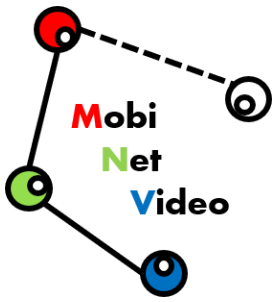


MobiNetVideo Newsletters

#7 – September 2021

TEC2017-88169-R MobiNetVideo (2018-2020-2021)

Visual Analysis for Practical Deployment of Cooperative Mobile Camera Networks

http://www-vpu.eps.uam.es/MobiNetVideo/

Extension progress report

During the nine months of extension that we were granted, the work was focused on developing applications within WP4 in order to show the results from the project in a more application-oriented approach, targeting to foster technology transfer. In this line, on 24th June, MobiNetVideo Industry Day was held (see below). Several publications were produced and send for review: currently we are still waiting for the completion of the review process in different international journals (the web site will be updated even after the project extension, when they are finally published). From the three PhD thesis working within the project framework, two of them, originally planned for completion during 2021, will be finally defended around summer 2022 (extensions were granted due to the effects of the pandemia).
The web site was maintained, and three deliverables were published: D3.2 "Technologies for mobile camera networks" (February 2021), D4 "Deployment and application scenarios" (September 2021), and D5.v4 "Results Report" (September 2021)

Industry day

The MobiNetVideo Industrial Day was held on June 24th, 2021. Although it was planned as a face-to-face event, finally it was held on-line in order to accommodate the suggestions of the attending companies.
Different applications and demonstrators developed within WP4 were presented:

- Continual learning for object detection
- Places detection in ego-centric video
- Contributions to the re-identification of objects
- Multi-Object Tracking and Applications

- Real-time camera operation and tracking for the streaming of teaching activities
- Computer Vision Applications for Event Cameras

After the presentatios, there was a discussion session with the industry representatives in order to get feedback, mainly with respect to suggestions for enhancements and new applications

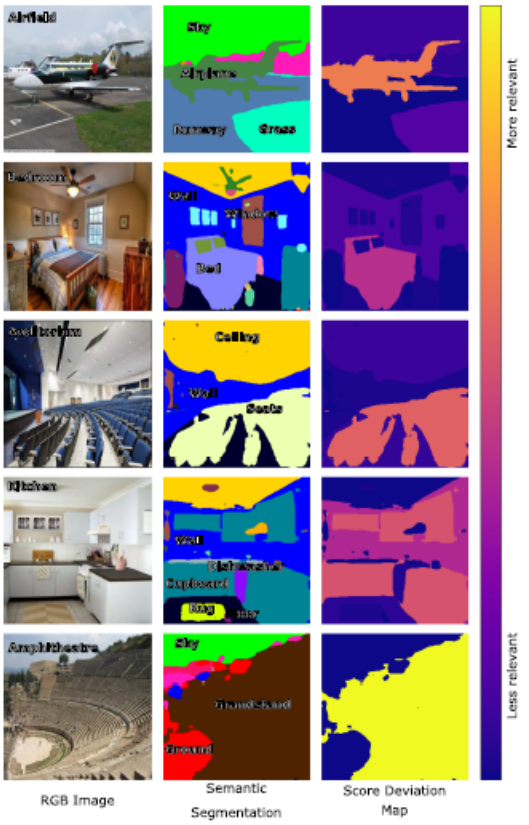
The Industry participation included people from the following companies: Thales Aelnia Space, Nokia Bell Labs Madrid, Treelogic, Vaelsys, Dive Tech, and Capgemini.

Seventh semester results

Conferences

Alejandro López-Cifuentes, Marcos Escudero-Viñolo, Andrija Gajic, Jesús Bescós, **Visualizing the Effect of Semantic Classes in the Attribution of Scene Recognition Models**, Proc. of Workshop "Explainable Deep Learning-AI" – 25th International Conference on Pattern Recognition (ICPR2020)Virtual Conference, Jan. 2021), LNCS 12663, pp.115-129, Springer, 2021 (DOI 10.1007/978-3-030-68796-0_9).

