RottenCAM:
Freshtastic - Redefining Produce Inspection

Morabito Mattia, Pallotto Francesco

June 26, 2023

1 Motivations

In a world where the need for sustainable practices is becoming more pressing by the day, AI can be a turning point in their growing pursuit. In particular, by looking at both global and local trends we can see how natural resources aren't correctly exploited, leading to the production of large amounts of preventable waste.

RottenCAM is an AI model designed to address this challenge, in particular by helping correctly classify produce freshness. This will hopefully empower users to make more conscious decisions not only for their wallets but also for their health and more sustainable future.

However, this application does not only target the general public but also the whole supply chain in order to minimize waste and streamline processes.

All of this will be framed over the on-device learning approach which naturally shares some of these concerns such as energy efficiency and reliability while keeping data private.

From an individual's standpoint, having a device with full offline functionality brings practical advantages that go beyond the limitations imposed by bandwidth constraints in day-to-day life.

Offline functionality eliminates the reliance on external network infrastructure, thereby removing the impact of bandwidth limitations. In situations where the available bandwidth is insufficient, an offline-capable device remains fully functional, allowing users to perform critical tasks without disruption. This becomes particularly valuable in remote areas, where access to reliable internet connections may be scarce, or during periods of high network congestion when bandwidth is strained.

It is not just about dealing with situations where the internet connection may be slow or unreliable, but also about ensuring long battery life, an essential feature of a dependable device. Offline functionality directly contributes to this. By relying on offline processing, the device minimizes the need for constant communication with external servers or networks. This significantly reduces the energy consumption required for network-related tasks, with significant benefits to battery life. Without the drain of data transmission and reception, the RottenCAM can dedicate its resources to the tasks at hand, optimizing performance and preserving battery power for extended periods.

This text will provide an in-depth exploration of the RottenCAM application, highlighting its key features and functionalities while also discussing any ethical considerations and commercial uses .

2 Problem definition

The problem at hand is to create an image classifier such that:

$$\underline{\hat{y}} = f(\underline{\underline{X}}, \underline{\hat{\theta}}) \tag{1}$$

where $\underline{\hat{y}}$ is an $N_{class} \times 1$ output vector of probabilities that the $W \times H \times C$ input image $\underline{\underline{X}}$ belongs to a class, given the estimated model parameters $\underline{\hat{\theta}}$ and the model f. This model can output $N_{class} = 2n$ classes, where a single species of produce n_i can either be fresh or rotten.

> Assumption: Each fruit can be unequivocally classified as either fresh or rotten.

The challenge is taking an input image with dimension $W \times H \times C$ for which even if we assume that each pixel is a `byte` for a full resolution image in `QQVGA` we get a memory occupation of:

$$W \times H \times C = 160 \times 120 \times 3 = 57.60\, kB \tag{2}$$

Given this figure we wanted to reduce the image size to:

$$W \times H \times C = 96 \times 96 \times 3 = 27.65\, kB \tag{3}$$

This is more manageable considering the memory constraint on our board (*Arduino Nano 33 BLE*) is $256\, kB$ where all parts of our system must fit.

3 Technical aspects

We will now provide an overview of the technical aspects of RottenCAM starting from how the model was built.

3.1 Model

For this application we developed a scalable model based on simple Separable Convolution Blocks. Each one of these blocks contains:

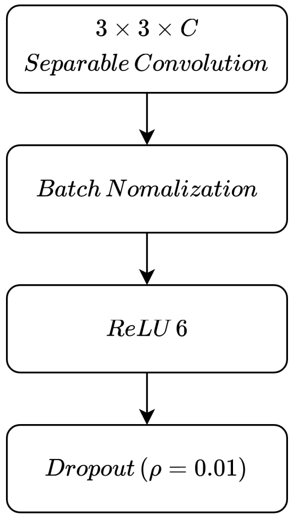


Figure 1: Custom Separable Convolution Block

Where C is the number of channels. $ReLU6$ was used to help the model converge to a quantizable level for the activations, while the dropout helps with overfitting. Each block has $strides = 2$ to scale down the feature maps.

