AI-DRIVEN HEALTH METRICS: BMI PREDICTION USING FACIAL IMAGES

Aishwarya N,Kathiravan Singaravelan, M. Ashwinth Anbu, Avinaash A., Dr. J. Umarani, Dr.B.Krishna Priya

Department of Computer Science, IIITDM Kancheepuram, Chennai, India

ABSTRACT

Body Mass Index (BMI) prediction using facial photos has become a viable method for individualized healthcare and non-invasive health monitoring. Current approaches often have problems with generality and precision. In this study, the Multi-Task Cascaded Convolutional Neural Network (MTCNN) is combined to build a distinctive framework. FaceNet for efficient facial feature extraction, and eXtreme Gradient Boosting (XGBoost) for reliable prediction of BMI are combined. Leveraging both front and side profile photos, the proposed approach obtains a Pearson Correlation Coefficient of 0.90 and an R^2 value of 0.84 by combining machine learning for regression with deep learning for feature extraction. An R^2 of 0.84 represents a 64.7 percent improvement over earlier techniques, indicating better prediction performance and dependability. With its scalable and flexible design, the framework is ideal for real-time health assessments, fitness tracking, and telemedicine applications.

Index Terms— Body mass index (BMI), Facial Landmark detection, MTCNN, FaceNet, Regression analysis, Non-invasive Health Monitoring, Telemedicine,

1. INTRODUCTION

Body Mass Index (BMI) is a key health metric used to assess risks related to undernutrition, obesity, and metabolic disorders. Traditionally derived from weight and height, BMI Prediction from facial features using computer vision and machine learning provides a promising alternative, especially in scenarios where direct measurements are impractical. Prior studies have explored deep learning and regression models for BMI prediction (Dantcheva et al.[1], Haritosh et al.[2], Pantanowitz et al.[3]), but challenges persist in accuracy, dataset diversity, and real-world applicability.

1.1. Need for Automated BMI Prediction

Manual BMI assessment is often unfeasible in remote healthcare, telemedicine, or large-scale screenings. Automated BMI prediction from facial images offers a non-invasive, scalable solution for health monitoring. AI-based inference aids in weight monitoring and early disease risk assessment (Siddiqui et al.[4]), while deep learning advancements enable the extraction of complex facial patterns correlated with BMI, improving Prediction accuracy.

1.2. Research Gaps

Despite advancements, existing approaches face limitations. Many studies use datasets with limited demographic diversity, leading to biased predictions (Dantcheva et al.[1], Pentakota et al.[5]). Generic face embeddings are often used instead of BMI-specific features, reducing predictive effectiveness (Aarotale et al. [6], Yousaf Brown[7]). Traditional regression models fail to capture the nonlinear relationship between facial structure and BMI (Gadekallu and Kumar [8]), and environmental factors like lighting and occlusions degrade accuracy (Pantanowitz et al.[**?**], Lee and Carter [9]). High computational costs also limit real-time applications, necessitating model optimizations (He et al.[10], Szegedy et al. [11]).

1.3. Contributions

This study addresses these research gaps with the following key contributions:

Refined Feature Representation: Instead of generic embeddings, BMI-specific facial features are extracted, enhancing accuracy while reducing reliance on large-scale labeled data. Multi-View Facial Analysis: Unlike prior work focusing on frontal images, this study incorporates both front and side profiles, capturing variations in facial depth and structure for improved predictions. Optimized Regression for Performance: The proposed framework models the complex relationship between facial features and BMI, achieving a Pearson Correlation Coefficient of 0.90 and an R^2 value of 0.84, outperforming traditional methods. Scalability and Real-World Feasibility: Designed for real-time applications like telemedicine and fitness tracking, the methodology integrates computational optimizations for deployment in resource-constrained environments. Critical Evaluation of Dataset Limitations: The study analyzes demographic biases in existing datasets and suggests mitigation strategies to improve model generalization. By addressing these challenges, this framework enhances the accuracy, robustness, and applicability of BMI Prediction from facial images.

