Proceedings of the Federated Conference on
Computer Science and Information Systems pp. 11–14
DOI: 10.15439/2020F146
ISSN 2300-5963 ACSIS, Vol. 21

A Social Robot-based Platform towards Automated Diet Tracking

Anastasios Alexiadis*, Andreas Triantafyllidis*, Dimosthenis Elmas,
Giorgos Gerovasilis, Konstantinos Votis, Dimitrios Tzovaras
Centre for Research and Technology Hellas
Information Technologies Institute (CERTH/ITI)
6km Charilaou-Thermi, Thessaloniki, Greece
Email: talex, atriand, dimoselmas, ggerovasilis, kvotis, Dimitrios.Tzovaras@iti.gr

***Abstract*—Diet tracking via self-reports or manual taking of meal photos might be difficult, time-consuming, and discouraging, especially for children, which limits the potential of long-term dietary assessment. We present the design and development of a proof of concept of an automated and unobtrusive system for diet tracking integrating: a) a social robot programmed to automatically capture photos of food and motivate children, b) a deep learning model based on Google Inception V3, applied for the use case of image-based fruit recognition, c) a RESTful microservice architecture deployed to deliver the model outcomes to a platform aiming at childhood obesity prevention. We illustrate the feasibility and virtue of this approach, towards the development of the next-generation computer-assisted systems for automated diet tracking.**

I. Introduction

CHILDHOOD obesity is a major public health challenge which is associated with the risk of developing serious life-threatening diseases [1], [2]. In this context, new computer-assisted technologies can provide useful means to monitor and manage childhood obesity, as well as influence health behaviour and lifestyle at early age [3], [4], [5].

Accurate and long-term diet tracking, is of great significance in childhood obesity prevention [6]. In this direction, computerised dietary assessment through food diaries and self-reports is a common approach [7], [8]. However, major problems in developed computerised tools are that they place a significant burden to the user, suffer from recall bias issues, and rely on technological literacy, often resulting to their early abandonment [9], [10]. Therefore, more unobtrusive and automated approaches are intensively required [11], especially in children, which may have difficulties in articulating their eating patterns.

In this work, we present the design and development of a social robot-based platform for automated food recognition, with the capability to further motivate children to adopt healthy diet habits. The platform employs a deep learning approach for fruit detection, based on camera images automatically captured by a commercially available social robot. The outcomes of the detection are delivered to a platform aiming at childhood obesity prevention, developed within the OCARIoT[1] project, via a service-oriented architecture. Overall, this work adopts and uniquely integrates enabling computer-assisted information and communication technologies, such as social robotics, deep learning and interoperable data communication interfaces, towards demonstrating the feasibility, usefulness and virtue of automated dietary assessment for the prevention of childhood obesity.

[1]https://ocariot.eu/

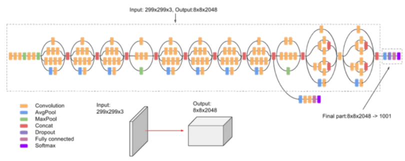

Fig. 1. Inception V3 (Source: https://cloud.google.com/tpu/docs/inception-v3-advanced)

II. Methods

A. Fruit Recognition Model

In a first step, our aim was to train and validate a fruit recognition model, based on the Google Inception V3 model [12], which is pre-trained on the ImageNet database. The Inception V3 model's architecture is shown in Figure 1.

The fruit recognition model classifies images into one of two classes—fruits and non-fruits. We gathered food images from the following sources: ImageNet, Food-101 Data-set[2], UEC Food 256 data-set[3] [13] and a data-set found in Zenodo[4]. The image-sets were split into two classes (fruits and non-fruits) and the pictures in the two classes were balanced. There was a total of 53884 fruit and not-fruit images. The images were cropped into 299x299 pixel chunks and horizontally flipped in a stochastic manner. We split the data-set 80%-20% stochastically.

