D4.2 – New Sensor integration report

Due date of deliverable: 30/11/2021

Responsible partner: SPXL

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

1. OVERVIEW OF THE DELIVERABLE

1.1. Scope

This report describes the explorative activities conducted during T4.2 on a select list of technologies considered strategic to improve the shelf system and the overall user journey in the micro-market.

1.2. Audience

This deliverable is public.

1.3. Summary

The principal areas addressed by this deliverable are:

- Technologies identified, based on the requirements as expressed in D2.1 and micro-market environment in D2.3.
- A set of prototyping and validation activity for each technology with related results.
- A testbed integration with related testing and validation. Based on such result a MVP is produced to be proposed and further validated in the pilots.

1.4. Structure

This deliverable is structured into 4 sections:

- Section 1 provides an executive summary of the document;
- Section 2 introduce the approach and the rationale behind it;
- Section 3-13 discusses the technologies explored;
- Section 14 offers some conclusions.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

2. INTRODUCTION

An integral part of the industrialization process is the exploration of the state of the art in research and industry fields. MIMEX consortium identified room of improvements on different aspects of the customer experience in the micro-market. To better serve the requirements and challenges coming from the deliverables of WP2, in the planning of T4.2 partners agreed to focus the effort on the overall micro-market. Moving the attention from the single shelf to the micro-market allow for a larger degree of solutions and innovations that result in improvements for customers and storeowners.

MIMEX partners explored the technologies within their competences and followed three main phases, that is reflected in the structure of this document:

1. Definition of the technologies and initial benchmark analysis
2. Lab prototyping and validation
3. Testbed integration and creation of an MVP and validation

2.1. Definition of the technologies and initial benchmark analysis

In this phase the selection of technologies happens, picking from the deliverable D2.1, but also D2.3. The rationale for a technology to be explored during T4.2 has been discussed in depth during scheduled meetings and can be summarized in the points following:

- benchmark with similar technologies, if applicable, to identify shortcomings or side effects of a technology in the retail context
- benchmark with similar applications, if applicable, to identify success stories and best practices
- estimated cost of instrumentation, to evaluate if infrastructure cost is sustainable for the experimentation and in general for the MIMEX product
- the added value to end-customer or shop managers
- any risk related to the approach, such as procurement time, availability and third party's dependencies

2.2. Lab prototyping and validation

Prototyping allows partners to acknowledge the technology better, learn how it works and produce a demonstration and a report, based on produced lab data, of the effectiveness of the technology for its usage in the micro-market. Validation may happen as a report describing the pros/cons of one or another approach but also a small physical setup (in the case of hardware artifacts) that can be used and monitored in a controlled environment. The outcome of the activities is collected as a set of KPIs. Thus, the consortium can decide if a scouted technology is eligible to be integrated in the testbed.

2.3. Testbed integration and creation of an MVP and validation

In this phase, mature technologies, validated in the laboratory, from an industrial compliance perspective and providing added value to the micro-market experience, are integrated into the testbed. The integration requires a few mandatory steps such as analysis, component interaction definition, and installation activities. Based on the type of technology, software integration may be needed to hook the new technology in the

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

MIMEX core software. Additionally, further steps related to shop management activities may be introduced, such as changing how a micro-market is managed or how some parts are functioning together.

Once these steps are ready, further validation and testing can be performed in the testbed, running automated integration tests and also performing test shopping sessions and collect data to benchmark and analyse. An assessment is made to demonstrate the value added of the technology and that the overall setup does not impact negatively on the micro-market functionalities from a customer and manager perspective. From the analysis of the testbed, the technologies can be integrated in the testbed and generally be considered an integral part of the micro-market as-a-product offering.

3. RFID

3.1. Introduction

Radio-frequency identification (RFID) uses electromagnetic fields to automatically identify, and track tags attached to shop products. An RFID system consists of a tiny radio transponder, a radio receiver and transmitter. This requires the attaching of passive tags to each product on a shelf with data incorporated into each tag. Each RFID tag can hold several data fields that include product numbers and information about each product/SKU (Stock Keeping Unit). In retail, RFID is primarily used for supply chain and loss prevention applications through the tagging and tracking of the movement of products inside a monitored space. As RFID adoption rate gains momentum, retailers are deploying more advanced RFID solutions to enhance their in-store customer experiences - tracking more complex customer behaviours which can lead to valuable insights into shopping habits. The pairing of SKU interactions with customer data in real-time has many benefits for micro-markets. For example, in a checkout zone, you can record product types as someone exits a store, which when combined with customer data like payment methods can determine whether buyers are returning shoppers. RFID tags can also be exploited to maintain and act on how customers move through a store. Decisions can be made by store owners about how to arrange store layouts and sharpen marketing tactics based on long-term movement analytics.

3.2. Testbed integration and MVP validation

The integration in the testbed requires a specific setup in the checkout area, bars are installed before the exit door and four antennas are connected based on the explorative work of the prototyping phase. A user virtual cart is aligned based on the data coming from RFID, using a process of reconciliation that update the product quantity on a cart item based on an history of events coming from the MIMEX core. Integrating the RFID item coming from the reader allow for a more robust cross-validation of incoming products. Considering however that the readings may come from different sources, it has less priority than other sources (like the smart shelf, product tracking system and direct user input). The testbed setup provides a tag registration procedure, where a batch of tags is collected as a raw list of tag serial numbers (used as unique identifiers). Manually the registered list of tags is mapped to a product category and its reference shelf identifier. Tracking that information allows to map effectively multiple event sources and provide an effective way to generate a reliable cart for the user. When a user exits the micro-market, the list of tags is removed to be ignored by consequent readings. In conclusion, the RFID technology can be considered promising for the micro-market use case, especially thanks to the flexibility in general of the radio frequency technology and the very cheap cost of the tags. The system, thanks to the great number of available "knobs" can be set in different ways to find the best configuration for the particular use case. However, this flexibility also negative aspects: indeed,

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

even if several tests have been carried out, in order to properly configure RFID and make it work in a real scenario additional analysis must be done. The lesson learned from the realized experiments are:

- With a transmission power of at least 0.5W for the RFID antennas, it is highly probable to have nearly 100% of object detection inside the user's bag/cart. Values under 0.5W do not provide a sufficient detection rate.
- Some kinds of objects, e.g., a bottle of shower gel, are more prone to not be detected because of the composition of the liquid inside. In those cases, a possible solution would be to put the RFID tag on the lid of the bottle. Moreover, if the lid is made of plastic or other non-metal materials, the RFID tag considered can be also a non-on-metal tag. In general, this can be considered as a rule for all the bottles that contain liquid.
- The rotation of the antennas toward the exit of the testbed helps in avoiding the detection of the objects in the shelves, but it cannot solve completely the problem. This very simple filtering technique can be used together with an RSSI filter.
- Filtering RFID tags on the shelves using an RSSI threshold is challenging. The experiments showed that a value around –80 dBm with the configuration considered can be effective. Depending on the configuration of the micro-market, several thresholds have to be tested to find the one that allows not detecting objects on the shelves while detecting objects in the users' bags/carts.
- If two users are too close when exiting from the micro-market ((e.g., 1 meter or 2 meters), distinguishing their carts is problematic.

As future work, we will use shielding material to increase the directionality of the antenna to point exactly toward the exit and detect only the objects of the first user. In particular, we considered ferrite absorber materials that have the capacity to absorb the electromagnetic waves, making the antenna more directional. Alternately, the market exit should be reconsidered, e.g., using an exit with an L-shape and a turnstile to allow only one person at a time to exit the market and cross the RFID gate.

4. Light Controller

4.1. Introduction

The design of the Light Controller is based in the functioning principles of the Digital Reader, aiming at a cost reduction in design and manufacturing processes for using similar technological bases. That is the reason why both designs share a lot of characteristics in common. As an example, both use the same type of microcontroller from the manufacturer ST due to the necessity of the Ethernet connection. To achieve the power aspect, in both cases is used the PoE technology, with the minor exception of the Light Controller using a standard class 0 instead of a standard class 6, which allows a greater power, in this case the 45 W minimum required to supply the led lamps. Due to the functioning of the lights being based on On/Off operation, meaning that they do not require to regulate the luminosity or any other parameter, it is only required from it to differentiate between two conditions dictated by the back-end. Therefore, 10 switches had been added to the design to control up to 10 led lamps (there were a total of 9 lamps in each rack).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

With this light controller, not only the proper illumination of the products but also the option to disable specific products in the shelf is achievable, just turning it off. Additionally, it allows to optimize the energy consumption, turning off each space in the shelf that runs out of products, lowering the energy consumed by the system.

4.2. Prototype and validation

- Provide description about prototype and validation, data and graph, photos
- Describe the conclusions and provide reason why we have not moved to next phase, if it is the case

The following block diagram represents all the elements that make up the design which has been proposed to implement the Light Controller. It shows that the entire operation of the system revolves around a CPU based on low-cost Risc V architecture that integrates all the peripherals required to manage the complete operation of the entire system. In the implementation of the design, once the technical and functional objectives have been met, cost and production feasibility have been prioritised over other conditioning factors, such as consumption or size, to name a few. For this reason, in the block diagram definition, a simple solution has been pursued that requires the fewest number of electronic components. The different elements of the block diagram that have been included in the design are described, in general terms, below. To speed up the proof of concept of the proposed technology, some evaluation boards have been used to reduce validation time by eliminating design and manufacturing tasks. The element of the system that creates the most uncertainty from an electronics development point of view is the configuration of the PD, which must meet very precise specifications within the standard so that everything fits together and works smoothly. As a previous experience of the possible contingencies that this part of the development can generate, we can make clear reference to the Digital Reader, where the element that generated the greatest problem in the fine-tuning of the designs was precisely the PoE technology. The evaluation board used to validate the power over Ethernet technology is the TPS2372-3EVM-757 from Texas Instrument, which is fully compatible with class 6 of the IEEE802.3bt standard, allowing up to 51W of power to be collected, which, in principle, would be more than enough to supply energy to the rest of the elements of the design that finally controls the LED lamps. Thanks to this prior validation of the design, all the elements of the system that have already been individually and jointly validated can be integrated into a single electronic design. In addition to working on a design that is known to be successful, it also reduces the commissioning time of the final design by having a prior knowledge base on the operation of the predefined system. Therefore, with all the data and metrics that have been extracted from this experience, it can be concluded that we have the optimum design for managing the operation of the LED lamps that are to be installed on each of the shelves that make up each MIMEX technology rack. In this way, it will be possible to activate or deactivate the lamps installed in the racks individually, being able to leave unlit those shelves where there are no products or which, for any other reason, it is not desired to keep illuminated.

4.3. Testbed integration and MVP validation

The roadmap for the next few weeks is described below, where the light controller is expected to be fully completed and when it is planned to be sent to be incorporated into the technology in the testbed that is being carried out in Trento. The following diagram shows the time evolution of the 4 milestones that have been set to complete the design and put it into operation. This diagram begins in November, where the proof of concept and the results described have been carried out, and evolves over the following weeks until December, when a definitive and integrated design will be available, incorporating all the elements described, ready to be tested in real operating conditions.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

5. AI Edge Accelerators

5.1. Introduction

The use of AI at the edge is driven by different needs:

- Sensitive and personal data that cannot be sent to the cloud for processing to ensure GDPR compliance.
- Latency involved in the cloud communication, processing and response can limit the overall scalability of the solution.
- Overall cost of the solution, the constant usage of cloud resources produces very high bills due to the constant data stream from sensors and cameras.

The “perimetral” context of the business logic is moved to the edge to provide low-latency, faster response time. Only a subset of the data generated by sensors is sent to the cloud after aggregating and filtering the data at the edge. This approach significantly saves the bandwidth and cloud storage costs for MIMEX adopters.

5.2. Prototype and validation

To compare an embed solution like Coral with a GPU solution for inference on object detection task required the setup of the card driver, the supporting libraries and some wrapper code (python based) to setup the environment [1]. We opted for the standard setup, which consume less power against the " maximum operating frequency " which uses the card at most of it’s capabilities with higher energy consumption and increased operational temperature of the device. The object detection model selected to benchmark is the Tensorflow based “Mobilenet SSD v2[2]” trained on the COCO dataset[3]. The basic setup demonstrates a good performance response, with an average of 25-30fps processing on a video stream of 640x480px. The overall performance applied to the retail use case is comparable to a GPU performance.

