Abstract

In today’s day and age, having the ability for people to identify the nutritional values which are about to be digested, is essential in having a balanced diet. In this paper a food recognition and calorie estimation application are being proposed whilst using a newly made “Maltese Food” dataset. For this research, the algorithm Mask Regional-Convolutional Neural Networks (Mask R-CNN) was used for both object detection and instance segmentation. A solution of how to augment datasets for food image recognition is also provided together with a process of automating the manual annotation process in similar scenarios. The dataset is made up of 60 manually taken images, 1,380 total augmented images and consists of six food classes, these being “Pastizzi”, “Qassatat”, “Qaghaq tal-ghasel”, “Gbejniet”, “Imqaret” and “Zalzett Malti”. The best accuracy which was achieved on this dataset for predicting the instances of each food item in the image was an IoU score of 87.13%, but the average performance was that of 80.45% IoU accuracy. As for the calorie estimation section, an average Root Mean Square Error (RMSE) of +-86.39 was given. In addition, the predicted calories were also compared to the real values which were achieved by weighing each item. In this area, an average RMSE score of +-102.65 was achieved, but the best-case scenario was that of +-75.