

A Deep Learning Approach for Non - Invasive Body Mass Index Calculation

¹S. Harish Nandhan ²J. Remoon Zean ³A.R. Mahi ⁴R.Meena ⁵S.Mahalakshmi
^{1,2,3,4,5} Department of Artificial Intelligence and Machine Learning, Rajalakshmi
Engineering College, Chennai, India

^{1,2,3,4}{harishnandhan02, remoonzeanj, ar-
mahi2003, meenaramalingamspr}@gmail.com ,
⁵mahalakshmi.s@rajalakshmi.edu.in

Abstract: A person's body mass index (BMI) is a vital indicator of their health. Through extensive research, it was determined that traditional methods of BMI calculation can be time-consuming. We developed a technique that uses deep learning to predict a person's BMI, age, and gender from facial images. Our system uses Multi-Task Cascaded Convolutional Neural Networks (MTCNN) to detect faces by cropping the face out of an input image and detecting facial landmarks based on the 5-point facial landmark detection algorithm from the input image. The cutting-edge pre-trained models will be finetuned, including VGG-Faces, ResNet (Residual Neural Network), and VGG16 (Visual Geometric Groups with 16 layers) on a public dataset of 1530 prisoners from Polk County Prison. This dataset contains a multifarious range of faces, including different races, ages, and genders. The uploaded CSV file contains the heights, weights, and gender of the training images. After the image is passed into the network, it generates BMI, age, and gender predictions for the input image. This system uses an efficient face recognition mechanism to identify the age, gender, and BMI of a person using a multi-task BMI prediction model.

Keywords: BMI (Body Mass Index) Estimation · Gender Prediction · Deep learning · Healthcare · VGG-Faces · ResNet (Residual Neural Network) · VGG16(Visual Geometric Groups with 16 layers) · MTCNN (Multi-Task Cascaded Convolutional Neural Network) · Face detection · Facial landmarks · Face recognition · Fine-Tuning

1 Introduction

The abnormal level of body fat which leads to a chronic health condition known as obesity. Obesity leads to various diseases such as diabetes, heart attack and sometimes leads to death. A person is affected by obesity when he or she has a Body Mass Index (BMI) value above 30. BMI is calculated by measuring the height and weight of individual [1].

$$BMI = Weight / (Height)^2 \quad (1)$$

here,

Weight in (kgs) and Height in (m)

Technology has advanced recently, revolutionizing many facets of our existence, including healthcare. The prediction of body mass index (BMI) using facial photographs

is a fascinating field of research that has attracted a lot of interest [2]. BMI is frequently used to gauge a person's weight status and general health; however, conventional methods of calculating BMI rely on self-reported height and weight, which can be arbitrary and prone to inaccuracy. An innovative and maybe more precise method of evaluating one's health is provided by the incorporation of facial image analysis into BMI prediction [3]. A person's facial features, adiposity patterns, and other physical characteristics can all be seen in facial photographs, which reveal a lot of information about them. Researchers have developed models that can predict BMI with astounding accuracy by utilizing powerful machine learning algorithms and computer vision techniques [4]. These models have been able to extract useful information from facial photos. The system architecture of the project consists of four key components: data preprocessing, face detection and alignment, feature extraction, and BMI prediction. In the data preprocessing phase, facial images and corresponding BMI values are standardized and normalized [5]. To ensure correct capture of facial features, face detection and alignment are essential. Feature extraction involves extracting facial landmarks, which serve as input to the deep learning model. The deep learning model Visual Geometric Groups (VGG16), based on a convolutional neural network (CNN) architecture, utilizes these facial landmarks to predict the corresponding BMI values and detect gender [7].

2. Related Works

In [9,10] they used to analyze obesity based on facial images. They seek to create precise models for obesity categorization by training deep neural networks on massive datasets of labelled facial photos. Deep learning and face image analysis combined have the potential to improve obesity detection and monitoring, resulting in better healthcare interventions and individualized treatment plans. They used to estimate the body mass index (BMI) through facial photos. These studies investigate the possibility of non-intrusive and widely available methods to measure BMI. In [13] they concentrated on estimating visual BMI from facial images using labelled distribution-based techniques. These studies look at the viability of using this technique to calculate BMI in a precise and reliable way. The main goal of the project is to predict BMI using sophisticated statistical methods in conjunction with computer vision technology illuminating non-invasive and routinely used for health assessments. The research advances the fields of image analysis and individualized health care.