The first layer is always a standard 2D convolution [1], while at the end of the last block we have a global average pooling and a final softmax fully connected layer.

This model also allows for 6 customizable parameters:

- n: the number of output classes;
- r: the resolution where the format is $W \times H$;
- α: the width multiplier, a float which applies a linear reduction on the number of channels of each filter (i.e $C = \alpha \cdot C_i$);
- β: the depth multiplier, an integer that determines how many blocks are present in the model;
- g: the channel gain, a float that specifies

how much the channels are increasing with each subsequent layer;

- *strides*: the stride for the initial 2D Convolution block.

The model naming reflects all the parameters as shown in the figure below.

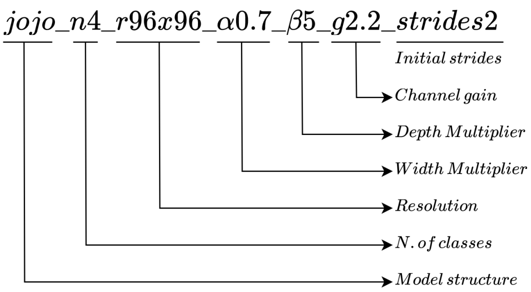


Figure 2: Model naming convention

At each layer (excluding the first one) the number of channels increases as:

$$C_i = \alpha \cdot C_{start} \cdot (i \cdot g) \tag{4}$$

Where i is the layer with $i \in [1, \beta]$, and C_{start} is the initial number of channels (fixed as 16 but configurable). This scaling mechanism is a way to taper the feature map size to optimize the memory demand of the model, in terms of both the number of parameters and the feature map size.

The final model has the structure shown in the next figure:

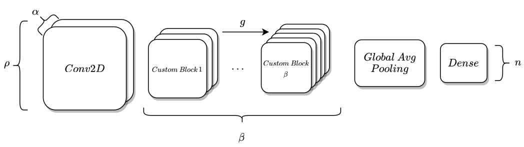


Figure 3: Model structure

Our final choice of model is the same as shown in the picture. We will report the theoretical calculations of the model size and MACS:

jojo_n4_r96x96_a0.7_b5_g2.2_strides2	
Number of parameters	29605
Max Consecutive Activations (kB)	39.744
MACs (M)	1.38
Number of layers	23

Table 1: Model Information

Moreover, by plotting the Number of Channels and the Feature Maps Resolution we can see the tapered effect:

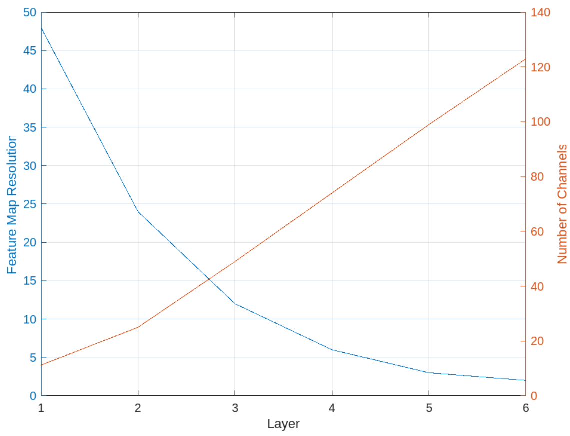


Figure 4: Number of Channels and the Feature Maps Resolution vs Layer

We can also see how the number of parameters, the feature map memory and the Number of MAC operations evolve over the layers:

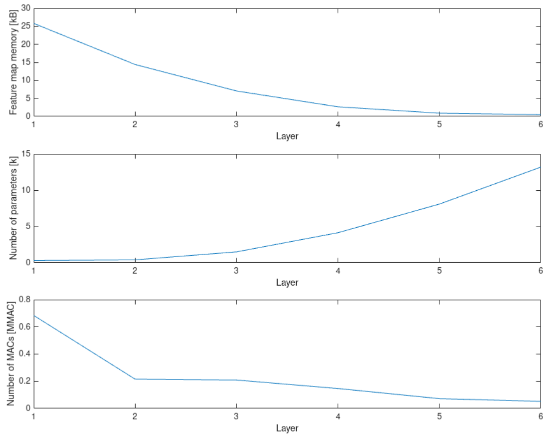


Figure 5: In order from top to bottom: Feature map memory (kB), Number of Parameters (k), Number of MAC operations (M)

