HAL open science

Real-Time Detection of Eating Activity in Elderly People with Dementia Using Face Alignment and Facial Landmarks

Mhamed Nour, Mickaël Gardoni, Jean Renaud, Serge Gauthier

To cite this version:

Mhamed Nour, Mickaël Gardoni, Jean Renaud, Serge Gauthier. Real-Time Detection of Eating Activity in Elderly People with Dementia Using Face Alignment and Facial Landmarks. 17th IFIP International Conference on Product Lifecycle Management (PLM), Jul 2020, Rapperswil, Switzerland. pp.671-679, 10.1007/978-3-030-62807-9_53 . hal-03753113

HAL Id: hal-03753113

https://inria.hal.science/hal-03753113v1

Submitted on 17 Aug 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

This document is the original author manuscript of a paper submitted to an IFIP conference proceedings or other IFIP publication by Springer Nature. As such, there may be some differences in the official published version of the paper. Such differences, if any, are usually due to reformatting during preparation for publication or minor corrections made by the author(s) during final proofreading of the publication manuscript.

Real-time detection of eating activity in elderly people with dementia using face alignment and facial landmarks

Mhamed Nour[1*], Mickaël Gardoni[2], Jean Renaud[2], Serge Gauthier[3]

[1] Laboratoire NUMERIX, Département du Génie de la Production Automatisée, GPA, ETS, Montréal CANADA

[2] Professeur associé au département du Génie de la Production Automatisée, GPA, ETS, Montréal CANADA

[3] Professeur au McGill University, Research Centre for Studies in Aging, Montréal, Canada.

mhamed.nour.1@ens.etsmtl.ca*
*Corresponding Author

Abstract. The aim of this research is to automatically detect the intake of meals for elderly people with dementia living alone by using the product lifecycle management concepts (PLM). We use it to create a service based on an Artificial Intelligence product which will reduce the need for a caregiver or a person as medical support, hence less travel (Co2, etc.), less time spent by a third party on a verification/monitoring task, greater autonomy and therefore improve quality of life. Thus, overall, for society, this service based on an AI product is a win-win approach for pollution and an increase in the value of the tasks of the caregiver and the person as a medical support.

The choice of appropriate AI assistive technology was done to satisfy both the elderly people with neurodegenerative disorders and the caregiver, to verify the ethical aspect, simplify design, optimize code and improve user friendly aspects. During all this process of design of the new service based on an AI product, the PLM concepts were fruitfully applied by involving the different experts concerned (medical, ethical, technological, etc.) and taking into account all the characteristics of the environment of the product from the beginning to the end.

Keywords. Automatic detection, human activity, eating detection, neurodegenerative disorders, assistive technology

1 Introduction.

Dementia and its related diseases strike very strongly [1]. With alarming statistics and a lack of curative and preventive solutions to date. However, the non-mediated approach improves patients' quality of life by allowing them to regain more autonomy, memory, etc. and thus to delay the progression of the symptoms of the disease.

The current challenge in treating patients with mental disabilities is whether the patient has eaten perfectly, especially if the patient lives alone. Another challenge is when assistance techniques to identify eating without ethical conflict, using these assisting technologies, how could one influence the person being followed to induce them to stop eating or encourage him to eat.

Recognizing human activities from video footage using assistive technology is a difficult task due to problems, such as background clutter, partial occlusion, scale changes, point of sight, lighting and appearance and accompanying noise. Many applications, including patient video monitoring systems, require a multiple activity recognition system.

In this article, we provide a detailed review of recent progress in research in the area of human activity classification [2-6]. Then include the ethical aspect in the classification because the majority of current proposals for recognition of activities are carried out only in laboratories.

The aim is to develop an automatic real-time meal-taking detection project in elderly people with early dementia living alone to ensure that they eat their meals normally. Our work will be distinguished by the practical application of the ethical aspect that will surely reduce the options but with a practical realization, will be better suited as a solution to the detection of meal-taking in patients at the beginning of dementia because there will be no object attached to the patients and with a respect of the autonomy and privacy of the patient.

