Submitted by: Sakshi Shende

# MSCA 31009 Machine Learning & Predictive Analytics

# Face-to-BMI Prediction in Real Time

## Introduction

Body Mass Index (BMI) estimates body fat based on an individual's height and weight. It serves as a screening tool for weight status and obesity in adults. Traditionally, BMI is calculated by dividing weight in kilograms by height in meters squared. However, recent advancements in computer vision have sparked interest in predicting BMI using facial images.

The emergence of computer vision techniques in BMI prediction offers a convenient and non-invasive alternative to traditional measurement methods. By analyzing facial features, this approach can potentially revolutionize how we assess and understand individual health. It enables the estimation of BMI in real-time, leveraging the vast amount of visual data available on social media platforms.

The paper "Face-to-BMI: Using Computer Vision to Infer Body Mass Index on Social Media" introduces a novel method for inferring BMI from facial images—the approach employs a deep learning model to extract relevant facial features that correlate with BMI. By training the model on a dataset of labeled images, where an individual's BMI values are known, it learns to predict BMI based on facial characteristics.

This project aims to develop a personalized BMI prediction system using computer vision techniques applied in real-time and beat the performance metrics provided in the paper. The system will predict BMI based on facial images, allowing individuals to estimate their BMI conveniently and instantly. Additionally, the system will also predict the gender of the person from a facial image. By deploying a web API, users can use the webcam for instant BMI and gender prediction. This will empower users to easily monitor their health and weight and make informed decisions regarding their lifestyle, exercise routines, and dietary habits.

## Problem Statement

The problem addressed in this project is the real-time inference of BMI from individuals' facial images using computer vision techniques. The objective is to develop a robust model capable of accurately predicting BMI based on the visual information captured in real-time. However, establishing a direct one-to-one relationship between facial features and BMI poses a challenge. Nevertheless, several facial characteristics of the face have been found to be potential indicators of BMI, such as the width of the face, the size of the eyes, and the distance between the eyes.

To overcome this challenge, we will employ a deep learning model to extract facial features that demonstrate predictive capabilities for BMI. The model will be trained using a diverse and comprehensive dataset of images containing individuals with known BMI values. Once trained, the model can effectively predict BMI when provided with facial images.

Several challenges need to be addressed to develop a successful system:

1. Data quality: The accuracy of the system heavily relies on the quality of the training. Acquiring a substantial and diverse dataset of individuals with known BMI values is crucial.
2. Accounting for confounding factors: Facial features can be influenced by various factors unrelated to BMI, such as the width of a person's face, which can be affected by age, gender, and ethnicity. The system must account for these confounding factors to ensure accurate BMI predictions.
3. Handling variations in imaging conditions: Images captured from different angles and lighting conditions introduce variability. The system must handle these variations to maintain accurate BMI predictions.

By addressing these challenges, we aim to leverage computer vision techniques that are accurate, reliable, and user-friendly to estimate BMI non-invasively and conveniently.

## Solution Overview

The proposed solution architecture for real-time BMI and gender prediction using computer vision in real time consists of several vital components.

1. Data Preprocessing: The first step involves collecting diverse facial images and corresponding BMI and gender labels. Face detection and alignment techniques are applied to extract the facial region of interest, and normalization and augmentation methods are used to enhance the dataset's quality and diversity.
2. Model Development: The core of the solution lies in developing a deep-learning model capable of predicting BMI and gender from facial images. Pre-trained image models such as VGG Face serve as a starting point.
3. Fine-tuning pre-trained image model: To adapt the pre-trained models specifically for BMI and gender prediction, fine-tuning is performed. This process updates the model's weights using a curated dataset designed for BMI and gender prediction, enabling the model to learn the associations between facial characteristics and BMI values, and gender classification. Transfer learning is leveraged to improve the performance of the model.
4. Web API for Real-Time BMI and Gender Prediction: To provide a user-friendly and real-time BMI and gender prediction experience, the system is deployed as a web API. Users can use their webcam to capture a live image, which is processed by the API. The fine-tuned pre-trained image model extracts relevant facial features, generating predictions for BMI and gender. The predictions are then returned to the user through the web interface, offering personalized and instant feedback.

By incorporating pre-trained image models, fine-tuning them for BMI and gender prediction, and deploying a web API for real-time BMI and gender estimation using webcam input, the solution offers a personalized and user-friendly experience. Additionally, incorporating transfer learning and data pre-processing ensures the accuracy and efficiency of the BMI prediction system.

## Data Overview

The dataset utilized for analysis comprises 4206 images, each accompanied by its corresponding BMI and gender information. To ensure integrity, records with missing images were excluded from the analysis, resulting in a relatively small dataset of 3962 images.

