NEXT GENERATION DATA INCUBATOR

EXPLORE PHASE
TECHNICAL SPECIFICATIONS

11/05/2023

CORE PARTNERS

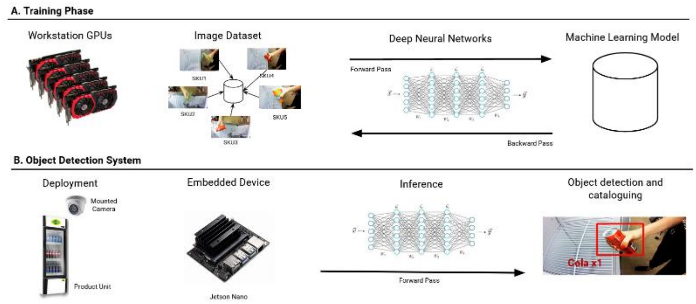
DATA PROVIDERS

This project has received funding from the European Union's H2020 research and innovation programme under Grant Agreement no 951981

1 Technical Specification Double-side Page

1. **TECHNICAL SCOPE:** The mock-up solution is suitable and correctly addresses the challenge/theme selected over the REACH dataset/s. The Big Data solution architecture proposed is adequate to tackle the data management issues associated to the solution in mind. "To what extent does the applications handle the data provided?"

ROSETTA will deliver an Edge AI system consisting of a Jetson Nano™ and a fixed camera module that can be installed on any product housing unit with the task of automatically detecting any product pickup using the camera feed alone. To this end, we will develop a complete, end-to-end trainable neural network that can perform object detection in real time. The network will be trained on GPU-powered desktop PCs and the resulting model will be able to run on a Jetson Nano™ device. The proposed system can be broken down to two distinct parts:



The **training subsystem** will use the image data provided by Migros to train an Artificial Neural Network (ANN) on desktop workstation GPUs. The network will be specifically designed to have enough capacity (trainable parameters) that it can accurately perform the task of object detection while running efficiently on a low-power low-compute device, like the Jetson Nano. The network will be fully parameterizable and fast to retrain to be able to support various data sources.

The **inference subsystem** will use the trained model for object detection on the Jetson Nano™ given images from the camera. It will operate solely on the resources provided by the device and will not require additional external resources. A stream of images will come from an installed camera on a product housing unit (e.g. refrigerator). The stream will be securely transferred to the device and fed to the neural network for object detection. Only information regarding the picked-up object will be extracted. The camera will only need to capture images of the customer's hands and item. This significantly minimizes concerns pertaining to customer privacy. Additionally, captured images will not have to leave the closed system that is formed between the camera and the Jetson Nano™. Further requirements concerning image streaming rate, communication protocols between camera and Jetson Nano™, additional product information will be finalized during the early stages of the design.

2. **SELECTION OF ALGORITHMS AND TOOLS:** The indicated Data Science approach, i.e. algorithms chosen, and Big Data architecture approach, i.e. tools chosen may successfully accomplish the required data governance, processing and analysis. A clear understanding of the used REACH dataset/s is demonstrated.

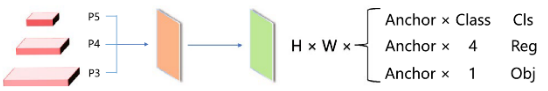
The dataset we will use will come from Migros, containing images of objects being grabbed from a housing unit. Our solution will be based on deep neural networks. We will adopt a Convolutional Neural Network architecture as they have shown great performance in image-based tasks. They are also fast to train and can be very compact for execution in constrained environments like the Jetson Nano™. They also do not require any advanced feature engineering as they can work directly in image space. Important identifying features of each object class are implicitly learned by the network during training. The network will be trained on a workstation using discrete GPUs. On the same workstation, the provided dataset will be curated, labelled, transformed and stored in a format ready for consumption by the training module. The software will be developed in Python using the PyTorch framework for tensor operations and OpenCV for image manipulation. PyTorch is also a great choice for its compatibility with the Jetson Nano™ SDKs. The network will be evaluated in terms of task performance and tuned with tools like Ray Tune. Python code will be statically analysed using Pylint.