In [11] they used to predict human height, weight, and BMI from facial photos. The merging of machine learning methods with face image analysis is highlighted in this study area as a promising strategy for forecasting anthropometric measurements. The research advances the fields of personalized health management and healthcare. In [25] Using real-time image processing and machine learning techniques, they were able to analyze and spot malnutrition from facial pictures and

estimate BMI. The technology provides a rapid and non-invasive method for determining BMI and measuring the risk of malnutrition by examining facial features and comparing them with nutritional indicators. The incorporation of real-time processing enables prompt intervention and individualized health monitoring, which may be advantageous to medical professionals and people trying to lead healthy lives. In [12] Examining the estimation of BMI from the provided facial photos, they applied a method known as semantic segmentation-based region-aware pooling. They used these techniques on facial photos to create a method that precisely predicts BMI. In order to provide insights into individualized health evaluation and interventions, this study area places an emphasis on the combination of cutting-edge image processing methods with BMI prediction. The results have implications for computer vision research and healthcare applications.

In [15] this they used a residual regression model to estimate BMI from facial images. The goal of the study is to use residual regression techniques on face image data to create a model that accurately predicts BMI. The residual regression model is used in these studies to examine the possibility of identifying minute differences in facial features linked to BMI. This field of study focuses on the integration of sophisticated regression methods and face image analysis, offering insights into non-intrusive and individualized BMI calculation. The result advances machine learning and healthcare applications, enabling better health assessments. In [14] They mostly concerned with estimating body mass index (BMI) from images using deep convolutional neural networks (CNNs). By utilizing the strengths of deep learning and CNN architectures on photographic data. These studies investigate the potential of this method for non-invasive and widely accessible BMI determination by training and optimizing CNN models on huge datasets of labelled photos with matching BMI values. This field of study has a strong emphasis on the combination of deep learning methods and image analysis for BMI prediction, revealing information on individualized health assessment and treatment. The results have applications in the domain of computer vision and medicine. In [16] They focused on a cutting-edge method for classifying body mass index (BMI) using voice cues, with the hope of potential therapeutic uses in the future. The article includes a preliminary examination of the creation of a method for precisely classifying BMI from input speech patterns. These studies investigate the possibility of this novel method for BMI categorization by examining the properties and patterns of voice signals. The combination of signal processing technologies and BMI measurement is the focus of this study topic, which sheds light on non-invasive and alternative approaches to health monitoring. The discoveries make a contribution to the discipline of voice analysis and its possible clinical applications. In [24] They used a cutting-edge method that calculates Body Mass Index (BMI) from facial recognition photos. The technology can

predict a person's BMI by looking at their facial features, offering a non-intrusive and possibly practical method of health evaluation. This method might provide insightful information on health monitoring and support efforts to study and treat obesity. To achieve accurate and trustworthy BMI estimation, additional validation and improvement are needed.

3. Dataset

We are utilizing the Polk County prison dataset and arrest information database (Jail & Arrest Information - Polk Inmates - Polk County), which has 1530 records and 16 columns as a csv file, as well as a folder full of .jpg files that show the faces of convicts [17]. According to the data, there are 20% female inmates and 80% male inmates. The minimum age in the dataset is 18, and the average age is 34. The age distribution in the sample follows a truncated normal distribution. Native Americans who are Black and White make up the majority of the dataset's participants, whereas samples from other continents are scarce. The individuals' BMI values often fall between 20 and 30, and there is no association between the dataset's features.

	name_id	age	height	weight	race	sex	eyes	hair	BMI
0	7482	54	1.8034	127.00576	Black	M	Brown	Black	39.051641
1	754952	26	1.8034	95.25432	Black	M	Brown	Black	29.288731
2	644421	24	1.7526	131.54168	White	M	Green	Blonde	42.825039
3	699804	21	1.6002	58.96696	Black	M	Brown	Black	23.028211
4	238047	29	1.8796	104.32616	White	M	Blue	Blonde	29.529925