Inception V3 is a Deep Convolutional Neural Network (ConvNet) designed for classifying images. Google states that the model has been shown to attain greater than 78.1% accuracy on the ImageNet data-set. We extended the model with the addition of the following layers:

- Average Pooling 2D with pool-size 8x8

[2]https://data.vision.ee.ethz.ch/cvl/datasets_extra/food-101/
[3]http://foodcam.mobi/dataset256.html
[4]DOI: doi.org/10.5281/zenodo.1310165

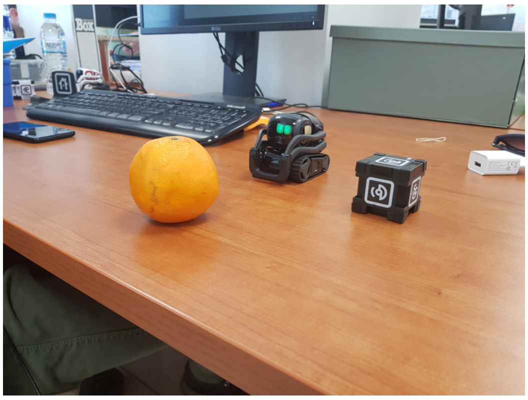

Fig. 2. Social Robot setup

- Dropout = 0.4
- Flatten
- Dense layer with 1 node, l2 regularization=0.0005, Sigmoid activation function, Xavier uniform initializer

The aim of these additions was to additionally train the model as a fruit classifier. We trained the model using Stochastic Gradient Descent, an optimisation algorithm with good performance over large data-sets. We utilized the following hyper-parameters:

- Initial Learning rate=0.01
- Momentum=0.9
- Decreased learning rate to 0.002 after epoch 15 and to 0.0004 at epoch 28
- Trained to binary cross-entropy loss 0.0018.

We used a multiple-crop (10-crop) strategy for classifying unknown images, where we produced 5 crops for each image to classify (upper left, upper right, lower left, lower right and centre), as well as the flipped versions of these crops. We classified each of these crops, for an image, and kept the one with the maximum value of the dependant variable (thus enabling us to identify a fruit in an image that also includes other unrelated objects).

B. Integration with a Social Robot

We further integrated our food recognition model with a commercially available social robot, aiming to apply the model outcomes in real-life and automate the process of food image recognition. The Anki Vector robot was adopted, which is equipped with a 720p (1280X720 pixels) High Definition camera and has additional interesting features, for example, it is easy to carry, it shows an engaging personality, e.g., showing feelings of happiness, sadness, anger, etc., through eye animations and movement with wheels, it can speak, display text/images on its display, and it is programmable via a Software Development Kit (SDK) which we have utilized.

We have developed an application that takes a picture using the Vector's built-in camera and passes it to the model for fruit classification. Upon detection of a fruit, the robot responds with speech, eye animations and movement, providing reward and motivation (e.g., "well done, fruits are good for your health", "congratulations for your choice", etc.). We have used a similar approach in our previous work with social robots targeting childhood obesity [14].

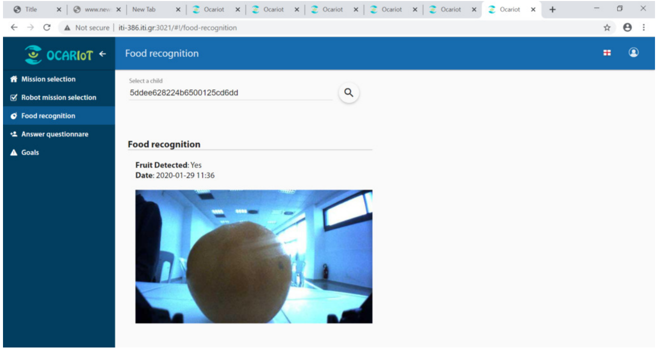

Fig. 3. Platform UI

To make the robot automatically move towards the food and capture an image, we implemented a method in which the user is required to place the robot's cube, as a reference object, in front of the food (Figure 2). The software we developed makes the robot to search for the reference cube and positions the robot towards facing the food, resulting overall to an automated food image recognition process with the help of a social robot, without requiring any significant manual effort by the child.

C. Software Architecture

Regarding the software architecture, two applications have been developed: An Angular application for the user interface and a Django server that provides an API for the application of the image recognition model and for controlling the robot. Both Angular and Django come equipped with a command line tool that can be used to quickly setup an application. This facilitated the rapid implementation of our prototype.

