CONVOLUTIONAL NEURAL NETWORKS FOR FOOD CALORIE ESTIMATION

LEONA RAMY

FACULTY OF COMPUTING AND INFORMATICS
UNIVERSITI MALAYSIA SABAH
2022

UMS
UNIVERSITI MALAYSIA SABAH

CONVOLUTIONAL NEURAL NETWORKS FOR FOOD CALORIE ESTIMATION

LEONA RAMY

THESIS SUBMITTED IN PARTIAL FULFILMENT FOR THE DEGREE OF BACHELOR OF COMPUTER SCIENCE WITH HONOURS (NETWORK ENGINEERING)

FACULTY OF COMPUTING AND INFORMATICS
UNIVERSITI MALAYSIA SABAH
2022

NAME	**:**	LEONA RAMY
MATRIC NUMBER	**:**	BI18110067
TITLE	**:**	CONVOLUTIONAL NEURAL NETWORKS FOR FOOD CALORIE ESTIMATION SYSTEM
DEGREE	**:**	BACHELOR OF COMPUTER SCIENCE WITH HONOURS (NETWORK ENGINEERING)
VIVA'S DATE	**:**	25TH JANUARY 2022

CERTIFIED BY;

1. **SUPERVISOR**

Signature

ERVIN GUBIN MOUNG, DR

DECLARATION

I hereby declare that the material in this thesis is my own except for quotations, equations, summaries and references, which have been duly acknowledged.

22nd FEBRUARY 2022

LEONA RAMY

BI18110067

ACKNOWLEDGEMENT

Foremost, I would like to express my gratitude to God for the guidance throughout my Final Year Project journey. I can finish my final year because of His strength for me. I would also thank my family in Christ for taking care of me keep supporting me physically and spiritually. I get experienced wonderful things throughout my study in UMS together with them, thank you so much. Besides, I would also like to thank my supervisor Dr Ervin Gubin Moung for the continuous support of my Final Year Project for his patience, enthusiasm and immense knowledge. He is always there when I need something regarding my project. He is a committed person and responsible for his supervision of students. His guidance helped me in all the time of research and writing this thesis. My sincere thanks also go to both Dr James Mountstephens and Associate Professor Dr Ag Asri Hj Ag Ibrahim as my examiner and panel for giving me chances to improve my Final Year Project report. Last but not least, I would like to thank my family and friends for always supporting motivate me. I would not achieve anything without them.

ABSTRACT

Understanding calorie intake is important as it influences an individual's health, and the amount of calories in food tells people how much potential energy they contain. In recent years, the food calorie estimation system has become popular due to calorie intake rose in most of the world because there has been a decrease in the quality of diet. However, a problem statement has been highlighted where existing researches did not include mass estimation procedures for their calorie estimation system. The number of calories estimated for each food item was constant. The purpose of this paper is to implement a food calorie system that utilized a mass estimation procedure, to test and evaluate the food calorie estimation to validate its effectiveness and to develop a GUI of the proposed system. The application of Deep Learning has been increasing dramatically in various fields, including food calorie estimation. In this paper, we proposed to build a system that can estimate the calorie content of food using the DL algorithm. Future works of this project is to fulfil all the objectives targeted.

ABSTRAK

Memahami pengambilan kalori adalah penting kerana ia mempengaruhi kesihatan individu, dan bilangan kalori dalam makanan memberitahu orang ramai berapa banyak tenaga berpotensi yang terkandung di dalamnya. Dalam beberapa tahun kebelakangan ini, sistem anggaran kalori makanan telah menjadi popular disebabkan pengambilan kalori meningkat di serata dunia kerana terdapat penurunan dalam kualiti pemakanan dalam kalangan masyarakat. Walau bagaimanapun, terdapat satu penyataan masalah di mana penyelidikan sedia ada tidak mempunyai prosedur anggaran berat untuk sistem anggaran kalori mereka. Anggaran bilangan kalori untuk setiap item makanan adalah malar. Tujuan kertas kerja ini adalah untuk melaksanakan sistem kalori makanan yang menggunakan prosedur anggaran jisim. Hal ini demikian kerana ingin menguji dan menilai anggaran kalori makanan untuk mengesahkan keberkesanannya dan untuk membangunkan GUI sistem. Aplikasi 'Deep Learning' telah meningkat secara mendadak dalam pelbagai bidang, termasuk dalam anggaran kalori makanan. Dalam kertas kerja ini, kami mencadangkan untuk membina sistem yang boleh menganggar kandungan kalori makanan menggunakan algoritma 'Deep Learning'. Kerja masa depan projek ini adalah untuk memenuhi semua objektif yang disasarkan.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

1.1 Introduction

Understanding calorie intake is important as it influences an individual's health, and the amount of calories in food tells people how much potential energy they contain. In nutrition, calories refer to the energy people get from food and drink and their point in physical activity. Everyone requires different amounts of energy each day, depending on age, size, gender, and activity level. In recent years, the food calorie estimation system has become popular due to calorie intake rose in most of the world because there has been a decrease in the quality of diet (Mozaffarian, 2017). Diet has become one of the top risk factors for poor health, and weight gain has been typically framed as a problem of excess calorie intake.

