
Creating Value
Leveraging Knowledge

Style transfer is the technique which allows us to apply the style of one image to the content of another image. Here we use image representations derived from Convolutional Neural Networks optimised for object recognition, which extract certain features from the image helpful for style transfer.

Objectives

- Convolution Model for feature Extraction.
- Extraction of feature level representations from image.
- Portability on android environment.

Content Image

Style Image

Output Image

Train Dataset

Imagenet

- The ImageNet dataset is a very large collection of human annotated photographs designed by academics for developing computer vision algorithms.

- More than 14 million images in the dataset
  - ➢ One thousand classes.
  - ➢ 1 million images that have bounding box annotations.

- Set of test images with manual annotation.

Source:Imagenet dataset preview

Test Dataset

National Gallery Art (Style Images)

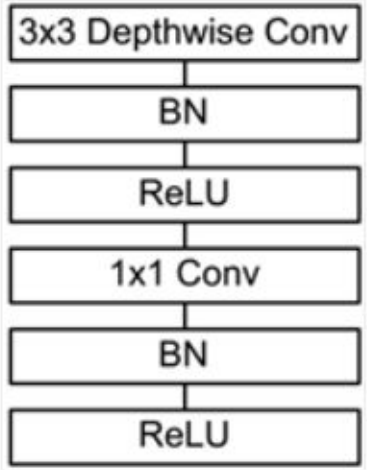
- Open source dataset for style images.
- Images of factual art objects.

Source:National Gallery of Art

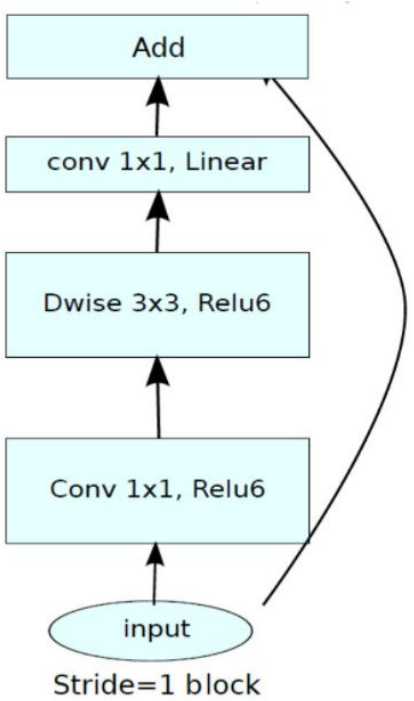
- Open source images taken from internet or local as a content image.

Model comparison

- Mobile models have been built on increasingly more efficient building blocks.
- MobileNetV1 introduced depthwise separable convolutions as an efficient replacement over traditional convolution layers.
- MobileNetV2 introduced the linear bottleneck and inverted residual structure in order to make even more efficiency.



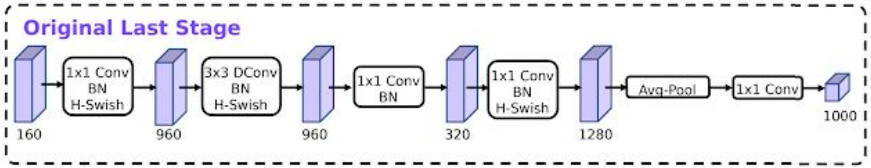
Source:MobileNet V1 Model

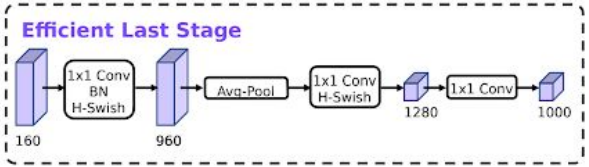


Source:MobileNet V2 Model

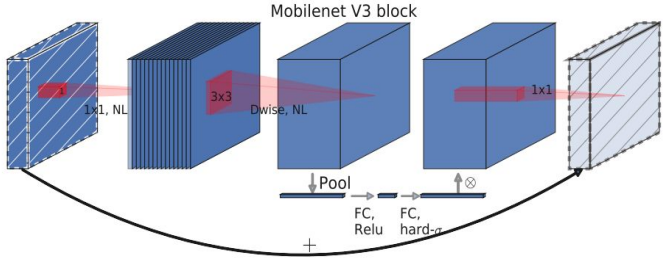
MobileNet-V3

- This architecture uses combination of:
  1.Depth-wise separable convolution.
  2.Linear bottleneck and inverted residual structure.
  3.Squeeze and excitation into the bottleneck structure.
- This architecture consists of 11 Bottleneck structures.





