Introduction:

Body mass index (BMI) is a measure of body fat based on height and weight that applies to adult men and women. It is calculated by dividing weight in kilograms by height in meters squared. BMI is a simple and inexpensive way to estimate body fat and is widely used as a screening tool for obesity and its health risks, such as diabetes, cardiovascular disease, and some cancers. It is used to classify people as underweight, normal weight, overweight, or obese. A BMI of 18.5 to 24.9 is considered normal weight. A BMI of 25 to 29.9 is considered overweight. A BMI of 30 or higher is considered obese.

However, BMI has some limitations as it does not account for muscle mass, body shape, or fat distribution. Moreover, BMI may not be accurate for some populations, such as athletes, children, or older adults. In recent years, there has been a growing trend of using BMI to predict health risks. This is because BMI has been shown to be a good predictor of cardiovascular disease, type 2 diabetes, and some types of cancer.

One study found that people with a BMI of 30 or higher were twice as likely to develop heart disease as people with a BMI of 25 or lower. Another study found that people with a BMI of 30 or higher were three times as likely to develop type 2 diabetes as people with a BMI of 25 or lower.

These studies suggest that BMI is a useful tool for identifying people who may be at risk for health problems related to obesity. However, it is important to remember that BMI is not a perfect measure of body fat and should not be used as the sole basis for making health decisions. Therefore, researchers have been exploring alternative methods to estimate BMI using computer vision techniques.

Computer vision is a field of artificial intelligence that deals with the extraction of meaningful information from digital images or videos. Computer vision algorithms can be used to identify objects, track motion, and even recognize faces. In recent years, there has been a growing interest in using computer vision to predict BMI. This is because computer vision can be used to extract information about a person's body shape and size, which can then be used to calculate BMI.

One of the challenges of using computer vision to predict BMI is that it can be difficult to accurately extract information about a person's body shape and size from images. This is because images can be distorted by factors such as lighting, pose, and background.

The use of computer vision to predict BMI has several potential benefits. First, it can be used to identify people who may be at risk for health problems related to obesity. Second, it can be used to track progress towards a healthy weight. Third, it can be used to make informed decisions about health.

The use of computer vision to predict BMI is still in its early stages, but it has the potential to be a valuable tool for improving public health. In recent years, several studies have shown that computer vision algorithms can be used to predict BMI with a high degree of accuracy.

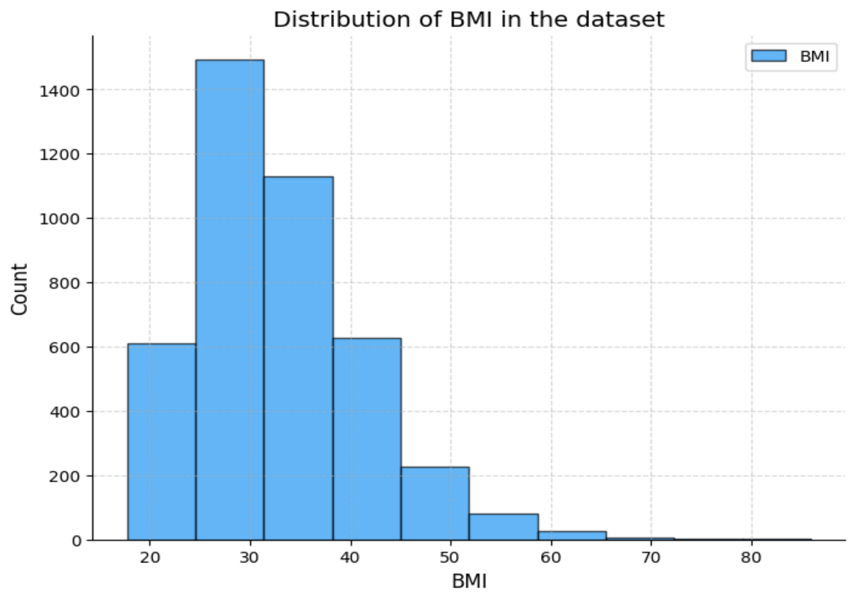
In this paper, I present a novel approach to infer a person's BMI in real time. I used pre-trained Haar cascade classifier for frontal faces and ResNet50 to detect BMI of the person. Deep learning is a branch of machine learning that uses artificial neural networks to learn complex patterns from large amounts of data. Neural networks are composed of layers of interconnected nodes that perform mathematical operations on the input data and pass the output to the next layer. The network can learn the optimal weights for each node by adjusting them based on the error between the predicted output and the actual output.