The performance of Convolutional Neural Networks for image classification has vastly and steadily increased during the last years. This success goes hand in hand with the need to explain and understand their decisions: opening the black box. The problem of attribution specifically deals with the characterization of the response of Convolutional Neural Networks by identifying the input features responsible for the model's decision. Among all attribution methods, perturbationbased methods are an important family based on measuring the effect of perturbations applied to the input image in the model's output. In this paper, we discuss the limitations of existing approaches and propose a novel perturbation-based attribution method guided by semantic segmentation. Our method inhibits specific image areas according to their assigned semantic label. Hereby, perturbations are link up with a semantic meaning and a complete attribution map is obtained for all image pixels. In addition, we propose a particularization of the proposed method to the scene recognition task which, differently than image classification, requires multi-focus attribution models. The proposed

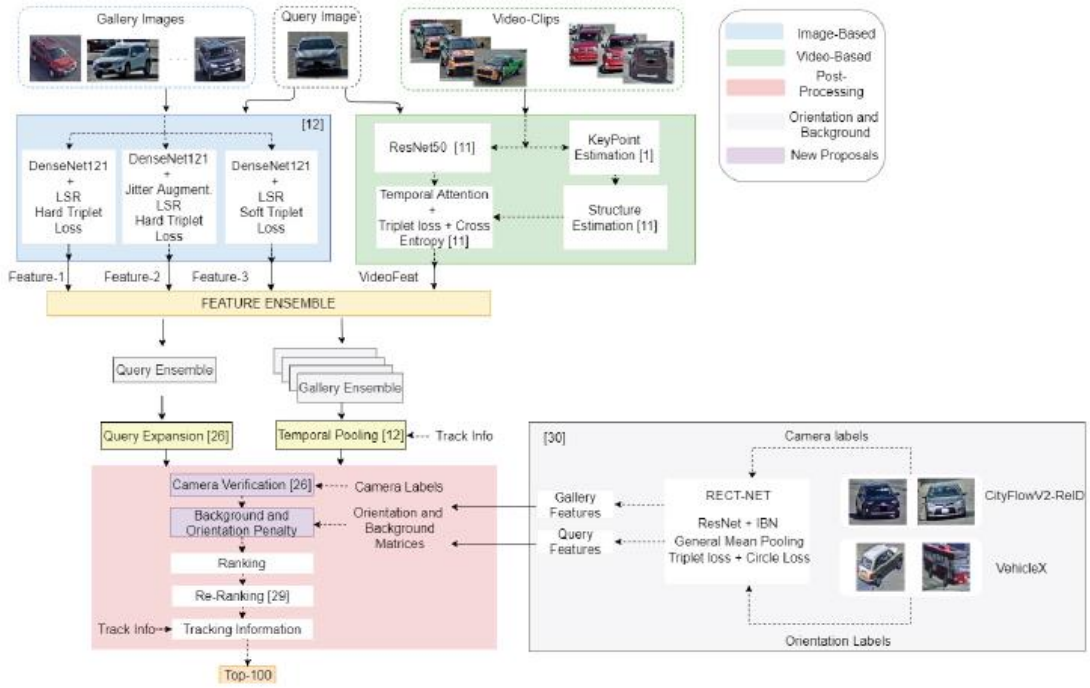


semantic-guided attribution method enables us to delve deeper into scene recognition interpretability by obtaining for each scene class the sets of relevant, irrelevant and distracting semantic labels. Experimental results suggest that the method can boost research by increasing the understanding of Convolutional Neural Networks while uncovering datasets biases which may have been inadvertently included during the harvest and annotation processes. All the code, data and supplementary results are available at http://www-vpu.eps.uam.es/publications/SemanticEffectSceneRecognition/

Marta Fernández, Paula Moral, Álvaro García-Martín, José M. Martínez, "Vehicle Re-Identification based on Ensembling Deep Learning Features including a Synthetic Training Dataset, Orientation and Background Features, and Camera Verification", Proc. of IEEE Int. Conf. on Computer Vision and Pattern Recognition (CVPR2021), Virtual Conference, Jun. 2021 (DOI 10.1109/CVPRW53098.2021.00459.

