Identification of malnutrition and prediction of BMI from facial images using real-time image processing and machine learning

Dhanamjayulu C¹ | Nizhal U N¹ | Praveen Kumar Reddy Maddikunta² | Thippa Reddy Gadekallu² | Celestine Iwendi³ | Chuliang Wei⁴,⁵ | Qin Xin⁶

¹ School of Electrical Engineering, VIT - Vellore, Tamilnadu, India
² School of Information Technology and Engineering, VIT - Vellore, Tamilnadu, India
³ Coal City University, Nigeria
⁴ Department of Electronic Engineering, Shantou University, Guangdong, China
⁵ Key Laboratory of Digital Signal and Image Processing of Guangdong Province, Shantou University, Guangdong, China
⁶ Science and Technology, University of the Faroe Islands, Faroe Islands

Correspondence
Chuliang Wei, Department of Electronic Engineering, Shantou University, Guangdong 515063, China. Email: chwei@stu.edu.cn

Abstract
Human faces contain useful information that can be used in the identification of age, gender, weight etc. Among these biometrics, body mass index (BMI) and body weight are good indicators of a healthy person. Motivated by the recent health science studies, this work investigates ways to identify malnutrition affected people and obese people by analyzing body weight and BMI from facial images by proposing a regression method based on the 50-layers Residual network architecture. For face detection, Multi-task Cascaded Convolutional Neural Networks have been employed. A system is created to evaluate BMI along with age and gender from human facial real-time images. Malnutrition and obesity are commonly determined with the help of BMI. In the previous works, height, weight, and BMI estimation through automatic means have predominantly focused on full-body images and videos of humans. The usage of facial images for estimating such traits have been given less importance. In order to facilitate the analysis, the dataset is cleaned along with metadata containing information about the persons height, weight, age, and gender. Gender-based analysis is performed for the prediction of BMI. Finally, an email containing the persons picture along with their details is sent to the concerned health officer.

1 | INTRODUCTION

In this modern world, various social networks like Facebook, Instagram, and Snapchat have various features, such as the exchange of photos, job searching, dating, and blogging. With digital cameras becoming increasingly popular, more and more people worldwide are recording their lives and posting the records on social media networks via photos or videos [1, 2]. Information on identity, the temperament of an individual, demeanor, as well as attributes such as pupil colour, gender, height, weight, age, etc are contained in human faces [3]. The focus has been focused primarily on facial recognition from the perspective of biometric data. Individual identification can be done through such biometric information [4]. Weight and BMI are pertinent indicators for health conditions [5] and excessive weight can be associated to obesity, diabetes, and cardiovascular diseases while lack of proper weight can lead to Malnutrition related diseases [6–8]. In this context, the method presented in this paper contributes to image-based automated self-diagnostic which is the current trend of neural networks in the medical field [9–11]. The goal of this work is to investigate the feasibility of body weight and BMI analysis from the visual appearance of real-time photographs of the human face. The proposed method is useful in establishing the relation between the characteristics of the human face and the body, such as body height and weight [12]. The widely used body fat Indicator the BMI (body mass index) is calculated as

\[ BMI = \frac{W\text{(kg)}}{H\text{ (m)}^2}, \]

where \( W \) and \( H \) are the weight and height, respectively.
The literature shows that BMI is one of the major risk factors for several diseases for employed peoples. The over BMI can cause cancers like colon cancer, breast cancer, and thyroid cancer etc for all genders [13]. BMI is considered a risk factor for myocardial infarction in the field of medicine. In patients, it is often known as a risk factor for dysfunctional angina. The risk of type 2 diabetes and cardiovascular disease (CVD) can be stratified using BMI. Since BMI is closely linked to some diseases, for personal health monitoring and medical research, BMI is considered to be significant. In this project, BMI is particularly used to identify people affected by Malnutrition and Obesity and give immediate attention to them. For the measurement of BMI, special devices are used in person. Through this work, BMI can be conveniently monitored and predicted automatically from images of peoples Daily Life [14, 15]. Through our work, medical researchers will benefit greatly from obtaining BMI data from social networks, which can provide large populations with a range of sources for health monitoring [16].