2 Artificial Intelligence Modeling.

Eating Activity Detection is part of the Human activity recognition (HAR) which aims [7] to recognize activities from a series of observations on the actions of subjects and the environmental conditions. The vision-based HAR research is the basis of many applications including video surveillance, health care, and human-computer interaction (HCI).

The challenges of the HAR domain are:

- the intraclass variation and interclass similarity,
- the recognition under real-world settings:
    - complex and various backgrounds.
    - multisubject interactions and group activities.

The paper reviews a list of papers in the Domain. They don’t take into account the ethical aspect of the assistive technology utilization and the respect the patient privacy.

The objective of this paper is the provide an approach for the Eating activity detection using assistive technology without the violation of the ethical aspect.

2.1 Assistive Technology for person with dementia review and ethical related Issues.

Paper [8] lists a series of AT methods to detect eating activity but all of them are still at the testing level in the LAB and they don't take into account the ethical aspect. The following are methods to detect eating using Assistive technology:

.

- Dietary Assessment Methods include food frequency questionnaires with limitations: forget, underestimate or overestimate portion sizes.
- Automated Dietary Monitoring includes oral sensors, wearable systems, on-body sensing approaches, crowdsourcing techniques, neckband wearable.
- Sensing with Objects, places and Artifacts include dining table, smart surface table, camera on dining ceiling light, sensor-embedded fork & interface mobile applications.
- Acoustic Sensing for Eating Detection include sound detection, bathroom sounds recorded with microphone.
- Recognizing Eating with On-body Inertial Sensing include small wearable accelerometers and gyroscope.
- Identifying Daily Routines and Patterns include the collection of data of many mobile phones over certain period on time.
- Techniques for Estimating Ground Truth in Real World Settings include the use of statistical machine learning for interference & modeling.

The objective of this paper is to use an assistive technology that respect the ethical aspect of the patient.

2.2 Artificial Intelligence and Machine Learning.

Human activity recognition is a challenging problem [9] given the large number of observations produced each second, and the lack of a clear way to relate accelerometer data to known movements. Deep learning methods such as recurrent neural networks and one-dimensional convolutional neural networks or CNNs have been shown to provide state-of-the-art results on challenging activity recognition tasks with little or no data feature engineering.

The following steps have been followed in order to build our Deep learning eating detection Model:

Step 1 — Data Pre-processing
Step 2 — Separating Your Training and Testing Datasets
Step 3 — Transforming the Data
Step 4 — Building the Artificial Neural Network
Step 5 — Running Predictions on the Test Set

Step 6 — Checking the Confusion Matrix
Step 7 — Making a Single Prediction
Step 8 — Improving the Model Accuracy
Step 9 — Adding Dropout Regularization to Fight Over-Fitting
Step 10 — Hyperparameter Tuning

2.3 Product Lifecycle Management concepts.

Product lifecycle management (PLM) [10] is the process of managing the entire lifecycle of a product from its conception, through design and manufacture, to service, and disposal. PLM integrates people, data, processes, and business systems and provides a product information backbone for companies and their extended enterprise. It can be represented as shown in Figure 1.
The conception of the Eating Activity Detection product is using the following steps of the PLM steps (Figure 1) to build our solution (building a prototype that grow up with the time):

- Manage and Collaborate during the whole project.
- Plan (Innovate and Specify the technology to be used for this project)
- Define (Develop and Validate the application)
- Build (Produce and Deliver. In the testing steps now.)
- Support (Service and Sustain)