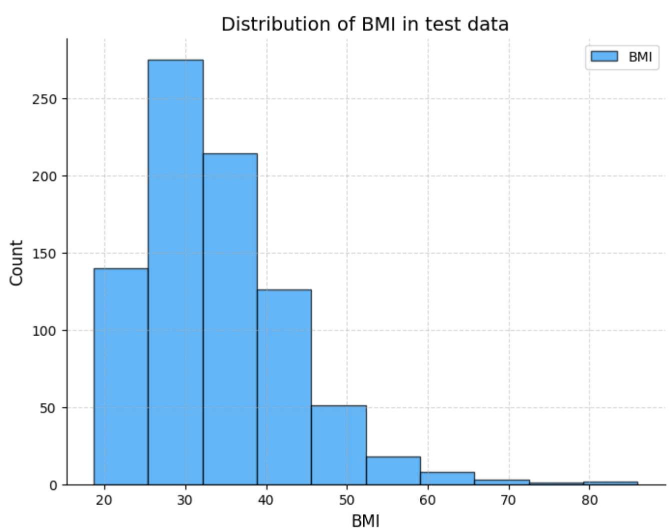
Methodology:

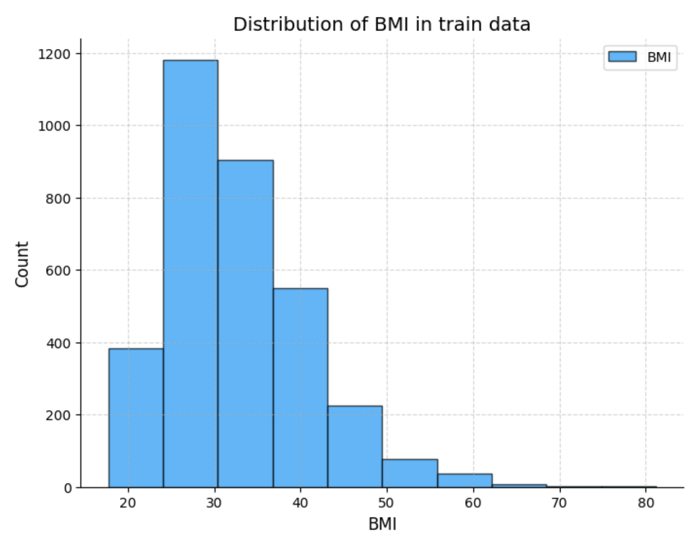
This section outlines the techniques used for building model. The dataset used in this research consists of labeled images depicting individuals under varying image qualities. The dataset comprises a total of 4000 images, each accompanied by attributes such as gender and body mass index (BMI). To evaluate the performance of the developed models, the dataset was split into training and testing sets, with 3210 images assigned for training purposes and 752 images reserved for testing.

- **Data Augmentation**: Given the limited size of the dataset, data augmentation techniques were employed to expand the available training data. By utilizing the ImageDataGenerator tool, a range of augmentation operations was applied to generate additional images. These operations included adjustments to the zoom level, modifications to image height, rotation transformations, and alterations to brightness levels. This augmentation process aimed to enhance the model's ability to adapt to real-world scenarios, as camera feed quality can vary due to environmental factors.
- **Validation Set:** To monitor the performance of the models during the training process and prevent overfitting, a separate validation set was created. This validation set consisted of 642 images from the original dataset. Throughout training iterations, model performance was assessed using this set, enabling the detection of potential issues such as overfitting or underfitting.
- **Preprocessing Techniques:** Prior to model training, appropriate preprocessing techniques were applied to ensure data compatibility with the selected models. These preprocessing steps aimed to enhance the quality and suitability of the images for subsequent analysis. Techniques such as normalization, resizing, and noise reduction were employed to standardize image attributes and mitigate potential variations.

