Face-to-BMI: Live BMI Readings

-Pathik Rupwate

Introduction

The Body-Mass-Index is defined as the body mass divided by the square of the body height. BMI is an estimate of body fat and a good gauge of your risk for diseases that can occur with more body fat. The higher your BMI, the higher your risk for certain diseases such as heart disease, high blood pressure, type 2 diabetes, gallstones, breathing problems, and certain cancers. Being over normal weight is becoming more and more common in America.

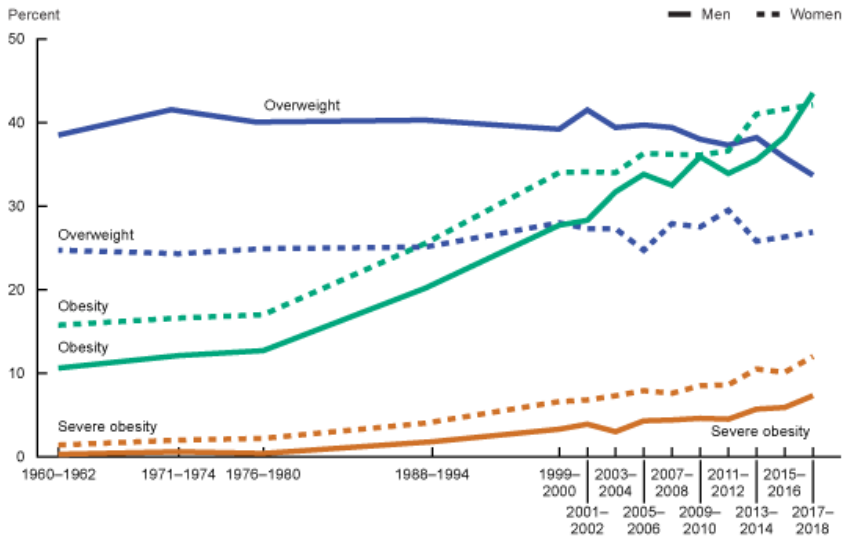

Fig 1: % of above-normal weight groups in America from 1980-2018

Data

We have the BMI Images dataset which contains a total of 4206 faces with corresponding gender and BMI information. Of these, seven were in the underweight range ($16 < BMI \leq 18.5$), 680 were normal ($18.5 < BMI \leq 25$), 1151 were overweight ($25 < BMI \leq 30$), 941 were moderately obese ($30 < BMI \leq 35$),

681 were severely obese (35 < BMI ≤ 40) and 746 were very severely obese (40 < BMI). The dataset contained 2438 males and 1768 females.

	Unnamed: 0	bmi	gender	is_training	name
0	0	34.207396	Male	1	img_0.bmp
1	1	26.453720	Male	1	img_1.bmp
2	2	34.967561	Female	1	img_2.bmp
3	3	22.044766	Female	1	img_3.bmp
4	4	37.758789	Female	1	img_4.bmp

Fig 2: BMI Dataset

We then map Male to 1 and Female to 0. We read in the images and split the data into train and test datasets with 20% test size.

Pre-trained Models

In this project we have used VGG-16 and VGGFace2 resnet50 models.

VGG-16 is a convolutional neural network. In VGG16 there are thirteen convolutional layers, five Max Pooling layers, and three Dense layers which sum up to 21 layers but it has only sixteen weight layers i.e., learnable parameters layer. The pretrained network can classify images into 1000 object categories, such as keyboard, mouse, pencil, and many animals. As a result, the network has learned rich feature representations for a wide range of images. The network has an image input size of 224-by-224

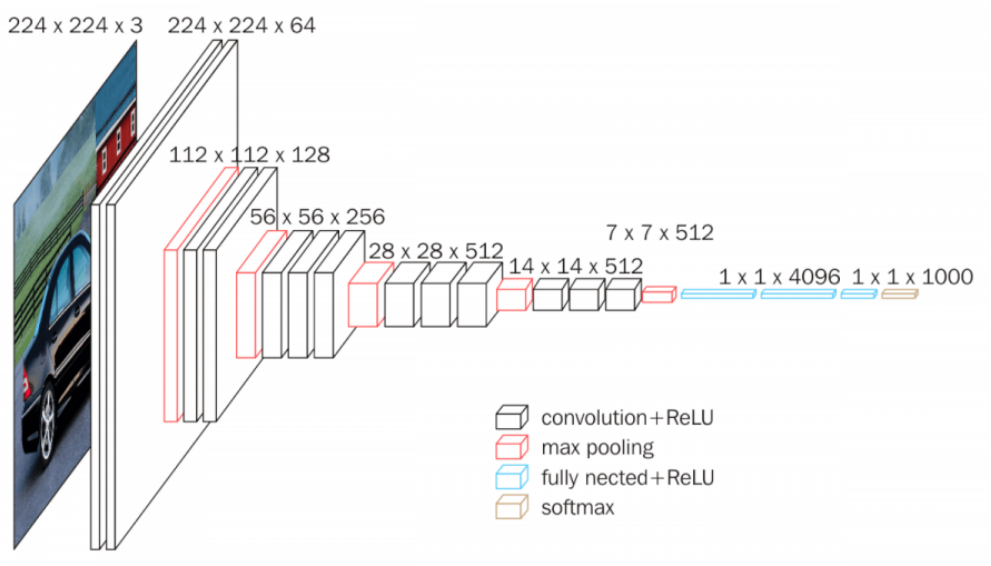

Fig 3: VGG-16 Model Representation

VGGFace2 is a large-scale face recognition dataset. Images are downloaded from Google Image Search and have large variations in pose, age, illumination, ethnicity and profession. The VGGFace2 dataset is made of around 3.31 million images divided into 9131 classes, each representing a different person identity.

