

Document downloaded from:

<http://hdl.handle.net/10251/178906>

This paper must be cited as:

Rincón Arango, JA.; Julian Inglada, VJ.; Carrascosa Casamayor, C. (2020). Towards the edge intelligence: Robot assistant for the detection and classification of human emotions. Springer. 31-41. https://doi.org/10.1007/978-3-030-51999-5_3



The final publication is available at

https://doi.org/10.1007/978-3-030-51999-5_3

Copyright Springer

Additional Information

Towards the edge intelligence: Robot assistant for the detection and classification of human emotions

J. A. Rincon, V. Julian, C. Carrascosa

Universitat Politècnica de València. VRAIN. Valencian Research Institute for Artificial Intelligence {jrincon, vinglada, carrasco@dsic.upv.es}

Keywords: EDGE AI, Assistant Robot, Emotions, Elderly

Abstract. Deep learning is being introduced more and more in our society. Nowadays, there are very few applications that do not use deep learning as a classification tool. One of the main application areas is focused on improving people's life quality, allowing to create personal assistants with canned benefits. More recently, with the proliferation of mobile computing and the emergence of the Internet of Things (IoT), billions of mobile and IoT devices are connected to the Internet. This allows the generation of millions of bytes of information about sensors, images, sounds, etc . Driven by this trend, there is an urgent need to push the IoT frontiers to the edge of the network, in order to decrease this massive sending of information to large exchanges for analysis. As a result of this trend, a new discipline has emerged: edge intelligence or edge AI, a widely recognised and promising solution that attracts with special interest to the community of researchers in artificial intelligence. We adapted edge AI to classify human emotions. Results show how edge AI-based emotion classification can greatly benefit in the field of cognitive assistants for the elderly or people living alone.

1 Introduction

Recent developments in artificial intelligent (AI) technologies have played an important role in several areas of knowledge such as medicine [1], robotics [2] or autonomous cars [3]. One of this late developments is the remote use of large computing units with high-performance servers, capable of executing millions of calculations per second (this is what we know today as *Cloud*). Among the recognised cloud services we can find Amazon Web Services (AWS)¹, which introduced its Elastic Compute Cloud; Microsoft Azure², announced as *Azure*, and Google Cloud Platform³. However, so much computing power has some drawbacks, such as the space they need and their high energy consumption.

¹ <https://aws.amazon.com/es/>

² <https://azure.microsoft.com/es-es/>

³ <https://cloud.google.com/>

In this regard, evolution of electronics and microelectronics has made it possible to create smaller, energy-efficient devices that are capable of accessing sensors or actuators, and with an internet connection, in other words, the Internet of Things (IoT) [4].

Using IoT devices we can connect everyday objects such as kitchen appliances, vehicles, thermostats, baby monitors, of the Internet through integrated devices. In this way, we can achieve a fluid communication between people, processes and things. However, sometimes, it is necessary to make a more intelligent interaction between the devices we have counted and the users. Nowadays, this is done by sending massively the data to the servers, which process, learn and classify these data. According to Cisco, about 50 billion IoT devices will be connected by 2020, but all these volumes of data are passive if they cannot be analysed or interpreted.

This high volume of data handling generates a series of problems such as latency in the reception of the results of the analysed data, the high energy consumption that these processing centers require and the space that they occupy. For these reasons, in the last few years a change in the IoT paradigm has been initiated, where it is no longer necessary to send data to servers for analysis, but it is now possible to perform an analysis or classification within these devices. This represents a great advance since these new devices have a low power consumption, are smaller and are equipped with structures capable of supporting deep learning models. These new elements are known as Edge AI devices, and they are being used in a massive way in applications of wearable devices such as [5], cyber-threats [6], implementation of machine learning algorithms in low memory devices [7], and assistant robot [8].

Personal assistants are perhaps some of the most interesting applications in which Deep Learning is used. This is mainly because they are in continuous interaction with humans. Allowing these assistants to learn their tastes, needs, perceiving how these needs affect their emotional states.