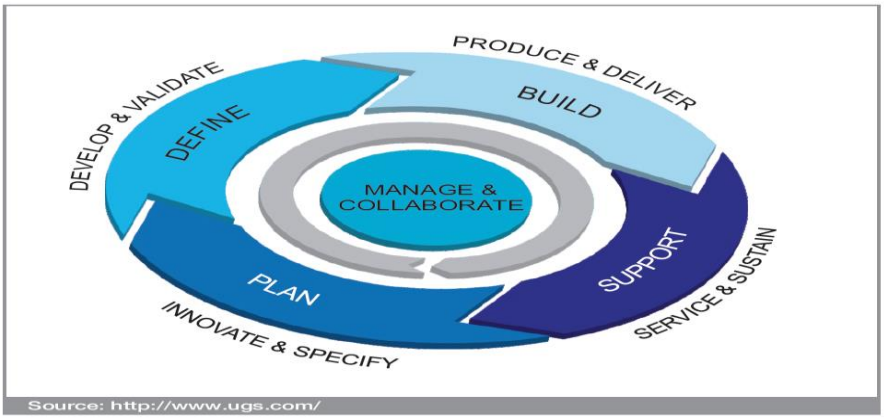

Fig 1, The PLM processes

3 Artificial Intelligence Modeling of the Real-time eating activity detection.

TensorFlow is an open-source software library for machine learning. It works efficiently with computation involving arrays; so, it's a great choice for the model. Furthermore, TensorFlow allows for the execution of code on either CPU or GPU, which is a useful feature especially when you're working with a massive dataset.

3.1 Facial landmarks detection.

Facial landmarks [11] are used to localize and represent salient regions of the face, such as: Eyes, Eyebrows, Nose, Mouth, Jawline. Facial landmarks have been successfully applied to face alignment, head pose estimation, face swapping, blink detection and much more.

The shape predictor algorithm [11] implemented in the dlib library comes from Kazemi and Sullivan's 2014 CVPR paper, The paper (Algo) addresses the problem of Face Alignment for a single image. The paper shows an ensemble of regression trees can be used to estimate the face's landmark positions directly from a sparse subset of pixel intensities, achieving super-Realtime performance with high quality predictions.

Our solution will use the facial landmarks detection to detection positions of the patient's face and store these positions in a database. In order to respect the patient's privacy, no pictures of the patients will be stored.

Figure 2 shows the Visualizing of the 68 facial landmark coordinates from the iBUG 300-W dataset. If the execution time of the solution is not acceptable, we may use another version of the facial landmarks that will use just for instance 5 facial landmark coordinates from the iBUG 300-W dataset.

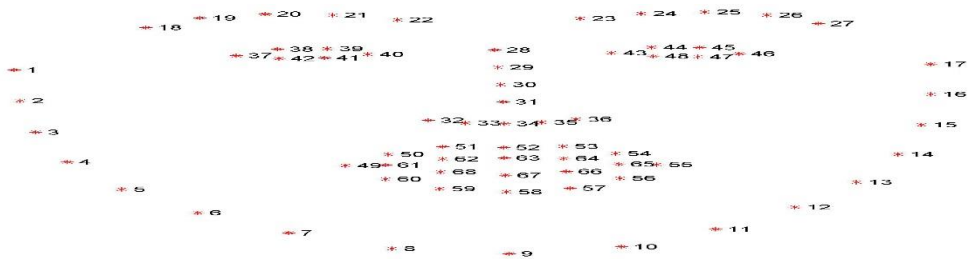

Fig. 2. Visualizing the 68 facial landmark coordinates from the iBUG 300-W dataset. [11].

3.2 Face Alignment

In order to have accurate data while recording the positions of multiple points of the face, a new procedure has been added to align the face of the patient. The face alignment procedure will be used during the data collection to train the model and during the track procedure.

Figure 3 shows the original picture on the right and the aligned picture on the left, the two eyes are aligned.