Several aspects influence the inspiration for this research. Firstly, the weight and fat of the body may be intuitively observed from human vision by humans from face images of other humans [17–19]. Without any trouble, human eyes can easily observe the rise in body fat. Secondly, machine learning methods can be employed to identify facial features like width of the cheek to chin, the ratio of width to upper face height, the ratio of perimeter to the area, size of the eye, lower face to face height ratio and mean eyebrow height, which are measures of obesity and malnutrition and is associated with BMIs [20, 21]. The recent health science research analysis says that humans conclude that a computational approach to the study of body BMI from human face images should be investigated [22]. Considering the close connection between BMI and some diseases, BMI is important for personal health monitoring and medical research. Generally, BMI is measured in person with special devices. For convenient monitoring, this work explores an automatic BMI prediction approach from people’s daily life photos. Nowadays, after the widespread usage of smartphones, a lot of photos are being uploaded and circulated in on the internet everyday. These photos can be utilized for assessment of the health condition of the person in the photo through image processing and machine learning techniques [23–25]. Our work can be of great benefit to medical researchers to access BMI data from social networks, which may provide lots of sources for health monitoring in large populations. People affected by malnutrition and obesity can be easily identified by applying our model on their photos and predicting their BMI. The results obtained can be sent to the concerned person through mail.

2 | LITERATURE SURVEY

Body weight and BMI have received an increased attention in the recent years. The fact that abnormal BMI could cause various diseases has increased research of contactless and low-cost estimation of BMI.

Previous works on estimating human body weight and BMI have utilized full body images. There are a few studies working on estimating human body weight or BMI from body related data, such as body measurements, 3-dimensional (3D) body data and RGB-D body images. Velardo et al. [26] studied the body weight directly from anthropometric data (body measurements) collected by NHANES. A polynomial regression model was employed to analyze the anthropometric data. Cao et al. [27] investigated the use of true measurements of the body (provided by CAESAR 1D database) for the prediction of certain soft biometrics, such as gender and weight. Detailed definitions about many different measurements of the anthropometric feature were included in their work. Velardo et al. [28] studied the weight estimation from 3D human body data by the same anthropometric features as in their previous work. Velardo et al. [29] estimated the weight of a person within 4% error using 2D and 3D data extracted from a low-cost Kinect RGB-D camera output. Nguyen et al. [30] proposed a weight estimator based on single RGB-D images, which utilized the visual colour cues, depth, and gender information. Nahavandi et al. [31] presented a skeleton-free Kinect system to estimate BMI of human bodies. Recently, Pfitzner et al. [32] described the estimation of the body weight of a person who is in front of an RGB-D camera with three different poses: lying, standing and walking.

To the best of our knowledge, the only work related to automated face-based estimation of BMI is a study by Wen and Guo [33], based on the MORPH-II dataset, which obtained mean absolute errors (MAEs) for BMI in the range from 2.65–4.29 for different ethnic categories. The study explored handcrafted features for BMI-estimation and specifically in the method the face was detected, normalized, and an active shape model was fitted, based on which, geometry and ratio features were extracted (cheekbone to jaw width, width to upper facial height ratio, perimeter to area ratio, eye size, lower face to face height ratio, face width to lower face height ratio and mean of eyebrow height), normalized and finally subjected to support vector regression. We note that the BMI-annotation of MORPH-II has not been made public.

Some works studied gender or body shape from body images or 3D scanners. Wu et al. [34] explored gender classification from unconstrained and articulated human body images. Cao et al. [35] developed a method based solely on metrological information from facial landmarks of 2D face images for gender prediction and demonstrated that the geometric features achieve comparable performance as appearance features in gender prediction. Gonzalez-Sosa et al. [36] studied gender estimation based on information deduced jointly from face and body and presented two score-level-fusion schemes of the face and body-based features which outperformed the two individual modalities in most cases. Balan et al. [37] studied the markerless human shape and pose capture from multi-camera video sequences using a richly detailed graphics model of 3D human shape. Their approaches required multi-camera video sequences for 3D model reconstruction. Lu et al. [38] collected anthropometric data by 3D whole body scanners, which consist of four sets of laser beams and CCD cameras.