This data processing in most cases, is used to perform simple actions in the systems. However, some of these actions could be performed within each of the devices, for this it is necessary that these devices have the ability to use AI techniques as automatic learning models. These models would help to detect patterns in the lower layers of the system, thus avoiding the massive sending of information to the upper layers. This would reduce response times, as well as the massive sending of information to the same point.

These devices capable of performing these actions at a low level, is what we know today as Edge-AI. Edge-AI enables the creation of intelligent solutions in real-time using deep learning techniques. These solutions must have a number of key features, such as energy efficiency, low cost and a balance between precision and energy consumption. Currently, deep learning techniques are conventionally deployed in centralised computing environments. However, these applications have some limitations such as costs generated by energy consumption, or latency in the network due to massive data sending. To address these limitations, edge computing, often referred to as *Artificial Edge Intelligence*, has been introduced,

in which calculations are performed locally from data acquired from various devices or sensors.

The challenges of meeting the requirements for implementing Edge-AI are to ensure high accuracy of the algorithms while having low power consumption. However, this would not be possible without the latest hardware innovations, including central processing units (CPUs), graphics processing units (GPUs), application-specific integrated circuits (ASICs) and system-on-a-chip (SoC) accelerators, which have made edge AI possible. Thanks to these advances there are applications such as the one presented by [9], in which they detect apples in real-time using Edge-AI or Smart Parking using Edge-AI [10].

This paper presents an assistant robot based on Edge-AI technology, which incorporates two devices capable of classifying emotional states and classifying the physical activities performed by the user.

2 Proposed Robot Assistant

The proposed edge AI-based IoT application is a low-cost assistant robot with the ability to classify emotions and recommend physical activities to improve the quality of life of users. The Figure 1 shows the diagram of our proposal, which is divided into two parts: detection of emotions and classification of physical activities.

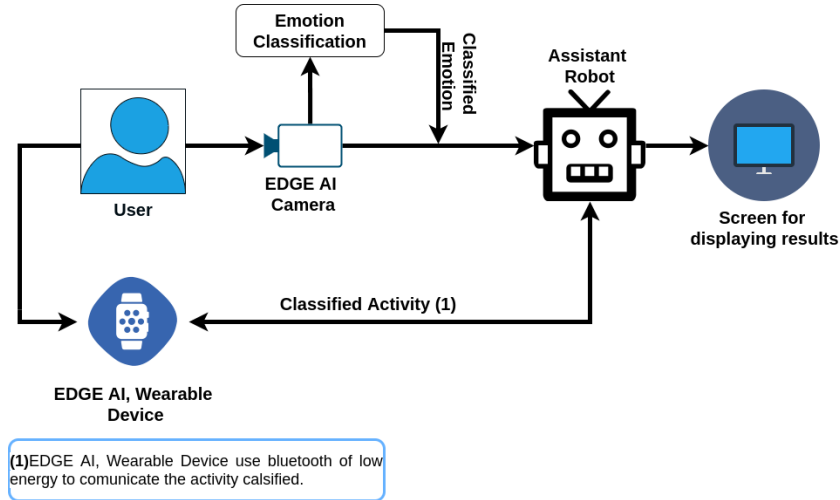


Fig. 1: Diagram of the proposed system.

The first one is in charge of classifying emotions and interaction with the user and the second one is in charge of detecting physical activity. The processes for

the classification of emotions and physical activities are carried out using EDGE-AI devices. Deep learning models are embedded in each device for classification process. The hardware and software used are described below.

To classify the emotions we use a M5StickV K210 A development system (Figure 2) which uses AI technologies embedded on a K210 chip ⁴. The Kendryte K210 is a chip system (SoC) with integrated machine vision and machine learning. This chip system employs advanced ultra-low processing with the help of a 64-bit dual-core processor. Also, it is equipped with a high performance hardware accelerator of the convolutional neural network (CNN) or KPU. The integrated machine vision function allows to create applications for object detection, image classification or face detection and recognition, providing the size and coordinates of the target in real time. With the K210's built-in microphone, it is possible to detect the orientation of the sound source, sound field images, voice alarm and voice recognition. All these features make this chip the ideal system for creating low-cost personal assistants.



