Face BMI Prediction Using Deep Neural Networks

By

Kishor Mannur

Supervisor: Utku Pamuksuz

A Course Project

Submitted to the University of Chicago in partial fulfillment

of the requirements for the degree of

Master of Science in Applied Data Science

Physical Sciences Division

May 2023

Abstract

Body Mass Index (BMI) is a vital health indicator widely used for assessing weight status and associated health risks. This paper introduces an innovative approach to accurately detect BMI from facial images by leveraging a combination of tuned deep-learning models. The primary objective is to surpass the current metrics published in research journals, achieving higher accuracy in BMI prediction. The methodology involves utilizing state-of-the-art architectures such as VGGFace, VGG16, ResNet50, DenseNet121, and others, each fine-tuned through multiple iterations to optimize their performance. A diverse dataset of facial images, encompassing a broad representation of age, and gender, is employed, accompanied by corresponding ground truth BMI labels. Preprocessing techniques, including image augmentation and normalization, are applied to enhance model robustness and generalization. Evaluation is conducted using various performance metrics, and the results are compared against benchmarks established by prior research. Preliminary findings demonstrate promising improvements over existing methods, highlighting the potential of the ensemble approach to advance BMI detection technology. The implications of this research extend to healthcare, wellness programs, and personalized health management, with the potential to improve BMI assessment accuracy and enhance overall health outcomes for individuals and populations at large.

Executive Summary

This paper presents an innovative approach to accurately detect Body Mass Index (BMI) from facial images by utilizing a combination of tuned deep learning models. The objective is to surpass current metrics published in research journals and achieve higher accuracy in BMI prediction. The methodology involves employing state-of-the-art models, including VGGFace, VGG16, ResNet50, DenseNet121, and others, which have undergone multiple iterations of tuning to optimize their performance.

A diverse dataset of facial images, covering a wide range of ages, gender, and ground truth BMI labels are used. Preprocessing techniques such as image augmentation and normalization are applied to enhance the models' robustness and generalization. Evaluation is conducted using various performance metrics, and the results are compared against benchmarks established by prior research. Preliminary findings show promising improvements over existing methods, highlighting the potential of the ensemble approach in advancing BMI detection technology. The implications of this research extend to healthcare, wellness programs, and personalized health management, with the potential to enhance BMI assessment accuracy and improve health outcomes for individuals and populations.

Introduction

Body Mass Index (BMI) is a widely used health indicator for assessing weight status and associated health risks. It plays a crucial role in determining the overall well-being of individuals and populations. Traditionally, BMI is calculated based on height and weight measurements. However, advancements in computer vision and deep learning techniques have opened new possibilities for BMI detection using facial images. Analyzing BMI from facial images can provide a non-invasive and convenient approach for assessing weight status.

Problem Statement

While BMI detection from facial images holds significant potential, achieving high accuracy in prediction remains challenging. Existing approaches often rely on simplistic features or do not fully leverage the capabilities of deep learning models. Consequently, there is a need to develop a more sophisticated and accurate BMI detection system using advanced deep-learning techniques.

The primary objective of this research is to surpass the current metrics published in research journals by employing a combination of tuned deep-learning models. By leveraging the strengths of state-of-the-art architectures, such as VGGFace, VGG16, ResNet50, DenseNet121, and others, we aim to improve the accuracy of BMI detection from facial images. Multiple iterations of tuning these models will be performed to optimize their performance specifically for BMI prediction. The proposed approach will utilize a diverse dataset of facial images, encompassing a wide range of ages and gender, to ensure robustness and generalization.

By addressing the limitations of existing methods and leveraging the power of deep learning models, we aim to advance the field of BMI detection from facial images. The results of this research have the potential to contribute to improved healthcare, wellness programs, and personalized health management by providing more accurate BMI assessments.

Analysis Goals

The primary goal of this research is to develop an advanced BMI detection system using tuned deep-learning models for the accurate prediction of BMI from facial images. To achieve this, the following analysis goals and scope have been defined:

Select and employ state-of-the-art deep learning models, including VGGFace, VGG16, ResNet50, DenseNet121, and others, known for their effectiveness in image recognition tasks. Perform multiple iterations of tuning to optimize the models specifically for BMI prediction. Fine-tune the models' hyperparameters, such as learning rate, batch size, and regularization techniques, to enhance their performance. Collect a diverse dataset of facial images, covering a broad range of age groups, gender, and ethnicity. Ensure sufficient representation from different demographics to enhance the generalization capability of the models. Curate the dataset with accurate ground truth BMI labels for each facial image.