We propose a deep residual neural network based approach to analyze BMI from human face images using image processing and machine learning. Full body images are not required for this
approach. This method estimates BMI along with the persons age and gender with more accuracy. People with abnormal BMI are declared as patients and their information is sent to through email using SMTP protocol. Multi task cascaded convolutional neural networks have been used for the accurate detection of the persons face by eliminating the other background details. MTCNN has also proved to be accurate in the location of the eyes, nose, mouth and other features in the image of the person. Apart from these, dataset cleaning and analysis have been done for better understanding of the face to BMI dataset. The proposed method has been evaluated and its MAE, root mean square error (RMSE) and correlation values have been compared with other methods.

2.1 Contribution

For directly estimating BMI from human face images, we contribute the following:

1. A ResNet based method has been proposed for estimation of BMI, age and gender from the face along with MTCNN for face detection. The trained residual neural networks take the face images as input and estimate the BMI, age and gender through regression pattern classification.
2. The output of the neural network model which is the input face image along with a bounding box with the estimated BMI, age and gender values are sent through mail using a simple mail transfer protocol (SMTP). The Multipurpose Internet Mail Extensions (MIME) is used to support text in character sets other than ASCII, as well as attachments of audio, video, images, and application programs.
3. The face to BMI dataset is pre-processed before giving it as input to the ResNet model. The null values are removed, the height column is converted to integer datatype and the BMI values were derived from both height and weight column. Data analysis has been performed on the dataset. The distribution of BMI values with respect to different factors like age, gender and race have been graphically represented.

The article is organized as follows: Section 2 explains Literature Survey; Section 3 presents the face detection and BMI estimation algorithm along with the block diagram, flowchart, and modes of operation. Section 4 shows the downloaded and cleaned face to BMI dataset. Section 5 describes the performance measures used to evaluate the model. Section 6 provides the detailed experimental results of the prediction and detection algorithms. The detailed comparative analysis is presented in Section 7. Finally, the conclusion shows Section 8.

3 | PROPOSED APPROACH

The proposed method to identify Malnutrition and obese children from human faces. The proposed system does not require the full body real image of a person. Face detection is done with the Multi-task Cascaded Convolutional Neural Networks on pictures with single/multiple faces. BMI, age, and gender are estimated from a persons face using residual neural networks. The problems of BMI, age, and gender estimation are posed as three separate regression pattern classification problems. The dataset of facial images taken from the internet along with their metadata containing information like gender, age, and BMI. The trained model is tested on images of both males and females. BMI is used as a measure to identify Malnutrition and obesity. The cases are classified into three categories: normal, malnutrition, and obese. The affected persons photo, BMI, age, and gender details are sent to the concerned health officer through email.

The model built is capable of detecting multiple faces accurately and estimating the BMI, Age, and Gender for each person whose image has been given as input to the model. The model performance is divided into three steps:

1. Pre-processing the input image:
   (a) The required features are formed by first loading the image, then resize it into 224 × 224 pixels for input to the training model, and finally convert it to a pixel array [26].
   (b) Map the BMI, age, and gender labels from the meta-data which is an excel file containing detailed information of each image to be trained.
   (c) A random sample from the train and valid dataset to build the generator for model fitting
2. Face detection using MTCNN:
   (a) First, for aligning the images, we need to pre-process the images to be trained by cropping and storing the detected faces.
   (b) Before applying the model, we need to detect the bounded faces from the input photo and then apply the training model for predicting BMI.
3. BMI prediction using neural network:
   (a) The apply transfer learning with the ResNet50 backbone.
   (b) Multi-task learning is applied by training the model to learn three tasks together

The input image is taken and cascaded convolution neural networks are used to detect the face in the image. The ResNet50 backbone is used for feature extraction. The backbone is loaded with pre-trained weights from the 1000-class ImageNet classification task [27] used for transfer learning. The extracted features are then applied to all the three heads, the BMI, Age and Gender head. These Prediction heads are two fully connected layers. The dataset is divided into training and testing datasets, the model is trained in the training mode and the performance of the trained model is evaluated in the testing mode as shown in the block diagram in Figure 1. During the training mode,

1. Face detection is done on the training dataset using Multi-task Cascaded Convolutional Neural Networks.
2. The image is resized to 224 × 224 pixels.
3. The detected face is sent as input to the neural network.
4. The images are trained using the ResNet architecture for identification of facial features useful in the estimation of BMI, age and gender values.
5. The network is trained end-to-end with the classical back-propagation algorithm.
6. The Caffe deep learning framework is used for implementation [28]. During the testing mode
(a) The input image is taken from the testing dataset.
(b) The real-time data image is given to the trained model.
(c) The BMI, age and gender values are evaluated through regression pattern classification.
(d) The accuracy, MAE, and AUC are calculated.
(e) The obtained results are sent through email. The flowchart representing the preprocessing, data analysis, training, and testing code is shown in Figure 2.