Fig. 2: Development system M5StickV.

The classification of physical activity is based on an Arduino Nano 33 BLE (Figure 3) with an embedded 9-axis inertial sensor, suitable for assisted mobility devices (such as canes or walkers), or physical therapy devices (such as dumbbells or medicine balls).

The Figure 4 shows the prototype in charge of carrying out the classification of physical activities.

Furthermore, this board can work with deep learning to use light tensorflow models in order to classify the activities performed by the user. This way the

⁴ <https://kendryte.com/>

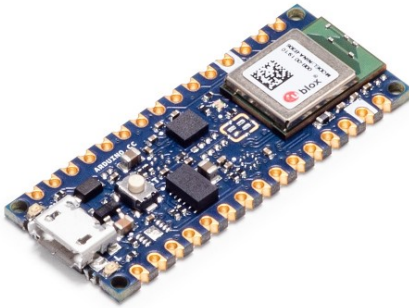


Fig. 3: Arduino Nano 33 Ble developer board.

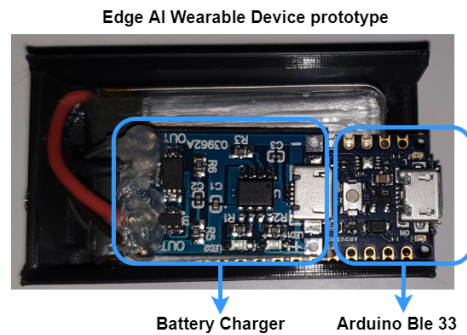


Fig. 4: Prototype Physical Activity Classifier.

robot tracks the activity and checks if the user has completed the exercise. The classification of the exercises is supported by a database built with exercises for boosting balance, posture, and core strength for elderly population. The database was built with 30 people, 15 men and 15 women between 35 and 50 years old. Each participant was asked to perform a total of 10 sets of 5 repetitions for each exercise. The robot was designed to fit into the environment of an elderly person in a lonely condition, meaning a small robot that can be placed on a desk. The robot has a camera that enables it to detect and identify the face in order to classify the user's emotional state. To increase the acceptability of the assistant by the user, the robot was given a pleasant appearance with a neutral look, ensuring that its physical appearance does not interfere with the human-assistant interaction. In addition, an information display device was added, which has an LCD screen using a M5Stack ⁵ on which the face of the assistant can be shown, as well as the activities to be performed and relevant messages.

⁵ <https://m5stack.com>

The Figure 5 shows the prototype of the assistant robot, with the different parts.

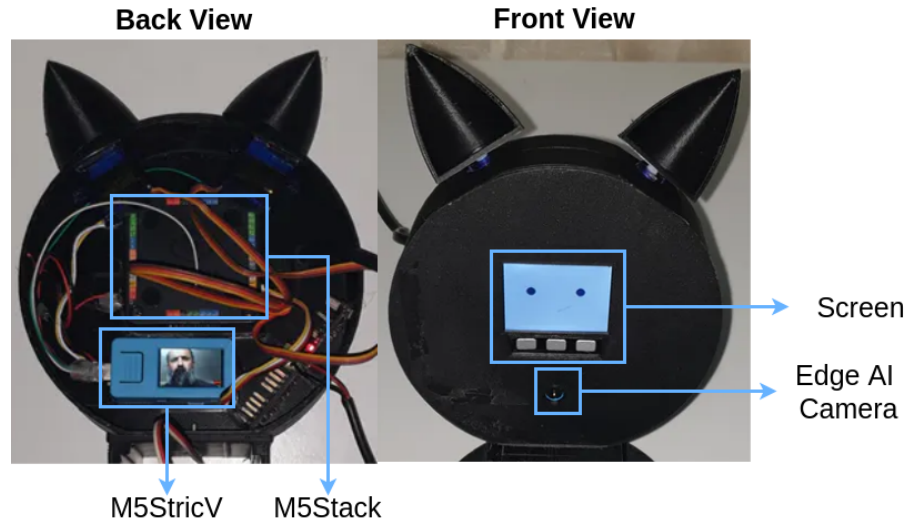


Fig. 5: Prototype of the assistant robot.