3.2 Deployment

The model was then quantized to `int8` with both input and output layer set to `int8`. The `tflite` file was then uploaded to Edge Impulse which gave us an estimate for the memory usage and processing time:

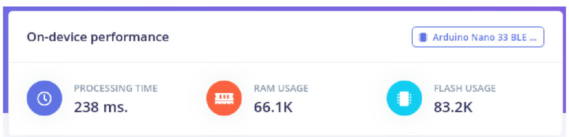


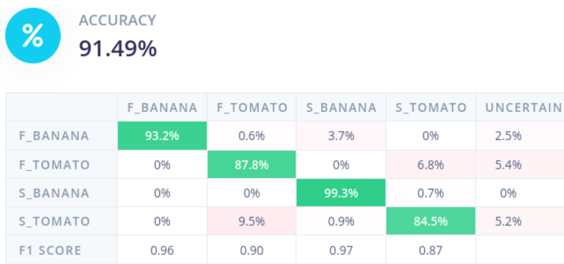
Figure 6: Edge Impulse profiling report

The computation time was then checked on the device which gave us:

- Image acquisition time: $T_{IMG} \approx 2500ms$;
- Inference time: $T_{INF} \approx 250ms$.

The bottleneck is the image acquisition time, which may be optimized by low-level optimizations of the camera library, which are outside the scope of this course.

Before showing the confusion matrix we need to specify that we have 4 classes to show the potential of the model beyond binary classification. However, in an industrial application, we do not need to worry about this sort of confusion under the assumption that each fruit/vegetable is sorted separately. The final accuracy metrics (extracted from edge impulse with confidence threshold $\theta = 0.6$) are shown below:


ACCURACY
91.49%

	F_BANANA	F_TOMATO	S_BANANA	S_TOMATO	UNCERTAIN
F_BANANA	93.2%	0.6%	3.7%	0%	2.5%
F_TOMATO	0%	87.8%	0%	6.8%	5.4%
S_BANANA	0%	0%	99.3%	0.7%	0%
S_TOMATO	0%	9.5%	0.9%	84.5%	5.2%
F1 SCORE	0.96	0.90	0.97	0.87	

Figure 7: Edge Impulse testing report (True labels along the y-axis, Expected labels along the x-axis)

While the false negatives do not represent an issue (lower diagonals of the confusion matrix), since we are not throwing away good produce, the false positives are an issue and must be minimized. In particular, in our case we would throw away fresh produce incorrectly identified as stale.

The overall average false positive rate is:

$$FPR = \frac{FP}{FP + TN} = \frac{16}{16 + 279} \approx 5,4\% \tag{5}$$

3.3 Dataset

The dataset is made up of images of two types of produce that serve as an example of a wider range RottenCAM could work with: bananas and tomatoes. It was compiled from multiple sources to enhance the generalization capabilities and robustness of the model. Namely:

Source 1 (≈ 80%): Spoiled and fresh fruit inspection dataset [2]

Source 2 (≈ 10%): Personal photos in various supermarkets, (Milan, Messina, Macerata, Genova)

Source 3 (≈ 10%): Image scraping under Creative Common License.

Figure 8: Locations where the personal images were acquired

All images were then resized and augmented via the online tool *Roboflow* with the following settings:

- 50% probability of horizontal flip and 50% probability of vertical flip to cope with various camera orientations;
- Random rotation of between $-10°$ and $+10°$ and random shear of between $-8°$ to $+8°$ horizontally and $-8°$ to $+8°$ vertically to cope with various viewing angles;
- Random brightness adjustment of between -15% and $+15\%$ percent to compensate for different exposure times;
- Random Gaussian blur of between 0 and 0.25 pixels to account for slightly out-of-focus images.