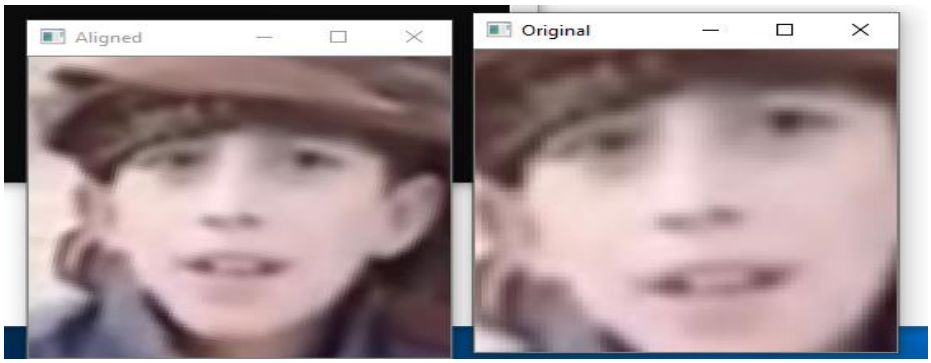

Fig. 3 shows the aligned procedure of the picture.

3.3 Data Collection.

Detecting the Jawlines and store their coordinate on csv files during the eating activity for many times. To verify the patient privacy, there is no video recording at all. For now, only two activities will be stored (eating and speaking), more activities will be added in the future. Figure 4. shows the corresponding plots of these eating activities.

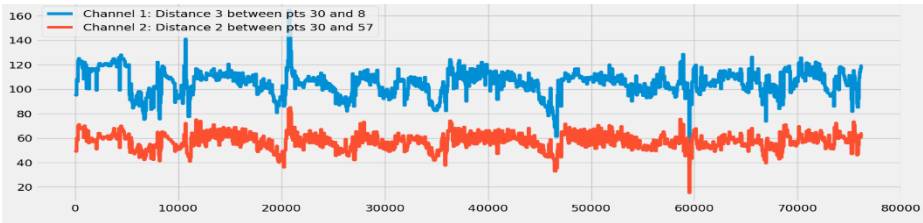

Fig. 4. the corresponding plots of these eating activities.

In order to distinguish between eating and speaking, we will record these two profiles and train the model to distinguish them. Figure 5 shows the patterns of these two activities. These patterns are repeated for many times and recoded using a data collection script.

Fig. 5. The plots of these eating and speaking activities.

The eating activity form a specific pattern. Figure 6 Shows the Coordinates of the facial points and the content of the csv file created by the Python Script. There are many similar csv files saved for the data training process.

	A	B	C
1	x_value	total_1	total_2
2			
3	1	99.08078	56.0803
4			
5	2	99.08078	56.0803

Fig. 6. Coordinates of the facial points and the content of the csv file created by the Python Script.

The dataset creation involved a python program for the creation of the csv files and another python program to plot the data. To respect patient privacy, no personal pictures or videos are recorded in this project.

3.4 Data Training

Datasets (training data) has been created which involves collecting data and labeling it. Having a high-quality labeled data is the key to develop good ML solutions.

The goal of the training is to build a classifier to be able to distinguish between these types of HAR activities. The fast Fourier transform (FFT) was used to obtain frequency characteristics.

After running FFT on the data, the Principal Component Analysis (PCA) was used to reduce the number of dimensions.

The windows size was 40, the number of trials used for training was 30, the rest was used for validation, the number of the PCA Components was 25 and the number of dimensions in the IMU Data was 6.

Four classes have been defined for this project: classes = ['Eating', 'Speaking', 'Smile', 'relaxing'].

The processed data becomes a set of 40-dimensional functionalities. An SVM (which is a Supervised Machine Learning algorithm) classifier was built using these feature sets in order to use an algorithm for classification. The SVM uses a kernel trick to transform data in order to found an optimal boundary between possible outputs.

The two models classifier.pkl and pca.pkl are the output of the train data program.

3.5 Real time activity tracking

A qualified classifier was used to track activities in real time. Data is collected in real time and a classification program is run continuously every 10 seconds.