3.1 Residual neural networks

The 50-layer Residual Network architecture is used in this project. Residual links overcome challenges encountered in deep neural networks such as disappearing or exploding gradients and degradation is useful and has been implicitly important for training very deep convolutionary neural networks [33] for other techniques such as initialization strategies, better optimizers, skip ties, information transfer. ResNets are known to improve the training speed while increasing the accuracy of object classification, object detection, and segmentation. The typical structure of a ResNet module is shown in Figure 3. As shown in the figure, when the model decreases in accuracy with increasing layers, the input layers bypass the hidden layers to the output layers. Sometimes deep neural networks perform better at the shallow layers than at the deep layers which are opposite
of what is expected. During such cases, residual networks are used so that no compromise is made on the models accuracy. When residual blocks are used along with the mapping of identities, the model can be expressed by the following equation:

\[ Y_{l+1} = Y_l + F(Y_l, X_l) \].  

(2)

Where the network input and output of the \( l^{th} \) unit are \( Y_{l+1} \) and \( Y_l \), respectively. The residual function is represented by \( F \) with \( Y_l \) and \( X_l \) as its parameters. The architecture defines the characteristics of residual neural networks based on sequentially stacked residual blocks. The original conv-BN-ReLU is used for the order of batch normalization, activation, and convolution in residual blocks. The two consecutive 3\( \times \)3 convolutions basic blocks are used in this work. The activation function used is a rectified linear unit.

3.2 | Simple Mail Transfer Protocol

Internet is becoming more popular nowadays and e-mail has emerged as one of the most important services. SMTP is one of the most popular methods to switch the e-mail from one person to the next [34]. While sending a mail, SMTP is used as a push protocol used for sending emails on the sender side whereas, in order to retrieve such addresses, IMAP (internet message access protocol) or POP (post office protocol) are used on the receiver side. SMTP is the protocol of an application layer: A TCP connection is opened on the client's side the person who wants to send the mail to the SMTP server and then sends the mail over the link. The SMTP server is always kept in listening mode. The SMTP mechanism initiates a port link (25) as soon as a client-side TCP connection is triggered. The client process can deliver the mail instantly to the receiver by successfully establishing the TCP link.

4 | DATASET ANALYSIS

The data used for training was scraped from the web by JavaScript using python. This small dataset consists of 1530 records representing each person in the training real-time images and 16 columns consisting of detailed information about each person. The data was combined with meta-data. Table 1 shows the original data downloaded from the internet before the cleaning process. The data was pre-processed. Height and weight were used to derive BMI. The dataset was cleaned, and the number format for gender was created (1 for male, 0 for female). The first five rows of the preprocessed dataset (head of the dataset) on which the training and testing process takes place are shown in Table 4. The height column in the original dataset is of string datatype representing the height in feet and inches. For the convenience of training the model to give better results and to derive the BMI column, the string in the height column is parsed into feet and inches column each of integer datatype. This process can be seen in Table 2. Then some incorrect inches present in the dataset with value greater than 12 were replaced by their proper values. Then the feet column was multiplied by 12 and added with the inches column to create a height column of integer datatype which represents the height only in inches. Then the height in inches was converted into meters and the weight in pounds was converted into kg by using the following formulas:

\[ \text{Meter} = \text{inches} \times 2.54/100, \]  

(3)

\[ \text{Kilogram} = \text{pounds} \times 0.453592; \]  

(4)

The BMI column was derived using the formula,

\[ \text{BMI} = \text{weight}/\text{Height}^2. \]  

(5)
A number format was created for gender since the neural network model uses regression pattern classification technique (Male = 1, Female = 0). These processes are indicated in Table 3.

1. As shown in Figure 5, age is a nearly truncated normal distribution with a minimum of 18 and an average of 34.
2. 80% of the data has real-time images of males as shown in Figure 4, thus displaying the gender imbalance in the data.
3. As shown in Figure 7, race is dominated by American Black and White people with less number of Asian samples.
4. The BMI graph is normally distributed, with a mean BMI of 26 as shown in Figure 6.
5. There is no correlation observed between BMI, age, gender, and race as shown in Figures 8–10.