All this yielded a 3x increase in the size of the dataset.

	Vanilla	**Augmented**
Training Set	4130	12390
Validation Set	814	814
Testing Set	564	564

Table 2: Dataset size

The following is an example of the augmented images, in which the rotation and blur are particularly evident:

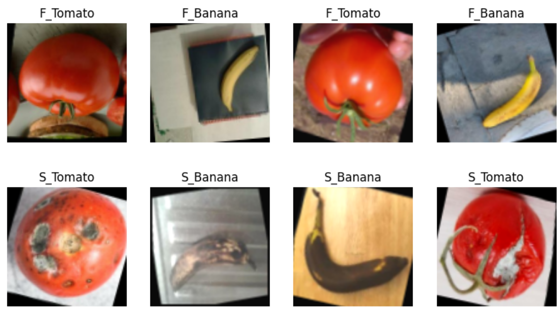


Figure 9: Examples of augmented images with their respective labels

The final Training-Validation-Testing split of our dataset and class distribution are:

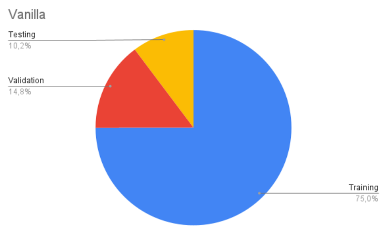


Figure 10: Vanilla dataset composition

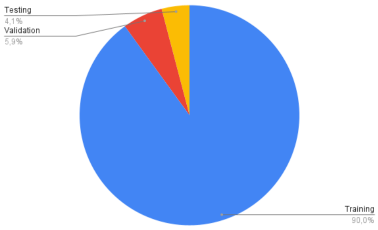


Figure 11: Dataset percentage for each class

As we can see, the dataset is unbalanced toward bananas, due to the availability of the fruit samples.

4 Ethics

4.1 Job Displacement

Work holds significance beyond providing income, contributing to a sense of dignity, purpose, and sense of accomplishment. Human replacement in the workplace on behalf of AI raises intricate ethical implications which require careful examination. Job displacement can have significant consequences for individuals and communities, including financial insecurity and difficulties in finding alternative employment. Ensuring the well-being and support of the affected workforce is a primary concern in managing the ethical implications of AI-driven job displacement.

While aknowledging the societal and psychological impacts that the loss of employment due to AI-driven job displacement can have, it is our belief that Artificial Intelligence can also create exciting new opportunities and transform industries. With careful consideration and collaboration between stakeholders, proactive policies and measures can be implemented, aiming to manage the highlited ethical ramifications and to ensure a just and inclusive transition to new work models.

4.2 Data Policies

Any new device capable of handling potentially sensitive or private information must be designed with the utmost care and attention to detail. As RottenCAM is intended to operate in a wide range of scenarios, its functionality is tailored to the diverse needs and contexts in which it may be deployed, with a strong focus on usability and privacy.

The latter is an area where on-device processing becomes crucial. While pictures of produce may not seem like an obvious privacy concern, mishandling of such images could have potentially harmful consequences.

Firstly, these pictures could capture more than just the targeted fruit or vegetable; they could inadvertently include personal information like faces or valuable belongings, especially since domestic usage is one of the primary applications of the RottenCAM. Another, less dramatic but still significant threat avoided by offline data processing is the potential misuse of a consumer's purchasing habits. If this information fell into the wrong hands, users could face issues like data profiling, data brokerage, and invasive targeted marketing. The problem also applies to the workplace. The presence of cameras could be perceived as unwelcome surveillance by workers of whichever sector, and the knowledge that any captured images would not end up in the wrong hands could provide piece of mind and trust towards employers. Moreover, given the spread of misinformation, we want to avoid the possibility of a malicious use of pictures of rotten produce to harm competitors.