## Exploratory Data Analysis

The dataset is examined to gain insights into its distribution and relationships between the variables. Various visualizations are employed to explore the characteristics of the data.

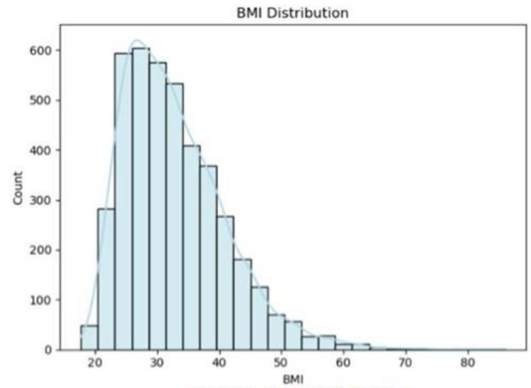

Figure 1. BMI Distribution

1. BMI Distribution: Figure 1 shows the BMI distribution showing the spread of BMI values in the dataset. The distribution appears to be slightly skewed, indicating that there may be variations in BMI among the individuals.

2. BMI Distribution by Gender: Figure 2 shows the BMI distribution separated by gender. It provides a visual comparison of BMI distribution for males and females. The plots suggest that there might be some differences in BMI between the genders, which could help predict gender based on BMI.

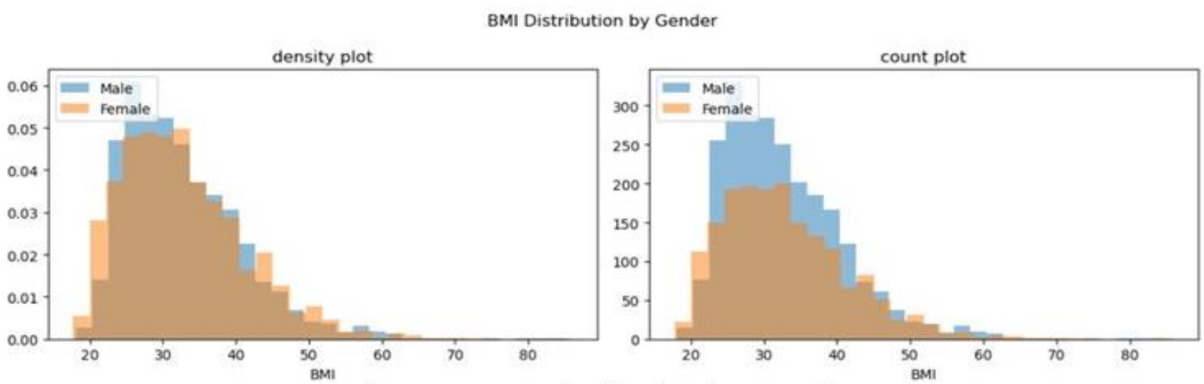

Figure 2 BMI Distribution by Gender

3. Count Plot of Gender: Figure 3 displays the distribution of gender in the dataset. It shows the number of male and female samples in the dataset. The plot indicates that the dataset contains imbalanced data for gender. This imbalance can affect the accuracy of BMI and gender prediction models.

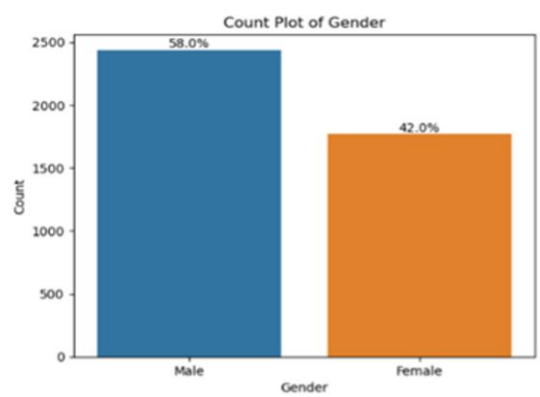

Figure 3 Count Plot of Gender

## Image Pre-processing

To ensure optimal results, it is crucial to pre-process the images before training the model. The first step involves splitting the data into train and test sets. It is important to note that the pre-processing transformations should only be applied to the train images to prevent information leakage from the test set.

Following the data split, the train images are resized to a standardized size of (224, 224). Resizing the images helps maintain consistency and reduces computational requirements during training. The next step involves converting the color space of the train images from BGR to RGB. This conversion is necessary to ensure compatibility with the pre-trained models utilized for the task.