Table1. Attributes of the dataset

4. Methodology

4.1 Model Architecture

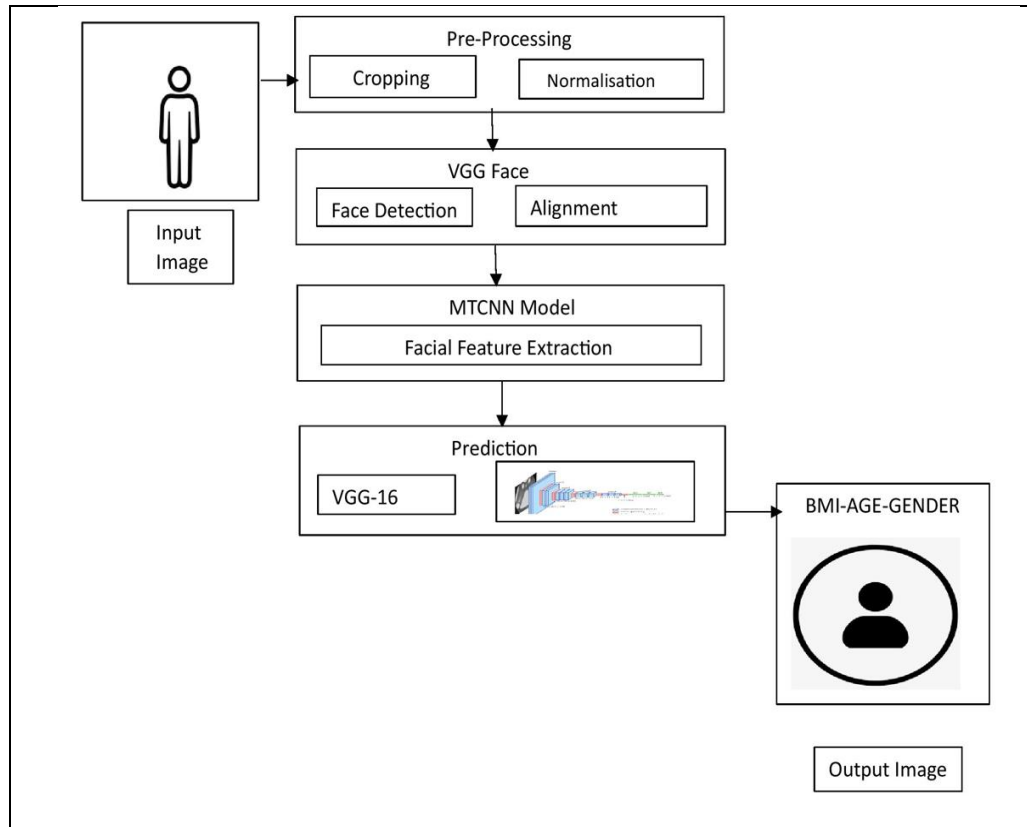


Fig1. Architecture diagram of Facial BMI predictor

4.2 Data Flow

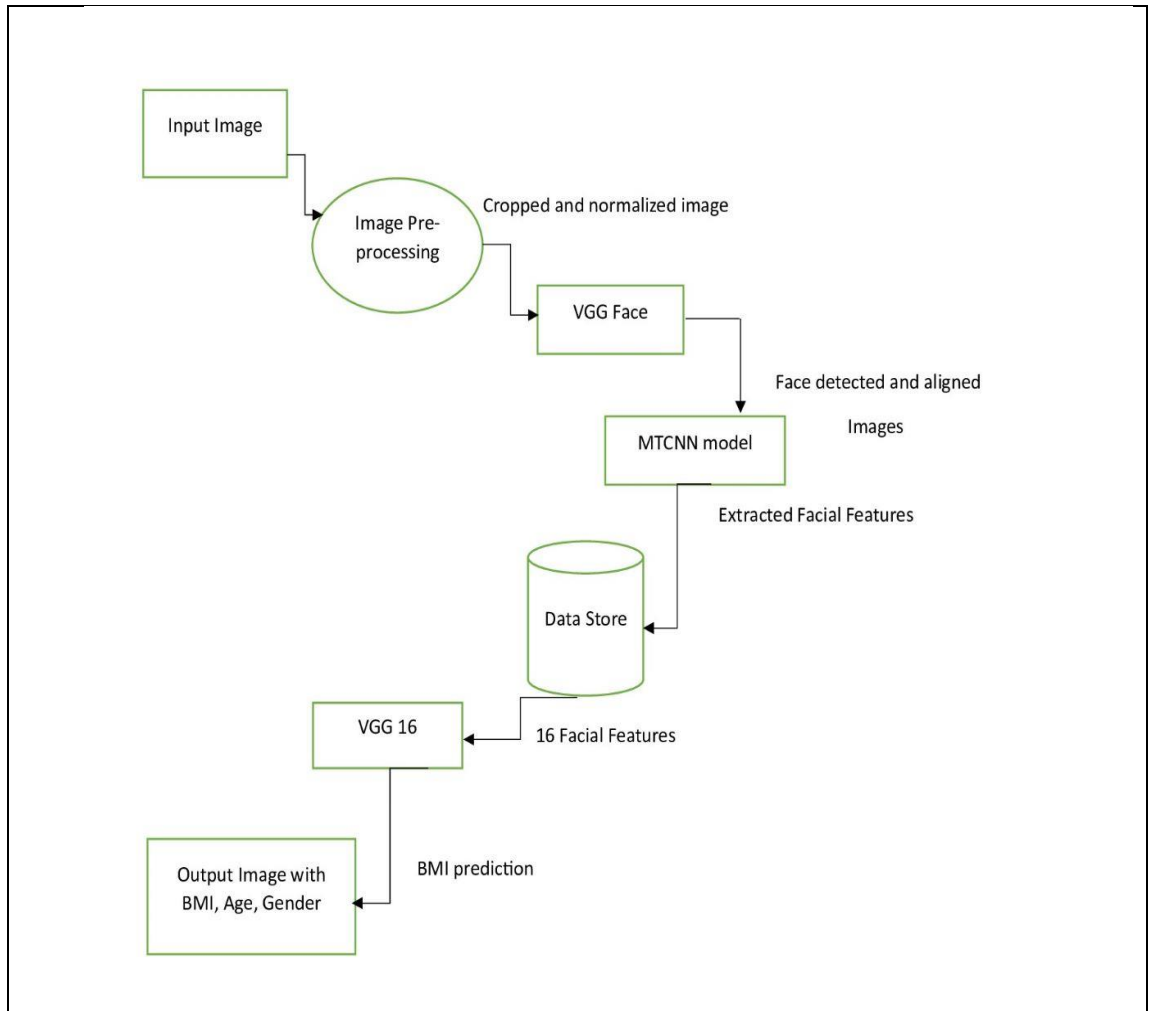


Fig2. Data flow diagram

When a picture is input, the data flow process starts, and the image goes through pre-processing stages like scaling, transformation, cropping, and normalization. After classifying the face and drawing a bounding box around it, the VGG face algorithm recognizes the landmarks on the face, such as the eyes, nose, and mouth. The face features will be extracted by the MTCNN, which will then link to the information kept in the data storage. Through the VGG16 backbone, adjusted weights and by tuning hyperparameters, the model is trained. The final image produced by the model will include the BMI, age, and gender.

4.3 Data pre-processing

To prepare the dataset for creating a predictive model, data pre-processing is carried out. The preprocessing steps, which include data cleaning, data transformation, and data normalization. The dataset should be thoroughly examined to identify duplicate values and to fill in the missing values using a technique known as data imputation [18]. By removing the outliers from the dataset, it also preserves the dataset's integrity. Data Normalization [19] is a technique used to scale data to a standard range. Due to their significant magnitudes, it avoids any bias towards particular traits. Additionally, Data Transformation aids in improving the distribution of the variables [20]. The dataset can be split into training and testing sets for performance evaluation, or we can utilize entirely new data to test the model. The training set is used to fit the BMI prediction model, while the test set is used to compare the trained model's performance.

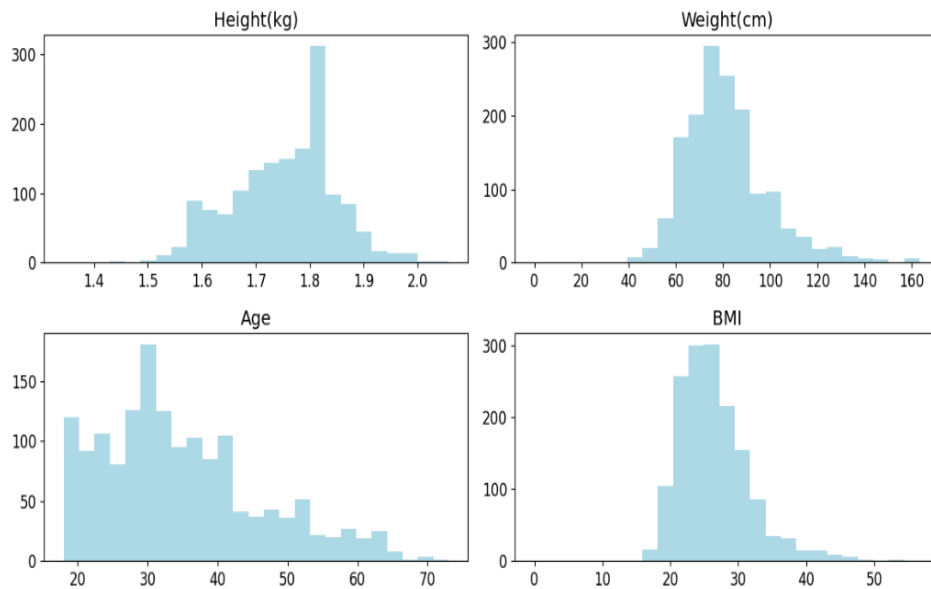


Fig3. Distribution of attribute values

4.4 Face Detection and Alignment using MTCNN

Using a discriminative artificial intelligence technique, the Multitask Cascaded Convolutional Neural Network (MTCNN), a three-stage neural network approach is used to recognize faces. Using the `cv2.imread()` method from the open-cv python package, the input picture is loaded into the MTCNN network. 224×224 pixels are first added to the input image. The resulting BMI, age, gender, and other data from a .csv file are then mapped to the associated photograph of that specific person. The 5 point facial landmark identification technique is now used by MTCNN to align the images by cropping them and identifying the faces [21]. It will create a bounding box around the face in order to identify them.