The combination of data augmentation, the creation of a validation set, and application of appropriate preprocessing techniques facilitated a comprehensive exploration and processing of the labeled image dataset. Below is the distribution of the images.







Model Building:

A thorough exploration and assessment of several widely used pre-trained models, namely VGG, VGG-Face, and ResNetV2 has been don. The evaluation criterion employed to gauge their suitability was the root mean square error (RMSE) value, which served as a measure of the models' performance.

To ensure a judicious approach, the top layers of these pre-trained models were preserved in their original state, as they encapsulated high-level features and representations. By doing so, the inherent capabilities of these layers were leveraged, allowing for the extraction of meaningful and discriminative features.

However, to adapt the models to the specific requirements of the research, a fine-tuning strategy was employed. This process involved the modification of the lower layers of the pre-trained models, making them amenable to the specific task at hand. By selectively adjusting the architecture layers, the models were optimized to enhance their performance in the context of the research objectives.

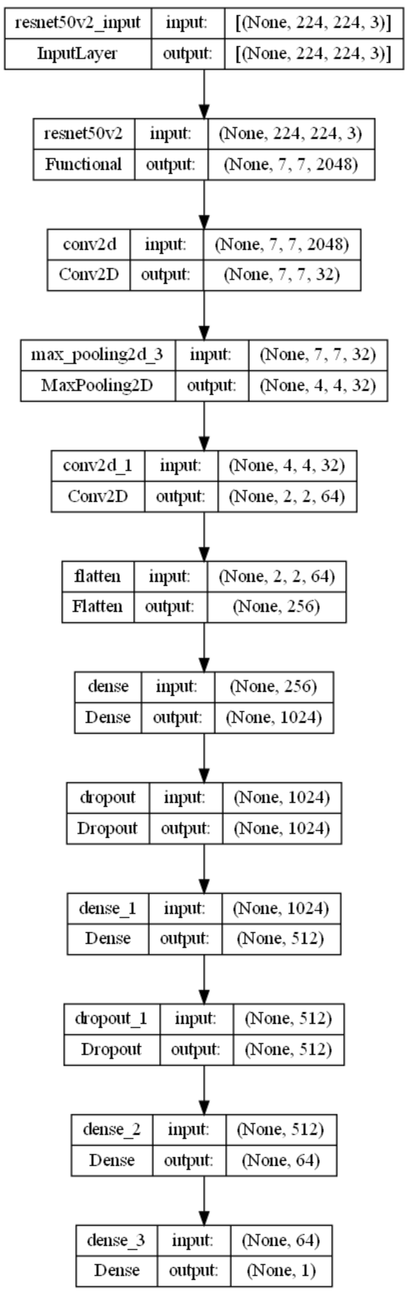
The careful fine-tuning of the architecture layers ensured that the pre-trained models were not only capable of capturing and leveraging high-level features but also had the flexibility to learn and refine lower-level representations. This approach aimed to strike a balance between utilizing the knowledge and expertise encapsulated in the pre-trained layers.

The BMI prediction model for regression was designed using various layers to effectively learn from the input data. The architecture includes two Conv2D layers, one MaxPooling2D layer, three Dense layers, two Dropout layers, and an output Dense layer with a 'relu' activation function.

To handle shape mismatches between layers, zero padding was applied. Zero padding helps maintain consistent dimensions throughout the network and ensures compatibility between layers.

The model was compiled using the Adam optimizer, a popular optimization algorithm known for its efficiency in handling large-scale problems. The mean squared error (MSE) was chosen as the loss metric, which measures the average squared difference between the predicted and actual BMI values. Minimizing the MSE loss helps optimize the model's ability to accurately predict BMI.