Regarding the integration of the image recognition model and the social robot, we adopted a REST microservice architecture, a new architectural style that structures an application into a set of small, independently deployable microservices, as opposed to traditional monolithic approaches. The microservice (tagged 'Food Tracking') can store the pictures of the fruit meals in a database in order to create a data-set that could be used to update the image recognition model. When the robot takes a picture of the fruit meal, the image recognition model is applied in order to identify the presence of a healthy food (certain varieties of fruit in this case). When the output of the model is available, the image is sent to the backend software and then the classification result and the image are correlated to the user's id. After the backend has received the image and the result, the platform's "dashboard" application is updated and the user can browse an updated version of their profile.

III. Results

We measured 99.68% accuracy on the validation set comprising a total of 10655 fruit and not-fruit images from the

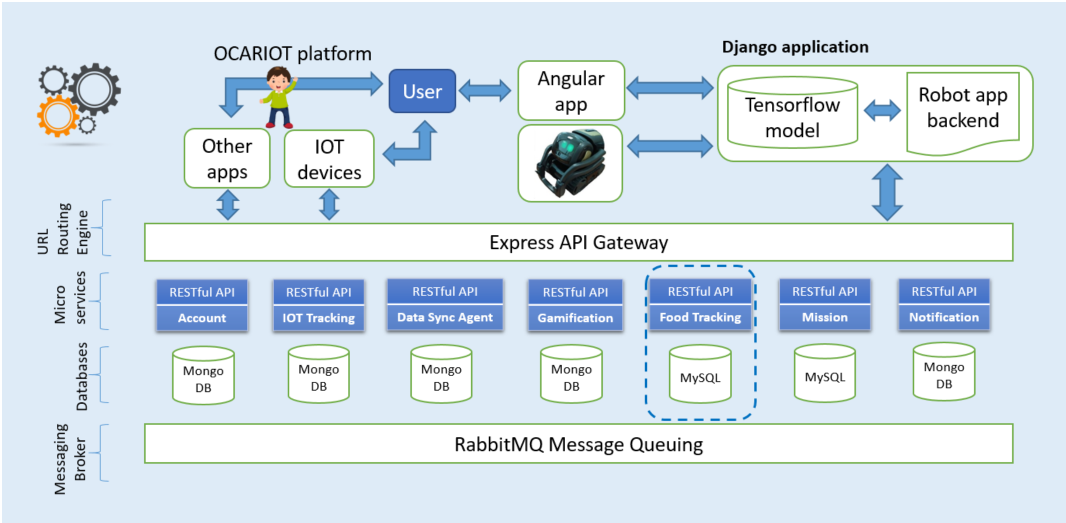

Fig. 4. Fruit Recognizing Social Robot integrated to OCARIoT

Fig. 5. Recognized as fruit

Fig. 6. Not recognized as fruit

combined data-set mentioned in section II-A, which shows that fruit recognition through images is an accurate method, and it could potentially replace tedious self-reports or surveys for fruit consumption.

The Angular and Django applications for controlling the functionality of the social robot and utilising the fruit-recognition model, were integrated within the OCARIoT software platform for childhood obesity prevention. A demonstration of our integrated system is shown in a YouTube video[5]. Figure 4 shows the architecture of the OCARIoT Platform with the social robot application that recognizes fruits integrated. In Figures 5 and 6 we show correct and incorrect classifications on fruits from the integrated fruit-recognition social robot system. We have observed that the distance to the target fruit is the most important indicator for classification accuracy. The robot managed to identify fruits from a distance $\sim$ 20 cm. The apple in Figure 6 was further from the threshold distance.

The Express Gateway, an open source API Gateway which is based on the Express.js framework, is used to redirect the client's requests to the respective microservices through URL routing. The RabbitMQ message broker is adopted in order to manage the message queues maintained between the microservices, thereby facilitating their communication (for example the food tracking microservice can use account information derived from the account microservice through RabbitMQ). RabbitMQ is an open source broker which allows transport-level security, through the use of the Advanced Message Queuing Protocol (AMQP), and fast communication over the Transmission Control Protocol (TCP).

In order for the robot apps to be executed successfully, the robot must be connected (via WiFi) to the same network with the computer executing the robot SDK. Both the robot's application and the Tensorflow model can be accessed by an API built using the Django web framework.