2.1 Software Description

This section describes each of the different steps carried out for the classification of emotions and the classification of activities. It describes the different tools used to carry out these tasks, as well as the division of the database and the different parameters that make up each of the networks used.

2.2 Emotion Classification

We propose an assistant robot with the ability to recognize emotional states. This ability provides the robot with information to interact more naturally with the user's and make some kind of recommendation based on their emotional state. To classify the emotions, a dataset was built in which the images were separated by classes, with each class representing an emotion. Seven emotions were taken into account to classify: fear, angry, disgust, happy, neutral, sad and surprised. We use the Karolinska Directed Emotional Faces (KDEF) database, containing 4900 images of human facial expression, as a source to build our dataset. The images show the faces of 70 amateur actors (35 men and 35 women) showing 7 different emotional expressions. Each expression is seen from 5 different angles. The selection criteria for the actors were: Age between 20 and 30 years and no

beard or moustache. During the photo session participants were asked to use some or no makeup and no earrings or glasses. To perform the experiments we divided the data set into three parts, 80% for training, 10% for testing and 10% for validation.

In order to carry out the qualification of emotions using EDGE devices, it is necessary to carry out a series of steps. These steps are necessary, since, the models obtained after the training using Keras [11] are not compatible with the K210 chip. These steps are divided into, (i) obtaining the H5 model, (ii) transformation to tensorflow lite and finally (iii) obtaining the Kmodel model. The network used to obtain the H5 model is a *Mobilenet* network [12]. This network was designed for mobile and integrated vision applications. MobileNets are based on a streamlined architecture that uses deep-separable convolutions to build deep and lightweight neural networks. We present two simple global hyperparameters that efficiently exchange between latency and accuracy. The Figure 6 shows the normalized confusion matrix of our experiments.

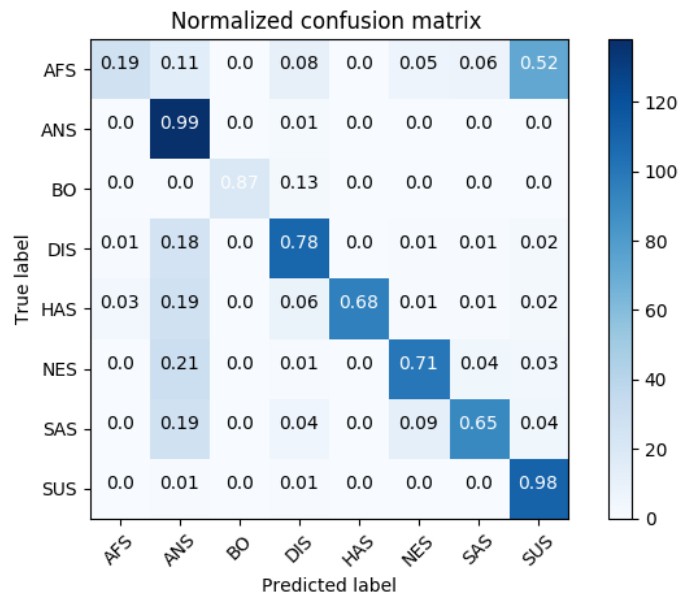


Fig. 6: Normalized confusion matrix of Emotion detection.

Once obtained from the trained model of this network, the next step is to transform it into TensorFlow Lite⁶. TensorFlow Lite has a set of tools that allows the develop to run TensorFlow models on mobile, embedded and IoT devices. The

⁶ <https://www.tensorflow.org/lite>

last step is the conversion of the tensorflow lite to K-Model, this transformation is performed using Sipeed tools ⁷.

2.3 Activity Classification

To carry out the classification of physical activities, a sequential network was used. To perform the classification of physical activities, a sequential network was used. The model configuration parameters are shown in the next JASON code (1), which specifies the density per layer and the activation rule for that layer.

Training results using this configuration can be seen below, in Figure 7 can be seen the Loss of the training and validation process.