4.3 Food Safety and waste

While other issues may inspire some ethical problems, talking about increasing food safety and reducing food waste isn't controversial. This topic is also included in the Sustainable development goals under target 12.3:

> By 2030, halve per capita global food waste at the retail and consumer levels and reduce food losses along production and supply chains, including post-harvest losses. [3]

However, given the fact we have a non-negligible FPR, we must take into account that some fresh vegetables may be thrown away as rotten. We are confident that expanding the available dataset and implementing adaptive learning will help reduce this impact significantly. In particular given the scalable design of the model we can assume that a bigger implementation may perform better.

Figure 12: UN infographic on food waste [5]

While also falling in the policies of reuse, reduce, recycle [4].

5 Market

This project aims to provide a pervasive market presence, ranging from personal use to industrial application. We will discuss each instance separately.

5.1 Private use

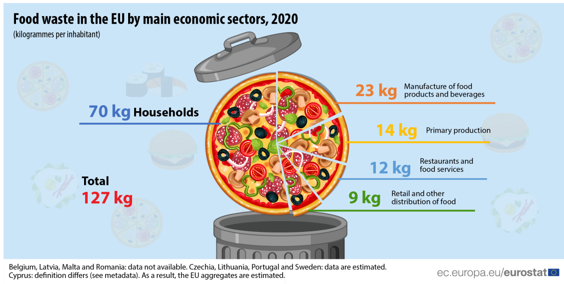


Figure 13: Food waste in the EU by main economic sectors [6]

In the context of private individuals, the importance of food in our daily lives cannot be overstated. Ensuring access to fresh and nutritious food is crucial for maintaining a healthy diet, although it can often be a challenging endeavor during the shopping process.

In this regard, the integration of AI algorithms can play a pivotal role. By providing users with a tool that possesses the ability to accurately assess the freshness of produce, we aim to accomplish the following objectives:

5.1.1 Reducing Food Waste

Based on reliable data collected by reputable agencies mentioned in [6] and [7], it has been observed that food waste amounts to an average of approximately 127 kilograms per capita annually. While it is acknowledged that not all of this waste is specifically attributed to produce, countries such as Italy exhibit a higher prevalence of waste concerning fresh fruits and vegetables [7].

The objective of our tool is to give customers the capability to accurately identify deteriorating produce. This application will serve as a guide for consumers, enabling them to make more informed purchasing decisions and promoting safer consumption practices. It is worth noting that a significant proportion of food waste, exceeding 50%, originates from private households [6], making the mitigation of this issue an urgent priority.

Potential applications for this model also include the implementation of smart fridges equipped with cameras capable of detecting signs of decay and providing timely alerts to users. In this context, the device can be configured to provide a pre-established rate of classifications per day, without incessantly taking pictures, which would be unnecessarily power consuming and could easily be perceived as a suspicious behavior.

Additionally, smart shopping assistants can be developed to provide on-demand classification of food items based on their freshness, thereby mitigating food safety risks. Furthermore, these assistants can offer users valuable information regarding the shelf life of different produce, aiding in reducing food waste and enhancing overall efficiency in the management of food resources.

5.1.2 Providing Savings

The reduction of food waste significantly influences shopping savings. Overbuying, which is a major contributing factor, can be mitigated by adopting conscious practices concerning portion sizes and purchasing only necessary quantities of food. By avoiding the purchase of excessive amounts, individuals can minimize waste and generate tangible savings by refraining from discarding unused food for which they have already incurred costs.

Nonetheless, achieving effective food waste reduction requires proper execution, as the task can otherwise become time-consuming and inefficient. In this context, our model aims to eliminate uncertainties associated with assessing produce freshness, offering reliable aid to individuals in making informed decisions regarding the quality and suitability of produce.

5.1.3 Improving Food Safety

The freshness of produce plays a pivotal role in ensuring food safety due to several key factors.

Firstly, fresh produce reduces the likelihood of harboring harmful microorganisms, including bacteria, viruses, and parasites. As fruits and vegetable age, their natural defenses against microbial contamination progressively weaken. This creates a favorable environment for the proliferation of pathogens, potentially leading to foodborne illnesses.