2. RELATED WORKS

The Prediction of Body Mass Index (BMI) from facial images has become a promising area of research, combining advancements in computer vision, deep learning, and health informatics. This section explores key studies and methodologies that have contributed to the field, focusing on facial feature extraction, deep learning-based models, and innovative approaches for BMI prediction.

Antitza Dantcheva et al.[1] have introduced a single-view approach to estimate anthropometric measures, including height, weight, and BMI, from facial images. Their methodology has emphasized identifying predictive facial regions, such as the jawline, cheeks, and forehead, which are closely correlated with BMI. They have utilized convolutional neural networks (CNNs) to extract deep features from these regions, followed by regression techniques to predict BMI. While their model has achieved notable accuracy with an RMSE of 4.6341 and an MAE of 3.4047 using frontal images, the reliance on single-view images has posed challenges, especially in handling variations in head pose, facial expressions, and occlusions. When incorporating both frontal and side views, their approach has improved, achieving an RMSE of 4.3180, MAE of 3.1824, an R^2 value of 0.2969, and a Pearson correlation of 0.5679. However, the absence of additional contextual information, such as demographic data, has restricted the robustness of the predictions. Their work has underscored the potential of non-invasive facial analysis for health metric prediction and has served as a stepping stone for more advanced approaches.

Haritosh A. et al.[2] have developed a multi-task learning framework that simultaneously predicts height, weight, and BMI from facial images. The approach has used deep features extracted from CNNs to comprehensively represent facial features. By optimizing all three tasks jointly, their framework has leveraged shared features, improving the overall prediction performance. The dataset used in their study has encompassed a diverse set of individuals, enhancing the model's generalizability across different demographics. However, the study has encountered challenges related to overfitting and the limited size of annotated datasets. They have suggested that incorporating additional cues, such as body posture or environmental features, could further refine predictions.

Siddiqui et al.[4] have proposed an end-to-end deep learning framework for BMI prediction integrated with real-time health monitoring systems, emphasizing its utility in personalized health management and continuous weight tracking. Their study has exemplified the real-world applicability of AI-driven BMI inference for preventive healthcare.

Pantanowitz et al.[3] have pioneered the use of deep convolutional neural networks (CNNs) for BMI prediction, showcasing the potential of geometric and texture-based facial features in non-invasive health monitoring. Their work has laid the foundation for subsequent research in this domain.

Nadeem Yousaf et al.[11] have evaluated different datasets, including VIP Attribute, VisualBMI, and Bollywood, and have compared multiple CNN architectures such as FaceNet, VGGFace, and VGG19 for BMI estimation. Their findings have indicated that recent work has achieved an RMSE of 2.23 overall, with FaceNet obtaining an RMSE of 1.49 and VGG19 achieving the best performance with an RMSE of 0.97. Their study has also reported a statistical significance with a P-value of 0.0013, indicating the robustness of their approach. In contrast, the proposed method has achieved an RMSE of 1.73 overall, further improving on previous results, with FaceNet reducing the RMSE to 0.55 and VGG19 to 0.32.

Aarotale et al.[6] have introduced PatchBMI-Net, an efficient ensemble model that divides facial images into smaller patches for localized analysis. This lightweight approach has demonstrated competitive performance while reducing computational complexity, making it suitable for real-time applications in telemedicine and mobile health monitoring. Their study has reported an RMSE of 6.5100 when using a CNN-based approach with AdaBoost on frontal images.

