Studies in Computational Intelligence 890

Diego Oliva
Salvador Hinojosa *Editors*

Applications of Hybrid Metaheuristic Algorithms for Image Processing

Springer

Diego Oliva · Salvador Hinojosa
Editors

Applications of Hybrid Metaheuristic Algorithms for Image Processing

Springer

Contents

Hybrid Metaheuristics and Other Image Processing Tasks

Health Applications

Intelligent System for the Visual Support of Caloric Intake of Food in Inhabitants of a Smart City Using a Deep Learning Model

José Mejía, Alberto Ochoa-Zezzatti, Roberto Contreras-Masse and Gilberto Rivera

Abstract In our days, a recurrent problem in the Smart Cities of all Latin America will be the degenerative illnesses linked to food, one of the main reasons for the lack of organization of each citizen, is not being able to adequately determine the caloric intake, a proposal for a solution , is the development of an application that has the capacity to be able by means of the recognition of patterns and a deep learning model, to be able to specify what percentage of the nutritional value of each meal is covered by the associated quantities, one of the advantages of using a food repository is that solutions based on deep learning are not ready-made solutions. A development process is necessary to acquire an adequate set of instances and to customize the intelligent system. The latter includes the customization of the user interface, as well as the way in which the system retrieves and processes the feeding scenarios later. The resulting scenarios can be shown to the user in different ways, and/or retrieved cases can be adapted to be reused later. This research is about an intelligent model for decision making based on deep learning to solve the existing problem in the planning of food distribution in the population of a Smart City, for this first, we mentioned the need for intelligent systems in the processes of decision-making, where they are necessary due to the limitations associated with conventional human decision-making processes, among them: human experience is very scarce with respect to being able to calculate in a correct way the caloric value of food intake and we must to consider that citizens in a smart city are tired of the burden of physical or mental work, in addition to human beings forget the crucial details of a problem, and many times are inconsistent in their daily decisions. Complexity and investment of the time necessary to make food decisions tend to be complex for health as well as the high frequency of decision making found in the distribution to supermarkets, which mostly supply the food of a population with tendency to increase of individuals of the groutier type when they eat their food late at night. We use an image repository from DataWorld to our research (https://data.world/).

Keywords Making-decision · Deep learning · Caloric intake of food

J. Mejía · A. Ochoa-Zezzatti (✉) · R. Contreras-Masse · G. Rivera
Universidad Autónoma de Ciudad Juárez, Juárez, Mexico
e-mail: alberto.ochoa@uacj.mx

D. Oliva and S. Hinojosa (eds.), *Applications of Hybrid Metaheuristic Algorithms for Image Processing*, Studies in Computational Intelligence 890,
https://doi.org/10.1007/978-3-030-40977-7_19

1 Introduction

Nowadays, Childhood morbid obesity in Mexico is a health problem characterized by excess body fat in children. This level of adiposity is a determining factor for their general health and well-being, making them prone to suffer other secondary pathologies and associated with social aspects such as discrimination in schools due to their physical condition. According to the data published by the World Health Organization (WHO), the prevalence of this disease internationally is increasing. In 2027 there were about 87 million overweight children, most of them in developing countries, especially in Latin America, leading Mexico.

As in adults, obesity requires several complementary tests for its diagnosis. One of the most commonly used scales is known as body mass index or BMI.2 also known as the Quetelet index, it takes the weight and height of the person as a reference, assessing its proportionality based on the mathematical formula: mass divided by the height squared. However, the weighting of the resulting figure varies in the case of children. Unlike in adults, the body mass index has to be moved to a percentile table, taking into account two essential criteria: the age and sex of the child in question. According to the same, it is established that children with a body mass index between 85 and 95 on that scale are overweight, in the present investigation is to propose a means to deal with the health consequences, through the implementation of an intelligent application that allows to indicate in visual form, the caloric intake of each child in their day to day.

In recent times, Deep Learning has evolved computer vision algorithms [1] and specially the segmentation algorithms used to determine the of various objects of interest within an image [2–4]. In particular, the segmentation of food has been revised in several works of literature to develop systems to help users measure their calorie in-take. In [5] it is considered a system based on deep learning for recognize the food and determine portion sizes, also the system is implemented on a mobile device. In [6] it is proposed an assistant calorie measurement system using deep convolutional neural networks that can run on smart-phones, allowing the user to take a picture of the food and measure the amount of calorie intake automatically. The method shows high accuracy on single food portions. In [7] it is proposed a CNN-based food image recognition system addressing the problem of food recognition in images. In [8] was developed a deep learning-based visual food recognition and adapted it to the design of a food recognition system employing computing-based service to overcome some inherent problems of traditional mobile cloud computing paradigm.