Secondly, fresh produce typically exhibits an extended shelf life before experiencing degradation and spoilage. During this period, the risk of microbial growth and contamination by foodborne pathogens remains relatively low.

Overall, the freshness of produce profoundly impacts food safety. Its role encompasses reducing the likelihood of microbial growth, extending shelf life while lowering risks of contamination, and ensuring that only high-quality produce reaches consumers. By prioritizing produce freshness, the consumption of compromised or unsafe food can be minimized, achieving health benefits and promoting a culture of food safety.

5.2 Retail

5.2.1 Quality control

The capabilities of the device make it a valuable tool with applications extending far beyond the individual household or single buyer. RottenCAM demonstrates potential for various commercial uses, including the retail sector. Its ability to accurately recognize the freshness or deterioration of produce positions it as a powerful instrument for quality control.

Through the classification of images captured by its camera module, RottenCAM could be a valuable asset in the assessment and selection of produce. One notable advantage is the potential reduction of human bias in quality control. The device's objectivity and consistency could significantly diminish the inherent subjectivity associated with manual inspections. By trusting the reliable identification of visual characteristics of fresh and stale produce, the quality control process can be streamlined, resulting in improved efficiency and ease of operation.

In fact, according to *AgShift*, a California-based startup, their platform utilizing computer vision algorithms to evaluate the quality of agricultural commodities has the potential to reduce the time required for quality control by up to 60% [8]. This reduction in time can have a positive impact on overall productivity and operational efficiency within the retail industry.

5.2.2 Supplier control

The device could grant retailers a competitive edge over suppliers, by equipping them with expedited quality control capabilities, thereby enabling greater control over the supply chain.

An efficient assessment of the quality of received produce could allow for the identification of subpar products in incoming shipments. Consequently, this would facilitate the establishment of supplier accountability for the quality of their deliveries, discouraging the shipment of low-quality produce. Simultaneously, it ensures the maintenance of consistent quality levels and adherence to a retailer's standards.

In this application, constant or periodic operation of the device would be unnecessary, and it could be configured to classify on an on-demand basis, thus sparing power consumption.

It is important to note that these advantages would not compromise retailer-supplier relationships. The device's improved objectivity in quality assessments would allow for constructive and evidence-based feedback to suppliers, which translates into enhanced performance evaluations and negotiations. Additionally, this objectivity would contribute to the reduction of quality disputes, as assessments and opinions could be substantiated by objective data.

Furthermore, the device would empower retailers to optimize their supply chain operations. By utilizing its assessments, identifying suppliers that consistently provide superior service would be easy, enabling retailers to prioritize sourcing decisions accordingly. This would result in an efficient and reliable supply chain, benefiting both the retailers and their customers.

Finally, if the device's data were diligently and ethically logged, it could offer valuable insights into supplier performance. Retailers could leverage this data to identify trends, patterns, and potential areas for improvement, further strengthening their position concerning suppliers.

5.2.3 Inventory management & Customer satisfaction

Thanks to an enhanced ability to identify stale or low-quality produce, retailers could gain insights into the demand for specific types of produce, enabling them to take immediate action such as requesting replacements from suppliers or adjusting order quantities. This would ensure a consistent supply of fresh produce while minimizing the presence of unwanted items in stock.

Consequently, this would result in a notable reduction of food waste within the inventory. The accurate detection of stale or spoiled produce would enable retailers to swiftly respond by implementing measures such as immediate sale or donation, thus minimizing the quantity of wasted food and associated costs. An example of how this could work is the replacement of traditional FIFO (First-In-First-out) practices with ones that consider the lifetime of the products in stock, minimizing the likelihood of unsold items and potential losses due to spoilage. This has the added benefit of catering to a consumer's needs and well-being.

Encountering subpar or decaying items seriously impacts the shopping experience, discouraging clients from returning. However, proactive rotation or removal of items nearing expiration ensures that customers are always presented with the freshest options, enhancing customer satisfaction and reducing the likelihood of selling stale or spoiled products.