The data consist mainly of frontal images. A variety of factors influencing the images includes illumination, speech, quality of the image and resolution. Embellishment of the photos along with the inclusion of lipstick, cosmetic surgery, beard, and moustache provide further complications. Metadata related to the subject were obtained. In the metadata, a number format for gender was created. Since we are applying the given data on a regression pattern classification model, if the predicted gender value is close to 1, then the person in the image is male and if the gender value is close to 0, the person in the image is female. The height column in the original data is of String data type. It is therefore divided into two separate columns named inches and feet which are of Integer datatype. The index column in the cleaned dataset represents the name of the image containing the person representing the particular row data. BMI for each row is derived from height and weight using Equation (1).

5. PERFORMANCE MEASURES

**MAE:** After considering all errors in a set of predictions, MAE calculates the average magnitude value without considering their direction as given in Eqn.6. It represents the average of the absolute distinctions between the absolute differences the predicted observations and actual values obtained where all individual distinctions are equivalent to each other, weightage in the
given data.

$$MAE = \frac{1}{n} \sum_{j=1}^{n} |y_j - \hat{y}_j|.$$  \hspace{1cm} (6)

**RMSE:** RMSE is a law of quadratic scoring that often calculates the average error magnitude just like MAE as given in Equation (7). The only difference is it calculates the average of square differences between prediction and actual observation is the square root.

$$RMSE = \sqrt{\frac{1}{n} \sum_{j=1}^{n} (y_j - \hat{y}_j)^2}. \hspace{1cm} (7)$$

**AUC:** Area Under the ROC Curve: Machine Learning needs to evaluate the outcomes of each model to understand its accuracy and usability which is an essential task. When it creates a classification type model and the AUC–ROC Curve can be relied on. When the output of the multi-class classification issue needs to be checked or visualized which assigns the results obtained to different classes, it can use the ROC (Receiver Operating Characteristics) curve of AUC (Area under the Curve) technique. Equations (8)–(10) represent the AUC technique.

$$TPR / \text{Recall} / \text{Sensitivity} = \frac{TP}{(TP + FN)}, \hspace{1cm} (8)$$

$$\text{Specificity} = \frac{TN}{(TN + FP)}, \hspace{1cm} (9)$$

$$FPR = 1 - \text{Specificity} = \frac{FP}{(TN + FP)}. \hspace{1cm} (10)$$

If the AUC is near to 1, it means that the model is an excellent model having a strong separability measurement while a poor model that has the worst amount of separability will have an AUC near to 0. If the AUC is near 0, then the model is predicting the opposite of the expected output. It calculates 0s as 1s and 1s as 0s. If AUC is 0.5, then there is no class separation potential for the model at all. By analogy, higher AUC means that the proposed model is very good at differentiating between patients who are affected by a disease and patients who are not affected by that particular disease.

## 6 RESULTS

In this section, the results obtained after training the model can be discussed. Dense layers follow the convolution neural blocks in our CNN network to output the prediction. Since we are predicting three values (BMI, age, and gender), it is necessary to build three separate models for each prediction head. This is will increase the maintenance efforts since the three models need to be trained and serialized separately as shown in Figure 11.

Since we are going to apply all three prediction heads on the same image can be shared the same backbone for the three different prediction heads, and thus only one model is needed for the training process. A class FacePrediction was built and trained for the implementation of the model. For this function, a Deep Relationship Network with mutual convolutional and task-specific fully linked layers with matrix priors is used. The number of trained samples is 1227 and the number of tested samples is 316 with 18 columns each. In the evaluation dataset, the ratio of the number of samples to the training dataset is approximately 1:4. During the training process, the validation loss improved from infinity to 12.4 in epoch 1 and from 12.4 to 4.98 in epoch 2.

### 6.1 Face detection

Multiple faces were detected within an image and the bounding box for each face is drawn with the help of cascaded convolution neural networks. The process of cropping the image is shown in Figure 12. Figure 13 displays the pixel values of the cropped image along with the confidence factor. Confidence scores are a critical component of face detection and comparison systems. MTCNN makes predictions of whether there exists a face in an image or how accurately a face matches with another face in another image, with a corresponding level.
of confidence score which represents the probable accuracy in the prediction [36]. The key points represent the position (or pixel values) of the left eye, right eye, nose, left and right part of the mouth.