2 Proposal Methodology

According to the tests, it is feasible to use a mobile application, to be able to determine the caloric consumption of each child, using real-time photographs of each meal made during the day and considering a nutritional pyramid, the most convenient time of the

day is when it begins, since together to establish specific parameters of food intake should be considered caloric intake and its representation in the population pyramid. This research tries to specify an alternative route to food supplements, among other advantages. The thought of not knowing how many calories are associated with our food is of great importance, it is because of them that Artificial Intelligence has a series of tools at hand to achieve this problem, which in this case is technology, information, processing images and some mathematics. It should never be considered that the qualitative aspects are not important and having a healthy figure allows children to interact in a better way, that is why each of the components of the intelligent application, all of them linked to the objective of this project, consists of a mobile information system for geo-referenced location of food intake in children with morbid infantile obesity (SIGOMI) on levels of insecurity in the correct calculation of caloric intake at each moment of the day, which is presented in detail below , this research includes the analysis by means of deep learning to determine the danger of a continuous excess of caloric intakes higher than those established by the WHO for each population group in a border society and to avoid a contingency occurring during a specific time associated with bad eating habits. We carried out an exhaustive analysis of other similar investigations, the only similar context is explained in [9], where the insecurity of a group of children who perform higher than normal intakes and in fast-paced sites is calculated, by means of the delivery of products food in different places with random programming, but this research does not consider real statistics in time and the perspective to suggest a scenario of caloric intake associated with age, height and current weight and with a diet appropriate for their population group, a proposal model of this research is shown in Fig. 1.

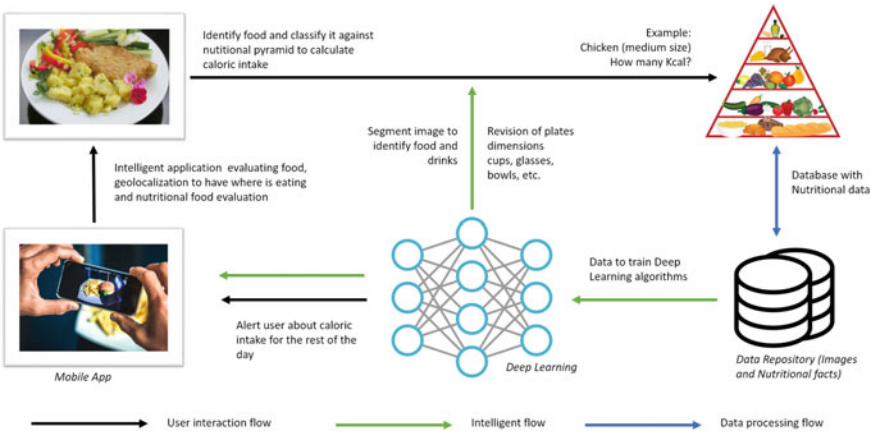

Fig. 1 Proposal model associated with our research

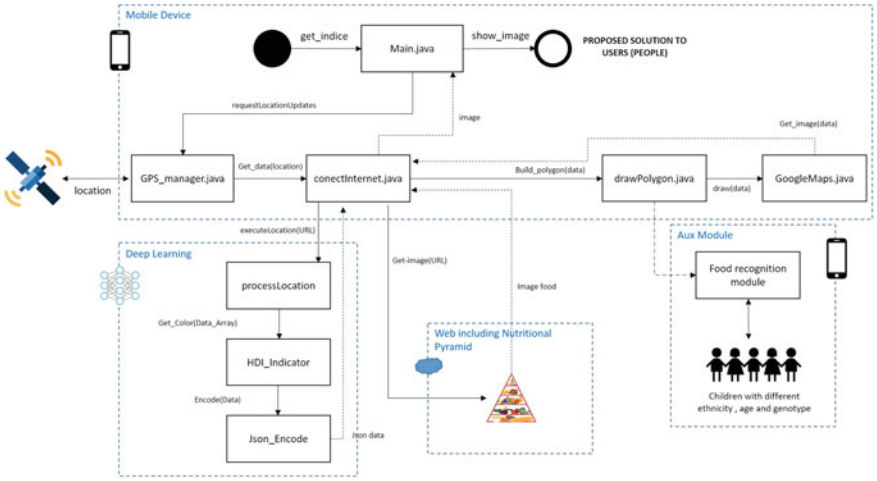

Fig. 2 Modular description of the components of our mobile computing application developed. *Source* Own design