Vehicle re-identification has the objective of finding a specific vehicle among different vehicle crops captured by multiple cameras placed at multiple intersections. Among the different difficulties, high intra-class variability and high inter-class similarity can be highlighted. Moreover, the resolution of the images can be different, which also means a challenge in the re-identification task. Intending to face these problems, we use as baseline our previous work based on obtaining different deep learning features and ensembling them to get a single, stable and robust feature vector. It also includes post-processing techniques that explode all the information provided by the CityFlowV2-ReID dataset, including a re-ranking step. Then, in this paper, several newly included improvements are described. Background and orientation similarity matrices are added to the system to reduce bias towards these characteristics. Furthermore, we take into account the camera labels to penalize the gallery images that share camera with the query image. Additionally, to improve the training step, a synthetic dataset is added to the original one

---



Master thesis

Semantic segmentation in 2D videogames, Javier Montalvo Rodrigo (advisor: Alvaro García Martín), Trabajo Fin de Máster (Master Thesis), Máster Universitario en Deep Learning for Audio and Video Signal Processing, Univ. Autónoma de Madrid, Jun. 2021.

Abstract: This Master Thesis focuses on applying semantic segmentation, a computer vision technique, with the objective of improving the performance of deep-learning reinforcement models, and in particular, the performance over the original Super Mario Bros videogame. While humans can play a stage from a videogame like Super Mario Bros, and quickly identify from the elements in the screen what object is the character they are playing with, what are enemies and what elements are obstacles, this is not the case for neural networks, as they require a certain training to understand what is displayed in the screen. Using semantic segmentation, we can heavily simplify the frames from the videogame, and reduce visual information of elements in the screen to class and location, which is the most relevant information required to complete the game. In this work, a synthetic dataset generator that simulates frames from the Super Mario Bros videogame has been developed. This dataset has been used to train

semantic segmentation deep-learning models which have been incorporated to a deep reinforcement learning algorithm with the objective of improving the performance of it. We have found that applying semantic segmentation as a frame processing method can actually help reinforcement learning models to train more efficiently and with better generalization. These results also suggest that there could be other computer vision techniques, like object detection or tracking, that could be found useful to help with the training of reinforcement learning algorithms, and they could be an interesting topic for future research.

Application for the demonstration of the automatic registration of transited spaces for contact tracing of infectious diseases using video signals from life-logging cameras, Daniel de Alcalá Valcárcel (advisor: Marcos Escudero-Viñolo), Trabajo Fin de Máster (Master Thesis), Máster Universitario en Deep Learning for Audio and Video Signal Processing, Univ. Autónoma de Madrid, Jun. 2021.

Abstract: In recent years having a register of the places that a person has visited is a very powerful tool for several tasks. It can even be used for tracing COVID-19 contacts, a ubiquitous topic in these times, or to model routines of a person to teach a machine. Previous work on this is scarce so they do not achieve accurate results. Ego-topo was the first work to reach a hopeful performance. The purpose of this thesis is to extend their approach, improve the results, and finally deploy it to a real application. In the first place, a study has been made of the systems that have led to the current Ego-topo. Next, the theoretical and algorithmic details of the Ego-topo system were delved into, with the aim of understanding its strengths and weaknesses. Later, several points of the system were modified in order to achieve added functionality, including enabling the use of external videos and the construction of combined graphs between different users. These combined graphs allowed the development of an application to detect contacts between two users in a domestic environment, with the aim of infectious diseases tracing. Before the application development, some weaknesses of the base system that harm its results were improved. The new changes were evaluated rst subjectively and then objectively with the creation of semi-automatic region annotations. Finally, with these improvements, the development of the application's graphical interface was carried out. These are initial steps in this topic which provide a wide eld of research and applications. Future works could transfer this illness tracking to bigger scale environments, for example an airport, and they may reach better results beneting from other features like the sound of different areas.