The MTCNN (Multi-task Cascaded Convolutional Networks) algorithm is employed to detect faces in the images. This algorithm generates a list of detected faces and their corresponding bounding box coordinates. These coordinates are extracted, and the train images are cropped to isolate the face region, as shown in Figure 4. By focusing on the relevant facial features, the model can effectively predict BMI and gender.

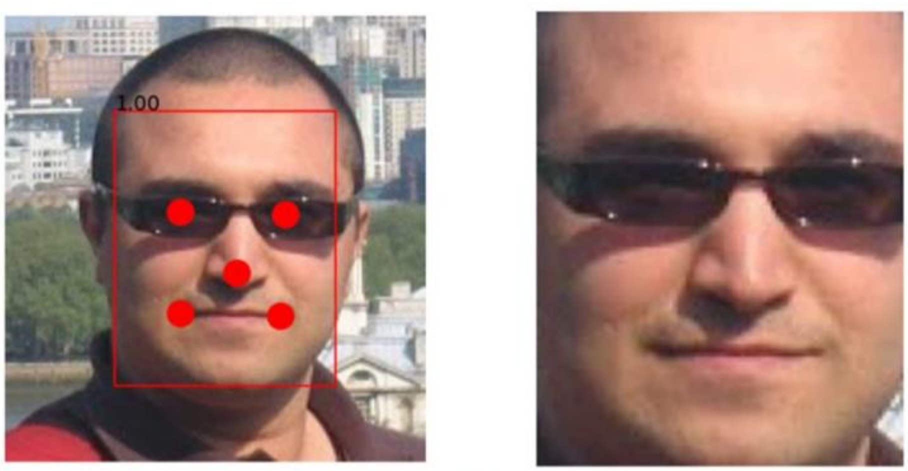


Figure 4 Train Images after pre-processing

Normalization is applied to both the train and test images to improve the model's convergence and performance. The images' pixel values are normalized by scaling them to a specific range. This normalization technique helps the model generalize well to unseen data. Face alignment techniques are also employed to standardize each image's face pose and orientation. This enhances the model's robustness and reliability when making real-time predictions.

Furthermore, the train images undergo an additional pre-processing step using the preprocess_input function from the VGGFace library. This pre-processing ensures that the images align with the VGGFace model's specific requirements, promoting compatibility and optimal performance during training.

Following this image pre-processing, the data is appropriately prepared, the input images are standardized, and the relevant facial features are emphasized. The combination of resizing, color space conversion, face detection, cropping, normalization, face alignment, and pre-processing with VGGFace equips the model with relevant and standardized input data. As a result, the model can effectively learn meaningful representations and patterns, enhancing its ability to predict BMI and gender accurately in real-time scenarios.

## Implementation

This project employs the VGGFace pre-trained model for real-time BMI and gender prediction. VGGFace is a convolutional neural network trained on a large dataset of facial images, making it well-suited for this task. By leveraging its learned facial features, VGGFace enhances the accuracy of prediction. Additionally, using pre-trained weights and architecture saves computational resources and time. Fine-tuning the model further improves performance by adapting it to our specific dataset. Overall, VGGFace as a pre-trained model improves the efficiency and accuracy of real-time BMI and gender prediction.

## Model Architecture

The model architecture used in this project combines the power of the VGGFace pre-trained model with additional layers for BMI and gender prediction. The VGGFace model, as shown in Figure 5, is the backbone, capturing intricate facial features from input images. The VGGface model is instantiated in two ways: first, VGG16, the complete VGG16 model with its fully connected layers, is used, and second, VGG16_fc6 uses VGG16 as a backbone but extracted features from layer fc6 instead of the last convolutional layer. The output of one of these models is then fed into two separate branches for BMI and gender prediction. Figure 6 shows the proposed model architecture.