2.1 *Design of the Mobile Application*

A decisive aspect of our research is the geo-reference of the location where food is consumed, this to determine the amount of associated caloric consumption, using our mobile application associated with Ubiquitous Computation it is possible to consider a decisive aspect that is the calories that are associated with the menus of the food stores in the locations, and determine if there is a reference proposed by our multi-agent system to a particular dish, that can help the correct selection of the caloric intake for the sample of children with morbid obesity. Using a geo-referencing model it is possible to determine the precise location of the diverse locations where the caloric management will be carried out, and therefore establish a criterion associated with the caloric intake of each food photographed and check if it meets the standards of previous diners using our multi-agent system, each image is stored in a food repository and linked to the nutritional pyramid acquired for each user, therefore there is little chance that an error will be generated, because there are different revisions to the caloric consumption of each user, in our case of each child with morbid obesity, as it is considered in each module of Fig. 2.

2.2 *Considerations of Our Problems and Their Impact on Society*

According to recent studies by the Pan American Health Organization, 17% of the infancy population in Ciudad Juarez—a frontier society—has characteristics of child morbidity, so that, in the wake of nutritional disorder, there is a problem of social

Fig. 3 A morbid obesity in children is characterized by anxiety problems, especially in single children. *Source* Cambridge News

psychology, characterized by the anxiety of having to eat, whenever there is a situation of stress, as can be seen in Fig. 3, this problem of children psychology generates a parallel problem that is the social stigma of not being left in the children stage, something that even leads to suicide in children, boys and girls.

In our days, Mexico now the tenth largest population with 128.7 million has the second place in the ranking of overweight population this affects the health system, as can be seen in Fig. 4, as the population grows in age, the Food has a deteriorating effect on their health and generates that a greater percentage of children and young people are affected by this paradigmatic change of diet, the high caloric intake of foods with saturated fats and types of salt harmful to health begin to affect the population as it is characterized a pattern of food that includes soft drinks with a refill system and ice cream and sweets with high concentrations of sugar.

The Organization for Economic Co-operation and Development (OECD), composed of 47 societies of high and medium economy, has counts of statistics of overweight in the target population of this study, as can be seen in Fig. 5, and is decisive for the consideration of this study that the young population has the highest morbid obesity growth in children, having food problems associated with caloric intake and food from commercial chains, which in many cases do not take care of caloric intake, nor present in detail such consumption.

If a determinant public policy associated with the decrease in the population of children with morbid obesity is not considered, it will not be possible to detail which nutritional aspects to consider could be decisive for the objective population, that is why determining a date established by the same OECD for the 2027 the critical route to follow in order to achieve the reduction of the caloric intake in overweight

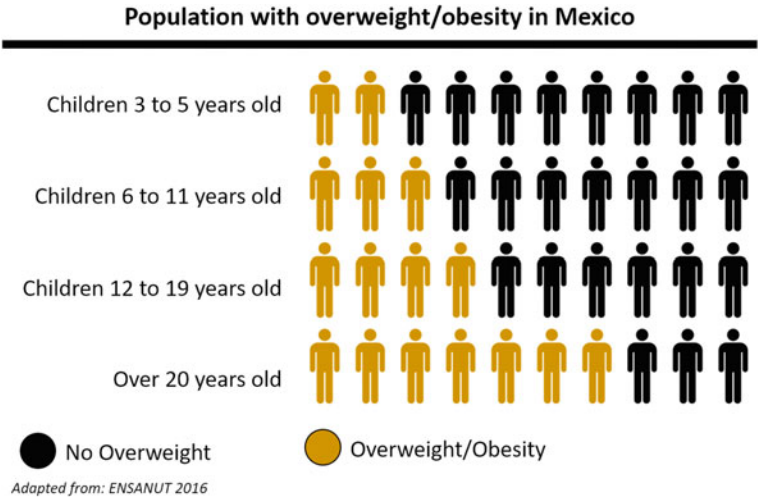

Fig. 4 Proportion of children with morbid childhood obesity, in Mexico. As the population gets old, the proportion of overweight increases