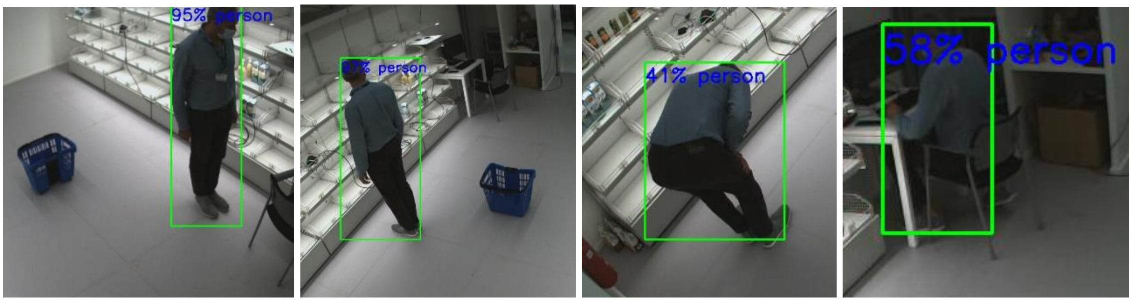

Figure 1 – Example of person detection in the testbed

Considering only the runtime part of the inference is not enough, the capability to train a model in an effective way is key to the adoption of such technology. Luckily considering Coral an integrated product from

---

[1] https://coral.ai/docs/accelerator/get-started/

[2] https://tfhub.dev/tensorflow/ssd_mobilenet_v2/2

[3] https://cocodataset.org/#home

This project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement **No 965486**

Google it leverages on the Tensorflow ecosystem to provide portability of the models between runtime architectures. To ensure a training is efficient and leverage on existing graph knowledge we are going to use a transfer learning approach to instruct the model recognizing specific situations and side cases that are common to our micro-market. Tensorflow Hub is a source of pre-trained model addressing different ML problems, including object detection. We search for the "lite" version[4], a particular format that can be easily converted to a Coral compatible format. The difference from a standard model is that information is packed in a way the Coral board can interpret with its resource constrains.

5.3. Testbed integration and MVP validation

The technology demonstrated to be flexible and cheap but there is no metrics about degradation of hardware overall and would require a dedicated design to use it as module on chip or in an integrated board. Considering the shortage of GPUs that characterized 2020/2021, this technology could be a viable alternative in the setup of the micro-market. The trade-off is that the TPU technology is owned and produced by Google exclusively with no alternative provider on the market. Similarly, AI on GPUs is mostly driven by NVIDIA but horizontal adaptation layer exists and a wide array of competitors (eg. AMD RocM [5]) can limit the risk of selecting a unique hardware technology from a single provider. As a conclusion, the Coral technology provide a rich ecosystem, good portability of trained models between GPUs and TPUs and good acceleration performances at a fraction of the price of standard GPUs. It is a technology to watch but still not mature enough to substitute a GPU in the edge.

6. Fridge Unit

6.1. Introduction

One of the potential functionality or service that a micro-market shop, as MIMEX, can successfully implement is related to the displaying of refrigerated products. Along the solution design and development path (T2.1-System requirements and technological scouting and T2.3-configurable environment), the consortium partners evaluated as a useful and exploitable feature the inclusion of a fridge unit in the MIMEX shop, that can give a benefit to this new kind of retail. Considering that the MIMEX solution will be also portable, for a mobile retail concept (container-like or movable shops), the fridge unit could represent a strong opportunity for the success of the whole solution, in terms of attraction for retailers and above all for potential users of cold beverages, fresh meal for fast lunch, that can appreciate the application of this format very much. Therefore, although the use of this technology was not foreseen within the project (in fact the MIMEX project foresees the inclusion of fresh products but not specifically products that need to be stored at a cold temperature), the consortium partners decided to deepen the possible integration of the weighing shelf system (already developed for the MIMEX racks) into a refrigerated unit. Indeed, Cefla has purchased a refrigerated unit on which a specific smart shelf has been studied and is under development and then it will be tested within the Trento testbed for the final integration in a MIMEX shop. Any other refrigerated unit for the others Pilot will be on charge of whom design, set up and install the MIMEX shop (T5.2 and 5.3). Although

[4] https://tfhub.dev/s?deployment-format=lite

[5] https://www.amd.com/en/graphics/servers-solutions-rocm-ml

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

the technology of the refrigerated sections is known, a series of tests and prototypes is becoming necessary in order to have the same reliability obtained in the racks currently installed in Trento.

6.2. Prototype and validation

After a scouting of reliable suppliers of refrigerated units, Cefla choose OSCARTIELLE spa based in Bergamo (Italy).

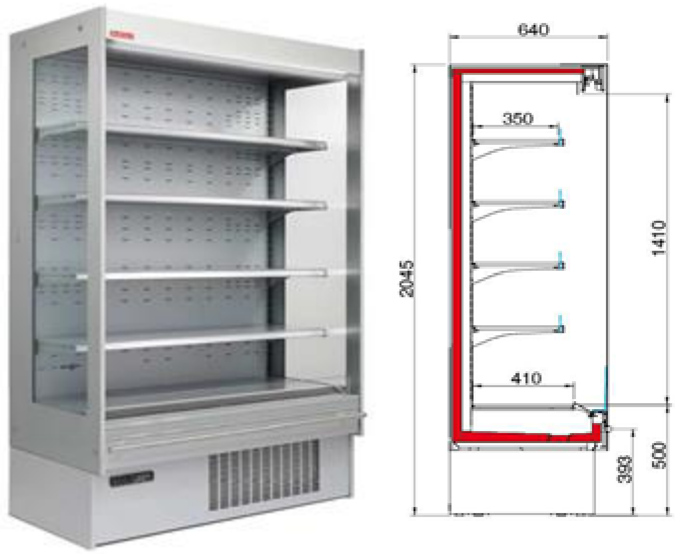

Figure 2 - View of the refrigerated unit

In figure 6-B a refrigerated unit hosted in a supermarket is represented and it is similar to the one that will be used in MIMEX test, but the one chosen for MIMEX is without the door in order to avoid misreading errors due to the vibrations in case of opening or closing or issues in the object recognition by the camera. The device is structurally different from the rack developed for the weighing shelf system actually in an ongoing testing phase in the Trento testbed. Basically, the rack is composed by three columns, every shelf level is composed by two shelves composed each one by a lower frame and an upper frame, joined by the load cell. Due to the components included inside, the refrigerated unit instead is composed by two columns, every shelf level is composed by two shelves but only one shared lower frame and the two upper frame. It's mandatory to ensure that the two shelves positioned on the same lower frame are able to avoid misreading error. Further, the positioning of the camera for the object detection has to be defined, as soon as the refrigerated unit will arrive in CEFLA.