Contributions to the re-identification of objects, Marta Fernández de Barrio (advisor: Alvaro García Martín), Trabajo Fin de Máster (Master Thesis), Máster

Universitario en Ingeniería de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2021.

Abstract: Vehicle re-identification has the objective of finding a specific vehicle among different vehicle crops captured by multiple cameras placed at multiple intersections. The relevance of this technology has been increasing during the last years due to its contribution to the functioning of smart cities and Intelligent Transport Systems. Among the different difficulties, high intra-class variability and high inter-class similarity can be highlighted. Moreover, the resolution of the images can be different, which also means a challenge in the re-identification task. Different bias towards several characteristics such as orientation and background can also be a drawback. Specifically, vehicles with similar orientation, background, type or colour may have a higher probability of being identified as the same ID. Furthermore, there is a limited number of labelled images, which makes it difficult to train the networks. Intending to face these problems, we use as baseline a previous work based on obtaining different deep learning features and ensembling them to get a single, stable and robust feature vector. Specifically, the system is composed of two main blocks: The image-based block and the video-based block, which provide, in total, four feature vectors for each image. Three of them contain information about the appearance and the last one includes structure information. Then, they are concatenated leading to a single vector. It also includes post-processing techniques that take advantage of all the information provided by the CityFlowV2-ReID dataset. Firstly, query expansion and temporal pooling are carried out to improve the feature vectors, and, secondly, a re-ranking step and the tracking information are included to improve the initial ranking. Then, in this thesis, several newly included improvements are described. First of all, background and orientation features are obtained for each sample in order to create distance matrices between them. Then, these matrices are used with the aim of reducing bias towards these characteristics. Furthermore, we take into account the camera labels to penalize the gallery images that share the camera with the query image. Finally, a synthetic dataset is added to the original training set so that the number of training samples is increased and, then, the training step is improved.

Aprendizaje y corrección de errores en sistemas de seguimiento basados en redes convolucionales siamesas (**Learning and correction of errors in tracking systems based on Siamese Convolutional Networks**), Álvaro Iglesias Arias (advisor: Marcos Escudero-Viñolo). Trabajo Fin de Máster (Master Thesis), Máster Universitario en Deep Learning for Audio and Video Signal Processing, Univ. Autónoma de Madrid, Jul. 2021.

Abstract: This Master's Thesis aims to study and refine trackers based on Siamese neural networks in the face of classic challenges in the field of object tracking, such as occlusions or distractors (presence on the scene of object with the same appearance as the target object). Since the explosion of deep learning techniques a few years ago and the increase in complexity and size of the available datasets there is no field of engineering that has not been affected by this. One of the most representative examples of this changes can be found in the field of the computer vision. There are so much different applications in this ambit each one with its one challenges. From all these applications one of the most challenging is object tracking due to the diversity of possible situations that can be encountered. Because of this tracker systems must be capable of adapt to multiple situations with ease In this project two different algorithms have been proposed and implemented in Matlab, both based in the technique known as backtracking (or backward tracking). Their objective is to mitigate the problems that arise in multiple trackers in presence of two different events: target occlusions and the appearance of distractors. These systems work as modules that can be attached to the end of the siamese neural networks that conforms the core of the tracker and use their output to refine the results obtained. This system-module approach makes this work flexible and open possibilities of improvements in multiple trackers that conform the state of the art. Also, to validate the results obtained a dataset with labeled occlusions at frame level has been created from a subset of videos obtained from other datasets. These videos and their events have also been tagged at video level. Lastly, results have been checked for each one of these events using our own dataset. Improvements and failures of each category have been analysed and commented also the problems found.

Real-time camera operation and tracking for the streaming of teaching activities, Javier Vinuesa Solana (advisor: Jesús Bescós Cano), Trabajo Fin de Máster (Master Thesis), Máster Universitario en Deep Learning for Audio and Video Signal Processing, Univ. Autónoma de Madrid, Jul. 2021.