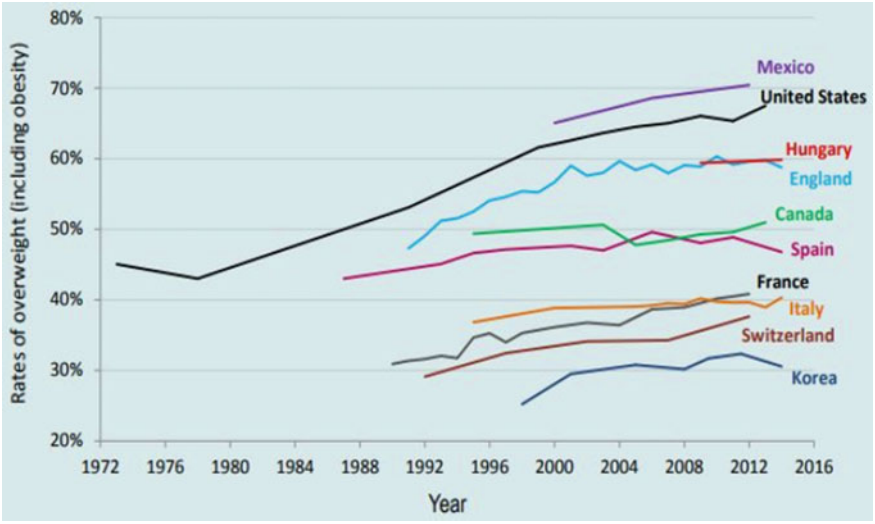

Fig. 5 Comparison of OECD countries with percentages of childhood obesity for at least one year prior to 13 years of age. *Source* OECD Statistics

children should be expensive, but not before establishing a particular nutritional management system associated with the size and weight indicated for each child and a child. specific exercise system to be carried out by the children of each of the 32 societies in Mexico, which considers socio-cultural and socio-economic aspects of each region.

2.3 Food Recognition Module

In this section more details of the food recognition stage are given, since it is a critical component of the proposed system (Fig. 2). The model has an interchange with recognition of food in the input image and the rest of data of our system. The proposed food recognition stage consists of a deep learning architecture. An input image is feed to the system and is subsequently segmented in several groups which are then classified as type of food or background. The architecture is based on Mask Regional Convolutional Neural Network (R-CNN) [10] which extends R-CNN [11] by adding an object mask to the existing branch for bounding box recognition.

Mask R-CNN is a deep learning architecture to solve instance segmentation, it uses the module for classification of fast-RCNN but with an extra module to create masks. Mask R-CNN adopts an identical first stage of Region Proposal Network, however, it adds a binary mask for each stage to produce binary mask from the input image. In this architecture, each module of the network has its own assigned loss, allowing the network to generate masks for every class without competition among classes. Figure 6, shows the network architecture. In this paper it is proposed a new architecture for the mask module in mask R-CNN, since the proposed network is focused on food recognition only, the network is simplifies to make it more easy to train and to have less computational burden as compared with for example ResNeXt [12] type networks of depth 50 or 101 layers which is a common choice used in for the

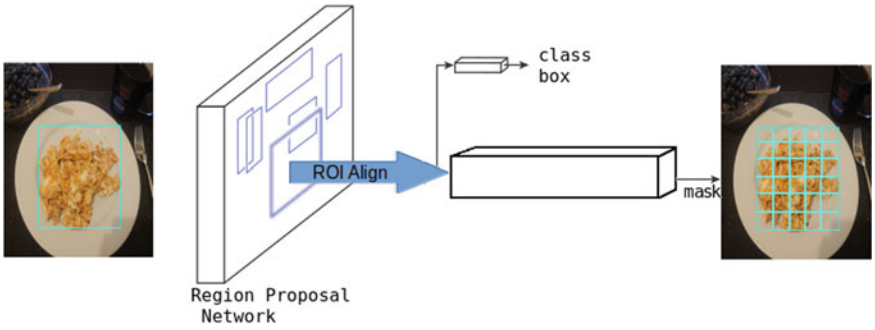

Fig. 6 Simplified diagram of the mask R-CNN network used for classification/segmentation of images. The network generates three types of output, the box where the food is, a binary mask delimiting the food, and the types (classes) of food found