6.3. Testbed integration and MVP validation

If the tests and the final assembly will give a positive outcome, a refrigerated unit is expected to be tested in Trento approximately within February '22. This activity follows its own development plan, independent of the deadlines of the deliverables expected for the project. An indicative roadmap for the development and the implementation of the refrigerated unit is in figure 6-C.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

7. ULTRA-WIDE BAND

7.1. Introduction

UWB is a radio technology that can use low energy levels for short-range, high-bandwidth communications over a large portion of the radio spectrum. UWB has traditional applications in non-cooperative radar imaging and has been exploited for data collection, precision locating, and tracking applications in the retail space. FBK has developed a device to measure the proximity of people and/or objects, called Janus (see picture below). It provides feedback signals (e.g., a blinking LED or vibration) when two people or objects are within a pre-defined range. Janus exploits standard BLE for detecting proximity and UWB to provide precise distance measurements with accuracy of tens of cm. Janus could be exploited to identify the proximity of an individual to an object in the store, e.g., shelf, door, etc, assuming the shelves are equipped with UWB anchors.

- Advantages:
  - *Positional accuracy*: 5 to 10cm accuracy;
  - *Range*: 100m line of sight, indoor 25m;
  - Availability: already present on high end Samsung and Apple smartphones, likely arriving on more phones in the near future;
  - Information capacity: none, only distance;
  - *High configurable*: in our configuration we obtain a data every 2s in case two devices are in range.
- Disadvantages:
  - *Cost*: 80 euro per device, anchors (on the shelf) and beacons (users).
  - *Usage*: Must attach the device to the user, by putting it in a pocket, on a lanyard, or on a shopping bag/cart. In the future, when UWB is available on more smartphones, the user's own device may be used;
  - *Dimensions*: Large ~15x10x5cm

7.2. Prototype and validation

To understand the possible advantage of using UWB inside the MIMEX experience, we performed two different tests. In the following, we describe each of them and the results obtained.

7.3. Testbed integration and MVP validation

The technology has worked well inside the testbed and its well-known accuracy has been observed in all the experiments. The possible integration of UWB inside the micro-market would be the installation of the devices on the shelves, for instance considering putting one device every two shelves. This system would allow recognizing whether a user is in front of a specific rack with very high accuracy. However, this implementation assumes that the user's smartphone is equipped with a UWB chip, which unfortunately is still not very common as only the last flagship smartphones have UWB inside. The alternative would consist of providing users a device at the entrance of the micro-market and collecting it at the exit. Clearly, this solution is impracticable for a real implementation of the micro-market. Moreover, the high cost of the devices (80 euro each) makes this technology very expensive. In conclusion, even if UWB is a promising technology it is still not ready to be implemented in this project.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

8. NEURAL NETWORK BASED PEOPLE TRACKING

8.1. Introduction

Object Tracking is a topic of Computer Vision which involves the process of tracking an object across a sequence of frames. If the object is a person usually it is referred to as Person Tracking. It has a special name because of (i) the specific interest in the task, (ii) the use of particularly developed detection modes and (iii) the characteristic articulated shape with respect to generic objects (iv) face detection can be used. People tracking task is strongly related to People Detection: it starts with all possible detections in a frame and gives them an ID. In subsequent frames the tracker has to carry forward a person's ID. If the person has moved away from the field of view then that ID is dropped, while if a new person appears then it starts off with a new ID. This is a challenging task since two people close to each other could appear similar, therefore causing the model to switch IDs, people may get hidden behind someone else or may disappear and reappear in later frames.

8.2. 2D object tracking

Early works [6][7] formulate instances association as a graph-based optimization problem following the "tracking-by-detection" paradigm, where a node represents a detection and an edge encodes the probability that two nodes are connected. The combination of visual and motion cues to represent each node often involves expensive calculations. The so-called simultaneous detection and tracking methods alleviate complex association strategies to link detected boxes through time. One of the first proposed methods using deep learning for single object tracking is GOTURN[8]. A model is trained on a dataset consisting of videos with labelled target frames. The objective of the model is to simply track a given object from the given image crop. To achieve this, they use a two-frame CNN architecture which uses both the current and the previous frame to accurately regress on to the object. Another class of object trackers uses Long Short Term Memory (LSTM) networks along with convolutional neural networks for the task of visual object tracking. LSTM networks are suitable for visual object tracking as they perform well at learning historical patterns. One of these methods is called Recurrent YOLO, or ROLO[9]. The method proposes to concatenate high-level visual features produced by convolutional networks with region information, specifically the authors use YOLO to collect rich and robust visual features, as well as preliminary location inferences; and them they use LSTM in the next stage as it is spatially deep and appropriate for sequence processing.