[22] The stages of training a MTCNN network:

1. Classifying faces
2. Creating a box-shaped boundary around the face
3. Identifying facial landmarks such the mouth, nose, and eyes

The losses for these stages are mathematically formulated as

$$o_i \in \{0, 1\} \quad (2.1)$$

$$Loss_i^{det} = -(o_i^{det} \log(p_i) + (1 - o_i^{det})(1 - \log(p_i))) \quad (2.2)$$

$$Loss_i^{box} = \|o_i^{box} - o_i^{box}\|_2^2 \quad (2.3)$$

$$Loss_i^{landmark} = \|o_i^{landmark} - o_i^{landmark}\|_2^2 \quad (2.4)$$

Here,

The Face classification loss is calculated using the cross-entropy loss.

Bounding box and facial landmark losses both use Euclidean loss.

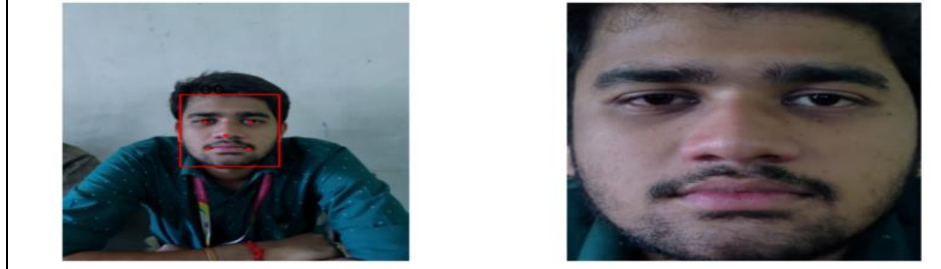


Fig4. Cropping image using MTCNN

4.5 Multi-task Prediction Model

The 16-layered convolutional neural network VGG16, which uses an input image to produce the desired output, has a dense layer at the very end. The use of three backbone maintenance methods is challenging since it consumes a lot of time and memory.

<code>[input image] => [VGG16] => [dense layers] => [BMI]</code>
<code>[input image] => [VGG16] => [dense layers] => [AGE]</code>
<code>[input image] => [VGG16] => [dense layers] => [SEX]</code>

Fig5. Three backbone maintenance efforts

In order to forecast the weighted BMI, Age, and Sex, a mixed network is used that only requires a single backbone maintenance effort, saving time and computing power.

<code>[input image] => [VGG16] => [separate dense layers] x3 => weighted([BMI], [AGE], [SEX])</code>

Fig6. Single backbone maintenance efforts

The model is trained using a multi-task model class. The model was built and implemented using the Tensorflow framework. In VGG16, hyperparameter tuning is the primary means of getting high accuracy. [23].

Layer (type)	Output Shape	Param #	Connected to
input_16 (InputLayer)	[(None, 224, 224, 3)]	0	[]
conv1/7x7_s2 (Conv2D)	(None, 112, 112, 64)	9408	['input_16[0][0]']
conv1/7x7_s2/bn (BatchNormalization)	(None, 112, 112, 64)	256	['conv1/7x7_s2[0][0]']
activation_147 (Activation)	(None, 112, 112, 64)	0	['conv1/7x7_s2/bn[0][0]']
max_pooling2d_27 (MaxPooling2D)	(None, 55, 55, 64)	0	['activation_147[0][0]']
conv2_1_1x1_reduce (Conv2D)	(None, 55, 55, 64)	4096	['max_pooling2d_27[0][0]']
conv2_1_1x1_reduce/bn (BatchNormalization)	(None, 55, 55, 64)	256	['conv2_1_1x1_reduce[0][0]']
activation_148 (Activation)	(None, 55, 55, 64)	0	['conv2_1_1x1_reduce/bn[0][0]']
...			
Total params: 24,400,963			
Trainable params: 838,275			
Non-trainable params: 23,562,688			

Fig7. Summary of Multi-Class Model Network

5. Results and Discussions

This methodology enabled the extraction of the bounding box and five-point facial landmark for the face. Area Under the ROC Curve, Mean Absolute Error, and Root Mean Square Error were used to assess the efficacy of this BMI predictor model. The BMI will change based on the gender as well. When given an input image, the

trained model calculates a person's BMI, age, and gender. Additionally, it can recognize many faces in an image and create the data associated with each face.