In order to track the model's performance and capture the best performing version, checkpointing was implemented. Checkpointing allows the model to save the weights and architecture at each epoch, and the best model based on validation metrics, such as MSE, is preserved.




```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 resnet50v2 (Functional)     (None, 7, 7, 2048)        23564800

 conv2d (Conv2D)             (None, 7, 7, 32)          262176

 max_pooling2d_3 (MaxPooling  (None, 4, 4, 32)         0
 2D)

 conv2d_1 (Conv2D)           (None, 2, 2, 64)          18496

 flatten (Flatten)           (None, 256)               0

 dense (Dense)               (None, 1024)              263168

 dropout (Dropout)           (None, 1024)              0

 dense_1 (Dense)             (None, 512)               524800

 dropout_1 (Dropout)         (None, 512)               0

 dense_2 (Dense)             (None, 64)                32832

 dense_3 (Dense)             (None, 1)                 65

=================================================================
Total params: 24,666,337
Trainable params: 5,570,273
Non-trainable params: 19,096,064
```

Fig 1: Model architecture and summary

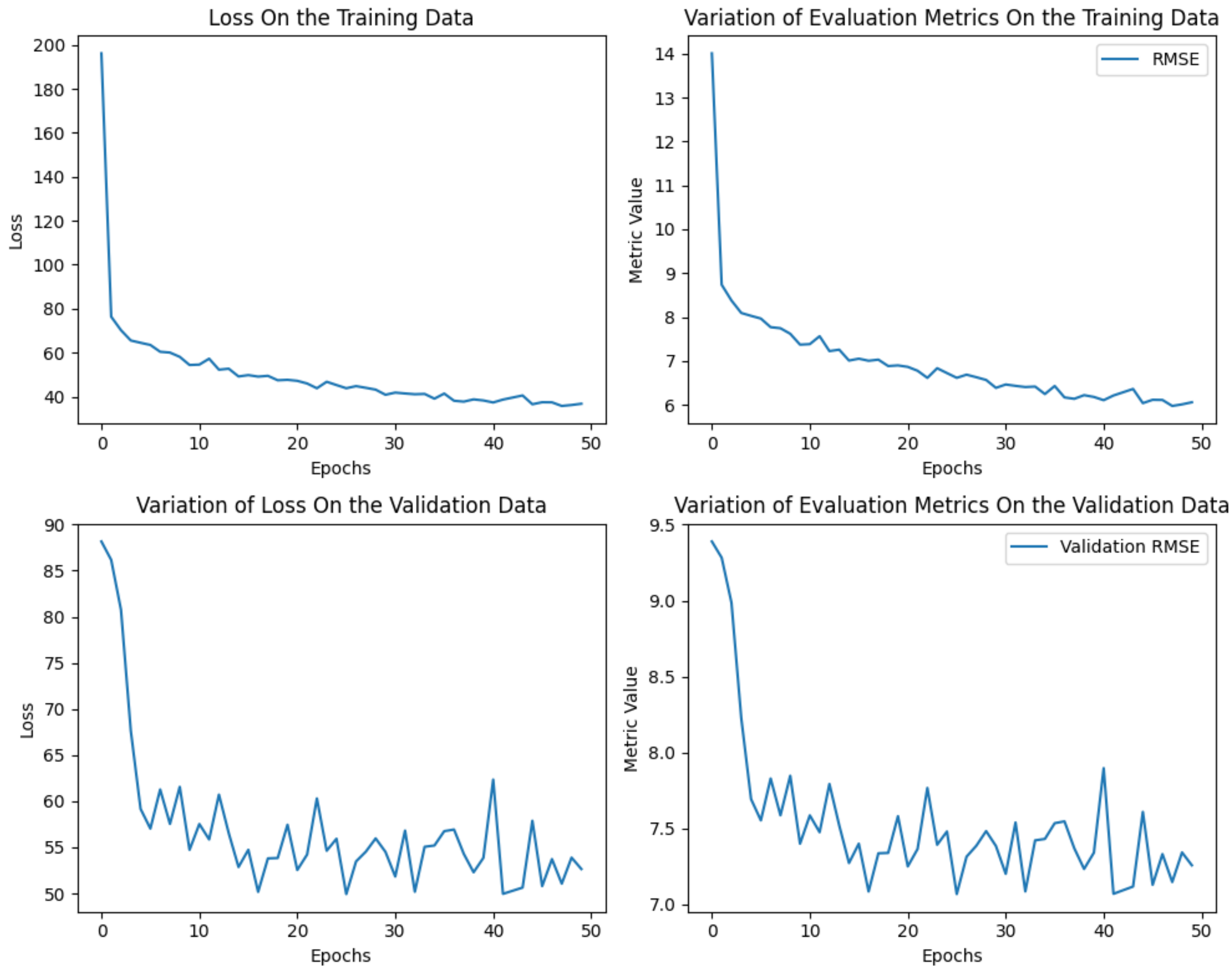
ResNet50 model exhibited a notably faster convergence rate in comparison to the other models under consideration. This observation implies that the ResNet50 model required fewer iterations to reach a satisfactory level of training performance.