Abstract: The primary driving force of this work comes from the objective of offering students the opportunity to attend a remote event from home or anywhere in the world in real-time. The main objective of this work is to build a real-time tracker to follow the movements of the lecturer. After that we will build a framework to handle a PTZ (Pan Tilt and Zoom) camera based on the lecturer movements. That is, if the lecturer goes to the left, the camera will turn to the left. To tackle this project, we will follow a project developed by Gebrehiwot, A. which involved building a real-time tracker. The problem of this tracker is that was implemented on Ubuntu and running with a very complex CNN which

required the use a good GPU on our computer. As Gebrehiwot, A. rightly points out at the end of his report, not everyone has an Ubuntu partition or a GPU on their computers, so we started moving the real time tracker to Windows. To achieve this objective, we used Anaconda Windows which made our work much easier. After that we implemented a lightweight backbone of the tracker allowing us to run it on computers with a fewer processing power. Once that all this process was done, we put into practice the mentioned framework for handling the movement of the PTZ camera. This framework uses the implemented lightweight tracker to follow the lecturer moves and depending on these movements the camera will pan and tilt automatically. We tested this framework on streaming platforms like YouTube proving that can greatly improve the quality of online classes. Finally, we draw conclusions from the work done and propose future work to improve the framework.

Continual Learning for Object Detection, Carlos Jiménez Muñoz (advisor: Álvaro García Martín). Trabajo Fin de Máster (Master Thesis), Máster Universitario en Deep Learning for Audio and Video Signal Processing, Univ. Autónoma de Madrid, Sept. 2021.

Abstract: This Master Thesis is contextualized in the context of the need to count containers in different cities in Spain for garbage collection. In order to perform this counting automatically, to save time and resources, it is proposed to use the technique of object detection on images captured from a car. At the research proposal level, the hypothesis is as follows: "it is possible to create a model/system that takes advantage of the knowledge acquired from the cubes of a city to facilitate learning in new cities, without signicantly worsening its results in the previously known city". To achieve this, one has to deal with a problem derived directly from the training algorithm of a neural network, the Catastrophic Forgetting. This problem implies that, once a model with certain weights is trained to solve a new task (e.g., recognition of a new class) its results will be worse on the previously known task/class. In order to try to solve this problem, a current called Continual Learning (also known as Incremental Learning or Lifelong Learning) arises within the area of Artificial Intelligence, which introduces techniques based on reproduction, regularization or isolation of parameters, among others, to try to mitigate the effect of Catastrophic Forgetting in the results of the model, retaining the knowledge prior to the retraining of the model with the new data. Specifically, in this project we have worked to achieve a system that is capable of detecting containers of new cities without losing accuracy in the detection of previously known cities. To this end, we have carried out annotation tasks, adaptation of a GitHub repository where the EfficientDet model is used for object detection, measurement and mitigation

of catastrophic forgetting and finally creation of an interface to facilitate the use of the application to a potential user.

Graduate thesis

Aprendizaje continuo mediante redes convolucionales (**Continuos learning using convolutional networks**), Anselmo Velázquez Pazos (advisor: Álvaro García Martín), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Feb. 2021.

Abstract: The main focus of this final degree project is the study and use of algorithms or continuous learning techniques for different sets of data and images. First, there will be a brief introduction to the basic concepts of Artificial Intelligence, convolutional neural networks and their application in the world of Deep Learning, since understanding and knowing this type of network is important for continual learning. Second, the datasets used and their given attributes are described. There is a set of own images provided by the Video Processing and Understanding Lab (VPU) of the UAM Escuela Politécnica Superior. In addition, a brief introduction will be made to the development enviroment that has been used to carry out this work and to the algorithms used in the different experiments. Third, the evaluation criteria of the results obtained will be explained, as well as their evolution during the use of different algorithms for continual learning. Different experiments will be carried out in which conclusions and results will be drawn, thanks to which the importance and power of continual learning will be seen. Finally, visual tests will be provided to demonstrate the results obtained and avenues of investigation that could be carried out in the future will be proposed.