Apply preprocessing techniques, such as image augmentation and normalization, to enhance the robustness and generalization of the models. Standardize the images to reduce variations in lighting conditions, pose, and facial expressions. Implement appropriate data augmentation techniques to increase the dataset size and introduce variability.

Evaluate the trained models using various performance metrics, including RMSE, MAE, accuracy, precision, recall, and F1 score. Compare the results obtained from the ensemble of tuned models against benchmarks established by prior research and existing methods. Assess the models' robustness by conducting cross-validation and testing on independent datasets.

Scope:

The scope of this research is limited to BMI detection from facial images using tuned deep-learning models. The focus is on improving accuracy and surpassing the current metrics published in research journals. The analysis will not cover other methods of BMI estimation, such as traditional height and weight measurements or alternative techniques not involving facial images.

The findings and outcomes of this research have the potential to contribute to the field of BMI assessment and have implications in healthcare, wellness programs, and personalized health management. The improved accuracy in BMI detection from facial images can aid in early identification of potential health risks and enable targeted interventions for individuals and populations.

Background

Data

The provided data consist of image files in .bmp format and a CSV file with related information to the images. However, there aren’t as many image files in the folder as there are entries in the CSV file. The provided data file consists of 4206 entries. It consists of five columns, each containing different types of information. The columns and their descriptions are as follows:

BMI: A float64 column representing the Body Mass Index (BMI) values of individuals. BMI is a numerical value calculated based on an individual's weight and height, and it serves as an indicator of weight status.

Gender: An object (string) column indicating the gender of the individuals. It provides information about the male or female categorization of the data points.

is_training: An integer column that denotes whether the corresponding entry is part of the training set or not. It serves as a binary flag, where a value of 1 indicates that the entry is included in the training set.

Name: An object (string) column containing the names or identifiers of the individuals. It provides a unique identifier for each data point.

This dataset can potentially be used for training and evaluating machine learning models for BMI prediction based on gender and other features. The information provided in the dataset can contribute to understanding the relationship between BMI, gender, and other variables, as well as developing accurate BMI detection models.

Methodology:

Data Collection and Preparation:

- Obtain facial image dataset in .bmp format and associated CSV file with BMI, gender, is_training, and name columns.
- Preprocess the data by extracting relevant information and organizing it for analysis.

Deep Learning Model Selection:

- Select state-of-the-art deep learning models known for their effectiveness in image recognition tasks, such as VGGFace, VGG16, ResNet50, DenseNet121, and others.
- Consider the models' architecture, complexity, and computational requirements.

Model Tuning and Hyperparameter Optimization:

- Implement multiple iterations of tuning to optimize the selected deep-learning models specifically for BMI prediction.
- Fine-tune the models' hyperparameters, including learning rate, batch size, regularization techniques, and optimizer algorithms.

Dataset Augmentation and Preprocessing:

- Normalize the facial images to reduce variations in lighting conditions, pose, and facial expressions.
- Implement appropriate data augmentation techniques may include rotation, translation, scaling, flipping, and adding noise to the images.

Model Training and Evaluation:

- Train the tuned deep learning models on the augmented and preprocessed training dataset. Monitor and record performance metrics during the training process, such as loss, accuracy, and validation metrics.
- Evaluate the trained models using various performance metrics, including RMSE, MAE, accuracy, precision, recall, and F1 score.
- Compare the results obtained from the ensemble of tuned models against benchmarks established by prior research and existing methods.

Result Analysis and Interpretation:

- Analyze and interpret the performance metrics obtained from the trained models.
- Identify the best-performing model(s) based on the evaluation metrics and benchmarks.
- Examine the models' strengths, limitations, and areas of improvement.

Limitations and Future Work:

- Discuss any limitations or constraints encountered during the research process.
- Propose potential areas for future research and improvement in BMI detection from facial images.

Conclusion:

- Summarize the key findings, outcomes, and contributions of the research project.

Feature Engineering:

Image Data Augmentation: Image Data Augmentation increases dataset diversity by applying transformations to images. Parameters like rescaling, rotation, shifting, shearing, zooming, and flipping are used with the ImageDataGenerator class.

Early Stopping: Early Stopping prevents overfitting by monitoring a chosen metric, like validation loss. Training stops if the metric doesn't improve for a specified number of epochs. Best weights can be restored based on minimum validation loss.

Model Training: The model is trained for a fixed number of epochs, with early stopping based on validation loss. Training history, including loss and metrics, is stored for analysis. Early stopping helps find the optimal training duration, preventing overfitting.