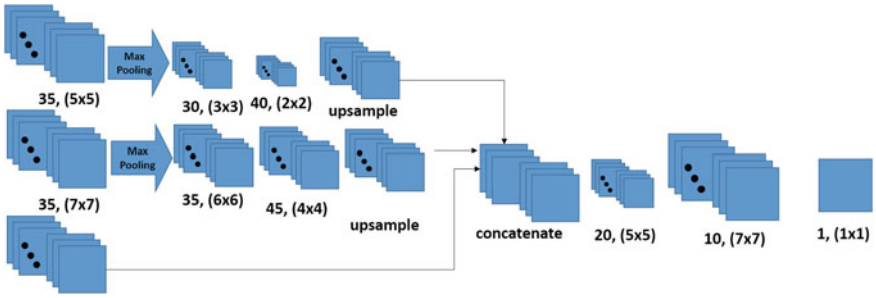

Fig. 7 The proposed architecture for the mask module. It consists of three channel of convolutional layers which are concatenated to have a rich feature map with information at different scales

module implementation in mask R-CNN implementations [10]. Thus the proposed architecture consists of three channels of convolutional networks, each channel is aimed to select different sizes of features: a large size channel with kernels of size 11×11, a medium size channel with output filter of size 4×4, and finally a small size channel with output filters of 2×2. The outputs of the different channels are concatenated in a feature map which contains information at different scales of the input image, note that for the concatenation the output of the medium and small channels are up sampled so that the output of the three channels have equal size. Subsequently the output is feed to a pair of deconvolution layers to recover the size of the original image and finally a convolution layer of 1×1 filter size to have a gray level image with the masks of food regions. Each mask of the output image has a gray level that indicates the class of the region that the mask encloses. The details of the described architecture are shown in Fig. 7.

In Fig. 8 shows several images used to train the network. To increment the transmission velocity to the server and reduce the computational cost, the input images were reduced to a fixed size of 128×128 pixels. This initial dataset was processed with data augmentation techniques [13, 14] to have a final data set of 4000 images, for this purpose each image in the initial data set was processed with translations, rotations, and scale changes. Also each image was segmented manually in background and the several types of food. The network was trained with a mean square error loss function, the implementation was programmed using the Keras framework [15] on a computer with 2.5 GHz Intel Core i7-4710HQ and a Nvidia GeForce GTX 980M (4 GB GDDR5 RAM) Graphics card.

Besides this, it is convenient to indicate that the image recognition module is still being improved for its correct implementation in the cell phone but due to the compact form of the architecture of our research it is considered that it is very viable that it can be implemented for any type of mobile system.

Fig. 8 Sample images from the input data. As can be seen, the dishes can consist of the same type of food, such as meat, chicken; or a combination like salads

3 Design of Experiments

To know the operation of the proposed model, we used the estimation of the sum of the food separately and finally the caloric combinations of a food that involves combinations of them. These proposals of food menus competed with each other in the category of proposal for the improvement of nutrition and therefore of the reduction of child morbid obesity and were evaluated by the intelligent application developed through their representatives in the deep learning model. For the estimation of the matrix of the selection of a menu according to the caloric intake, 30 runs of the experiment were conducted which were carried out under the same conditions, and for the group of finalist proposals, a design of experiments according to the attributes of each proposal to obtain a better estimate of the final classification. In each run, 350,000 function evaluations were performed. The average over the 30 runs was calculated for each proposal. Then, the average ranking was obtained to determine the 24 best proposals that will compete in the final ranking. Three measures are calculated from the 30 runs: mean, median and interquartile range. The interquartile range has an amplitude of 50% of the whole value of the median (second quartile Q2), which is calculated through the Q1 lower quartile (first quartile) and upper quartile (Q3 quartile third quartile). In descriptive statistics, a quartile is any of the three values that divide the ordered data into four equal parts, so that each part represents 1/4th of the population of the sample. The difference between the upper and lower quartiles is called the interquartile range. In the results section, the estimation of our approach to the analysis of food menu proposals for children with a new eating disorder is presented.

3.1 Experimentation

In order to be able similar, the most efficient arrangement of individuals associated with the food in a social network, we developed an atmosphere able to store the data of each one of the representing individuals of each society—including their food—, this with the purpose of distributing of an optimal form to each one of the evaluated

Table 1 Orthogonal array

A	B	C	D	E	F	G	H	Color
H	H	H	H	H	H	H	L	1
H	H	H	H	H	H	L	H	2
H	H	H	H	H	L	H	H	3
H	H	H	H	L	H	H	H	3
...	...	...	...	...	...	...	...	...