Segmentación objeto-fondo mediante redes convolucionales (**Object-background segmentation using convolutional networks**), Raúl Arcos Serrano (advisor: Álvaro García Martín), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jun. 2021.

Abstract: Currently, techniques based on deep learning or Deep Learning have achieved really good performances in a large number of different tasks in which they have been applied. Regarding the data to be used, there is a great availability of different sets that will allow the training of the models. These data sets may or may not be labeled, which is a very expensive task and influences the type of learning as well as the efficiency of the model, being those that have more

efficient labels then employing supervised learning. Therefore, the objective of this Final Degree Project is the implementation of a Deep Learning system based on convolutional neural networks, which has the ability to solve an image classification problem. To do this, we are going to focus on the evaluation of a specific case: semantic segmentation. However, we are going to investigate further in this technique to be able to differentiate between a specific class and the rest of the image, thus allowing us to differentiate between a specific object and the background. The first part of the project will contain the state of the art, where we will see the most general concepts related to neural networks. This section will allow us to put ourselves in the appropriate context in order to better understand what we have done throughout the work. Next, we will explain how our design will work in terms of the development environment, the networks used to train our model, the steps followed to obtain the results and how these steps have been modified from the base code. Finally, we will evaluate the experiments carried out as well as the results obtained in order to draw conclusions about the work carried out and, based on these, propose a series of ideas as future work.

Ejecución de redes neuronales en móviles Android con aceleración hardware mediante Keras y Tensorflow (**Running Neural Netowrks in Android Mobile Phones qith hardware acelleration with Keras and Tensorflow**), Ángel Fragua Baeza (advisor: Miguel Ángel García), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería Informática, Univ. Autónoma de Madrid, Jun. 2021.

Abstract: The Neural Networks have evolved in every way possible, since its first appearance in 1943 by Warren McCulloch. These early networks were too simple and the had plenty of limitations, but after a decade it arises a new model named Perceptron. This new Neural Network brought a new idea of learning, but shortly after there were some discoveries about its limitations, like the impossibility of learning the pattern of the XOR gate. It was that moment when the first "AI Winter" occurred. This problem was solved by the Backpropagation mechanism. Since that moment until now, there has been plenty "AI Winters", but with effort and time everyone has been solved. One of the biggest problems of the Neural Networks that nowadays is still a thing, is its high computational cost. However, there are some options that let you use this high demanding Networks with minimal resources with high efficiency. This is the reason why this project will make a full study about Neural Networks in mobile devices, taking advantage of all the tools that TensorFlow Lite provides. This paper will make use of some pretrained Networks obtained from Keras, that can be divided in two main groups, the more consolidated models including VGG16 and ResNet50, and the models focused on mobile devices like MobileNet and EfficientNet. The Neural

Networks are trained based on a very well-known dataset named ImageNet, which objective is to be able to classify those images or some similar ones once trained. These Networks will be converted to its corresponding TensorFlow Lite models, applying them some optimizers in those cases that they are possible. Once obtained the TensorFlow Lite model, it will be an exhaustive study about the impact of applying different hardware and software acceleration on the image inference, using its specific delegates like the GPU and NNAPI. Also, it will be introduced these new TensorFlow Lite models into a real Android application, with the corresponding image pre-processing required for each model. This will close the complete cycle about using Neural Networks within a mobile device from the start.

Detección de colisiones mediante procesado de video (**Collisions detection using video processing**), Sergio Avello Largo (advisor: Álvaro García Martín), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2021.

