SinBMI: Estimating BMI from a Single Image

Tzvi Tal Noy, Ido Sagi, Nurit Spingarn Eliezer
Signal and Image Processing Laboratory (SIPL)
Andrew and Erna Viterbi Faculty of Electrical & Computer Engineering
Technion – Israel Institute of Technology, Haifa, Israel
https://sipl.ece.technion.ac.il
Email: {tzvinoy2, idosagi12, nurit.spini}@gmail.com

***Abstract*—Body mass index is a quantitative measure that relates an individual's weight to their height, providing a standardized indicator of body composition. This metric has been extensively validated and is considered a reliable biometric tool in medical evaluations. However, traditional methods to calculate BMI require direct measurements or self-reported data, which can be time-consuming and subject to error. Therefore, there is great interest in non-apparatus and remote BMI measurement approaches. In this paper, we introduce SinBMI, a novel deep learning model that estimates human BMI from a single image. SinBMI is more than an order of magnitude lighter than existing single-image BMI estimation models, making it highly suitable for mobile deployment. SinBMI achieves state-of-the-art performance on the largest dataset to date, with an MAPE of 9.38% and an MAE of 2.92.**

I. Introduction

As health awareness grows globally, BMI has become a widely used metric for assessing individual health, and identifying risks associated with underweight and overweight conditions. The common way to calculate BMI is according to:

$$BMI = \frac{weight[Kg]}{height[m]^2}. \tag{1}$$

However, obtaining accurate height and weight measurements can be time-consuming and prone to errors. Obtaining accurate physical measurements is challenging in many real-world scenarios, such as in remote areas or busy clinics. Inaccurate measurements may lead to misclassifications of health risks, potentially delaying early intervention for conditions like obesity, cardiovascular disease, or malnutrition. These challenges motivate the development of alternative approaches for rapid and reliable BMI estimation.

II. Related work

With the growing success of deep learning methods, recent studies have shown the ability to estimate the BMI from 2D images. For example, [1] employed convolutional models on multiple images of the same person but required pose alignment and was constrained by limited data. Similarly, [2] focused on facial images with the same limitation. Other approaches have tackled single-image BMI estimation but relied heavily on feature extraction, increasing memory and computation requirements. Furthermore, [3] used anthropometric, 3D, and statistical features, requiring multiple pre-trained models, while [4], the most relevant to our work, used two models: one for deep features and another for anthropometric features, combining them for Gaussian Process regression, which is resource-intensive. In contrast, SinBMI surpasses current models and achieves state-of-the-art performance, while enabling fast BMI estimation on mobile devices.

III. Proposed Solution

In this paper, we demonstrate the efficacy of an end-to-end lightweight deep learning model for direct BMI extraction from a single 2D input image. SinBMI architecture consists of an EfficientNet-B2 backbone [5], renowned for its efficiency in mobile applications. Extensive testing with different backbones revealed that EfficientNet-B2 offered the best trade-off between accuracy and resource efficiency. The backbone is followed by a seven-layer perceptron (MLP), where each layer performs a linear transformation and applies a GELU activation, iteratively reducing the 1000-dimensional feature vector to a scalar BMI prediction (see Fig. 1).

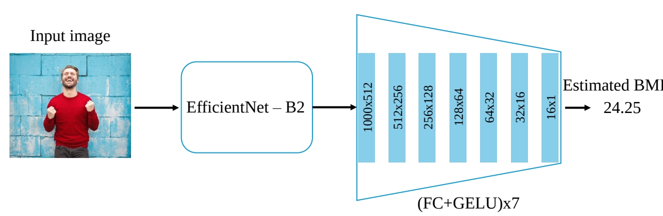

Fig. 1: **SinBMI architecture.** We used EfficientNet-B2 followed by an MLP that gradually reduced dimensionality to a scalar BMI prediction.

We construct our dataset by merging the two largest publicly available BMI datasets. The first, 2D-Image-to-BMI [4], includes 4,189 images (1,477 males, 2,712 females) sourced from Reddit posts about dieting. The second, Visual-Body-to-BMI [6], contains 5,900 images from 2,950 individuals. Both provide frontal-view RGB images in free poses with random backgrounds, along with reported height and weight for BMI calculation. Figure 2 illustrates the BMI distribution, which closely follows global trends [7]. However, the dataset is imbalanced, which may introduce bias into the learning process, favoring overrepresented BMI ranges.

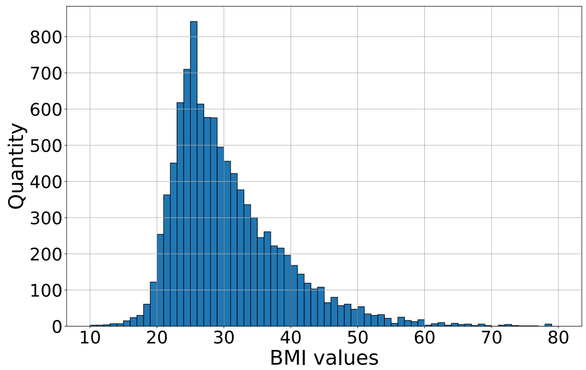

Fig. 2: **Distribution of BMI values.** Note the significantly reduced image counts at the extreme BMI ranges (BMI < 20 and BMI > 40).