It serves two purposes for this project:

1. For the alignment, extraction, and cropping of the facial features in an image. Before training the model, preprocessing of each image for cropping and extracting the faces by focusing on the facial part is taken place. For the next step of training the images, the cropped faces are stored separately.

2. Enable prediction for multiple persons in the same image. When multiple faces are present in an image, all the faces will be detected from the input image and the same preprocessing step will be applied for each image.

### 6.2 Prediction results

The trained model class Face Prediction can be used in three different ways for prediction purposes.

1. **Predict from a single image**: As shown in Figures 14 and 15, the model can be applied to an image containing a single person to get the BMI, age, and gender details. Figure 14 contains the image of a male and Figure 15 contains the image of a female.

2. **Predict multiple faces**: As shown in Figure 16, the model can be applied to images with multiple males and females to detect faces and predict for all faces.

3. **Predict from a directory**: The trained model can be used to predict from a directory and output as a pandas data frame.

### 6.3 Send output through email

The output obtained is sent through email. smtplib is used for this purpose. The email contains an attachment of the persons real-time image as shown in Figure 17. In the real-time image of the person, the BMI, age, and gender details are also given. A python is a powerful tool that offers a native library to send emails and no external library is required to import any. Smtplib is the native library of python which creates a session object for the SMTP client via which it sends emails to any legitimate
The performance metrics of the ResNet model

| Ref. No.       | BMI (MAE) | BMI (COR) | Age (MAE) | Gender (AUC) | BMI (RMSE) | Age (RMSE) |
|---------------|-----------|-----------|-----------|--------------|------------|------------|
| ResNet50      | 5.02      | 0.325     | 7.164     | 0.998        | 4.56       | 5.66       |
| Vgg16 [39]    | 6.131     | 0.199     | 9.573     | 0.99         | 4.99       | 6.04       |
| Vgg16 seqMT [39] | 7.074    | 0.285     | 8.895     | 0.988        | 5.21       | 7.02       |

6.4 The tensorboard results for BMI loss obtained during model training

While training the model, the validation loss and the BMI loss decrease during each epoch. As the neural network model trains, the decrease in loss values has been graphically represented in the below Figure 18. These graphs were obtained through tensorboard. Before training, the validation loss is at infinity and as the model trains it goes below 5.

7 COMPARISON

The previous approaches have focused on the estimation of BMI from full-body images. Some of the works used anthropometric features and some other works included deep neural networks for estimation of BMI from full-body images. Most of the photos available on the internet are nowadays only face images. After the introduction of front cameras on mobile phones, people have started to take photos of themselves on their own. Such photos include only the face part and not the whole body. Taking the above fact into consideration, we have developed a model that uses MTCNN for face detection and residual neural networks for predicting BMI from human facial images. MTCNN has proven to be better than other algorithms when it comes to detecting multiple faces. It also displays the confidence factor during face detection which proves the accuracy of the algorithm in detecting faces. Residual neural networks overcome the shortcomings like degrading performance which is encountered in other types of deep neural networks and thus increases the accuracy of the predicting model. The model has been compared to predict BMI from face images with two other methods. One is the VGG16 method and the other is a sequential multi-task VGG method. As input, the methods involve specific frontal face images. For comparison, we took 1543 real-time images that contain images of the frontal face and then crop the pictures of the face. The 1543 images are divided into collections of training and evaluations, which contain 1227 and 316 images, respectively. The comparison of the results is shown in Table 5.

It can be shown that the approach suggested is outperforming the other two methods in BMI and gender prediction but performs slightly less than the sequential multitask method in age prediction. The performance measures used for comparison are the MAE and area under the curve (AUC). The BMI correlation factor has also been evaluated for all the 3 models. Our model has achieved an MAE of 5.02 for BMI, a BMI correlation of 0.325; the MAE for age is 7.164 and the area under the curve value for gender was 0.998. Much better results could be achieved with a better dataset that has an equal distribution of males and females and a wide range of age groups from small children to old people with different BMI values. The performance of a neural network model mainly depends on the dataset used to train the model. Thus, by obtaining a better dataset, our model will be able to achieve lower error values and higher accuracy and precision.