---

[6] Jerome Berclaz, Francois Fleuret, Engin Turetken, and Pascal Fua. Multiple object tracking using k-shortest paths optimization.TPAMI, 33(9):1806–1819, 2011

[7] Li Zhang, Yuan Li, and Ramakant Nevatia. Global data association for multi-object tracking using network flows. CVPR, 2008

[8] David Held, Sebastian Thrun, Silvio Savarese, "Learning to Track at 100 FPS with Deep Regression Networks", ECCV 2016 - CODE https://github.com/davheld/GOTURN

[9] G. Ning et al., "Spatially supervised recurrent convolutional neural networks for visual object tracking", IEEE International Symposium on Circuits and Systems (ISCAS), 2017, pp. 1-4, doi: 10.1109/ISCAS.2017.8050867

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

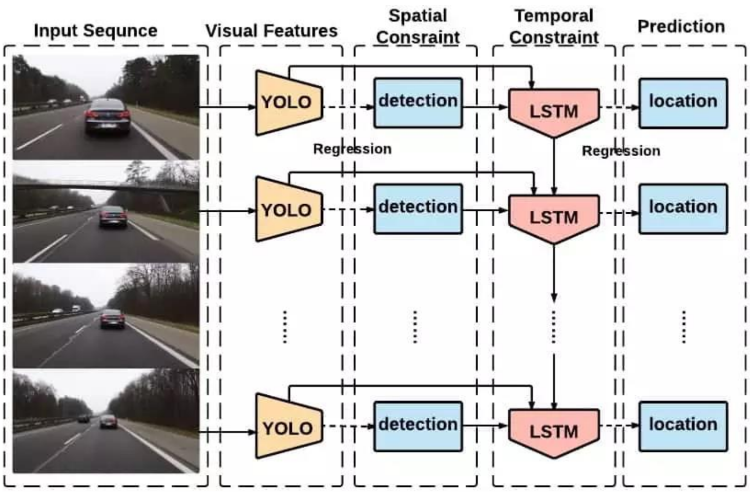

Figure 3 – LSTM use to localize detected object

In contrast, an emerging keyword of the same period in the tracking task is SORT (Simple Online and Realtime Tracking) which is a pragmatic approach to multiple objects tracking with a focus on simple and effective algorithms to associate objects for real-time applications[10] . The paper presents a simple tracking framework that focuses on frame-to-frame prediction and association. The authors showed that the tracking quality is highly dependent on detection performance and by capitalising on modern developments in detection, state-of-the-art tracking quality is achieved with only traditional tracking methods, such as the Kalman Filter and Hungarian algorithm, leading to best performance with respect to both speed and accuracy. SORT is proposed as a baseline for subsequent algorithms. The drawback with approaches such as Optical Flow, Kalman filtering, Kanade-Lucas-Tomasi (KLT) feature tracker, mean shift tracking is that they can't predict the abrupt motion and direction changes, which is sometimes the case of people. An accurate tracker needs to consider the appearance of the objects to be tracked in order to focus on re-identification to connect detections or to handle long term occlusion. In "Simple online and realtime tracking with a deep association metric" [11] is presented DeepSORT, an extension of SORT that incorporates appearance information through a pre-trained association metric: it uses the learned information to track the people's trajectory until they leave the camera frame. More recent state-of-the-art trackers [12] [13] learn a deep network to predict the displacement (motion) of instances based on both visual and geometric features, significantly outperforming the simpler SORT. The consideration of SORT, that a better motion model is the key to improve the local linking accuracy, suggests that particular care must be taken when choosing the detector, which must be state-of-the-art. As Deep Learning has led to higher performance on detection with respect to traditional techniques, one of these methods should be used. YOLO v3, for example, is one of them, Faster R-CNN another.

---

[10] A. Bewley, Z. Ge, L. Ott, F. Ramos and B. Upcroft, "Simple online and realtime tracking," 2016 IEEE International Conference on Image Processing (ICIP), 2016, pp. 3464-3468, doi: CODE https://github.com/ab- ewley/sort

[11] N. Wojke, A. Bewley and D. Paulus, "Simple online and realtime tracking with a deep association metric", IEEE International Conference on Image Processing (ICIP), 2017, pp. 3645-3649, doi: 10.1109/ICIP.2017.8296962 - CODE https://github.com/nwojke/deep_sort

[12] Xingyi Zhou, Vladlen Koltun, and Philipp Krahenb. "Tracking objects as points". ECCV, 2020

[13] Philipp Bergmann, Tim Meinhardt, and Laura Leal-Taixe. "Tracking without bells and whistles". In ICCV, 2019

Very recently in "FairMOT: On the Fairness of Detection and Re-identification in Multiple Object Tracking" [14] is proposed the following architecture, which is divided into Object Detection Branch and Re-Identification Embedding Branch.

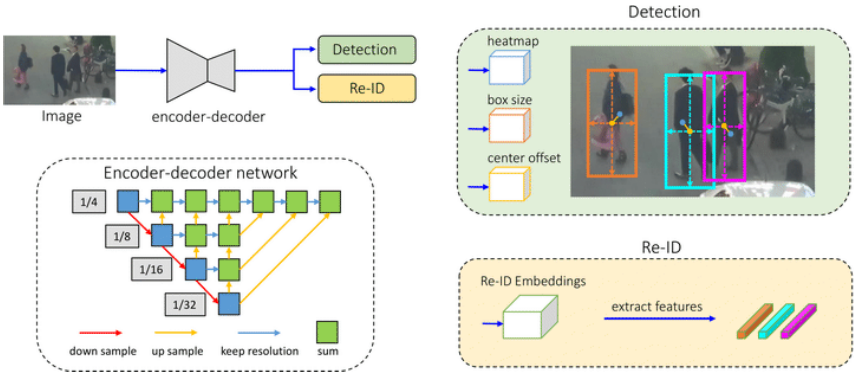

Figure 4 – Relation of object detection and embeddings in Re-Identification

The Object Detection Branch considers the heat map, the bounding boxes center offsets, and their size. The Identity Embedding Branch is responsible for generating the vector representation of the image patch that corresponds to the predicted bounding box, encoded into a 128-dimensional vector. From the first frame are predicted boxes and their corresponding Re-ID vectors. Then the Re-IDs and bounding boxes for the subsequent frame are also generated and all the Re-IDs are compared by means of a similarity function: if the similarity is high, they can be labeled as the same person from the previous frame. The network is trained with multiple losses, three for object detection and one for re-identification. CenterTrack 15 performs object detection and tracking simultaneously using center points. Each object is represented by a single point at the center of its bounding box identified by a CenterNet detector. It applies the detection model to a pair of images and detections from the prior frame in order to localize objects and predict their associations with the previous frame. SiamMOT16 is a region-based Siamese Multi-Object Tracking network that detects and associates object instances simultaneously. It includes a motion model that estimates the instance's movement between two frames such that detected instances are associated. It combines the region-based detection network Faster-RCNN with an implicit motion model (IMM) and an explicit motion model (EMM). Differently from CenterTrack, "Tracking objects as points " 17 implicitly infers the motion of instances with point-based features SiamMOT uses region-based features and develops template matching to estimate instance motion.