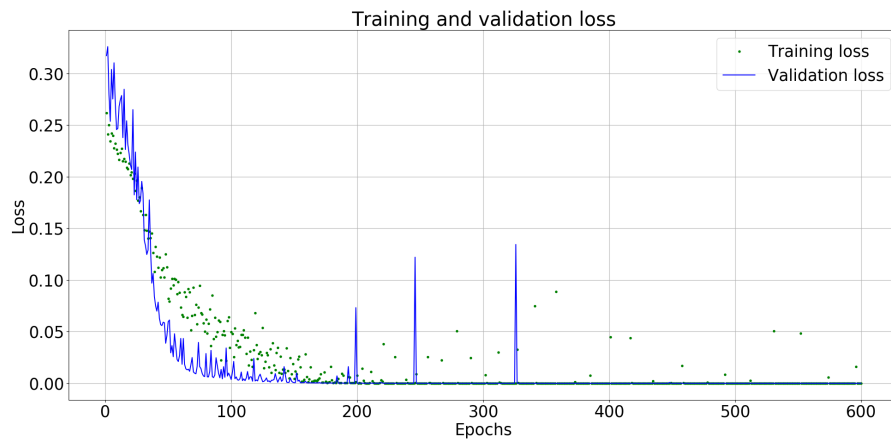


Fig. 7: Loss of Training and Validation process.

The Figure 8 shows the Mean Absolute Error, which allows us to determine the accuracy of our classification model. In the tests performed this error is very low, so we can determine that our classification of activities is correct.

3 Conclusions and future work

This paper presents the integration of two Edge AI technologies for the classification of emotions and the classification of physical activities for humans. The proposed system uses low-cost, low-power sensor systems. In this way, the system developed does not need to send the information obtained to servers for

⁷ <https://www.sipeed.com>

```

{
  'build_input_shape': (None, 714),
  'layers': [{ 'class_name': 'Dense',
    'config': { 'activation': 'relu',
      'activity_regularizer': None,
      'bias_constraint': None,
      'bias_initializer': { 'class_name': 'Zeros',
        'config': {}},
      'bias_regularizer': None,
      'dtype': 'float32',
      'kernel_constraint': None,
      'kernel_initializer': { 'class_name': 'GlorotUniform',
        'config': {'seed': None}},
      'kernel_regularizer': None,
      'name': 'dense',
      'trainable': True,
      'units': 50,
      'use_bias': True}},
    { 'class_name': 'Dense',
      'config': { 'activation': 'relu',
        'activity_regularizer': None,
        'bias_constraint': None,
        'bias_initializer': { 'class_name': 'Zeros',
          'config': {}},
        'bias_regularizer': None,
        'dtype': 'float32',
        'kernel_constraint': None,
        'kernel_initializer': { 'class_name': 'GlorotUniform',
          'config': {'seed': None}},
        'kernel_regularizer': None,
        'name': 'dense_1',
        'trainable': True,
        'units': 15,
        'use_bias': True}},
    { 'class_name': 'Dense',
      'config': { 'activation': 'softmax',
        'activity_regularizer': None,
        'bias_constraint': None,
        'bias_initializer': { 'class_name': 'Zeros',
          'config': {}},
        'bias_regularizer': None,
        'dtype': 'float32',
        'kernel_constraint': None,
        'kernel_initializer': { 'class_name': 'GlorotUniform',
          'config': {'seed': None}},
        'kernel_regularizer': None,
        'name': 'dense_2',
        'trainable': True,
        'units': 2,
        'use_bias': True}}}
  'name': 'sequential'
}

```

Listing 1: Structure of the model for the classification of activities.