MobileNet-V3 Architecture

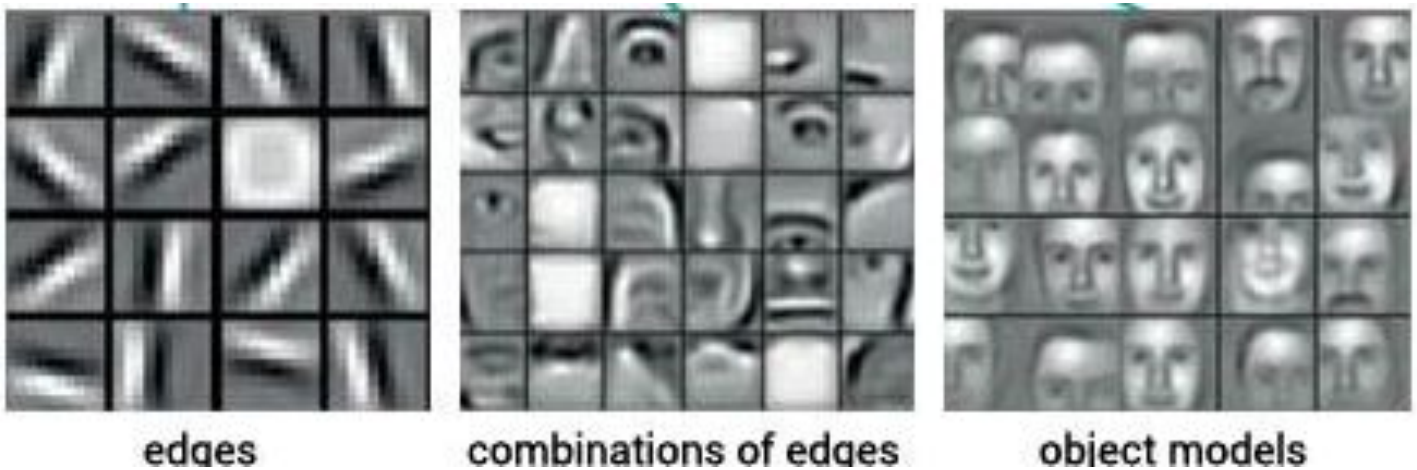


Input	Operator	exp size	#out	SE	NL	s
$224^2 \times 3$	conv2d, 3x3	-	16	-	HS	2
$112^2 \times 16$	bneck, 3x3	16	16	✓	RE	2
$56^2 \times 16$	bneck, 3x3	72	24	-	RE	2
$28^2 \times 24$	bneck, 3x3	88	24	-	RE	1
$28^2 \times 24$	bneck, 5x5	96	40	✓	HS	2
$14^2 \times 40$	bneck, 5x5	240	40	✓	HS	1
$14^2 \times 40$	bneck, 5x5	240	40	✓	HS	1
$14^2 \times 40$	bneck, 5x5	120	48	✓	HS	1
$14^2 \times 48$	bneck, 5x5	144	48	✓	HS	1
$14^2 \times 48$	bneck, 5x5	288	96	✓	HS	2
$7^2 \times 96$	bneck, 5x5	576	96	✓	HS	1
$7^2 \times 96$	bneck, 5x5	576	96	✓	HS	1
$7^2 \times 96$	conv2d, 1x1	-	576	✓	HS	1
$7^2 \times 576$	pool, 7x7	-	-	-	-	1
$1^2 \times 576$	conv2d 1x1, NBN	-	1024	-	HS	1
$1^2 \times 1024$	conv2d 1x1, NBN	-	k	-	-	1

Source:Searching for MobileNetV3

Methodology

- Use of Pre-trained MobileNet v3 model.
- Extraction of high level features for the content image from the deeper layers.
- Extraction of texture features for the style image from the shallow layers.
- Initialising the output image with random noise or content image.
- Appropriate loss function for the content image and the style image.