[14] Yifu Zhang, Chunyu Wang, Xinggang Wang, Wenjun Zeng, Wenyu Liu, "FairMOT: On the Fairness of Detection and Re-identification in Multiple Object Tracking", International Journal of Computer Vision, 129, 3069–3087 (2021) https://doi.org/10.1007/s11263-021-01513-4 https://arxiv.org/abs/2004.01888 - CODE https://github.com/ifzhang/FairMOT and https://paperswithcode.com/method/fairmot

[15] Xingyi Zhou, Vladlen Koltun, and Philipp Krähenbühl. Tracking objects as points. ECCV 2020 https://arxiv.org/abs/2004.01177 - CODE https://github.com/xingyizhou/CenterTrack

[16] Bing Shuai, Andrew Berneshawi, Xinyu Li, Davide Modolo, Joseph Tighe, "SiamMOT: Siamese Multi-Object Tracking", CVPR 2021 https://arxiv.org/abs/2105.11595 - CODE amazon-research/siam-mot: SiamMOT: Siamese Multi-Object Tracking and https://pythonrepo.com/repo/amazon-research-siam-mot-python-deep-learning

[17] Xingyi Zhou, Vladlen Koltun, and Philipp Krähenbühl. Tracking objects as points. ECCV 2020

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

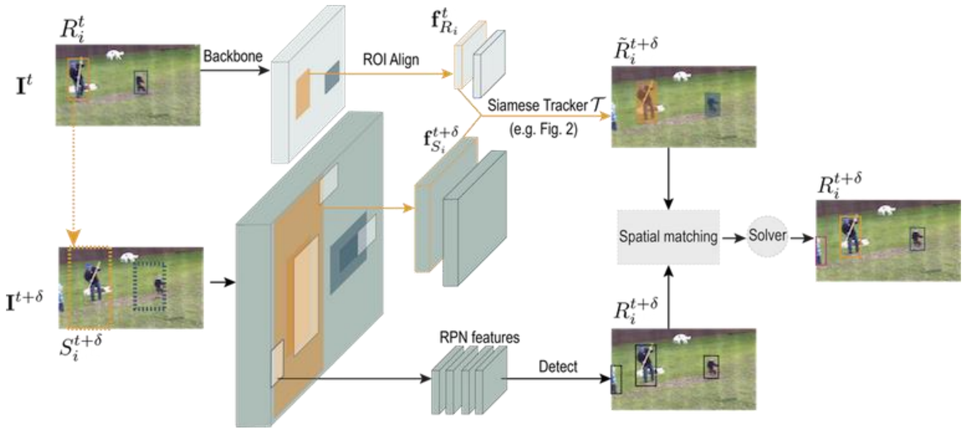

Figure 5 – CenterTrack architecture

8.3. 3D object detection and tracking

While people detection in single 2D images has improved considerably with the deep learning advent, relatively little progress has been made in multi-camera multi-people tracking algorithms. 3D trackers replace the object detection component in standard tracking systems with 3D detection from monocular images[18] or 3Dpoint clouds[19]. Tracking then uses an identity association model, for example, 3DT[20] detects 2D bounding boxes, estimates 3D motion, and uses depth and order cues for matching, while AB3D[21] combines a Kalman filter with accurate 3D detections (PointRCNN). The most challenging part of multi-view detection is gathering occupancy knowledge about people from multiple views. Before the surge of deep neural networks, modelling the correspondence across cameras was mostly done by probabilistic modelling of objects. Working in a multi-view environment the power of Deep Learning should be combined with occlusion reasoning instead of with 2d people detection in single images and putting them into correspondence to achieve 3D localization. This approach is introduced in "Deep Occlusion Reasoning for Multi-Camera Multi-Target Detection" [22] where is presented an architecture that combines Convolutional Neural Nets and Conditional Random Fields for the multi-camera multi-people tracking scenario. The model is end-to-end and the robustness is achieved thanks to high-order Conditional Random Field terms that model potential occlusions. More specifically, they reason on a discretized ground plane in which detections are represented by boolean variables. The CRF is defined as a sum of high-order terms whose values are computed by measuring the discrepancy between the predictions of a generative model that takes into account occlusions and those of a CNN that can infer that certain image patches look like specific body parts. The method outputs probabilities of presence on the ground plane for each temporal frame which are linked

[18] Ren, J., Chen, X., Liu, J., Sun, W., Pang, J., Yan, Q., Tai, Y.W., Xu, L.: "Accurate single stage detector using recurrent rolling convolution", CVPR 2017

[19] Shi, S., Wang, X., Li, H.: "PointRCNN: 3D object proposal generation and detection from point cloud", CVPR 2019

[20] Hu, H.N., Cai, Q.Z., Wang, D., Lin, J., Sun, M., Krhenbhl, P., Darrell, T., Yu, F., "Joint monocular 3D vehicle detection and tracking", ICCV 2019

[21] Weng, X., Kitani, K.: "A baseline for 3d multi-object tracking", arXiv:1907.03961 (2019)

[22] Pierre Baqué, François Fleuret, and Pascal Fua. "Deep Occlusion Reasoning for Multi-Camera Multi-Target Detection", International Conference on Computer Vision, pages 271–279, 2017 https://arxiv.org/abs/1704.05775 - CODE: https://github.com/pierrebaque/DeepOcclusion

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

into full trajectories using a known flow-based approach. ” Bringing Generalization to Deep Multi-view Detection” [23] aims to predict a bird's eye view occupancy map of people from multiple camera views.