Abstract: This Bachelor Thesis consists of two quite different parts. In the first one, a technology with so much potential as accident detection by video analysis is presented. The possible applications where its use would be very useful and the role it could play in the near future are discussed. Then, the basic concepts on which this science is based are introduced, such as Machine Learning, Artificial Vision, artificial neural networks, or object detection. In this way, even if the reader is not an expert in the field, he/she will be able to get an idea of the context in which accident detection is included and will be able to understand the content of the following chapters. The second part is somewhat more technical, and we focus on a real case of accident detection around a vehicle. Based on a model that had already been developed using the Faster R-CNN object detector and a VGG neural network, we conducted an experiment by modifying the object detector (with EfficientDet) and feature extraction modules (with VGG16, AlexNet, DenseNet, ResNeXt). The algorithm is trained and tested with a dataset of clips in which accidents occur around a vehicle and is able to generate the probability of accident that exists for each of the frames. From this probability and after a complicated processing, the results of Area Under the Curve and Tiempo al Accidente are obtained to measure the performance of the experiment. Finally, all the results obtained are compared and analyzed to find the best model structure depending on the task to be conducted.

Adaptación de algoritmos de acuerdo con información contextual extraída automáticamente (**Algorithms adaptation based on automatically extracted contextual information**), Javier Santos Gimeno (advisor: Marcos Escudero-

Viñolo), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Jul. 2021.

Abstract: Since the advent of deep learning-based solutions, the performance of computer vision tasks has increased exponentially. Setting up an object detector to operate correctly in different scenarios has become a difficult task. Therefore, it is essential for an object detector to have a correct configuration of confidence or decision thresholds. To achieve this correct configuration of decision thresholds, we study in this work the feasibility of using a semantic segmentator, which provides the necessary contextual information. The semantic segmentator provides semantic information for each image of a video, associating each pixel a label of the class to which it belongs. To do so, we start by performing a state-of-the-art study of the basic concepts of object detectors and semantic segmenters. Then, we describe the different models of object detectors and semantic segmentators that have been used for the development of the work. Subsequently, the working environments used and the data sets used, both in the work and in the detection and segmentation methods, are explained. Next, the development of the work is explained, including the object detector, the semantic segmentator and the characterization and classification of the detections. Finally, the results obtained in the experiments are evaluated and analyzed. For this purpose, a comparison of the results obtained by different combinations of detector and segmentator is carried out. Based on these results, objective conclusions are drawn and possible future work is proposed.

Desarrollo de redes neuronales en dispositivos móviles mediante PyTorch Mobile (**Development of Neural Netowrks in Movile Deveices using PyTorch Mobile**), David García Espinoza (advisor: Miguel Ángel García), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Sept. 2021.

Abstract: Today, Deep learning is being a key tool for the development of new technologies and the advancement of humanity. For a few years due to the progress of hardware and the possibility of handling large amounts of data (Big Data), this technology is widely implemented in computers and PCs, allowing them to perform tasks that would not be possible without Deep learning. Despite this, some devices still do not have the necessary facilities to handle neural networks with ease. An example of this are mobile devices and smartphones. This Final Degree Project aims to study the use of neural networks in these devices and compare them with hardware that is already used to their application. To carry out this task, it will first be necessary to do a state-of-the-art analysis in which the internal functioning of neural networks will be detailed.

We will delve into the architecture of convolutional neural networks as it is of greater interest for the work. Later, we will detail the architectures of the networks that are used. We will talk about the database that we will use and follow the steps to be able to adapt neural networks to mobile devices. In the next chapter, the necessary tests will be carried out to be able to compare the use of neural networks in some devices or others. Specifically, a PC and an emulated mobile device and other personnel will be used. Finally, we will conclude the work and we will see what are the possibilities of improvement in the different future works.

Análisis de estrategias de binarización para la aceleración de redes neuronales profundas (**Analysis of binarization strategies for the acceleration of deep neural netowrks**), Héctor Aparicio Herreros (advisor: Miguel Ángel García), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Sept. 2021.