8 CONCLUSION

In this work, the relationship between body weight and facial appearance is investigated and the BMI values from human face images are estimated. The MAEs, correlation values, and area under the curve values have been compared with other models and the better performance of our model validates its use in the estimation of BMI. More specifically, the process can be classified into three steps: A dataset is downloaded from the inter-
net and the images are preprocessed for face detection; neural networks are trained on the cropped faces to build a BMI estimation model; lastly based on the values estimated, the person is classified and his details are sent through email. To address the BMI estimation problem, a residual neural network is built and trained, and to facilitate this analysis, the face to BMI image dataset was collected and cleaned. There was no gender bias observed during BMI estimation but more work needs to be carried out for better results. The age, weight, and BMI estimator were motivated by the current need for self-diagnostic tools in the field of remote healthcare caused due to increasing population and degrading health, as well as for providing better security through soft biometrics categorization in security applications.

The Errors are within appropriate ranges of the three estimation tasks tested by multiple measurements. In certain cases, the proposed approach works better compared to the methods of body image analysis. In addition, the residual neural networks significantly outperform the Vgg16 feature and sequential multitask feature on BMI estimation. The main disadvantage of our project is the absence of a large dataset with wide ranges of BMI values. A larger dataset with a wide variety of images can make the training model learn better thus providing better results. It is promising to evaluate body weight or BMI from the 2D face images visually, based on all experimental data. In the future, we will explore the DNN-based method to improve the accuracy of this visual BMI analysis problem.