societies. One of the most interesting characteristics observed in this experiment was the diversity of the cultural patterns established by each community related with the food. The scenes structured associated with the agents cannot be reproduced in general, since they only represent a little while moment in the space and time of the different societies. These represent a unique form and innovating of adaptive behavior which solves a computational problem that it does not try to clustering the societies only with a factor associated with his external appearance (attributes of each society), trying to solve a computational problem that involves a complex change between the existing relations. The generated configurations can be metaphorically related to the knowledge of the behavior of the community with respect to an optimization problem (to select culturally 47 similar societies and their food, without being of the same quadrant associated with Nutritional pyramid). The main experiment consisted of detailing each one of the 568 diverse kind of food or their combinations, with 500 agents, and one condition of unemployment of 50 epochs, this allowed us to generate the best selection of each Quadrant and their possible location in a Diorama associated with food, which was obtained after comparing the different cultural and social similarities from each community, and to evaluate with Multiple Matching Model each one of them. The developed tool classified each one of the societies pertaining to each quadrant, with different wardrobe for societies that included linguistic identity and for societies only with cultural identity, this permit identifies changes in the time respect at other societies. The design of the experiment consists in an orthogonal array test, with the interactions between the variables: emotional control, ability to fight, intelligence, agility, force, resistance, social leadership, and speed. These variables are studied in a range of emotions associated with the food, another variable is the color of food represented by numbers (0–256). The orthogonal array is $L - N(2^8)$, in other words, 8 factors in N executions, N is defined by the combination of possible values of the 8 variables an the possible range of color (To see Table 1).

For measure the performance of the network we use accuracy for the recognition module in R-CNN and for the segmentation mask stage we used the Intersection over Union (IoU) which is a common metric [16, 17] used to evaluated segmentation accuracy. This metric computes the total of pixels in the intersection between the set of predicted pixels (A_{pred}) and ground truth pixels (A_{GT}) for each class, and divides by the number of pixels in their union, as is shown in Eq. 1.

$$\text{IoU} = \frac{A_{pred} \cap A_{GT}}{A_{pred} \cup A_{GT}} \tag{1}$$

4 Results

The so-called Somatotype is a crucial element in our research, because the caloric intake continues and the existing genetics determines bodies of a certain form, which require different treatments in the use of reduction of felling, and therefore the existing determination of the real possibility to continue in a process of caloric reduction, which will be effective in the near future, as it can be the youth, the seven different types of somatotypes can be visualized in Fig. 9.

One of the most relevant aspects of our research, is that if you adjust the caloric intake prior to adolescence, each of the individuals will not be the target of bullying in their youth, considering that they could reduce at least one or two sizes to the utilizer our proposal of application of mobile computing.

4.1 Results of the Segmentation Network

In Fig. 11 is shown the output from the segmentation on several images and Fig. 10 shows the output of both the segmentation combined with the classification output. The averages of IoU of the segmented regions for each type of food are shown in Table 2. Most classes of food have an average above of 0.9 of accuracy, except for liquids, water, coffee; this could be because of to the recipient used to contain the liquid, which could have a large variation in size, color and shape.

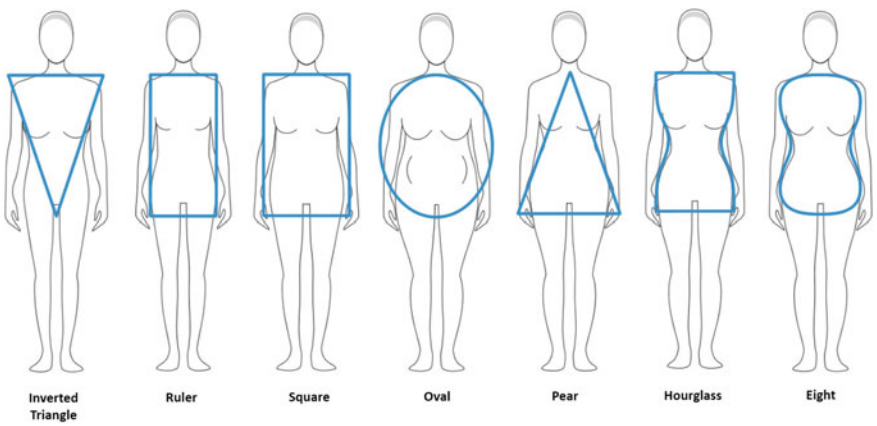

Fig. 9 A subsequent aspect of our research is the establishment of somatotype models in girls and therefore an aspect of the type of clothing, which can be designed for everyday use, as proposed in [18]