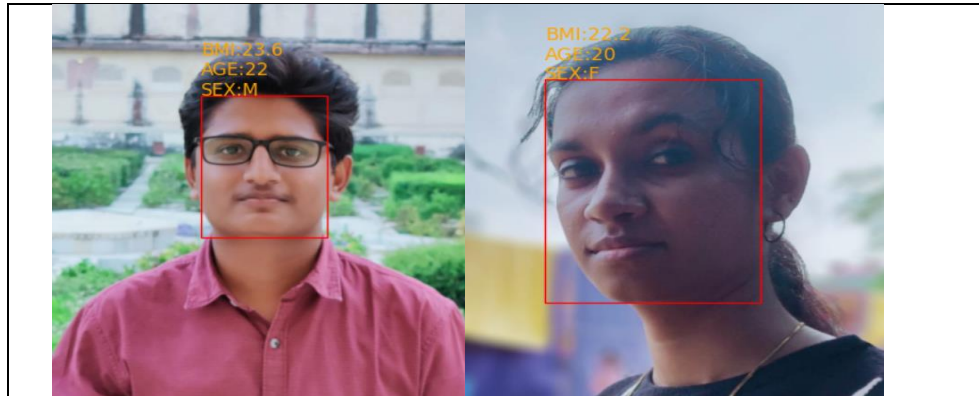


Fig8. Output Image with BMI, Age, Sex

The accuracy of the model is greatly influenced by the hyper parameter tweaking. The parameters, such as mode, model type, batch size, and number of epochs, must be explicitly defined.

The value loss during training decreases from infinite to a relatively low amount as the number of epochs rises. The accuracy of the model also improves when the loss goes down. The output produced by Tensorboard is used to diagrammatically show the loss decrease. Tables 2 and 3 include the performance metrics used to determine the model's effectiveness.

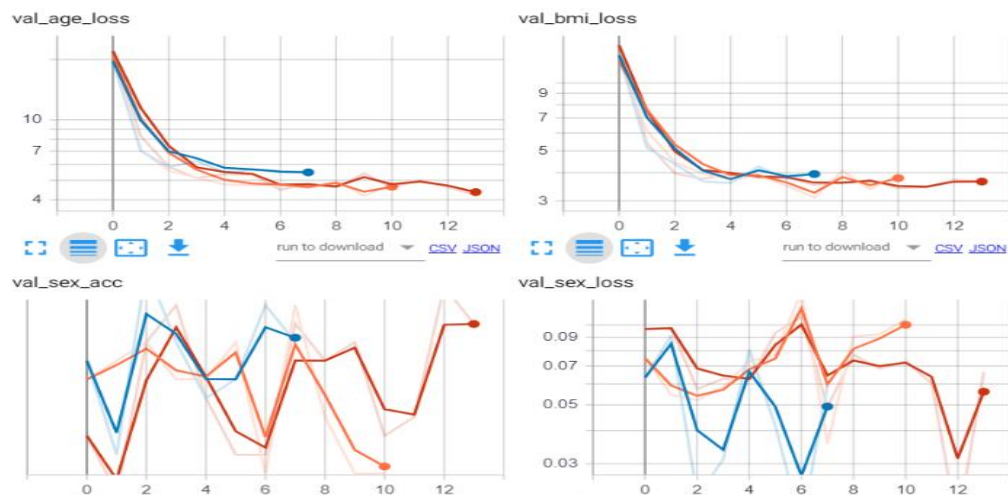


Fig9. Tensorboard Representation of Losses

MODEL	BMI(RMSE)	AGE(RMSE)	SEX(AUC)
Vgg16	4.56	5.66	0.99
Vhh16_fc6	4.99	6.04	0.99
Resnet50	5.21	7.02	0.99

Table2. Performance metrics – Root Mean Square Error and Area Under the ROC Curve

In comparison, VGG16 and Resnet50 performed better than any other algorithms. In face-based deep learning challenges, evaluating the gender-based BMI predictions is much more difficult, and our model makes accurate predictions. The accuracy will also be impacted by the gender-specific characteristics because the prediction uses distinct values for each gender. The model's accuracy significantly enhanced as a result of the training's increased epoch count and implicit declaration of the hyperparameter tuning. By training the model in a 16-layered VGG network utilizing every parameter, the model is able to associate the data with each recognized face.