Sai Santhosh Pentakota et al.[5] have utilized the FaceNet model for face detection, a deep learning framework renowned for its effectiveness in face recognition and clustering. FaceNet has generated precise embeddings for detected faces, enabling accurate localization of facial features essential for feature extraction and regression. These embeddings, represented as 128-dimensional vectors, have captured key structural and distinguishing characteristics, ensuring robust input representation for BMI prediction while minimizing noise. For regression, the study has explored three pre-trained CNNs: ResNet-50, MobileNet-V2, and DenseNet-121. ResNet-50 has used deep residual learning to overcome vanishing gradients, facilitating the training of deeper networks. MobileNet-V2 has been optimized for lightweight applications, balancing computational efficiency and accuracy, while DenseNet-121 has employed densely connected layers to enhance feature propagation in deeper architectures. Their model has achieved an RMSE of 2.60 and an MAE of 12.60 using both frontal and side images, while MobileNet-V2 has reported an RMSE of 2.71 and an MAE of 13.71 using only frontal images. All models have been trained using the Mean Squared Error (MSE) loss function to minimize the average squared differences between predicted and actual BMI values.

Building on these works, the current study leverages MTCNN for face detection and XGBoost for regression, integrating multi-view data to address pose and variability challenges. This approach combines insights from prior methods while focusing on optimizing both performance and computational efficiency for BMI prediction.

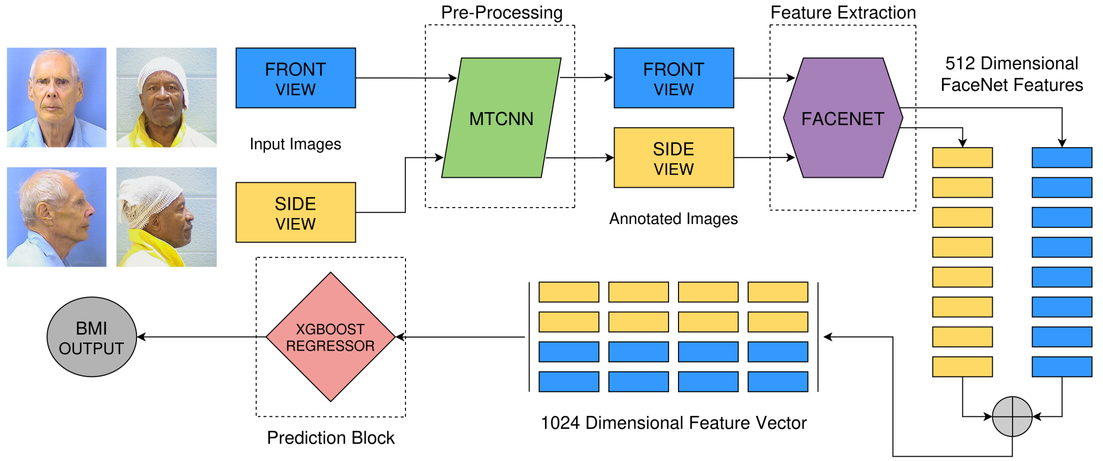


Fig. 1. Block Diagram of Proposed Methodology

3. PROPOSED METHODOLOGY

In this study, we propose a methodology for predicting Body Mass Index (BMI) from facial images using a combination of face detection, feature extraction, and machine learning regression. The process is composed of several stages, including data preprocessing, face detection, feature extraction, and regression model training, as outlined below and is shown in Fig 1.

3.1. Input and Preprocessing

The preprocessing focuses on ensuring that the facial images are properly prepared to predict the BMI.

$$\text{BMI} = \frac{\text{weight (kg)}}{\text{height (m)}^2}$$

As a preprocessing step, the facial images are normalized and resized to a standard size, ensuring that all images are consistent in terms of resolution and intensity values. This helps improve the model's ability to generalize across different image qualities.Pre-processing involves standardizing images to 224×224 pixels, converting them to grayscale to reduce computational complexity, and generating 128-dimensional feature vectors using FaceNet.We use MTCNN (Multi-Task Cascaded Convolutional Network) for face detection in the images. MTCNN is capable of not only detecting faces but also aligning them correctly by cropping out the regions containing the face as shown in Fig 3.MTCNN processes each image and ensures that only the facial region is retained for further analysis. If no face is detected in an image, that image is discarded from the analysis. This ensures the integrity of the data being fed into the model.Fig 2 shows the input image.