Model Outputs

Content Image

Style Image

Stylized Image

Model Outputs

Content Image

Style Image

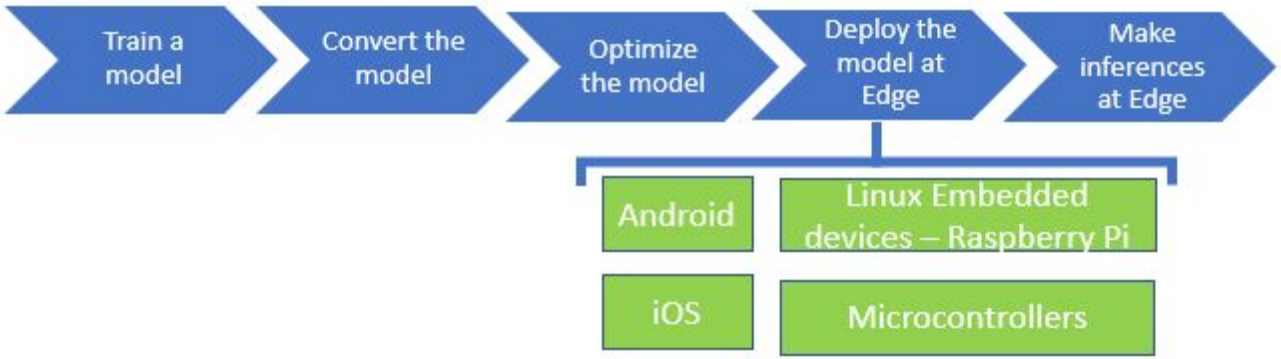
Styling Stepwise

Android Deployment

Requirements to interface at the Edge-

- **Light-weight**
- **Low Latency**
- **Optimal power consumption**
- **Pre-trained**
- **Optimization**