IV. Discussion

We presented the design and development of platform for automated diet tracking based on a programmable social robot, the application of a deep learning model for fruit recognition, and the integration of the model outcomes with a computerised system targeting childhood obesity prevention, through a REST microservice architecture. Our platform constitutes a proof-of-concept, demonstrating the integration of different enabling technologies, towards the development of the next-generation computer-assisted systems for automated diet assessment, which are also capable to motivate individuals to be more engaged with the acquisition of healthy diet habits. Through taking advantage of the programmable robot's in-built capabilities such as a camera, text-to-speech synthesis and

[5]https://www.youtube.com/watch?v=ZSH-WW-rBjY

eye animations, the predictive capabilities of deep learning, as well as an architecture which allows extensibility and interoperability with other software components, our aim was to develop a novel unobtrusive system requiring minimal user interactions.

The system developed could be particularly useful for children, which may face difficulties in self-reporting diet information due to issues related to recall or tedious repetitive user-to-system interactions. Furthermore, child interaction with other computerised systems such as mobile phone devices, would be likely to require parental consent, a good knowledge of using mobile apps, as well as manual taking of photos which may be of low quality. In this context, our system differentiates from previously systems which have been examined [15], [16], [17], and shows the high potential of the application of significantly more engaging and automated systems. In particular, social robot-assisted systems have been demonstrated to be highly attractive to children and useful [18], [19], which has been a motivation for following this approach.

Our future work involves the recognition of different categories of food through social robot-captured images, which would enable a more holistic approach in accurate dietary assessment. Furthermore, the addition of speech recognition in the system would enable dialogue interactions between the robot and the child, which could facilitate motivation of children to acquire healthy diet habits or improve system certainty (e.g., the social robot could ask the child about a meal in the case of robot's low certainty in detecting a specific food in an image, and receive a verbal response). Moreover, the longitudinal collection of all captured diet information would reinforce personalised data analytics, which could indicate behaviours requiring guidance and attention, or revealing potential risks. The deployment of computational models and decision support systems has shown promise in this direction [20]. Finally, a study with children should be conducted, enabling the evaluation of the platform in real-world scenarios, e.g. at home, or within educational sessions at school settings. Furthermore this will allow us to measure the real-world accuracy of the model on a set of images captured by the social robot in practical usage scenarios.

In conclusion, we regard social robots as valuable agents that can support humans in engagement with healthy behaviours. To this end, the work presented in this paper is a step towards automated dietary support of children by social robots.

Acknowledgements

The study was supported by the European Union‘s HORIZON 2020 Programme (2014-2020), under ID no 777082, and from the Brazilian Ministry of Science, Technology and Innovation through Rede Nacional de Ensino e Pesquisa (RNP) under OCARIoT

References

[1] M. Shields, M. S. Tremblay, S. Connor Gorber, and I. Janssen, "Abdominal obesity and cardiovascular disease risk factors within body mass index categories.," Heal. reports, vol. 23, no. 2, pp. 7–15, Jun. 2012.

[2] I. Vucenik and J. P. Stains, "Obesity and cancer risk: evidence, mechanisms, and recommendations," Ann. N. Y. Acad. Sci., vol. 1271, no. 1, pp. 37–43, Oct. 2012, doi: 10.1111/j.1749-6632.2012.06750.x.

[3] E. B. Tate et al., "mHealth approaches to child obesity prevention: successes, unique challenges, and next directions.," Transl. Behav. Med., vol. 3, no. 4, pp. 406–415, Dec. 2013, doi: 10.1007/s13142-013-0222-3.

[4] A. J. Smith, A. Skow, J. Bodurtha, and S. Kinra, "Health Information Technology in Screening and Treatment of Child Obesity: A Systematic Review," Pediatrics, vol. 131, no. 3, pp. e894–e902, Mar. 2013, doi: 10.1542/peds.2012-2011.

[5] P. W. C. Lau, E. Y. Lau, D. P. Wong, and L. Ransdell, "A Systematic review of information and communication technology-based interventions for promoting physical activity behavior change in children and adolescents," J. Med. Internet Res., vol. 13, no. 3, 2011, doi: 10.2196/jmir.1533.