At the edge (e.g. on a refrigerator unit), a camera will interface with a Jetson Nano™ to feed images in a controlled, predefined rate. The device will forward the images through the pre-trained model and detect any present objects. After a number of detections of the same object within the stream of images, the system will be able to accurately report a specific object grab from inside the unit. Additional interfacing between the camera and the embedded device will be done through native APIs and SDKs. Camera images will be securely streamed through direct wired connection or wirelessly. All information will be captured and processed inside the system of the camera and the Jetson Nano™.

This project has received funding from the European Union's H2020 research and innovation programme under Grant Agreement no 951981

3. **TECHNICAL SCALABILITY AND FLEXIBILITY OF THE SOLUTION:** The solution can truly cope with humongous and increasing datasets, potentially from diverse data providers, and is flexible it to adapt to other related domains.

Scalability and flexibility have been at the core of our design during the EXPLORE phase. The goal is to be adaptable to new challenges that will arise during the EXPERIMENT phase as well as diverse data providers. To this end we have designed a neural network architecture that can be trained end-to-end on desktop GPUs with reasonable resources and time. As shown in the Figure, the network will employ an architecture that simultaneously trains for the tasks of classification, regression and objectness which will allow it to adapt to diverse data providers beyond the challenge. The capacity of the network will also be chosen depending on the target hardware as well as the available training resources (time, compute). This will allow us to build AI models that train under specific constraints and are tailored to the target hardware capabilities and task performance characteristics. This is a similar approach to other state-of-the-art models that deploy versions of varying parameter capacity to support a range of performance requirements and hardware specifications.



4. **DATA GOVERNANCE AND LEGAL COMPLIANCE:** Data sharing challenges, data governance and legal compliance, must be observed. The proposed solution is compliant with the current data legislations concerning security and privacy (e.g. GDPR).

Access to the data for training will only be given through specialized secure channels dictated by the server infrastructure and agreed with the Data Provider. The REACH Big Data infrastructure will be our first choice during training. Any future labelling or augmentation of the data that might be required, will only use identifying information and features of the object that it contains. All other information will be treated as background and no personal or private information will be exposed. No identifying information of the customer will be required or stored for later use, thus fully complying with EU GDPR.

The requested system is approached as an Edge AI solution thus all information will be captured and processed inside the closed system of the camera and the Jetson Nano™ without external communication with a server. Regarding data acquisition (images of customers grabbing a product) the module will be able to operate as-is in a fully secure, offline fashion. Images will not require any customer identifying information therefore ensuring the privacy of the customers and complying with EU GDPR. Only the grabbed object will need to be (partially) visible. As part of our module, image data is transient. Images will be processed in real-time and only information of the detected objects will have to be kept in memory.

5. **QUALITY ASSURANCE AND RISK MANAGEMENT:** Feasible and credible quality process followed for the final product generation. The potential risks in all the phases of the project (design of the solution, development, testing, deployment…) are identified and convincing mitigation plans put in place.

The work will be carried out iteratively, aiming at producing an early MVP, and continuously improving it based on feedback through regular meetings with the Data Provider. Software development will follow Agile methodologies. Code quality will be validated with tools like, Farma-C, Pylint. Git version control will be used for issue tracking and milestones.

Our team is composed of highly skilled professionals with a strong research background. The proposed approach benefits from advancements in AI. However, during the EXPLORE phase, we identified the following risks and mitigation strategies:

Inadequate training data: neural network training and model performance is very dependent on the data used for training. There is also the case of missing and underrepresented occurrences within the dataset. As a mitigation plan, the quality of the data will be discussed with the Data Provider and if required we will incorporate open datasets.

Object occlusions: partially or fully occluded objects will pose a challenge to the object detection system as their identifiable features might be occluded. Partial occlusion will rarely be an issue for a deep learning approach. Full occlusion will be treated as special cases and reported as such.

Device performance: there is a risk that the Jetson Nano™ will not have enough processing power to properly utilize the trained model. This is an additional reason why our proposed architecture was designed to adapt to existing requirements.

This project has received funding from the European Union's H2020 research and innovation programme under Grant Agreement no 951981

REACH

NEXT GENERATION DATA INCUBATOR

This project has received funding from the European Union's H2020 research and innovation programme under Grant Agreement no 951981

CORE PARTNERS

DATA PROVIDERS