Modeling Frameworks:

VGG16 with Dropout and L2 Regularization: VGG16 is a pre-trained convolutional neural network (CNN) model that is commonly used for image classification. In this model, the base VGG16 model is used as a feature extractor, and additional layers are added on top for prediction. Dropout layers are used to prevent overfitting, and L2 regularization is applied to the fully connected layers for weight regularization.

ResNet50V2 with Custom Dense Layers: ResNet50V2 is another pre-trained CNN model known for its deep residual architecture. In this model, the ResNet50V2 base model is combined with additional layers for classification. The model includes convolutional, pooling, and dense layers. The initial layers of ResNet50V2 are frozen, while the last few layers are trainable.

DenseNet121 with Dropout: DenseNet121 is a pre-trained CNN model that emphasizes feature reuse and dense connections between layers. In this model, the DenseNet121 base model is used as a feature extractor, and additional dense layers are added for prediction. Dropout layers are incorporated to regularize the model and prevent overfitting.

VGG16 for Feature Extraction with Fine-tuning: This model utilizes VGG16 as a feature extractor rather than a complete model. The last convolutional layer is extracted and used as an output layer. The initial layers of VGG16 are frozen, while the later layers are fine-tuned by making them trainable. Additional fully connected layers are added on top for classification.

Findings

Evaluation Metrics for Various Model Configurations:

Model Configuration	RMSE	MAE
VGG16 (2 Dense Layers followed by Dropouts, Regularization	10.18	7.66
VGG16 (2 Dense Layers followed by Dropouts, Regularization, Data Augmentation	10.61	7.79
VGG16 (2 Dense Layers followed by Dropouts, Learning Rate Tuned)	10.65	8.11
DenseNet121	12.51	9.31
ResNet50	11.31	8.35

In this evaluation of various model configurations, the performance of different models was measured using the RMSE (Root Mean Square Error) and MAE (Mean Absolute Error) metrics. The first model, VGG16 with 2 Dense Layers followed by Dropouts and Regularization, achieved the lowest RMSE of 10.18 and the lowest MAE of 7.66. This indicates that it had the highest predictive accuracy among the models tested. The second variant of VGG16, which included Data Augmentation in addition to Dropouts and Regularization, had slightly higher RMSE and MAE values of 10.61 and 7.79 respectively. The third version of VGG16, with Learning Rate Tuning, also had similar performance with an RMSE of 10.65 and MAE of 8.11. DenseNet121 obtained higher scores with an RMSE of 12.51 and MAE of 9.31, while ResNet50 achieved an RMSE of 11.31 and MAE of 8.35. These results suggest that the VGG16 model with 2 Dense Layers followed by Dropouts and Regularization exhibited the best performance in terms of accuracy.

Conclusion

In my final deliverable for the project, I conducted a comprehensive evaluation of various model configurations, including VGG16 with different variations, DenseNet121, and ResNet50. This analysis has allowed me to gain a deeper understanding of their performance and choose the most suitable model for my specific task. The impact of my work can be measured in terms of improved predictive accuracy, enabling me to make more informed decisions and potentially reducing errors or costs associated with incorrect predictions. Going forward, my next steps involve deploying the chosen model in a production-like environment, monitoring its performance, and conducting further evaluation using real-world data to assess its practicality.

Furthermore, this project has opened possibilities for exploring advanced techniques and extensions, such as incorporating transfer learning or experimenting with different architectures, to enhance the performance of the model. Overall, this individual project has been a valuable learning experience, providing me with insights and skills in model evaluation and selection.

References:

- Enes Kocabey, Mustafa Camurcu, Ferda Ofli, Yusuf Aytar, Javier Marin, Antonio Torralba, Ingmar Weber (2017). Face-to-BMI: Using Computer Vision to Infer Body Mass Index on Social Media. Cornell University arXiv
- Chong Yen Fook; Lim Chee Chin; Vikneswaran Vijean; Lim Whey Teen; Hasimah Ali; Aimi Salihah Abdul Nasir (2021). Investigation on Body Mass Index Prediction

from Face Images. 2020 IEEE-EMBS Conference on Biomedical Engineering and Sciences (IECBES)

- P. Vishnu Raja; K. Sangeetha; D. Sanjay Kumar; A. Surya; D. Subhathra (2022). Prediction of human height, weight and BMI from face images using machine learning algorithms. AIP Conference Proceedings 2393, 020192 (2022)