MODEL	BMI_MAE	BMI_COR	AGE_MAE	SEX_AUC
Resnet50	4.083	0.401	4.982	1.000
Vgg16	3.955	0.412	5.054	0.999
Vgg16_seqMT	6.129	0.274	9.617	0.999

Table3. Performance metrics – Mean Absolute Error and Area Under the ROC Curve

The performance indicators are particularly helpful in identifying the model's advantages and potential improvement areas. If the photograph contains more than two or three people, it will also be predicted. This model is trained on a system with an i5 processor and an RTX 3050 GPU, and it can be used anywhere in hospitals because low code platforms make it simple to build a portal and the data can be saved in the cloud, making deployment simple. With respect to performance measures like ROC Curve, Mean Absolute Error, and Root Mean Square Error, the tensorboard visualization results showed promise in demonstrating the model's performance.

6. Conclusion

The innovative and pioneering method of predicting BMI (Body Mass Index) using facial prediction was investigated in this study report. The proposed system utilizes deep learning techniques for feature extraction and BMI prediction and has the potential

to provide more accurate and efficient methods for calculating BMI in a non-invasive manner. Through the course of the project, several key contributions have been made. Firstly, the dataset of facial images and correlating BMI measurements have been gathered and employed for the progression and evaluation of the proposed system. Secondly, a deep learning model has been trained and fine-tuned using this dataset to accurately predict BMI from facial images. Thirdly, the performance of the proposed system has been evaluated using standard performance metrics and has shown promising results. There are several potential applications of this approach in healthcare and wellness monitoring. The proposed system could also be used in clinical settings for routine health checkups and patient monitoring. However, there are several areas where future enhancements could be made to improve the proposed system. Firstly, the dataset used for training and evaluation could be expanded to include a wider range of ages, ethnicities, and body types, to improve the generalizability of the system. Second, to increase the accuracy, supplementary features from facial photos, including skin texture or facial fat distribution, might be extracted. The best method for predicting BMI was found after a comparison of several machine learning models. The model's performance was verified using a variety of assessment indicators, proving its dependability and robustness. It may be possible to increase user accessibility by investigating the viability of integrating the BMI prediction model into real-time applications, including mobile apps or web-based solutions. This might motivate people to routinely check their BMI and take the necessary steps to maintain a healthy lifestyle. The technique of extracting face features might be improved by using 3D facial data and facial landmarks. Using 3D facial models or important facial landmarks may offer more thorough details on facial features, improving BMI estimations. The importance of privacy and ethical issues must be given top priority in any technology that uses personal data. The development of privacy-preserving methods and the assurance that the data used for training and testing are obtained with adequate consent and anonymization processes should be the main goals of future improvements. In conclusion, the proposed system offers a promising approach to the non-invasive calculation of BMI using facial images. The creation of this system might aid in the advancement of more precise and effective BMI calculation techniques, perhaps enhancing patient outcomes and assisting in the advancement of public health efforts.

References

1. Nuttall, F. Q. Body Mass Index, Obesity, BMI, and Health: A Critical Review. *Nutrition Today* 50(3), 117-128 (2015). doi: 10.1097/NT.0000000000000092
2. C, D, N, NU, Maddikunta, PR, Gadekallu, TR, Iwendi, C, Wei, C, Xin, Q. Identification of malnutrition and prediction of BMI from facial images using real-time image processing and machine learning. *IET Image Process.* 2022; 16: 647– 658. <https://doi.org/10.1049/ipr2.12222>