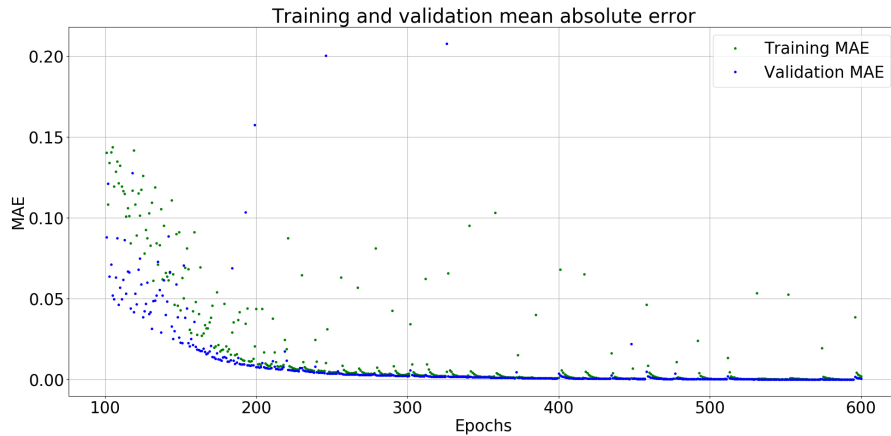


Fig. 8: Training and validation absolute error.

analysis, thus decreasing the latency in obtaining responses. This represents a breakthrough in the state of the art and tools used to make applications using Edge AI. Due to the combination of several deep learning models and the use of different low cost hardware devices. Future work will focus on the development of new tests with a larger number of users. These new tests will allow us to use the information obtained to improve our learning models for better recognition of different activities and emotions. We also focus our future research on the integration of speech recognition with these devices (Edge), so that users can use voice to control the robot.

Acknowledgements

This work was partly supported by the Generalitat Valenciana (PROMETEO/2018/002), by the Spanish Government (RTI2018-095390-B-C31) and Universitat Politècnica de Valencia Research Grant PAID-10-19.

References

1. Anthony Chang. The role of artificial intelligence in digital health. In *Digital Health Entrepreneurship*, pages 71–81. Springer, 2020.
2. Li Yang, Tony L Henthorne, and Babu George. Artificial intelligence and robotics technology in the hospitality industry: Current applications and future trends. In *Digital Transformation in Business and Society*, pages 211–228. Springer, 2020.
3. Hamid Khayyam, Bahman Javadi, Mahdi Jalili, and Reza N Jazar. Artificial intelligence and internet of things for autonomous vehicles. In *Nonlinear Approaches in Engineering Applications*, pages 39–68. Springer, 2020.

4. Fan Liang, Wei Yu, Xing Liu, David Griffith, and Nada Golmie. Towards edge-based deep learning in industrial internet of things. *IEEE Internet of Things Journal*, 2020.
5. Pradeep Baliganapalli Nagaraju, Adam Jamison Oliner, Brian Matthew Gilmore, Erick Anthony Dean, and Jiahua Wang. Data analytics in edge devices, January 9 2020. US Patent App. 16/573,745.
6. Mojtaba Eskandari, Zaffar Haider Janjua, Massimo Vecchio, and Fabio Antonelli. Passban ids: An intelligent anomaly based intrusion detection system for iot edge devices. *IEEE Internet of Things Journal*, 2020.
7. A Harish, Saksham Jhavar, BS Anisha, and P Ramakanth Kumar. Implementing machine learning on edge devices with limited working memory. In *Inventive Communication and Computational Technologies*, pages 1255–1261. Springer, 2020.
8. Jaime Andres Rincon, Alberto Martin, Ângelo Costa, Paulo Novais, Vicente Julián, and Carlos Carrascosa. Emir: An emotional intelligent robot assistant. In *AfCAI*, 2018.
9. Ruimin Ke, Yifan Zhuang, Ziyuan Pu, and Yinhai Wang. A smart, efficient, and reliable parking surveillance system with edge artificial intelligence on iot devices. *arXiv preprint arXiv:2001.00269*, 2020.
10. Vittorio Mazzia, Aleem Khaliq, Francesco Salvetti, and Marcello Chiaberge. Real-time apple detection system using embedded systems with hardware accelerators: An edge ai application. *IEEE Access*, 8:9102–9114, 2020.
11. François Chollet et al. Keras. <https://github.com/fchollet/keras>, 2015.
12. Andrew G Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, and Hartwig Adam. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861*, 2017.