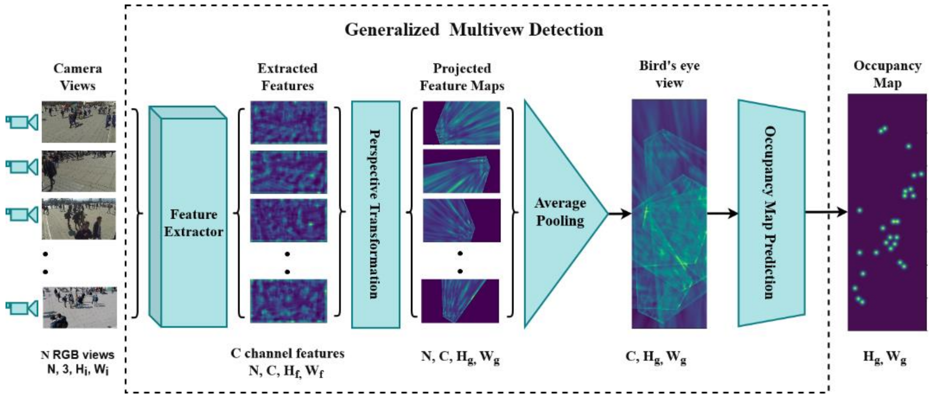

Figure 6 – Multicamera 3D tracker architecture

9. Door Control

9.1. Introduction

Access management is a well-known challenge in any access-restricted area and requires fine grained access control and constant monitoring to avoid uncontrolled or malicious activities inside a protected facility. The micro-market faces the same challenges, being at the same time a commercial activity opens to anybody and a context where access must be limited to address monitoring activities and management duties. The door access system requires two key features: to provide instant feedback of the status of a door (open or closed) and to be controlled remotely based on the business logic as defined by the requirements. In that perspective the right tool is an electronic relay system equipped also with digital input circuits. To ease the integration, it should also expose an API (Application Program Interface) for remote control served over the network.

Figure 42
Advantech controller

The market provides different solutions but not all met the minimum criteria of industrial strength, ease of communication and interaction. The product selected for testing has been the “WISE-4060/LAN 4DI/4Relay Modbus Ethernet I/O” from Advantech. The product exposes a RESTful web API for integration. It offers an HTML5 web interface to configure the module without the limitation of a platform or operation system. It has a wide degree of operating temperatures and includes diagnostics functionalities. It mounts 4 relay and 4 digital inputs that are singularly controllable.

[23] Bringing Generalization to Deep Multi-view Detection (2021) https://arxiv.org/abs/2109.12227] [CODE: https://github.com/jeetv/gmvd

This project has received funding from the European Union’s Horizon 2020 research and innovation program under grant agreement **No 965486**

9.2. Prototype and validation

As first setup a set of LEDs light and Arduino based output has been prepared to simulate the communication scenarios for opening a door (by turning on and off the lights) and sending signals on the input circuit to notify of status changes (sent intermittently by the Arduino wire up).

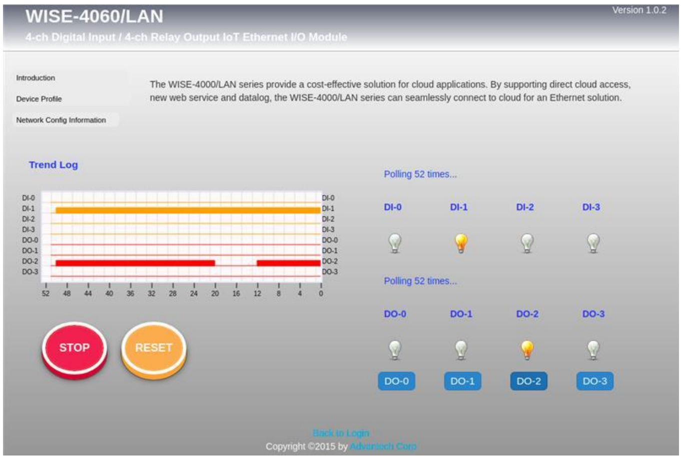

Figure 7 - WISE user interface

To validate the solution, a REST client has been configured to interact with the relay interface and monitor/interact with the connected components

9.3. Testbed integration and MVP validation

The module has been integrated to the testbed, connecting an electric door lock to the relay and the lock open/closed sensor to the input circuit. The overall integration has run smooth and with a single WISE device both the entrance and exit door can be monitored.

10. Temperature Covid19 TOTEM - Temperature Monitor, Hand Sanitizer and Mask Checking

10.1. Introduction

Device developed specifically for the MIMEX project to combine essentially two functionalities: channel to communicate with the user before entering the Micro-market and a guarantee for a minimum safety health compliance. It works as a door keeper and a granting access point for the customers after complying with the requirements related to the body temperature, mask recognition and hands sanitized. One of its characteristics is the portability so it is easily removed once it is not needed anymore, mainly because of the recovering situation of the Covid-19 crisis, but still useful to fight contagious diseases with similar characteristics like flu or common colds. The totem device works in a simple way: when it receives an external signal (QR code validation), it checks if it meets the requirements for entry: user with a facemask and body temperature below 37ºC. If it is correct, the user can use the hydroalcoholic gel dispenser. An external signal is sent back to the cloud. The latter, depending on the case, will display a different message: "entrance

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

allowed/denied" or "wait, too many people" and it will operate accordingly. Therefore, the totem is a co-creation between Métrica6 and Spindox. Métrica6 is in charge of the mechanical and electronic development of the totem device and Spindox developed for this device to work in the MIMEX ecosystem -cloud: backend, entrance Totem, door's controller and customer app. The totem is the result of the integration of the next devices:

- Proface X [TI]
- Smart Angel Gel Dispenser

The Proface X [TI] is a product developed by the company ZKTeco to work on the control access through computer vision technology while reducing the risk of infection and germs spreading during the recent global public health issue as well. The enables fast and accurate body temperature measurement and masked individual identification during facial and palm verification. Includes the next features:

- Ultra-large capacity of facial templates for 1:N verification: 30.000 (standard); max. 50.000 (optional).
- Huge capacity of palm templates; 1:N – 5.000 palm templates.
- Anti-spoofing algorithm against print attack (laser, colour and B/W photos), videos attack and 3D mask attack.
- Intelligent energy-saving design; precisely evaluation of the distance (up to 2.5m (8.2 ft)) between the user and device by a microwave detector before waking up the recognition terminal.
- 2MP starlight CMOS sensor camera with WDR function.
- 8" touch screen with 400 lux, which offers high visibility under strong and direct light.
- Wide range of working temperature (16 ºC ~ 35ºC; 60.8 ºF ~95ºF).
- 0.1s High Speed Temperature Detection, measurement distance of 30 - 120 cm.
- Temperature Measurement Accuracy: ±0.3 ºC (±0.6 ºF) with the distance of 80cm and 25ºC environment by the laboratory tested;
- Mask detection; facial verification available with masks.