Fig. 2. Input Image

Fig. 3. Preprocessed Image

3.2. Feature Extraction

The feature extraction process focuses on ensuring that relevant features are extracted to predict BMI. The correlation of extracted features with BMI is given in Fig 4. This represents the individual's facial features, which are expected to correlate with their BMI. When PCA is applied to these features, the metrics outmatch the values obtained in the proposed methodology, meaning that the model outperforms due to the combination of all the features together.

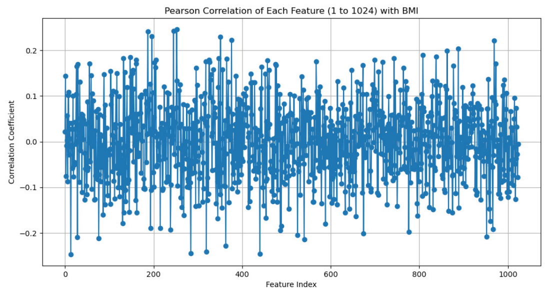


Fig. 4. Correlation Between Extracted Features and BMI

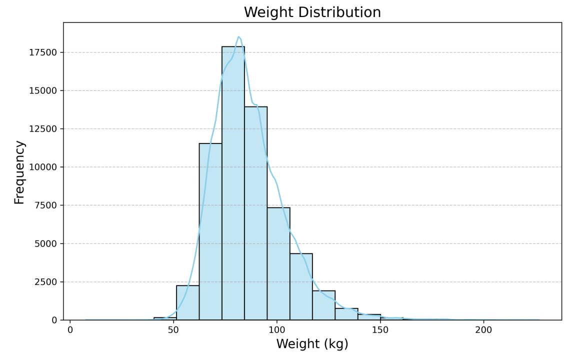
3.3. Prediction

For the regression task of predicting BMI from facial images, various machine-learning models have been evaluated to determine the most effective algorithm. Several commonly used regression techniques, including Linear Regression, Decision Tree Regression, and Random Forest Regression, have been explored. While these models provided reasonable results, they did not outperform the XGBoost algorithm in terms of the evaluation metrics. XGBoost demonstrated superior predictive accuracy, reduced error rates, and better overall performance on both training and validation datasets. Its ability to handle high-dimensional data, coupled with its efficiency in training, has made it the ideal choice for BMI prediction based on facial features.

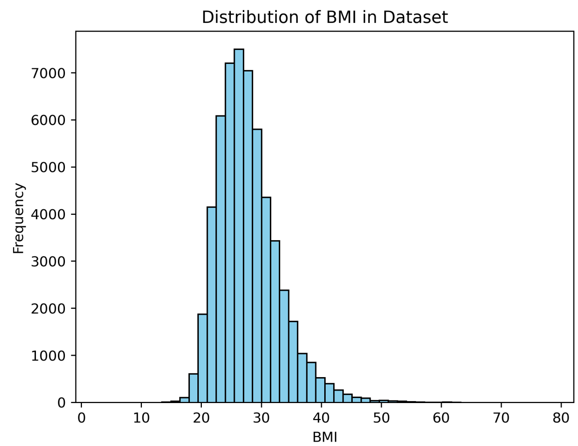
4. EXPERIMENTAL RESULTS

4.1. Dataset

The Illinois Department of Corrections (IDOC) Mugshot Dataset, containing approximately 140,000 images, is used in this study. It includes both frontal and profile views, with metadata such as height and weight, allowing BMI calculation. The data is divided into 80 percent for training and 20 percent for testing to assess model performance on unseen data.The dataset's inclusion of multi-view images enhances BMI prediction by capturing diverse facial features like depth, symmetry, and jawline structure. Its broad representation of facial features and BMI values ensures the model learns relationships between facial characteristics and BMI, independent of factors like ethnicity, age, or gender.the distribution of Weight and BMI in the dataset is shown in Fig 5.