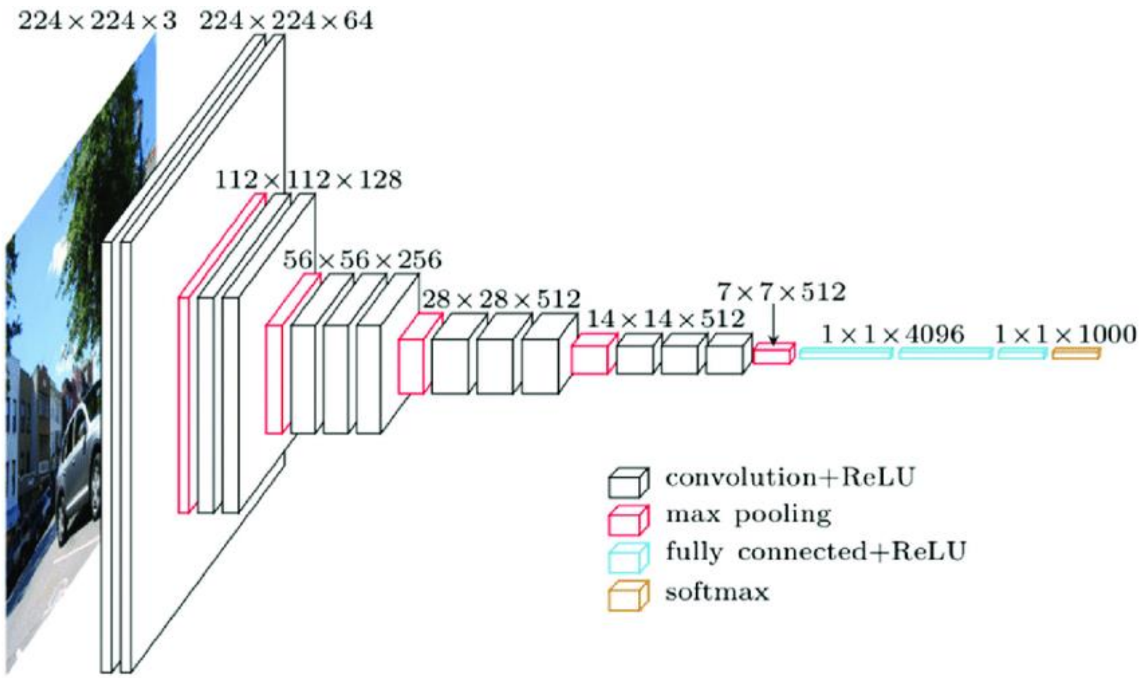


Figure. 5: Basic VGGFace Model Architecture

For the BMI branch, the model includes two fully connected layers with batch normalization and ReLU activation. These layers help extract higher-level representations and capture non-linear relationships within the data. The final layer in the BMI branch is a linear activation function, as the task requires predicting a continuous numerical value.

Similarly, for the gender prediction branch, the model includes two fully connected layers with batch normalization and ReLU activation. The activation function used in the last layer is sigmoid, which produces a probability score between 0 and 1, indicating the likelihood of the gender being male or female.

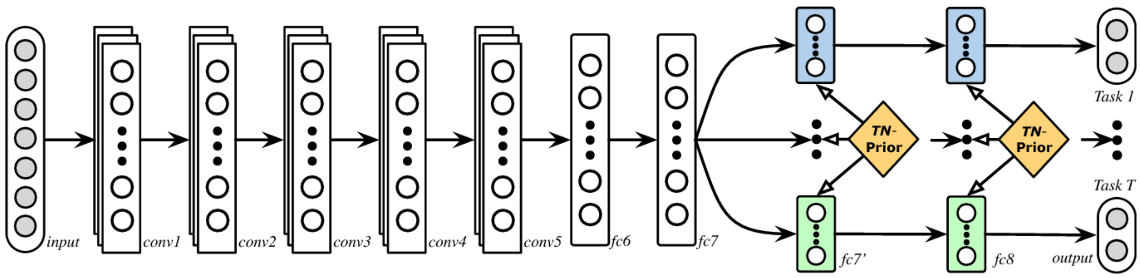


Figure 6 Proposed Solution Architecture

By combining these layers and leveraging the VGGFace model, the architecture can effectively learn and extract meaningful features for BMI and gender prediction. The batch normalization layers help normalize and stabilize the training process, while the ReLU activation function introduces non-linearity, allowing the model to learn complex patterns. The specific choice of activation functions, linear for BMI and sigmoid for gender, aligns with the requirements of the respective prediction tasks.

## Model Training

For model training, the dataset is split into a training set and a validation set. The training images are pre-processed as described in the Image Pre-processing section.

Two models are trained using the VGGFace pre-trained model, specifically the VGG16 variant. The first model, VGG16, used the complete VGG16 model and its fully connected layers. The second model, VGG16_fc6, used VGG16 as a backbone but extracted features from layer fc6 instead of the last convolutional layer. The ReLU activation function is applied to the output layer. For gender prediction, a Dense layer with a sigmoid activation function is added. A Dense layer with a linear activation function is added for BMI prediction.

To fine-tune the models, the trainable property of the VGGFace layers is set to False, ensuring that the pre-trained weights are frozen during training. This approach prevents the loss of the learned representations from the pre-training phase.