Figure 7 provides a live activity tracking with the above trained model and shows live tracking result. It also shows the position values, their diagram plots and eating activity detection.

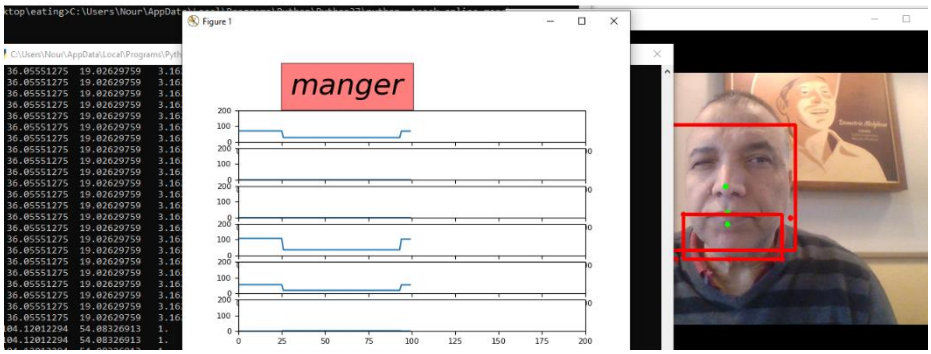

Fig. 7. Output of the track python script detecting eating activity.

3.6 Methodology validation

A validation process is currently installed in order to evaluate the prototype. This process need improvement with a configuration of parallel programming to speed up the execution of the real-time tracking activity. Then a real-time evaluation process can be done taking into account the variability of the PCA components.

4 Conclusions and Future works

This paper proposes a new approach for human activity detection specially the eating detection for elderly with data collection and tracking. At this stage optimizing performance and increasing classification accuracy is not the main goal of this project.

Most of the work in literature are still at the lab level and most of them use wiring assistive technology which may violate the ethical values of the patients. Our next step will be the optimization of the approach by adding a parallel computation.

References

1. Source Fondation pour la recherche médicale : https://www.frm.org/alzheimer/ampleur-maladie.html.
2. https://www.pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/
3. One Millisecond Face Alignment with an Ensemble of Regression Trees Vahid Kazemi and Josephine Sullivan KTH, Royal Institute of Technology Computer Vision and Active Perception Lab Teknikringen 14, Stockholm, Sweden {vahidk,sullivan}@csc.kth.se
4. Real Time Activity Recognition with accelerometer and gyroscope sensor (MYO) https://medium.com/softmind-engineering/real-time-activity-recognition-with-accelerometer-and-gyroscope-sensor-myo-663cda8536e6
5. https://github.com/hiukim/myo-activity-track?files=1
6. Belinda Bennett,a Fiona McDonald,a Elizabeth Beattie,b Terry Carney,c Ian Freckelton,d Ben Whitea & Lindy Willmotta Assistive technologies for people with dementia: ethical considerations, Bull World Health Organ 2017;95:749–755 | doi: http://dx.doi.org/10.2471/BLT.16.187484
7. Shugang Zhang,1 Zhiqiang Wei,1 Jie Nie,2 Lei Huang,1 Shuang Wang,1 and Zhen Li1, A Review on Human Activity Recognition Using Vision-Based Method, Hindawi, Journal of Healthcare Engineering, Volume 2017, Article ID 3090343, 31 pages https://doi.org/10.1155/2017/3090343
8. Tri Vu 1, Feng Lin 1, Nabil Alshurafa 2,3 and Wenyao Xu 1,* Wearable Food Intake Monitoring Technologies, A Comprehensive Review, Computers 2017, 6, 4; doi:10.3390/computers6010004
9. https://www.digitalocean.com/community/tutorials/how-to-build-a-deep-learning-model-to-predict-employee-retention-using-keras-and-tensorflow
10. Lakshminadh Javvadi, Product life cycle management An introduction, https://www.researchgate.net/publication/285769844