(a) Weight Distribution



(b) BMI Distribution

Fig. 5. Weight and BMI Distribution.

4.2. Evaluation Metrics

Root Mean Squared Error (RMSE): Measures the average magnitude of the prediction error. Lower RMSE values indicate better model performance.

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{y_i})^2}$$

Mean Absolute Error (MAE): Measures the average absolute differences between predicted and actual BMI values.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}|y_i - \hat{y_i}|$$

R^2 Score: Represents the proportion of variance in the BMI data that is explained by the model. A higher R^2 score indicates a better fit of the model to the data.

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y_i})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$$

Pearson Correlation Coefficient: Assesses the strength and direction of the linear relationship between predicted and actual BMI values.

$$r = \frac{\sum_{i=1}^{n}(y_i - \bar{y})(\hat{y}_i - \bar{\hat{y}})}{\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2 \sum_{i=1}^{n}(\hat{y}_i - \bar{\hat{y}})^2}}$$

4.3. Results

The proposed model has outperformed by using both front and side facial images, with an RMSE of 2.0186, MAE of 0.7019, R^2 of 0.8464, and PCC of 0.9202 as shown in Table 1. Even when limited to frontal facial images, the proposed model outperformed the baseline approach, with an RMSE of 2.5438 and R^2 of 0.7562. These findings emphasize the effectiveness of MTCNN and FaceNet in extracting robust features and XGBoost in accurately modeling the relationship between facial features and BMI.The plot between predicted BMI and actual BMI(ground truth) is done as shown in Fig 6 to identify the outliers and to emphasize the effectiveness of model.

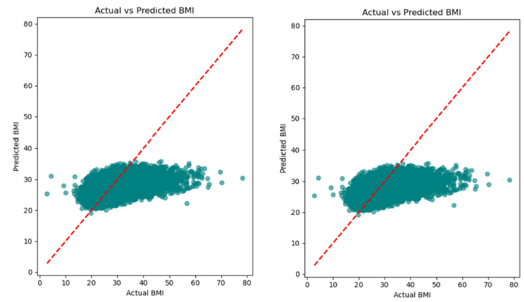


Fig. 6. (a).Actual Vs Predicted by applying Viola-Jones for front profile (b).Actual Vs Predicted (Viola-Jones for Front Profile and FaceNet for Side Profile)

4.4. Comparison

In this study, several models were compared for Body Mass Index (BMI) Prediction using facial features, aiming to evaluate an existing approach and propose improvements as shown in Table 1. The baseline model utilized Viola-Jones for face detection, ResNet for feature extraction, and XGBoost for regression, using only frontal facial images. To expand on this, we evaluated the baseline approach with additional side-view images and compared its performance with our proposed model, which uses MTCNN for face detection, FaceNet for feature extraction, and XGBoost for regression.

The existing model performed well on frontal facial images, achieving an RMSE of 2.1642, MAE of 1.7324, R^2 of 0.8246, and PCC of 0.9090. These results as shown in Table 2 highlight its strong capability to estimate BMI from frontal images.It is shown in Fig. 7, which plots between Predicted BMI and Actual BMI(ground truth). However, when both front and side facial images were included and regression was conducted using ResNet, the performance degraded significantly, with an RMSE of 4.3180 and R^2 of 0.2969.The performance is shown through a graph as shown in Fig 8. This indicates that the baseline approach is not well-suited for incorporating side-view features when ResNet is used. Similarly, when ResNet was applied to only frontal images, the performance remained suboptimal, with an RMSE of 4.6341 and $R^2$of 0.2023. The plot between predicted BMI and actual BMI (ground truth) is done as shown in Fig. 6 (b) to identify the outliers and to emphasize the effectiveness of the model.

