Conclusion

In this paper, a newly created data set for six common Maltese foods has been presented whilst also proposing a solution of how other food datasets could be augmented and automatically annotated, to reduce the manual labour involved in creating such a dataset. The “Maltese Food Dataset” itself is made up of 60 original images and 1,380 augmented images which are split as “test” and “train” at a ratio of 70/30, 70% training and 30% testing. It also includes three random configurations, and, in each configuration, the images are randomly split into the previously mentioned ratio. The food classes included are “Pastizzi”, “Qassatat”, “Qaghaq tal-ghasel”, “Gbejniet”, “Imqaret” and “Zalzett Malti”.

Apart from that, an application making use of Mask R-CNN was proposed to estimate the calorie values for each food of the six food classes. The algorithm was used to detect both the food classes and identify the region of the item at which it is located, whilst simple proportion and a two-euro coin (the reference object) was also used to identify the proportions of each food item in the dataset. The process was also repeated on the three configurations provided by the dataset.

Using the proposed dataset three results per metric where achieved. In identifying the area of the food items, the IoU values are 87.13%, 73.66% and 80.57%. When taking the average of these configs, the value is that of 80.45%. The accuracy for estimating the calorie value using R.M.S.E are +-80.36, +-79.30 and +- 99.50 for the actual predictions, averaging out at a value of +-86.39. On the other hand, the R.M.S.E for the real predictions are +-94.20, +-138.24 and +-75, averaging a value of +-102.65. With the gathered results, it is possible to answer the research questions which were submitted at the begging of the scientific study. With an average IoU value of 80.45% it could be stated that Maltese foods have the potential of being identified consistently just by using machine learning techniques and that a two-euro coin can be used as the reference object to identify proportions from the image. The proposed a question which stated, if it would be possible to achieve nutritional values from an image. Taking into consideration that the average RMSE for both real and actual scenarios are +-86.39 and +-102.6, this difference might not be extremely small, but since this is only an estimate and the RMSE is still quite low since 100 calories difference are not a lot in a single plate of multiple food items. Therefore, it is possible to use such techniques to extract such values from images.

In future research, improvement could be made of the proposed solutions in various ways. It would be interesting to provide images which are taken from multiple different angles rather than just a single top view perspective for starters. Another improvement might be to add a way of removing the reference object by training models capable of identifying different plate size. One might also opt to enhance the current research by adding more Maltese food classes to the proposed dataset or even training the model with more variations of the given images.