Furthermore, during the training process, approximately 5.5 million parameters were updated and fine-tuned. This parameter count encompasses both the parameters inherited from the pre-trained models and the newly learned parameters specific to the research task.

The superior convergence speed of the ResNet50 model suggests that its architecture and pre-trained weights were particularly well-suited for the specific problem at hand. This efficiency in convergence can be attributed to the ResNet50 model's ability to effectively capture and leverage complex patterns and hierarchical representations within the dataset.

Results:

During the evaluation process, several metrics, including loss and RMSE, were utilized to assess the performance of different models, namely VGG Face, ResNet50V3, and VGG. Among these models, ResNet50 demonstrated the lowest validation loss. However, it is crucial to consider that model performance can be affected by various factors, such as the complexity of the architecture, the dataset's size and diversity, and the chosen hyperparameter settings.To gain a more comprehensive understanding of the models' capabilities, further analysis and experimentation are essential. Specifically, conducting experiments on larger and more diverse datasets would provide valuable insights into the models' generalizability and robustness.



Plots showing the training and validation loss and RMSE of epochs

Model achieved a training rmse is 5.97 and a validation rmse is 7.06.

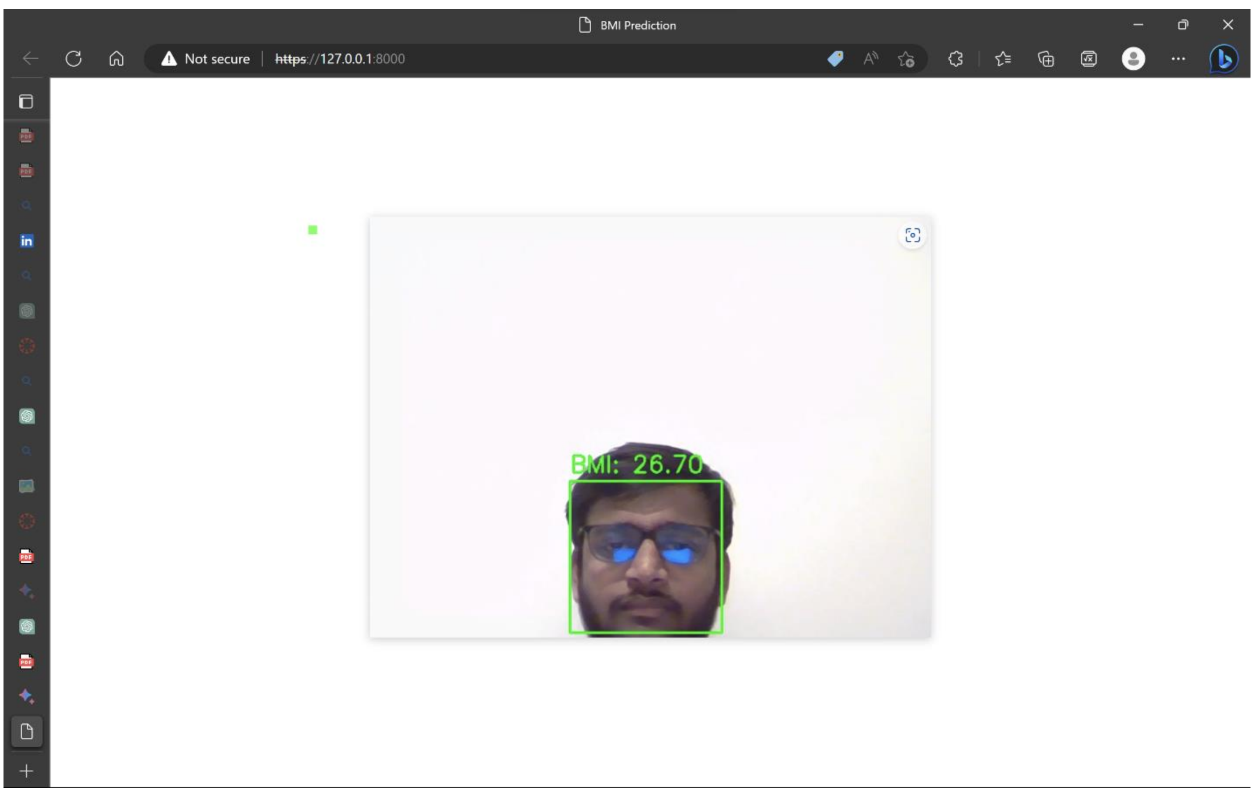
Deployment:

The fine-tuned BMI estimation model was successfully deployed as a web API utilizing the Flask framework. Flask, known for its lightweight and efficient nature, serves as an excellent platform for hosting machine learning models as web applications. By leveraging Flask, we have enabled real-time BMI prediction from facial images through a user-friendly web interface.

To facilitate local deployment, Flask web server and seamlessly integrated the BMI estimation model into the server using a RESTful API. Users now have the convenience of uploading their facial images via the web application and promptly receiving BMI predictions in response.

For accessibility and ease of use, the deployment code and comprehensive instructions available in a dedicated GitHub repository. This repository serves as a valuable resource for developers and interested parties seeking to deploy the BMI estimation model in their own environments.

The deployment of the BMI estimation model as a web API broadens its reach, offering the BMI prediction capability to a wider audience. Moreover, it opens possibilities for integration into various applications, including those related to fitness and health. This deployment represents an important step toward utilizing the BMI estimation model in diverse contexts and enhancing health-related platforms with its predictive capabilities.



Results

Conclusion:

In this study, we conducted research on the development and fine-tuning of deep learning models aimed at BMI estimation using facial images. Model exhibited promising performance, demonstrating high accuracy and low error metrics. BMI estimation models provided reliable predictions of BMI based on facial features.

The results highlight the potential of deep learning models in accurately estimating BMI from facial images. These models have valuable applications in various fields, including healthcare, fitness tracking, and body composition analysis. By leveraging deep learning techniques, we can extract meaningful insights and predictions from facial images, thereby advancing research in computer vision and machine learning.

However, it is crucial to acknowledge the limitations of our study. Our models were trained and evaluated on a specific dataset, and their performance may vary when applied to different populations or datasets with distinct characteristics. Therefore, it is essential to further validate the models using larger and more diverse datasets to assess their generalizability and robustness.

For future work, we propose deploying the model as a cloud application to enhance scalability and accessibility. One potential approach is to containerize the model using Docker and deploy it on cloud platforms like Heroku or Google Cloud. Alternatively, the model could be developed as a mobile application using Android Studio. Such deployments would enable easy access to the BMI estimation service from anywhere, facilitating broader adoption and utilization.

In summary, our study demonstrates the effectiveness of deep learning models for BMI estimation from facial images. While further validation is needed, the potential applications of these models in various domains are promising. By deploying the models as cloud or mobile applications, we can enhance their accessibility and usability, contributing to advancements in computer vision and machine learning research.