ORCID
Thippa Reddy Gadekallu https://orcid.org/0000-0003-0097-801X

REFERENCES
1. Zhu, Y., et al.: Still-to-video face matching using multiple geodesic flows. IEEE Trans. Inf. Forensics Secur. 11(12), 2866–2875 (2016)
2. Gadekallu, T.R., et al.: A novel PCA–whale optimization-based deep neural network model for classification of tomato plant diseases using GPU. J. Real-Time Image Process. 1–14 (2020)
3. Dantcheva, A., Elia, P., Ross, A.: What else does your biometric data reveal? A survey on soft biometrics. IEEE Trans. Inf. Forensics Secur. 11(3), 441–467 (2015)
4. Ahmed, U., Srivastava, G., Lin, J.C.W.: A deep reinforcement learning model for data sanitization in IoT networks. Comput. Netw. 107914 (2021)
5. Javed, A.R., et al.: Automated cognitive health assessment in smart homes using machine learning. Sustain. Cities Soc. 65, 102572 (2021)
6. Wu, J.M.T., et al.: Fuzzy high-utility pattern mining in parallel and distributed Hadoop framework. Inf. Sci. 553, 31–48 (2021)
7. Lin, J.C.W., et al.: Linguistic frequent pattern mining using a compressed structure. Appl. Intell. 1–18 (2021)
8. Deepa, N., et al.: An AI-based intelligent system for healthcare analysis using ridge-adaline stochastic gradient descent classifier. J. Supercomput. 77, 1998–2017 (2021)
9. Gadekallu, T.R., et al.: Deep neural networks to predict diabetic retinopathy. J. Ambient Intell. Humaniz. Comput. (2020)
10. Iwendi, C., et al.: The use of ensemble models for multiple class and binary class classification for improving intrusion detection systems. Sensors 20(9), 2559 (2020)
11. Garg, R., et al.: Using NMR-based metabolomics to evaluate postprandial urinary responses following consumption of minimally processed wheat bran or wheat aleurone by men and women. Nutrients 8(2), 96 (2016)
12. Rativa, D., Fernandes, B.J., Roque, A.: Height and weight estimation from anthropometric measurements using machine learning regressions. IEEE J. Transl. Eng. Health Med. 6, 1–9 (2018)
13. Renelhan, A.G., et al.: Body-mass index and incidence of cancer: A systematic review and meta-analysis of prospective observational studies. Lancet 371(9612), 569–578 (2008)
14. Jain, A.K., Ross, A., Pankanti, S.: Biometrics: A tool for information security. IEEE Trans. Inf. Forensics Secur. 1(2), 125–143 (2006)
15. Venkatraman, S., Alazab, M.: Use of data visualisation for zero-day malware detection. Secur. Commun. Netw. 2018, 1728303:1–1728303:13 (2018)
16. Nagar, A., Nandakumar, K., Jain, A.K.: Multi-biometric cryptosystems based on feature-level fusion. IEEE Trans. Inf. Forensics Secur. 7(1), 255–268 (2011)
17. Ranjan, R., Patel, V.M., Chellappa, R.: Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. IEEE Trans. Pattern Anal. Mach. Intell. 41(1), 121–135 (2017)
18. Vasan, D., et al.: Imcfn: Image-based malware classification using fine-tuned convolutional neural network architecture. Comput. Networks 171, 107138 (2020)
19. Bugeja, S., Garg, L., Audu, E.E.: A novel method of EEG data acquisition, feature extraction and feature space creation for early detection of epileptic seizures. In: 2016 38th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC), pp. 837–840. IEEE (2016)
20. Wang, H., Hu, J., Deng, W.: Face feature extraction: A complete review. IEEE Access 6, 6001–6039 (2017)
21. Zehra, W., et al.: Cross corpus multi-lingual speech emotion recognition using ensemble learning. Comput. Intell. Syst. 1–10 (2021)
22. Lu, Z., Jiang, X., Kot, A.: Deep coupled ResNet for low-resolution face recognition. IEEE Signal Process Lett. 25(4), 526–530 (2018)
23. Wang, J., Zhang, Y.: Median filtering forensics scheme for color images based on quaternion magnitude-phase CNN. Comput. Mater. Contin. 62(1), 99–112 (2020)
24. Xu, G., et al.: Image interpolation via Gaussian-sine interpolators with partition of unity. Comput. Mater. Contin. 62(1), 309–319 (2020)
25. Yun, H.J., Kim, K.J., Chun, J.C.: Shadow detection and removal from photo-realistic synthetic urban image using deep learning. Comput. Mater. Contin. 62(1), 459–472 (2020)
26. Cesar, I., et al.: Enhancing the performance of image preprocessing for classification and object detection. Politehnica i dizajn 8(01) (2020)
27. Russakovsky, O., et al.: Imagenet large scale visual recognition challenge. Int. J. Comput. Vision 115(3), 211–252 (2015)
28. Cengil, E., Çınar, A., Özbay, E.: Image classification with Caffe deep learning framework. In: 2017 International Conference on Computer Science and Engineering (UBMK), pp. 440–444. IEEE (2017)
29. Velardo, C., et al.: Building the space scale or how to weigh a person with no gravity. In: 2012 IEEE International Conference on Emerging Signal Processing Applications, pp. 67–70. IEEE (2012)
30. Nguyen, T.V., Feng, J., Yan, S.: Seeing human weight from a single RGB-D image. J. Comput. Sci. Technol. 29(5), 777–784 (2014)
31. Nahavandi, D., et al.: A skeleton-free Kinect system for body mass index assessment using deep neural networks. In: 2017 IEEE International Systems Engineering Symposium (ISSE), pp. 1–6. IEEE (2017)
32. Pitzner, C., May, S., Nüchter, A.: Body weight estimation for dose-finding and health monitoring of lying, standing and walking patients based on RGB-D data. Sensors 18(5), 107138 (2020)
33. Keeler, J.D., et al.: Residual activation neural network. Expert Syst. Appl. 9(1), III–III (1995)
34. Klensin, J., et al.: SMTP service extensions. RFC 2846, November (1995)
35. Cao, D., et al.: Can facial metrology predict gender? In: 2011 International Joint Conference on Biometrics (IJCB), pp. 1–8. IEEE (2011)
36. Xiang, J., Zhu, G.: Joint face detection and facial expression recognition with MTCNN. In: 2017 4th International Conference on
37. Balan, A.O., et al.: Detailed human shape and pose from images. In: 2007 IEEE Conference on Computer Vision and Pattern Recognition, pp. 1–8. IEEE (2007)

38. Lu, J.M., Wang, M.J.J.: Automated anthropometric data collection using 3D whole body scanners. Expert Syst. Appl. 35(1-2), 407–414 (2008)

39. Tammina, S.: Transfer learning using vgg16 with deep convolutional neural network for classifying images. Int. J. Sci. Res. Publ. 9(10), 143–150 (2019)

How to cite this article: 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. 2021;1–12.
https://doi.org/10.1049/ipr2.12222