In past research, it has been found that the authors did not deal with any mass estimation procedure. One way to estimate the volume of food is to use the object referencing approach, and there has been extensive research regarding volume estimation. For example, Poply and J. (2020) used tray/plate as a reference object, and the number of pixels associated with the reference object is counted to get the 'pixels per square inch' measure for that image. Once the mass is calculated, a calorie table lookup is performed to calculate the calorie contents of the detected food items. However, the system cannot handle food items with concave/convex structures since the only top-view of images has been used to grab the features.

Since artificial intelligence (AI) is introduced to the world, a subset of AI called Machine learning has revolutionized different fields in the past few decades. This ML is a subfield of computer science (CS). It is often referred to as predictive modelling and predictive analysis. Its aim and usage are to build new or leverage existing algorithms to learn from data, find patterns, and build generalized models that produce accurate predictions with new and unseen data. Besides that, it has appeared in the computer vision field and gain its popularity in many areas. Recently,

the application of DL has been increasing dramatically in various fields, including food calorie estimation. In this paper, we propose to build a system that can estimate the calorie content of food using the DL algorithm.

1.2 Problem Statements

Currently, there are many published reports on the food calorie estimation system. However, a problem statement has been highlighted where existing researches did not include mass estimation procedures for their calorie estimation system. The number of calories estimated for each food item was constant.

1.3 Hypothesis

Incorporating a convolutional neural network-based for food calorie estimation can increase the final performance evaluation of system.

1.4 Research Question

1. What is the efficient approach/method to use for volume estimation?

2. How effective is the proposed Convolutional Neural Network-based food calorie estimation algorithm?

3. What is the expected outcome of the proposed system?

1.5 Project Objectives

1. To implement a food calorie system that utilized a mass estimation procedure. (This objective mapped with Research Question 1)

2. To test and evaluate the food calorie estimation to validate its effectiveness. (This objective mapped with Research Question 2)

3. To develop a GUI of the proposed system. (This objective mapped with Research Question 3)

1.6 Project Scope

1. Image is taken with camera phone perpendicular to the surface, consists of reference object.

2. The calorie estimation system is web-based.

3. Focus only on the use of CNN as an approach.

CHAPTER 2

LITERATURE REVIEW

2.1 An Overview of Deep Learning

At a very high level, machine learning teaches a computer system how to make accurate predictions when fed data. Machine learning modes are typically trained to carry out useful tasks based on manually built features derived from raw data or features extracted from other basic machine learning models. Deep learning (DL) is a part of machine learning in artificial intelligence, or more specifically, deep learning is considered an evolution of machine learning. DL model uses hierarchical neural networks to analyse data continually with a similar logic structure.

The great success of DL has been widely applied used in many domains such as object detection, speech recognition, automatic machine translation, advertising, medical image analysis, etc. The most popular and successful type of DL model is convolutional neural networks (CNN) in image recognition. It is a more efficient and powerful way to learn good image representations and structured data. Subhi et al. (2019) confirmed that CNN exhibited considerably higher accuracy than other conventional methods. DL algorithms have quickly become a methodology choice for analysing food images now.

2.1.1 Deep Neural Networks

The original concept of deep learning is from studying artificial neural networks (ANNs), and ANNs have become an active research area during the past few decades. It is essential to utilize neurons to produce real-valued activations to construct standard neural networks (NNs), and NNs behave as expected by adjusting the weights. A traditional NN consists of many simple, connected processors called neurons, producing a sequence of real-valued activations. The research on the deep neural network has stirred a great deal of attention, and a series of exciting results are reported.

2.1.2 Convolutional Neural Networks

Convolutional Neural Networks (CNN) is one of the most famous learning algorithms in the computer vision field due to their good recognition capabilities. CNN can take in an input image, assign important learnable weights and biases to various aspects or objects in the picture, and differentiate one from the other. A convolutional neural network consists of a few different layers: the Convolutional layer, the Pooling layer, the activation layer, and the Fully Connected layer. The convolutional layer used info from the input data and output a feature map with the aid of kernels. The number of convolutional layers is varying from architecture to architecture. The first level of convolutional layers learns low-level features like bright and dark pixels.

In contrast, the second layer of convolutional layers may learn horizontal edges and vertical edges. The next level of the convolutional layer knows some more complex functions like ears and nose mouth. As layers increased, the neural network learns some even more complex tasks like gestures, objects, and characters. These feature maps are passed through a non-linear activation function that accelerates the CNN to understand complex processes. Lastly, one or more fully connected layers summarized this learnable information and gathered it into a softmax classifier. Softmax classifier gives the output probability of each class for the given input.