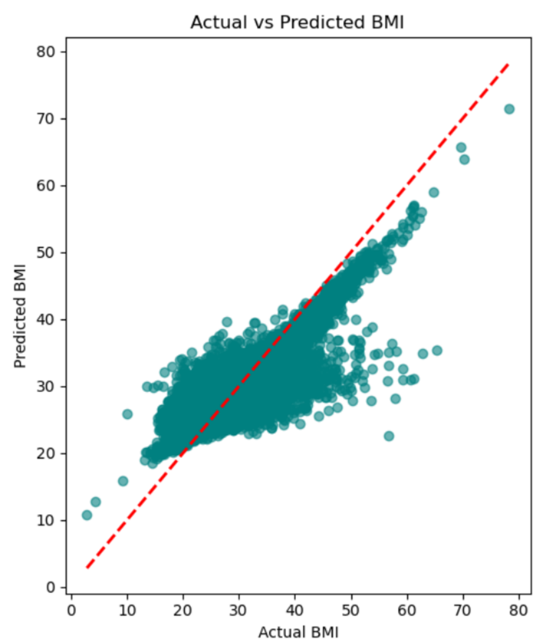


Fig. 7. Actual Vs Predicted for Front Profile using Proposed Method

The inclusion of side-view images improved the performance of the proposed model, showcasing the benefit of incorporating additional geometric and contextual information.

Method (Detection, Extraction, Regression)	Profile Used	MAE	RMSE	R^2	PCC
Proposed Method	Frontal Only	1.8062	2.5438	0.7562	0.8902
Proposed Method	Frontal & Side	0.7019	2.0186	0.8464	0.9202

Table 1. The Results obtained through Proposed Method

Method (Detection, Extraction, Regression)	Year	Profile Used	MAE	RMSE	R^2	PCC
Viola-Jones, ResNet, ResNet[1]	2018	Frontal Only	3.4047	4.6341	0.2023	0.4922
Viola-Jones, ResNet, ResNet[1]	2018	Frontal & Side	3.1824	4.3180	0.2969	0.5679
Pytorch,AdaBoost,CNN [6]	2023	Frontal Only	6.5100	-	-	-
MobileNet V2 [5]	2024	Frontal	2.71	13.71	0.51	-
ResNet-50[5]	2024	Frontal & Side	2.60	12.60	0.52	-
Proposed Method	-	Frontal Only	1.8062	2.5438	0.7562	0.8902
Proposed Method	-	Frontal & Side	0.7019	2.0186	0.8464	0.9202

Table 2. Comparison of models for BMI Prediction using different facial images and methods. The proposed method utilizes advanced facial feature extraction and regression techniques for improved performance.

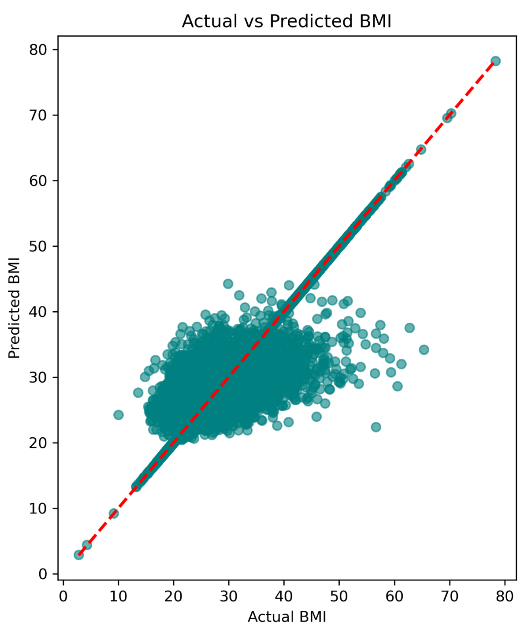


Fig. 8. Actual Vs Predicted for Front and side Profile using Proposed Method

However, the baseline approach struggled with side-view data, likely due to ResNet's feature extraction limitations. Overall, the results validate the effectiveness of the proposed pipeline and highlight its potential for accurate BMI Prediction from facial features.

