Faculty of Computers and Artificial Intelligence

Computer Science Department

2021/2022

CS 396 Selected Topics in CS-2

Research Project

Report Submitted for Fulfillment of the Requirements and ILO's for Selected Topics in CS-2 course for Fall 2021

Team ID No.

	ID	Name	Grade
1.	201900320	زكريا علي زكريا مصطفي	
2.	201900465	عبدالله فتحي سيد محمدعليوه	
3.	201900592	لؤي صالح احمد محمد حسين	
4.	201900486	علي حسن علي هريدي	
5.	201900634	محمد اشرف عبد الباري عبد الله	
6.	201900049	احمد عادل الامير احمد	
7.	201900633	محمد اشرف خليفة علي	

Delivered to:

Dr. Wessam El-Behaidy

Eng. Salma Doma

Eng. Ahmed Nady

Architecture used in the paper: -

1-The paper used a residual neural network(resnet34).

2-A ResNet can be called an upgraded version of the VGG architecture, with the difference between them being the skip connections used in ResNets. In the figure below, we can see the architecture of the VGG as well as the 34 layer ResNet.

4-You might wonder why we have taken padding as 'same' only for all Convolution layers. The reason behind this is, we have to maintain the shape of our input until we add it to the residue. If the shape of the input gets changed, we will get a Numpy Error saying- "Two arrays with different shapes cannot be added".

5-Some of the points to note in this convolution block are, the residue is not directly added to the output but is passed through a Convolution Layer

3-The image size 224 x 224 pixels.
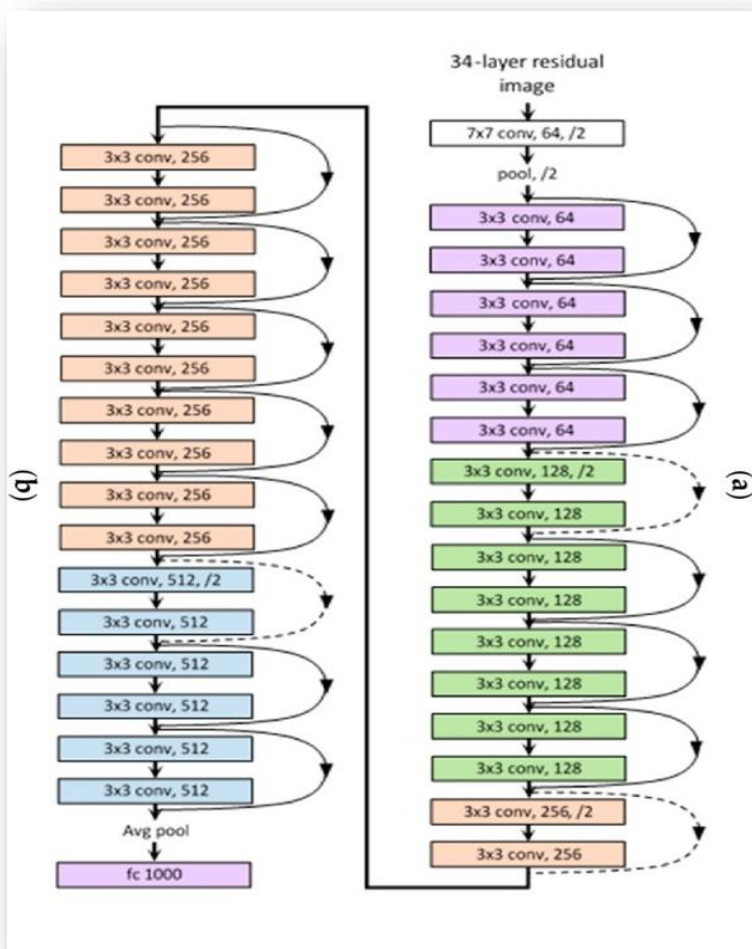
$$n_{out} = \left\lfloor \frac{n_{in} + 2p - k}{s} \right\rfloor + 1$$

n_{in}: number of input features

n_{out}: number of output features

k: convolution kernel size

p: convolution padding size

s: convolution stride size

Algorithm for Convolutional Block

```
X_skip = Input
```

```
Convolutional Layer (3X3) (Strides = 2) (Filters = f) (Padding = 'same') →(Input)
```

```
Batch Normalisation →(Input)
```

```
Relu Activation →(Input)
```

```
Convolutional Layer (3X3) (Filters = f) (Padding = 'same') →(Input)
```

```
Batch Normalisation →(Input)
```

```
Convolutional Layer (1X1) (Filters = f) (Strides = 2) →(X_skip)
```

```
Add (Input + X_skip)
```

```
Relu Activation
```

Algorithm for Identity Block

```
X_skip = Input
```

```
Convolutional Layer (3X3) (Padding='same') (Filters = f) →(Input)
```

```
Batch Normalisation →(Input)
```

```
Relu Activation →(Input)
```

```
Convolutional Layer (3X3) (Padding = 'same') (Filters = f) →(Input)
```

```
Batch Normalisation →(Input)
```

```
Add (Input + X_skip)
```

```
Relu Activation
```

