I. J. Education and Management Engineering, **2024, 2, 53-60**
Published Online on April 8, 2024 by MECS Press (http://www.mecs-press.org/)
DOI: 10.5815/ijeme.2024.02.05

Classification of Food Objects Using Deep Convolutional Neural Network Using Transfer Learning

Dipta Gomes*
University of Ulster, Belfast, Northern Ireland, UK
Email: diptagomes@gmail.com
ORCID ID: https://orcid.org/0000-0003-3019-6051
*Corresponding author

Received: 17 June, 2023; Revised: 04 August, 2023; Accepted: 17 September, 2023; Published: 08 April, 2024

Abstract: With the advancements of Deep Learning technologies, its application has broadened into the fields of food classification from image recognition using Convolutional Neural Network, since food ingredient classification is a very important aspect for eating habit recognition and also reducing food waste. This research is an addition to the previous research with a clear illustration for deep learning approaches and how to maximize the classification accuracy to get a more profound framework for food ingredient classification. A fine-tuned model based on the Xception Convolutional Neural Network model trained with transfer learning has been proposed with a promising accuracy of 95.20% which indicates a greater scope of accurately classifying food objects with Xception deep learning model. Higher rate of accuracy opens the door of further research of identifying various new types of food objects through a robust approach. The main contribution in the research is better fine-tuning features of food classification. The dataset used in this research is the Food-101 Dataset containing 101 classes of food object images in the dataset.

Index Terms: Food classification, Deep Learning, Convolutional Neural Network, Data Augmentation

1. Introduction

Classification of food had been a major challenge in the research field of Computer Vision and Object detection and automation of the ingredient detection had been the most awaited aspect of the field of Deep Learning. Applications such as Food Detection System, Calorie Measuring System and Diet Monitoring System had been important tools for everyday life for all levels of users around the world. Recent research on Image processing and detection of object from images had reached a breakthrough in different fields of research [1,2,3,4,5,6] like surveillance systems, medical imaging, remote sensing, emotion detection and object detection. Various research has shown machine learning techniques proved successful in detecting food objects from images successfully [1,7,8]. In most recent years, it was shown deep learning tools such as Convolutional Neural Networks (CNN) proved a very successful tools for high-efficiency detection model [9].

Object detection and classification from images is a challenging job due to the presence of low image quality, distortions, disturbances, intensity of light and small image dataset size. Comparisons among various Underwater Object detection from images has been represented by Dipta et al. [10], where various deep learning methods and conventional methods are compared, and a detailed evaluation of methods had been illustrated. Beside all these obstacles, food normally gets deformed and the shape and appearance of food changes with time as classification of food had always been a hard task for researchers and food cooked in different recipes give the images high variations and forms. In every food object detection system, it is required to be fast as well as accurate and needs to take consideration of the color, texture and quality of food in an image. A Real Time Sign Language Detection model had been put forward by P. P. Urmee et al. [11] where the Xception Model had been used on top of a proposed Bangla Sign Language dataset called BdSLInfinite. A convolutional neural network model (CNN) model had been proposed on top of Xception Model by P. P. Urmee et al. [11] to detect real time Bangla Sign Language. Here in the research work a dataset called BdSLInfinite dataset had been proposed for Bangla Sign Language for people with hearing difficulties. The model achieves an accuracy of 98.93% and a average response time of 48.53 minutes.

The dataset to be used for training is the Food101 dataset containing 101 food categories with over 100,000 images. Due to this high volume of images and classes of food, this dataset proved to be an excellent choice for this research.