Fig. 10 Segmentation results of the proposed network

5 Multivariable Analysis

The prediction of future events is a difficult task to perform, because it requires an extensive multivariable analysis, it is also impossible to perform on several topics [19]. There are several methods that have been used as an auxiliary tool for the construction of estimation models. In our case, in the review of the literature we have detected that there are not enough antecedents in the area. In this work, the use of ubiquitous computation, image processing and deep leanings combine to predict the behavior in an evaluation of caloric intake in children with morbid childhood obesity, including the proposal of public policies by the state government. Chihuahua and is very similar to that proposed in [20]. Our approach proposes a model that includes two main characteristics: behavior associated with caloric intake and cultural characteristics of food in Mexico [13]. The model incorporates historical information on the allocation of food from a food repository, which families have made throughout previous editions linked to children's food on a normal day. In addition, the model includes information on the intrinsic characteristics of the nutritional menu that represents each proposal of caloric intake.

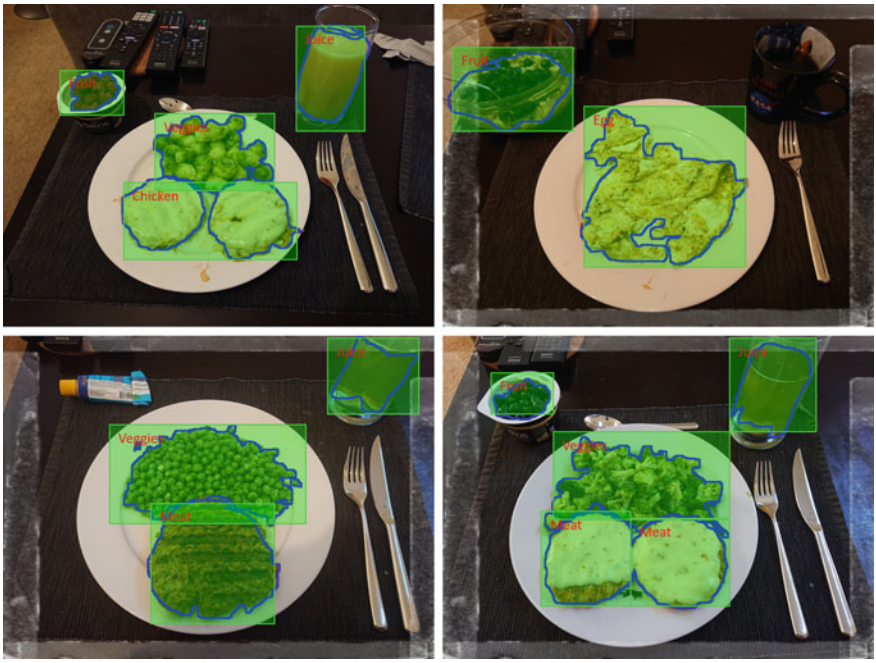

Fig. 11 Output images of the mask stage combined with the classification stage

Table 2 Measures of intersection over union (IoU) for each class

Class	IoU	Class	IoU	Class	IoU
Background	0.98484454	Egg	0.94114587	Vegetable	0.96853764
Juice	0.87009111	Coffee	0.81870388	Beans	0.90149943
Beef	0.8459461	Cereal	0.93650956	Fruit	0.97946966
Yogurt	0.97838868	Chicken	0.9701942	Cheese	0.98694365
Cream	0.99432157	Bread	0.95197437	Potato	0.9634425
Tortilla	0.9580226	Rice	0.94042164	Water	0.84217442

6 Conclusions and Future Research

The contribution of greater value of our research is related to a ubiquitous computer system capable of determining in real time the caloric intake of a child from the related food and immediately compare it with the population pyramid that should be covered as a function of his size, age, sex and physical complexion, as is proposed in [14]. Also, our proposed recognition module is compact and has an average IoU accuracy above of 90% for most types of food. A future research would be to analyze the proposals of nutritional menus from small companies such as the Faroe Islands, Guernsey, Jersey, Liechtenstein, Kosovo or Gibraltar, who face problems of various

food types including the obtaining of tempura and fresh fruits very different from countries with more than 1 million inhabitants and with extensive cultivation extensions. The OECD currently uses this type of innovative methods based on Artificial Intelligence to correctly characterize and evaluate the different opinions of different societies in the context of being able to listen to all existing voices, including those of minorities, something similar could be used in Ciudad Juáirez in where 37.14% of the population was not born in the state of Chihuahua - and because of this, the consumption of dishes from Zacatecas, Durango and Coahuila have a total and complete caloric intake, reaching 46.89% of the people who do not In the City, an approach based on these two major minorities could improve citizen participation.