Figure 8 - Proface X [TI] Device

Along with the purchase of the device, the manufacturer provides the software development kit, what allows the team to adapt the device to the necessities of the project in terms of communication and improve the usability of the features offered by the device.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

The Smart Angel Gel Dispenser is a product developed by Metrica6. Following the work plan to fight contact contagious diseases, specially Covid19, this advanced dispenser combines the easiness of use and portability while offering a sensation of safety where it is placed. Its use is compatible indoors and outdoors without risks of malfunctioning. In figure 46 is possible to see the customized version of the dispenser for the MIMEX project where the dimensions can also be appreciated (130x50x15 aprox.).

Figure 9 - Gel dispenser before the integration with the Proface X [TI]

10.2. Prototype and validation

For a valid functional prototype, we need the whole structure and the measuring device properly assembled to work with the stablished dimensions in order to assure the right operability towards the user (dimensions showed in Figure 57).

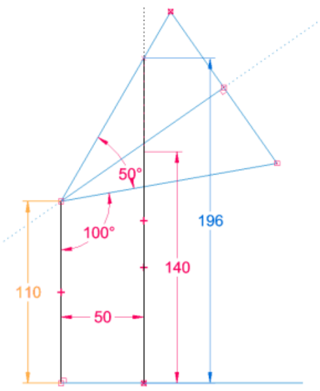

Figure 57 - Functioning dimensions of the Covid19 Totem.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

- Temperature measuring test and mask detection: ZKTeco, the manufacturer of the Proface X [TI], ensures a correct temperature and mask detection in the range between 30-120 cm long. It can detect, with a ± 0.3 ºC tolerance, the temperature in a range comprehended between 20-50 ºC. The test will start operating the device while pointing to 5 different people individually one after another. The device was tested under different environments to provide with enough results to give a reliable outcome of the tests. Summarized these are the most remarkable aspects to consider when operating the device. The device Proface X [TI], as a technology to act on the mask recognition and the temperature measurement, works as needed and expected. Only under extreme circumstances has showed difficulties to perform its task, but this can be easily avoided because of portability of the device (even after its integration with the rest of the technology).
- Hydroalcoholic Gel Density Test: It is important that the gel is kept in a controlled operating range, as this is because a variation in the gel, pointed out by a change in the density, has a negative impact in the performance or the quality of the product. Although the device has been tested with water, density 1g/ml, it is advised that the pump should always work with lower densities, such as that of the hydroalcoholic gel shown in the table below. This is because a higher density of the fluid would result in the obstruction of the system. To start with the study about the internal mechanism of the system, we stick and work with specific values, as shown in the following table (Table 4).
- Mechanical Test: The hydraulic pump is powerful enough to impulse the hydroalcoholic gel with the mentioned density. As a result, we reduce the working time of the pump to push the gel through the dispensation hole. If the solution is present in the pipe at all moment, the actioning time is reduced and the dispensation time is lower.
- IR sensor distance: The device counts with an infrared sensor in the front part where the gel is dispensed, so it easily recognizes when an object enters the programmed zone. This means, the sensor operates when a body interrupts its read range. To improve the accuracy of the activation, a test was performed to avoid accidental actionings, when it is not meant to use the device. If the distance at the end of the sensor is about 40 cm, it will be placed at 35 cm, which is the distance needed by the user to introduce the hand under the tap.
- Connectivity of the totem: To set ready the totem, the communications between the dispenser and the Proface X [TI] were studied. Through the relay, the Proface sends an order to the dispenser to "be ready" to perform the dispensation, being this the last action to perform on the user to ensure the healthy requirements (along with an optional final message on the screen). After this, a confirmation signal is sent by the totem to the MIMEX system to be used to grant/deny the access. Once the user passes the mask recognition and the temperature measurement, the IR sensor of the dispenser starts working, being ready to dispense the hydroalcoholic solution once the user put the hands under the dispensation hole. After the dispensation is performed, the IR sensor goes back to its "stand by" status.

10.3. Testbed integration and MVP validation

As a conclusion for this technology, the integration of the MIMEX device and ecosystem-cloud totem into TRENTO testbed cannot yet be justified until all phases of the roadmap described above (including the

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**

integration tests themselves in TRENTO testbed) have been completed. Integration will be validated if all tests are successful. It should be noted that, once the project is completed, and due to the evolution of the global pandemic in which we find ourselves, it is possible that the device will not have the same relevance and needs as at the beginning of the project. To this end, its functions can be enabled and disabled according to the needs required in each circumstance: user verification, access control, facemask recognition, body temperature, gel placement, etc.

11. CONCLUSIONS

This document describes the experiments and explorative activities on technologies supporting and improving the initial assumptions and functional stack for the micro-market. The overall task experience has provided the opportunity to demonstrate the feasibility of different technologies in an industrial context, providing evidence of strongness and weakness of the different approaches. The micro-market improves from the result of this task by adding functional components that solve concrete problems such as addressing improving the reliability of virtual carts, sanitary compliance in the COVID-19, strict and integrated access control over entrance and exit doors. Other activities open to upcoming technological trend such as UWB positioning which is a technology that started appearing in the most recent consumer phones. Energy and power reduced AI inference is already possible but limited on some extents due to limited availability of mature technologies and framework support. In depth analysis and benchmark of market and research approaches allow a better understanding of the technological panorama and provide support during the phases of analysis and technology selection. Based on the test and development of T4.2 we can use a data driven approach to evaluate the challenges and opportunities of mentioned technologies and better understand impacts on testbed and pilot adoption.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement **No 965486**