ResNet-34 Structure and Code

layer name	34-layer	50-layer	101-layer
conv1	$7 \times 7, 64$, stride 2		
	3×3 max pool, stride 2		
conv2_x	$\begin{bmatrix} 3 \times 3, 64 \\ 3 \times 3, 64 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 64 \\ 3 \times 3, 64 \\ 1 \times 1, 256 \end{bmatrix} \times 3$
conv3_x	$\begin{bmatrix} 3 \times 3, 128 \\ 3 \times 3, 128 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$	$\begin{bmatrix} 1 \times 1, 128 \\ 3 \times 3, 128 \\ 1 \times 1, 512 \end{bmatrix} \times 4$
conv4_x	$\begin{bmatrix} 3 \times 3, 256 \\ 3 \times 3, 256 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 6$	$\begin{bmatrix} 1 \times 1, 256 \\ 3 \times 3, 256 \\ 1 \times 1, 1024 \end{bmatrix} \times 23$
conv5_x	$\begin{bmatrix} 3 \times 3, 512 \\ 3 \times 3, 512 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$	$\begin{bmatrix} 1 \times 1, 512 \\ 3 \times 3, 512 \\ 1 \times 1, 2048 \end{bmatrix} \times 3$
	average pool, 2048-d fc		

Dataset details: -

1-this project use an Audience data set created to use to Predict age's Person,

2-the Dataset contain image of Person's Faces,

3-the Dataset Contain 18000 image between women, men, Childs, adults, youth, babies, elderly.

4-classify the dataset into 18 Class

5-the Dataset contain people of age range (1-100)

6-The images cover large variation in pose, facial expression, illumination, occlusion, resolution
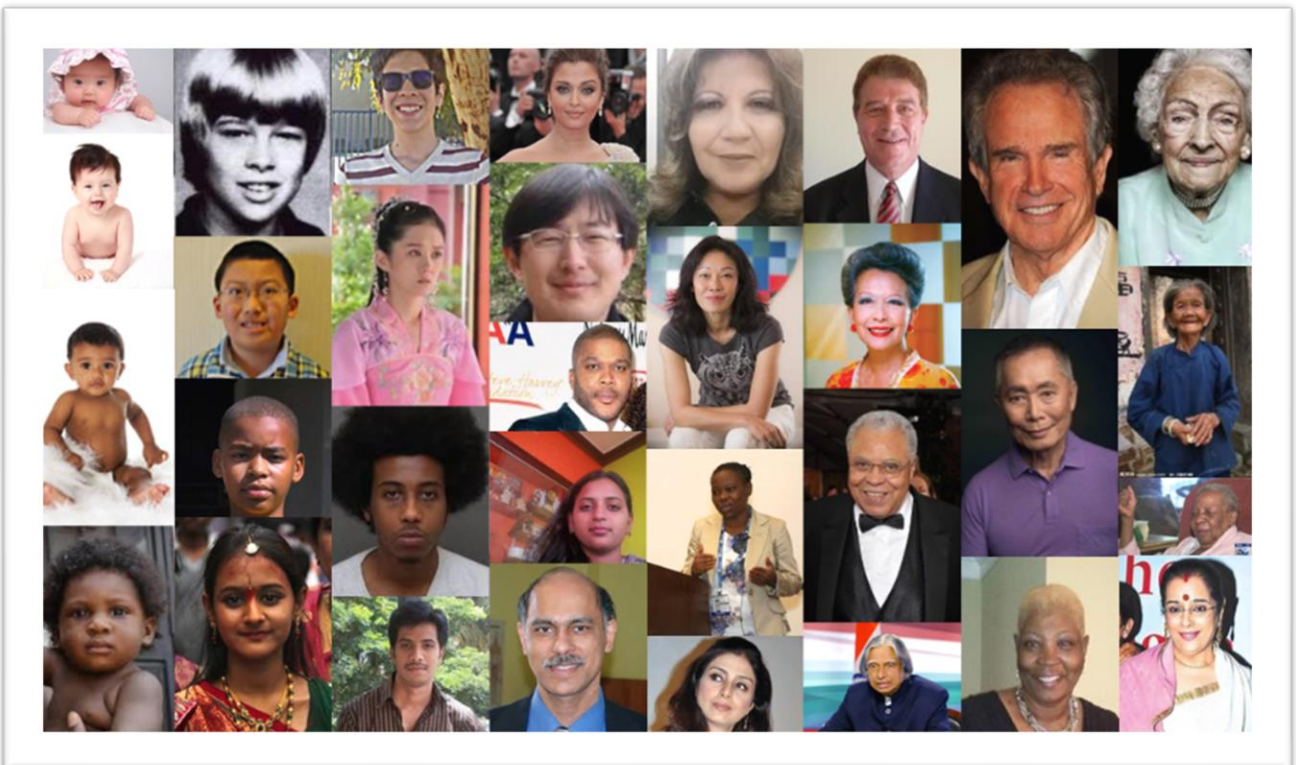
Labels:

The labels of each face image are embedded in the file name.

• [age] is an integer from 1 to 100, indicating the age

Implementation details: -

We have implemented the code using tensorflow library

& the tensorflow pipeline to preprocess the date then we used the famous resnet 34 architecture that have been published in 2015 then started building the model then we compiled it using the adam optimizer with initial learning rate of .0002, sparse categorical cross entropy , 50 epoch & verbose value of 2 to provide info about our model through its training process

(use the link provided here to see for yourself)

https://colab.research.google.com/drive/1bD_EjirKmbRHlP-C1AtohMoKj3tIm7G_?usp=sharing