[6] E. P. Abril, "Tracking Myself: Assessing the Contribution of Mobile Technologies for Self-Trackers of Weight, Diet, or Exercise," J. Health Commun., vol. 21, no. 6, pp. 638–646, Jun. 2016, doi: 10.1080/10810730.2016.1153756.

[7] A. G. Arens-Volland, L. Spassova, and T. Bohn, "Promising approaches of computer-supported dietary assessment and management-Current research status and available applications.," Int. J. Med. Inform., vol. 84, no. 12, pp. 997–1008, Dec. 2015, doi: 10.1016/j.ijmedinf.2015.08.006.

[8] A. H. Andrew, G. Borriello, and J. Fogarty, "Simplifying mobile phone food diaries," in Proceedings of the 2013 7th International Conference on Pervasive Computing Technologies for Healthcare and Workshops, PervasiveHealth 2013, 2013, pp. 260–263, doi: 10.4108/icst.pervasivehealth.2013.252101.

[9] S. M. Schembre et al., "Mobile Ecological Momentary Diet Assessment Methods for Behavioral Research: Systematic Review.," JMIR mHealth uHealth, vol. 6, no. 11, p. e11170, Nov. 2018, doi: 10.2196/11170.

[10] D. Lupton, "'I Just Want It to Be Done, Done, Done!' Food Tracking Apps, Affects, and Agential Capacities," Multimodal Technol. Interact., vol. 2, no. 2, p. 29, May 2018, doi: 10.3390/mti2020029.

[11] T. Prioleau, E. Moore Ii, and M. Ghovanloo, "Unobtrusive and Wearable Systems for Automatic Dietary Monitoring.," IEEE Trans. Biomed. Eng., vol. 64, no. 9, pp. 2075–2089, Sep. 2017, doi: 10.1109/TBME.2016.2631246.

[12] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* Las Vegas, NV, 2016, pp. 2818-2826, doi: 10.1109/CVPR.2016.308.

[13] Kawano Y., Yanai K. (2015) Automatic Expansion of a Food Image Dataset Leveraging Existing Categories with Domain Adaptation. In: Agapito L., Bronstein M., Rother C. (eds) Computer Vision - ECCV 2014 Workshops. ECCV 2014. Lecture Notes in Computer Science, vol 8927. Springer, Cham

[14] A. Triantafyllidis, A. Alexiadis, D. Elmas, K. Votis, D. Tzovaras, A social robot-based platform for prevention of childhood obesity, in: Proc. - 2019 IEEE 19th Int. Conf. Bioinforma. Bioeng. BIBE 2019, Institute of Electrical and Electronics Engineers Inc., 2019: pp. 914–917. doi:10.1109/BIBE.2019.00171.

[15] A. Myers et al., "Im2Calories: towards an automated mobile vision food diary," 2015, 10.1109/ICCV.2015.146.

[16] S. Mezgec and B. Koroušić Seljak, "NutriNet: A Deep Learning Food and Drink Image Recognition System for Dietary Assessment," Nutrients, vol. 9, no. 7, p. 657, Jun. 2017, doi: 10.3390/nu9070657.

[17] Y. Hswen, V. Murti, A. Vormawor, R. Bhattacharjee, and J. Naslund, "Virtual avatars, gaming, and social media: Designing a mobile health app to help children choose healthier food options," J. Mob. Technol. Med., vol. 2, no. 2, p. 8, 2013, doi: 10.7309/jmtm.2.2.3.

[18] O. Mubin, C. J. Stevens, S. Shahid, A. Al Mahmud, and J.-J. Dong, "A Review of the Applicability of Robots in Education," Technol. Educ. Learn., vol. 1, no. 1, pp. 1–7, 2013, doi: 10.2316/Journal.209.2013.1.209-0015.

[19] O. A. Blanson Henkemans et al., "Design and evaluation of a personal robot playing a self-management education game with children with diabetes type 1." 01-Jan-2017, doi: 10.1016/j.ijhcs.2017.06.001.

[20] A. Triantafyllidis et al., "Computerized decision support and machine learning applications for the prevention and treatment of childhood obesity: A systematic review of the literature," Artif. Intell. Med., vol. 104, p. 101844, Apr. 2020, doi: 10.1016/j.artmed.2020.101844.