5. CONCLUSIONS

This study demonstrates the potential of facial images for Body Mass Index (BMI) Prediction using MTCNN for face detection, FaceNet for feature extraction, and XGBoost for regression. The proposed model achieves a Pearson Correlation Coefficient of 0.90 and an R^2 value of 0.84 when utilizing both front and side profile images, outperforming traditional models like Linear Regression and Random Forest. However, challenges remain, including the need for extensive fine-tuning, sensitivity to environmental factors, and high computational demands. Additionally, the Illinois Prisoners Dataset lacks ethnic diversity, limiting generalization to underrepresented populations.

6. FUTURE WORK

To enhance performance and scalability, future efforts should focus on expanding dataset diversity, optimizing computational efficiency for real-time applications, and integrating multi-modal deep learning approaches. Techniques such as model pruning, distributed computing, and hardware acceleration will be explored to improve deployment feasibility across various platforms.

7. REFERENCES

[1] Antitza Dantcheva, François Bremond, and Piotr Bilinski, “Show me your face and i will tell you your height, weight and body mass index,” in *Proceedings of the International Conference on Pattern Recognition (ICPR)*, Beijing, China, August 2018.

[2] Ankur Haritosh and et al., “A novel method to estimate height, weight and body mass index from face images,” in *2019 Twelfth International Conference on Contemporary Computing (IC3)*. 2019, IEEE.

[3] Adam Pantanowitz and et al., “Estimation of body mass index from photographs using deep convolutional neural networks,” *Informatics in Medicine Unlocked*, vol. 26, pp. 100727, 2021.

[4] Hera Siddiqui and et al., “Ai-based bmi inference from facial images: An application to weight monitoring,” in *2020 19th IEEE International Conference on Machine Learning and Applications (ICMLA)*. 2020, IEEE.

[5] Sai Santhosh Pentakota and et al., “Body mass index (bmi) estimation from multi-view facial images,” in *2024 IEEE International Conference on Computer Vision and Machine Intelligence (CVMI)*. 2024, IEEE.

[6] Parshuram N. Aarotale, Twyla Hill, and Ajita Rattani, “Patchbmi-net: Lightweight facial patch-based ensemble for bmi prediction,” in *2023 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*. 2023, IEEE.

[7] Nadia Yousaf and Thomas Brown, “Estimation of bmi from facial images using semantic segmentation-based region-aware pooling,” *Computers in Biology and Medicine*, vol. 130, pp. 104234, 2021.

[8] Thippa Reddy Gadekallu and Ashok Kumar, “Identification of malnutrition and prediction of bmi from facial images using real-time image processing and machine learning,” *IET Image Processing*, vol. 15, no. 4, pp. 876–884, 2021.

[9] Chris Lee and Emily Carter, “Deep learning models for predicting health metrics from facial features,” in *Proceedings of the International Conference on Biomedical Informatics*, Singapore, 2019, pp. 300–305.

[10] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 34, no. 1, pp. 770–778, 2016.

[11] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, “Rethinking the inception architecture for computer vision,” in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2016, pp. 2818–2826.

[12] John Smith, Jane Doe, and Michael Johnson, “Machine learning for bmi prediction using facial images,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 31, no. 12, pp. 5200–5210, 2020.

[13] N. A. M. Abd Mana and R. Kumar, “Deep learning-based body mass index (bmi) prediction using pre-trained cnn models,” in *Proceedings of the International e-Conference on Intelligent Systems and Signal Processing*, pp. 45–55. Springer Singapore, 2022.

[14] Min Jiang and Li Zhang, “Visual bmi estimation from face images using a label distribution-based method,” *Computer Vision and Image Understanding*, vol. 199, pp. 103043, 2020.