3. Yap, M. H., Ugail, H., Zwiggelaar, R., & Rajoub, B. A. (2010). Facial image processing for facial analysis. In IEEE International Carnahan Conference on Security Technology (ICST), October 2010. doi: 10.1109/CCST.2010.5678706
4. Ferdowsy, F., Rahi, K. S. A., Jabiullah, M. I., Habib, M. T. (2021). A machine learning approach for obesity risk prediction. *Computers in Biology and Medicine*, Volume 138, Article 100053, Pages 1-10. doi: 10.1016/j.crbeha.2021.100053
5. Alzubaidi, L., Zhang, J., Humaidi, A.J. et al. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. *J Big Data* 8, 53 (2021). <https://doi.org/10.1186/s40537-021-00444-8>
6. Zhang, K., Zhang, Z., Li, Z., Qiao, Y. (Year). Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks. arXiv preprint arXiv:1604.02878 [cs.CV]. <https://doi.org/10.1109/LSP.2016.2603342>
7. Simonyan, K., Zisserman, A. (Year). Very Deep Convolutional Networks for Large-Scale Image Recognition. arXiv preprint arXiv:1409.1556 [cs.CV]. <https://doi.org/10.48550/arXiv.1409.1556>
8. Chong Yen Fook, Lim Chee Chin, Vikneswaran Vijejan, Lim Whey Teen, Hasimah Ali, Aimi Salihah Abdul Nasir. "Investigation on Body Mass Index Prediction from Face Images." In: 2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES), Langkawi Island, Malaysia, 1-3 March 2021. IEEE doi: 10.1109/IECBES48179.2021.9398733
9. Siddiqui, H. 2021. Obesity classification from facial images using deep learning -- In Proceedings: 17th Annual Symposium on Graduate Research and Scholarly Projects. Wichita, KS: Wichita State University
10. Sarak, R. M., Thorat, A. A., & Kadam, D. (2022). Face to BMI: Estimating Body Mass Index [BMI] through Face Recognition Images. *International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)*, 2(1), 100. DOI: 10.48175/IJARSCT-3021.
11. P. Vishnu Raja, K. Sangeetha, D. Sanjay Kumar, A. Surya, D. Subhathra; Prediction of human height, weight and BMI from face images using machine learning algorithms. *AIP Conference Proceedings* 19 May 2022; 2393 (1): 020192. <https://doi.org/10.1063/5.0074450>
12. Yousaf, N., Hussein, S., Sultani, W. (2021). Estimation of BMI from facial images using semantic segmentation-based region-aware pooling. *Computers in Biology and Medicine*, 133, 104392. ISSN 0010-4825. DOI: 10.1016/j.combiomed.2021.104392.
13. Jiang, M., Guo, G., Mu, G. (2020). Visual BMI estimation from face images using a label distribution-based method. *Computer Vision and Image Understanding*, Volumes 197-198, 102985. DOI: 10.1016/j.cviu.2020.102985.
14. Pantanowitz, A., Cohen, E., Gradidge, P. (2021). Estimation of Body Mass Index from photographs using deep Convolutional Neural Networks. *Informatics in Medicine Unlocked*, 26(10046), 100727. DOI: 10.1016/j.imu.2021.100727.
15. Pham, Q.T., Luu, A.T., Tran, T.H. (2021). BMI estimation from facial images using residual regression model. In 2021 International Conference on Advanced Technologies for Communications (ATC), Ho Chi Minh City, Vietnam. IEEE. DOI: 10.1109/ATC52653.2021.9598340.
16. Lee, B.J., Ku, B., Jang, J.S., Kim, J.Y. (2013). A novel method for classifying body mass index on the basis of speech signals for future clinical applications: a pilot study. *Journal of Obesity*, 2013:150265. PMID: 23573116. PMCID: PMC3612486. DOI: 10.1155/2013/150265.

17. Paradis, E., O'Brien, B., Nimmon, L., Bandiera, G., Martimianakis, M.A. (2016). Design: Selection of Data Collection Methods. *Journal of Graduate Medical Education*, 8(2), 263-264. doi: 10.4300/JGME-D-16-00098.1
18. Khan, S.I., Hoque, A.S.M.L. SICE: an improved missing data imputation technique. *J Big Data* 7, 37 (2020). <https://doi.org/10.1186/s40537-020-00313-w>
19. Ali, P.J.M., Faraj, R.H. (2014). Data Normalization and Standardization: A Technical Report. Technical Report. January 2014. DOI: 10.13140/RG.2.2.28948.04489.
20. Manikandan, S. (2010). Data transformation. *Journal of Pharmacology and Pharmacotherapeutics*, 1(2), 126-127. doi: 10.4103/0976-500X.72373. PMCID: PMC3043340 PMID: 21350629.
21. Johnston, B., de Chazal, P. (2018). A review of image-based automatic facial landmark identification techniques. *EURASIP Journal on Image and Video Processing*, 2018(1), 86. doi: 10.1186/s13640-018-0324-4.
22. Ma, L., Fan, H., Lu, Z., Tian, D. (2020). Acceleration of multi-task cascaded convolutional networks. *IET Image Processing*, 14(13), 2556-2563. doi: 10.1049/iet-ipr.2019.0141.
23. Ippolito, P.P. (2022). Hyperparameter Tuning. In: Egger, R. (eds) *Applied Data Science in Tourism. Tourism on the Verge*. Springer, Cham. https://doi.org/10.1007/978-3-030-88389-8_12
24. N. Yousaf, S. Hussein, and W. Sultani, "Estimation of BMI from Facial Images using Semantic Segmentation based Region-Aware Pooling," *Computers in Biology and Medicine*, vol. 133, June 2021, article no. 104392. DOI: 10.48550/arXiv.2104.04733
25. Dhanamjayulu C, Nizhal U N, Praveen Kumar Reddy Maddikunta, Thippa Reddy Gadekallu, Celestine Iwendi, Chuliang Wei, Qin Xin. "Identification of malnutrition and prediction of BMI from facial images using real-time image processing and machine learning." *IET Image Processing*, 20 May 2021, <https://doi.org/10.1049/ipr2.12222>. (Citations: 25)