The VGG-16 worked best with this dataset and is the one used for the final trained model.

Fine-tuning the Model

We load the pre-trained VGG16 model without the top layer and add a new fully connected layer for classification. We then add a Global Average Pooling layer, a dropout layer of 50%,a dense layer with gelu activation and a final Dense layer with gelu activation. To prevent overfitting, we added one dropout layer with a dropout of 50% into our model architecture. We also used Gaussian Error Linear Unit (Gelu) as an activation function; it combines the properties of the RELU activation function, Dropout, and Zoneout. Due to this, it tends to generalize better when there is more noise in the data so we used it in our models.

Evaluation

Mean absolute error of the final model is 7.88. This is pretty good and improves on the initial paper's findings.

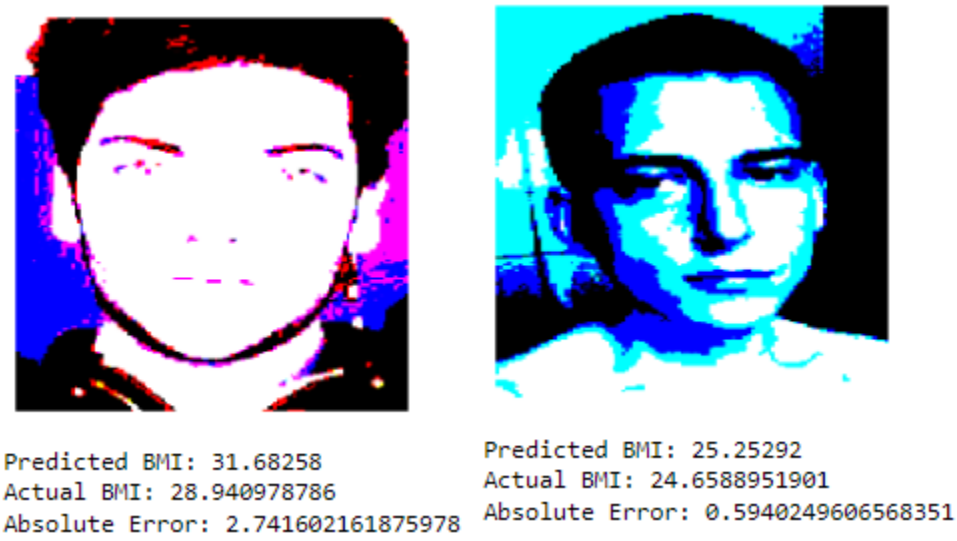

Fig 4: Predicted BMI of Test Images

BMI in Real Time with Camera Input

I used CV2 from Jupyter notebooks to read in image data from the webcam in live time and predict the person's BMI. The app shows the BMI in live time and saves the predicted BMI after the webcam video has ended.

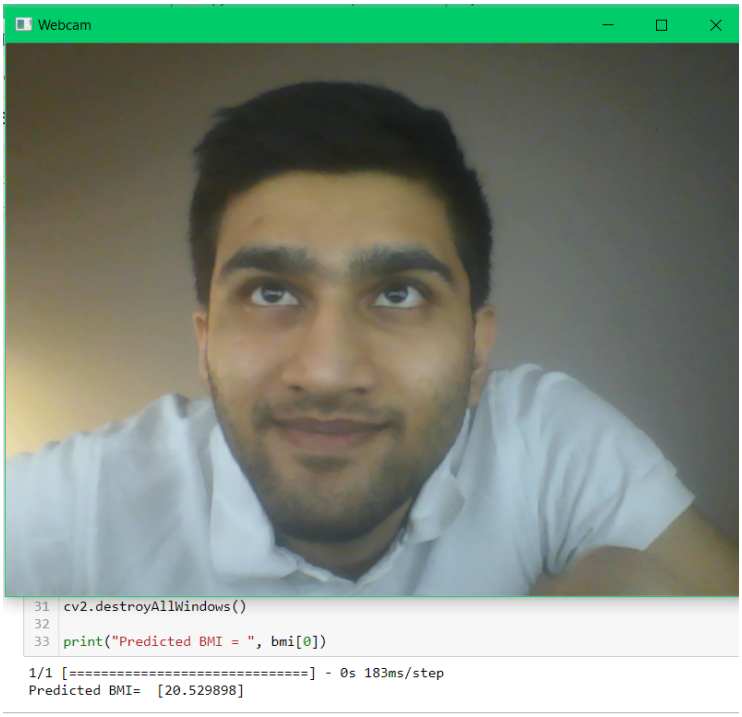

Fig 5: Predicted BMI using Real-time Camera Input

Future Improvements

Deploy flask on Google Cloud platform to make it the web API available on the internet.

Train with a larger dataset of faces.

Get more facial data from different regions around the world.

Apply more face recognition techniques to get better face images.

Apply more layers and fine-tune model further