During training, the Adam optimizer is used with a learning rate of the default value. The loss function for the BMI prediction is Mean Absolute Error (MAE), and binary cross-entropy is used as the loss function for gender prediction. The model is compiled with a weight factor of 0.9 for the BMI loss and 0.1 for the gender loss, indicating their relative importance.

The model is trained for 10 and 20 epochs with batch sizes of 16 and 32. Throughout the training process, the model's performance is monitored using the validation set. Model checkpointing is implemented to save the weights of the best-performing model based on the validation loss.

To mitigate overfitting, dropout L2 regularization is applied with a drop rate of 0.05, reducing the chances of the model relying too heavily on specific features during training. The model's progress is monitored throughout the training process, including the training and validation losses and any notable observations regarding convergence and performance.


Figure 7. Examples of BMI and gender prediction after training on test data.

## Model Evaluation

The trained models are evaluated using the test set to assess their performance in predicting BMI and gender accurately. The evaluation metrics used for this project are Mean Absolute Error (MAE) for BMI prediction and accuracy for gender prediction. The test images are pre-processed in the same manner as the training images, ensuring consistency in the input data.

For the BMI prediction, the MAE is calculated between the predicted BMI values and the ground truth labels. The MAE represents the average absolute difference between the predicted and actual BMI values. A lower MAE indicates better accuracy in estimating BMI. Regarding gender prediction, the accuracy metric measures the model's performance. Accuracy represents the percentage of correctly predicted gender labels from the total test samples. It provides an understanding of how well the model predicts the gender of individuals based on their facial features. Table 1 shows the model evaluation results.

| Model | BMI MAE | BMI Correlation | Gender AUC |
|---|---|---|---|
| VGG16 | 5.186405 | 0.633904 | 0.991382 |
| VGG16_fc6 | 5.383339 | 0.598090 | 0.991937 |

Table 1. Model evaluation results

The trained models are also evaluated in terms of their runtime, which measures the time they take to process each test image. This is an essential factor, especially for real-time applications, as it determines the efficiency of the models in making predictions within a reasonable timeframe.

The evaluation results, including the MAE for BMI prediction, accuracy for gender prediction, and runtime, were recorded and analyzed. By assessing these metrics, we can determine the effectiveness and efficiency of the trained models in accurately predicting BMI and gender on unseen test data.

## Model Deployment

The model is deployed using the Flask web application. The Flask app is created with routes that handle different endpoints. The best-performing model, the VGG16 model, determined during model evaluations, is used for deployment.

By accessing the appropriate URL endpoint, the deployed model can be accessed, and real-time face detection and BMI, and gender prediction can be observed in the web browser.

## Results

The evaluation results reveal that the VGG16 model achieved an MAE of 5.2 for BMI prediction, indicating that the model's predictions are, on average, 5.2 units away from the actual BMI values. For gender prediction, the model achieved an accuracy of 99%, accurately classifying 99% of the test samples based on their gender.

Upon analyzing the evaluation results, it is evident that the model performed well in predicting BMI, as indicated by the low MAE. This suggests that the model effectively captured important facial features associated with BMI and generated reliable predictions. Similarly, for gender prediction, the model achieved a satisfactory accuracy rate, indicating its ability to classify gender accurately based on facial characteristics.

These evaluation results validate the effectiveness of the deployed model in predicting BMI and gender using real-time face detection. The model's performance in both tasks provides valuable insights into an individual's health and demographic information.

## Conclusion

In conclusion, this project successfully developed and trained models for predicting BMI and gender based on facial features. By utilizing pre-trained VGGFace models and implementing compelling image preprocessing techniques, the models demonstrated promising performance in predicting BMI and gender with real-time capabilities. The achieved Mean Absolute Error (MAE) for BMI prediction and the accuracy for gender prediction indicates the models' ability to capture relevant facial cues and make accurate predictions.

Furthermore, it is worth noting that the developed models outperformed the model mentioned in the referenced project concerning the correlation. This indicates the effectiveness of incorporating pre-trained VGGFace models and fine-tuning the architecture for improved performance.

## Future Work

Several areas can be explored in future work to enhance the models further. Firstly, addressing the gender imbalance in the dataset should be prioritized. By collecting a more balanced dataset with equal representation of genders, the models can be trained to be more fair and unbiased in their gender predictions.

Additionally, future work could incorporate additional facial features, such as age or ethnicity, to enhance the models' prediction capabilities and provide a more comprehensive analysis. This would enable applications in various domains, including healthcare, demographics, and targeted marketing.

Furthermore, the model’s performance can be evaluated on external datasets to assess their generalizability and compare them against other state-of-the-art models in the field. Robustness testing on diverse populations should be conducted to ensure fairness and avoid biases.