Abstract: We could take as a fact the increasing importance that the Artificial Intelligence (AI) has acquired in many different fields of our daily lives and the environment that makes it possible. One of the key technologies regarding this growth is the Deep Learning. The development of this technique is based on logic structures similar to the organizational architecture of the nervous system of the human beings, where different layers of the process are specialized in the detection of explicit characteristics of the perceived objects. These organizational architectures are the ones we know as: artificial neural networks. Since its theoretical conception, developed by Warren McCulloch and Walter Pitts in 1943 [1] the neural networks have constantly evolved, what was a great advance for the IA, however, its investigation has suffered diverse periods of reduction of funds and interest due mainly to its high computational cost, the big amount of used memory and the execution times needed for its training. Improving the effectiveness as well as reducing costs and avoid the technical considerations previously mentioned are the main goals in order to optimize these networks, and many strategies have been used with this purpose, such as parallelization or binarization. In order to carry out this analysis, I first studied the status of the art of the Neural Networks, focusing in the convolutional and binarized networks, with the goal of knowing their functioning and objective. Then, I will explain the materials and tools used in order to analyze these models in a deeper way, going through its programming language (Python), framework (PyTorch) and work environment (Google Colaboratory). I will also study the functioning of the Neural Networks: Resnet18, VGG16 and the database Cifar10 and MNIST. The next step consisted on analyzing the models in which binarization was used, contrasting them with those in which it was not, taking as

the basis the following aspects: the classification of the accuracy of the dataset used in each case, the execution time and the losses of evaluation of the model. Finally, I will expose the main conclusions in the wake of the obtained results and the possibilities of the future work that could emerge thanks to this analysis.

Análisis de estrategias de cuantificación para la aceleración de redes neuronales profundas (**Analysis of quantification strategies for the acceleration of deep neural networks**), Jorge Carmona Blanco (advisor: Miguel Ángel García), Trabajo Fin de Grado (Graduate Thesis), Grado en Ingeniería de Tecnologías y Servicios de Telecomunicación, Univ. Autónoma de Madrid, Sept. 2021.

Abstract: Artificial Intelligence is one of the technologies that has grown the most in recent years. It is not a new tool, as it already emerged more than 50 years ago with a simple neural network model, the Perceptron. However, this model soon became obsolete due to the fact that it found certain limitations. After many "Winters of Artificial Intelligence", where this technology stayed stagnant, at present, not only have many of the limitations been overcome, but its development is expected to grow in the coming years and its impact in the economy and in the society begins to emerge. The three reasons for this rapid growth are: the increasing amount of data generated, thanks to the Internet and IoT devices; the great improvement in computing that allows training and using neural networks, which involve a high cost; and the growth in research on this technology. However, although devices have more and more computational resources, neural networks are also increasingly complex, and require higher memory consumption and a slower training and inference process (or that requires a greater computational load). To reduce the size of the network and the time it takes to perform the calculations for each of the layers, there are techniques that allow to optimize and improve the efficiency of the network. That is why, in this project, we will focus on understanding how one of these optimization techniques works, called quantization. To do this, the study begins by conducting an analysis of the state of the art of the basic concepts of Machine Learning and Deep Learning, as well as the quantization technique that we will use later. Subsequently, three pre-trained neural network models (ResNet, ResNext and MobileNetV2) from a Python library (PyTorch) will be implemented in Google's programming environment, called Google Collaboratory. These models will be tested and evaluated with the Imagenette database. Once a model has been launched in Google Colab, the quantization process will be carried out to optimize the network and make it more efficient. There are three quantization methods available in PyTorch: dynamic quantization, post-training static quantization, and quantization aware training. Therefore, the three methods will be studied and implemented on our pretrained networks. And finally, the

behavior of this technique on these models will be studied and the results obtained from quantized neural networks (with the three types of quantification) and unquantified neural networks will be compared. Based on these results, some conclusions will be reached, where the benefits and drawbacks of this optimization technique are collected, and the possible applications and improvements that are intended to be achieved in the future will be specified.

The MobiNetVideo (TEC2017-88169-R) project is supported by

GOBIERNO DE ESPAÑA MINISTERIO DE ECONOMÍA, INDUSTRIA Y COMPETITIVIDAD