References

1. I. Goodfellow, Y. Bengio, A. Courville, *Deep Learning* (MIT Press, 2016)
2. L.C. Chen, G. Papandreou, I. Kokkinos, K. Murphy, A.L. Yuille, Deeplab: semantic image segmentation with deep convolutional nets, atrous convolution, and fully connected CRFS. IEEE Trans. Pattern Anal. Mach. Intell. **40**(4), 834–848 (2018)
3. G. Lin, A. Milan, C. Shen, I.D. Reid, RefineNet: multi-path refinement networks for high-resolution semantic segmentation, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, vol. 1, no. 2 (2017), p. 5
4. H. Zhao, J. Shi, X. Qi, X. Wang, J. Jia, Pyramid scene parsing network, in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017), pp. 2881–2890
5. M. De Bonis, G. Amato, F. Falchi, C. Gennaro, P. Manghi, *Deep Learning Techniques for Visual Food Recognition on a Mobile App. MISSI* (2018), pp. 303–312
6. L. Herranz, W. Min, S. Jiang, Food recognition and recipe analysis: integrating visual content, context and external knowledge (2018). arXiv:1801.07239
7. W. Hui, M. Merler, R. Uceda-Sosa, J.R. Smith, Learning to make better mistakes: semantics-aware visual food recognition. ACM Multimed. 172–176 (2016)
8. K. Yanai, T. Maruyama, Y. Kawano, A cooking recipe recommendation system with visual recognition of food ingredients. IJIM **8**(2), 28–34 (2014)
9. C. Liu, Y. Cao, Y. Luo, G. Chen, V. Vokkarane, Y. Ma, Deepfood: Deep learning-based food image recognition for computer-aided dietary assessment, in *International Conference on Smart Homes and Health Telematics* (Springer, Cham, 2016), pp. 37–48
10. P. Kuhad, A. Yassine, S. Shimohammadi, Using distance estimation and deep learning to simplify calibration in food calorie measurement, in *2015 IEEE International Conference on Computational Intelligence and Virtual Environments for Measurement Systems and Applications (CIVEMSA)* (IEEE, 2015)
11. P. Pouladzadeh, P. Kuhad, S.V.B. Peddi, A. Yassine, S. Shirmohammadi, Food calorie measurement using deep learning neural network, in *IEEE International Instrumentation and Measurement Technology Conference Proceedings (I2MTC)*(2016), (pp. 1–6)
12. S. Xie, R. Girshick, P. Dollár, Z. Tu, K. He, Aggregated residual transformations for deep neural networks, in *IEEE Conference on Computer Vision and Pattern Recognition* (2017)
13. O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, in *International Conference on Medical Image Computing and Computer-Assisted Intervention* (Springer, Cham, 2015), pp. 234–241
14. X. Cui, V. Goel, B. Kingsbury, Data augmentation for deep neural network acoustic modeling. IEEE/ACM Trans. Audio Speech Lang. Process. (TASLP) **23**(9), 1469–1477 (2015)
15. F. Chollet, et al. Keras (2015)

16. G. Papandreou, L.C. Chen, K.P. Murphy, A.L. Yuille, Weakly-and semi-supervised learning of a deep convolutional network for semantic image segmentation, in *Proceedings of the IEEE International Conference on Computer Vision* (2015), pp. 1742–1750
17. G. Papandreou, L.C. Chen, K. Murphy, A.L. Yuille, Weakly-and semi-supervised learning of a DCNN for semantic image segmentation (2015). arXiv:1502.02734
18. C. Liu, Y. Cao, Y. Luo, G. Chen, V. Vokkarane, M. Yunsheng, P. Hou, A new deep learning-based food recognition system for dietary assessment on an edge computing service infrastructure. IEEE Trans. Serv. Comput. **11**(2), 249–261 (2018)
19. K. He, G. Gkioxari, P. Dollár, R. Girshick, Mask r-cnn, in *2017 IEEE International Conference on Computer Vision (ICCV)* (IEEE, 2017), pp. 2980–2988
20. S. Ren, K. He, R. Girshick, J. Sun, Faster r-cnn: towards real-time object detection with region proposal networks, in *Advances in Neural Information Processing Systems* (2015), pp. 91–99