2.2 An Overview of Volume Estimation and Its Relation to Food Calorie Estimation

Food is any substance ingested to provide an organism with nutritional help. Food typically comes from plants, animals, or fungi and contains essential nutrients, such as carbohydrates, fats, proteins, vitamins, or minerals. A calorie is a unit of energy. Calories in diet refer to the calories people gain from food and drink and their power in physical activity. Food calories are essential to the human body as it provides energy for them to operate. Several studies have confirmed that reducing the number of calories in the diet by approximately 30% can lead to significant health benefits (Holowko et al., 2019). Modern society faces increased health problems as the population's eating habits and the lack of healthy food consciousness have caused obesity and poor nutrition among young adults (Chong et al., 2019). About 2.8 million people worldwide die because they are overweight or have melancholic obesity every year (WHO, 2021). Malaysians have been facing obesity and eating disorders since 2000 (Lee et al., 2019) and Southeast Asians are known for having a higher number

of obese people in the world as they have unhealthy eating habits and lifestyles (Mamun et al., 2020). Thus, daily tracking of people's calorie intake is crucial as it helps to achieve weight loss goals and promote a healthy lifestyle. Since people are dependent on intelligent technologies, applying an application to monitor the individual eating habit helps in many aspects. Ege et al. (2017) stated that currently, many mobile applications for recording everyday meals had been released because of a rise in healthy thinking on eating. (Attokaren et al., 2017) over the last few decades, research has been focused on automatically recognizing the food and their nutritional information from images captured using computer vision and machine learning techniques. Based on the literature review, there are many outcomes of volume estimation in calorie estimation applications which are web-based and mobile-based systems.

Ruenin et al. (2020) reported that they searched for food in the divided food tray, then the type of food in each section is predicted, and the weight of each kind of food is reported. Food pictures of before and after consumption are taken. The system displays the weight and calories of food before and after consumption to compare the amount of food the elderly consume. The Faster R-CNN technique is used for detecting and classifying food. CNN technique which used InceptionResNetV2, is utilized for the training weight estimation model. Cropped images from the Food Detection and Classification part are imported to the weight estimation. Then, calories of food are estimated from a reference table using the rule of three with the calculated weight of the food picture.

Poply & J (2020) used tray/plate as an object referencing approach to calculate the surface area of food items. Pixels per square inch are estimated to get the number of pixels associated with that object using Mask R-CNN. Then, a lookup table in which each food item is associated with its' calories per square inch' metric is built. A calorie can be estimated using the food surface areas through a lookup table and mathematical calculation.

Raikwar & Jain (2018) proposed an SVM algorithm to identify the food item. Next, the calorie value of food items is provided using a calorie map. There is no volume estimation procedure in this paper.

Deshmuk et al. (2021) used a coin as a reference object to determine the size of the food, and the quantity of pixels is compared to a given dimension. Three types of shapes are used in the experiment, which are ellipsoid, pillar, and irregular. The

following structures have their equations. Next, the mass of food and calorie can be determined using mathematical calculation. The measurement error of each food item was less than 10%.

Balbin et al. (2019) utilized software development for their proposed system. The volume of food will be calculated by multiplying the area obtained by Graph Cut Image Segmentation from the depth obtained by Ultrasonic sensing. Then, the mass will be received by multiplying the volume with the density. The calorie measurement of the food item will be determined using mathematical calculation. The mean accuracy of the proposed system is 88.18%.

Chiang et al. (2019) proposed a food calorie and nutrition system that can analyze a food composition based on a provided image. The system is based on Mask R-CNN. The weight of an object is estimated through the idea of the recognized food. The number of pixels in the image is captured through the food mask. The calories and nutrients of the food in the picture can be estimated using food nutrition and calorie tables. The authors used the Ministry of Health and Welfare Nutrition Database as a reference. First, the food class and food weight are recognized by food calorie and nutrition analysis, and then the calories and nutrients of the food in the image are estimated. The average absolute error and relative error were 8.22 and 0.13, respectively.

Pouladzadeh et al. (2016) proposed Graph Cut segmentation combined with a deep neural network for food classification and recognition. The authors utilized two types of reference objects as volume estimation of food which was finger-based and calculated. Calorie computation is performed by calculating the size of the food item concerning the finger in the frame.

Surya Gunawan et al. (2018) utilized the Generalized Regression Neural Network (GRNN) to predict the food intake calorie from digital image input. A trained nutritionist estimates the calorie of food. The proposed system has a large prediction error due to a very large variation of the calorie needed to be predicted.