Tensorflowlite



Working-

- Select and Train a Model
- Convert the Model using Converter
- Save the Model

Tensorflowlite

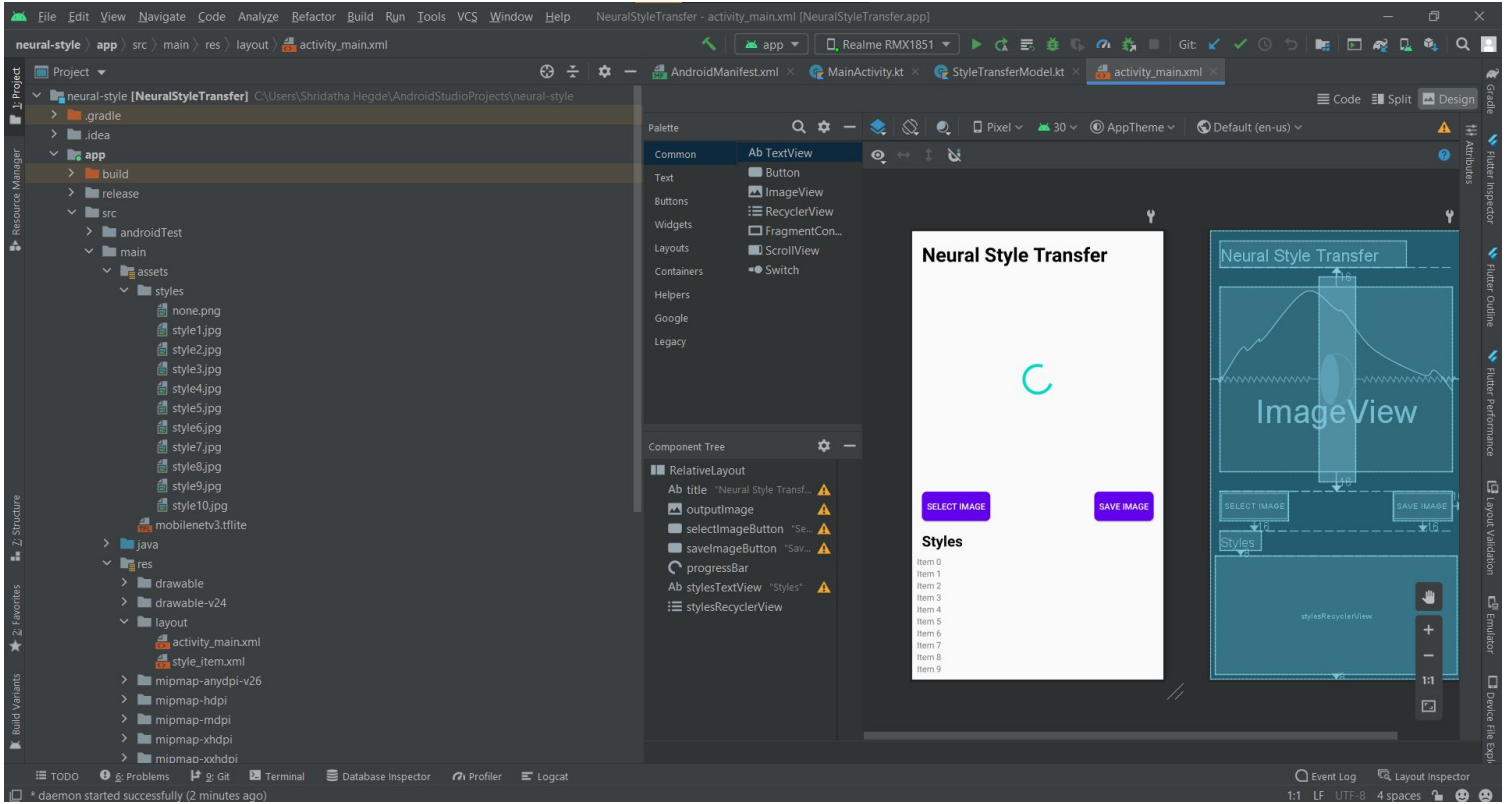
Need of Optimization of the Model

- Take less space on the Edge devices.
- Faster download time on networks with lower bandwidth.
- Occupy less Memory for the Model to make inferences faster.

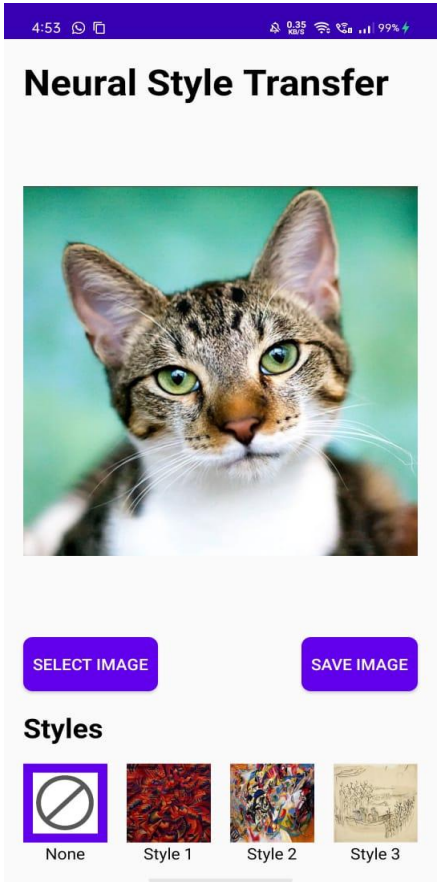
Optimization Methods

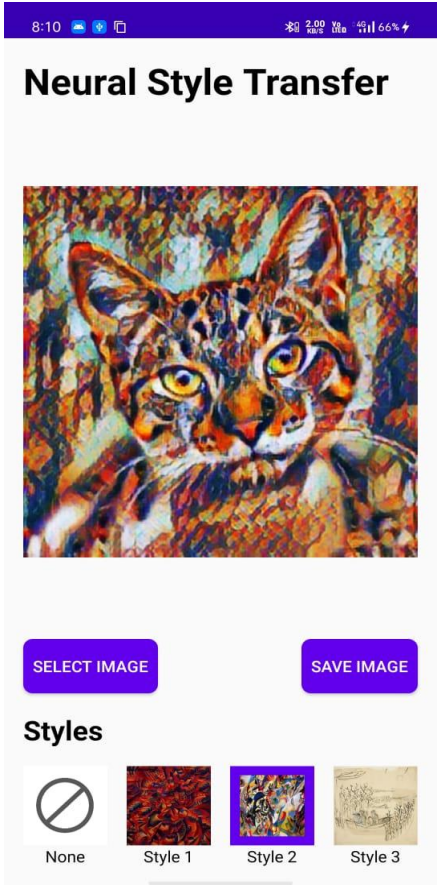
- Quantization
- Weight Pruning

Android Studio Implementation



Android Deployment





Thank you