We applied data augmentation [8] to compensate for this imbalance and to improve model generalization by introducing natural variations in image statistics. Transformations included horizontal flips, additive Gaussian noise, and rotations up to 10°. These augmentations expanded the dataset to 19,336 images.

IV. Results

In order to evaluate our method, we use mean average error (MAE) and mean average percentage error (MAPE), which are commonly used evaluation metrics [3], [4]. During training, we minimized the mean squared error (MSE) loss between the predicted BMI and the ground truth. Our training regime includes Adam optimizer and plateau scheduling for 40 epochs.

TABLE I: **BMI estimation results.** SinBMI outperforms [3] and [4] across all architectures, with optimal performance using EfficientNet-B2.

Method	Arch.	MAE	MAPE [%]	Params [M]
[3]	Multiple models	3.22	10.27	216.59
[4]	Multiple models	3.96	13.31	490.14
SinBMI	VIT	6.23	18.74	87.25
SinBMI	ResNet101	3.14	10.01	45.23
SinBMI	MobileNet	3.50	11.22	3.23
SinBMI	**EfficientNet-B2**	**2.92**	**9.38**	**9.79**

In Table I, we summarize the performance of SinBMI compared to [3] and [4]. Additionally, we include performance across alternative architectures. Our findings consistently show that SinBMI with EfficientNet-B2 outperforms all the methods compared.

For a more precise evaluation, SinBMI is assessed across BMI categories, widely recognized by medical and health authorities. These categories divide the BMI range into five groups: 'underweight' (<18.5), 'normal' (18.5–25), 'overweight' (25–30), 'obese' (30–40), and 'extremely obese' (>40). Figure 3 demonstrates strong performance within the BMI range of 18.5–40, which is also the most common in the general population. However, performance degrades at the edges of the BMI spectrum due to the limited data in the extreme categories, which makes it difficult for the model to generalize.

The graphs compare the performance with and without data augmentation according to the evaluation metrics- MAE and MAPE. According to both metrics, data augmentation significantly improves performance.

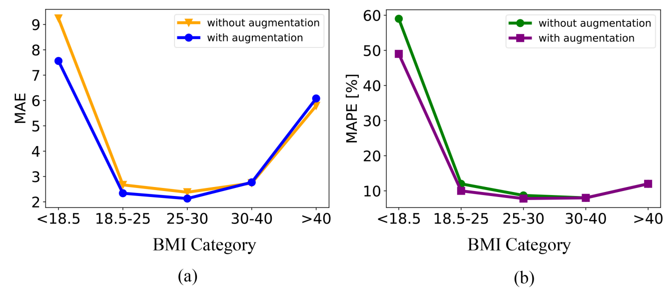

Fig. 3: **Performance evaluation.** Comparison of SinBMI performance across BMI categories with and without data augmentation. (a) MAE results, (b) MAPE results.

V. Live Demonstration

To demonstrate our system's efficiency and ease of use, we will stage a live demo in which participants stand approximately one meter from a standard webcam connected to a laptop. The webcam captures a single image that is processed in real time by our model. The estimated BMI is then immediately displayed on a dedicated screen, showcasing the speed and accuracy of our approach.

VI. Conclusions

In this paper, we propose SinBMI, an end-to-end deep learning architecture designed to estimate BMI from a single image. SinBMI introduces a novel, lightweight, and high-accuracy approach to BMI estimation, making it well-suited for deployment on mobile devices. Our model achieves state-of-the-art MAE and MAPE performance, demonstrating particularly high accuracy in the most common BMI categories.

References

[1] A. Pantanowitz, E. Cohen, P. Gradidge, N. J. Crowther, V. Aharonson, B. Rosman, and D. M. Rubin, "Estimation of body mass index from photographs using deep convolutional neural networks," *Informatics in Medicine Unlocked*, vol. 26, p. 100727, 2021.

[2] M. Jiang, G. Guo, and G. Mu, "Visual bmi estimation from face images using a label distribution based method," *Computer Vision and Image Understanding*, vol. 197, p. 102985, 2020.

[3] J. Huang, C. Shang, A. Xiong, Y. Pang, and Z. Jin, "Seeing health with eyes: Feature combination for image-based human bmi estimation," in *2021 IEEE International Conference on Multimedia and Expo (ICME)*. IEEE, 2021, pp. 1–6.

[4] Z. Jin, J. Huang, W. Wang, A. Xiong, and X. Tan, "Estimating human weight from a single image," *IEEE Transactions on Multimedia*, 2022.

[5] M. Tan and Q. Le, "Efficientnet: Rethinking model scaling for convolutional neural networks," in *International conference on machine learning*. PMLR, 2019, pp. 6105–6114.

[6] M. Jiang and G. Guo, "Body weight analysis from human body images," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 10, pp. 2676–2688, 2019.

[7] "Time trends in body mass index distribution in the general population in denmark from 1987 to 2021," https://ugeskriftet.dk/dmj/time-trends-body-mass-index-distribution-general-population-denmark-1987-2021.

[8] D. A. Van Dyk and X.-L. Meng, "The art of data augmentation," *Journal of Computational and Graphical Statistics*, vol. 10, no. 1, pp. 1–50, 2001.