Yogaswara et al. (2019) proposed a system that can calculate the food calorie content based on the volume of food using the Mask R-CNN method. For volume calculation, the plate is used as a calibrator for the area. The area of each food object detected, or segmented will be measured, then will be compared with the size of the detected plate. The volume of the thing will be multiplied with the object density constant to get the mass object. Then, calorie can be estimated by referring to the

calorie table from the Indonesian Ministry of Health Promotion brochure on "Healthy Lifestyle." performing a mathematical calculation for Calories of each class food object with respective mass can be obtained. The average accuracy of calorie calculation obtained is 97.48%.

Mittal et al. (2019) proposed a system that helps in calculating the number of nutrients using image processing and machine learning methods to recognize the food. A hole is used for calibration purposes to calculate the area of nutrition. Depth of fruit or vegetable is known from the height of the minor axis; then, multiplying the size and area will obtain the volume. Finally, the calories of food can be calculated using mathematical calculation. The average accuracy obtained is 88%.

Kohila & Meenakumari (2017) proposed a system that can predict a calorific value for mixed food using image processing. Morphological operations are used to identify/extract and calculate the shape of the image, volume, mass, and density. The standard nutritional value of fruits and vegetables is obtained from the health organization. Calorie value can be obtained by mathematical calculation by referring to the value of features extracted for calorie measurement and calories value for different samples.

Poply & Arul Jothi (2021) proposed CNN to carry out a segmentation a refined image segmentation. In this work, surface area estimation is carried out using reference objects which are plate/tray and coin sizes. The number of pixels is counted. The mathematical calculation is used to perform the volume and mass estimation. The preliminary information for height, density, calories are obtained from reliable online sources. Calories are estimated by multiplying the estimated mass with calories per 1 gram of the objects. The percentage accuracy obtained for calorie prediction is 93.06%.

Sadeq et al. (2018) proposed calorie estimation from food images using distance information. For solid food items, the distance between food and camera is estimated from the user's hand and device orientation posture. For liquid food items, an image is captured from a predefined distance from the fluid container. The distance information is used to find the centimeter to pixel ratio. K means clustering for segmentation of food portion, CNN for recognition food items are utilized in this paper. The mathematical calculation is used to perform both solid and liquid food items. The estimated volume of food items converted to mass and calories were

calculated using a nutrition database. The average standard error and relative standard error obtained are 7.46 calories and 4.46%, respectively.

Ando et al. (2019) proposed a mobile application for volume-based food calorie estimation using depth cameras. The CNN-based food segmentation method is used. 3D food objects are divided into many small pieces of an elongated rectangular parallelepiped; the volume of each piece is calculated and sum up. A reference object is used in this paper. Distance between the camera and the reference plane is estimated. For calorie estimation, the estimated food volume is applied to a regression equation of the calorie amount, the parameters of which are trained for each food category.

Table 2.1: Summary of Food Calorie Estimation Approaches

Author	**Volume Estimation approach/ method**	**Calorie Estimation approach/ method**	**Dataset/Data base**	**Result**
Ruenin et al. (2020)	Faster R-CNN (food detection and classification) CNN (weight estimation model)	Use reference table from a hospital	New dataset	Average value of MAPE: 16.9729
Poply & J, et al (2020)	Mask R-CNN (object detection) - Use reference object	Build a lookup table (source: FoodData central) consists of each food item associated with its 'calories per square inch' metric.	40 images randomly selected from UNIMIB 2016 Food Database.	Average accuracy of single food items: 95.45%

		Calorie is estimated using the lookup table and mathematical calculation		whole meals: 93.4%
Raikwar & Jain (2018)	SVM classifier - No volume estimation	Use calorie map	New dataset	Recognition accuracy: 90.66%
Deshmuk et al. (2021)	Faster R-CNN (for object detection) - Use reference object - Mathematical calculation	Refer to current value of food calorie - Use mathematical calculation	ECUSTFD	Measurement error: <10%
Balbin et al. (2019)	CNN (food type recognition) - Multiplying area from Graph Cut segmentation with the depth obtained	Refer to nutritional value from Department of Science and Technology – Food and Nutrition Research Institute - Perform mathem	New dataset	Mean accuracy: 88.18%

	from Ultrasonic sensing	atical calculation		
Chiang et al. (2019)	Mask R-CNN (detect and recognized food class and masks) - Number of pixels captured through the food masks - Mask prediction - Use mathematical calculation	Use food nutrition and calorie tables, Ministry of Health and Welfare Nutrition Database. - Mathematical calculation	New dataset combined with Food-256	Average Absolute error: 8.22 Relative error: 0.13
Pouladzadeh et al. (2016)	Deep CNN with Graph Cut segmentation (for food classification and recognition) Use reference object - Finger based and calculate	Mathematical calculation only	New dataset	Average standard error: 0.15