By utilizing data collected from the device's assessments, retailers can gain insights into the popularity and demand for specific types of produce: a product that is often found to be in subpar conditions is likely to be in lower demand than one that is always fresh. This information can be used to enhance the accuracy of demand forecasting, enabling retailers to adjust their inventory levels based on customer preferences and market trends. Furthermore, this process can be automated to optimize the replenishment of fresh stock with minimal manual intervention, ensuring efficient and timely restocking of inventory.

In general, maintaining high-quality standards directly translates into improved customer satisfaction. Specifically, visibly fresh items are the best received and play a key role in shaping a retailer's reputation:

> The degree of consumer satisfaction in five European countries France, Denmark, Finland, Portugal, and Switzerland found that product quality is the most important attribute of store image [9].

As customers value the transparency and trust that comes from knowing the freshness of the products they purchase, the improvements to quality control provided by RottenCAM could help build customer loyalty.

In the event of issues with the quality of a purchased item, the device's assessments could provide dissatisfied customers with valuable evidence to support their claims. This would allow retailers to address complaints promptly and take appropriate actions such as refunds, exchanges, or providing replacements, all the while building customer satisfaction and trust.

5.3 Industry

5.3.1 Food sorting and grading process

In the agricultural industry, after harvest, produce must be sorted and inspected for defects. A complex system of conveyor belts usually does this, however, the process of visual inspection is still carried out by humans.

While human supervision is still critical in this kind of process, many solutions have been proposed to integrate AI in industries such as airport security [10]. Given this interest, we believe that the workers and industry can benefit from AI assistance.

In this scope our application will aid the industry in converting to a semi-automatic system, improving processing times and making the process more reliable. We believe that a device-centered approach will provide faster response and classification times while preventing possible data leaks from the factory. Clearly, in this scenario the device's classifications would be needed at a fast and constant rate, making performance an important aspect.

Improving the accuracy of the process may also yield a reduced carbon footprint since it removes non complying produce before transportation and storage. Using tiny devices for this application will streamline the maintenance by replacing heavy machinery which is difficult to replace and require specific technical knowledge. Moreover, on-device learning can compensate for the possible seasonability of produce, adapting over time to provide more accurate predictions.

Such algorithms can be expanded to also encompass the food grading process where produce is sorted based on some visual characteristics. This application however is just a possible development and will need extensive cooperation between companies to build a cohesive dataset.

The implementation of rigorous quality control measures throughout the entire process of harvesting, packing, and distributing fresh produce significantly contributes to maintaining food safety. These measures involve meticulous inspection, sorting, and removal of damaged or spoiled items. Ensuring that only fresh and visually acceptable produce reaches consumers minimizes the likelihood of encountering food safety risks associated with deteriorated or contaminated items.

5.3.2 Federated Learning

This application may also be fitted to allow for a federated learning architecture where a common classification model is shared between companies which then update it, therefore contributing to higher quality produce reaching the retail sector. This in turn will improve customer satisfaction across the whole industrial chain.

In this context, we can imagine that a single company might have multiple conveyor belts for different kinds of produce. The proposed model can be simplified to classify a single species of produce and distributed across the floor plan.

The implementation is an example of Horizontal FL, where the sample space is different across species of products while the features that we can extract from an image are the same (i.e. mold, indentation, etc.). This allows for a centralized server to combine different models to enhance the predictions.

> Note on labeled samples: A worker provides the labeled information at deployment time to initialize the model. This helps with drift concerning the environment while later on the adaptiveness of the application will take care of the seasonability of the produce.

References

[1] Andrew G. Howard and Menglong Zhu and Bo Chen and Dmitry Kalenichenko and Weijun Wang and Tobias Weyand and Marco Andreetto and Hartwig Adam (2017), *MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications*, Addison-Wesley Professional.

[2] Pachon Suescun, Cesar Giovany; Pinzón Arenas, Javier Orlando; Jiménez-Moreno, Robinson (2020), "Spoiled and fresh fruit inspection dataset", Mendeley Data, V1, doi: 10.17632/6ps7gtp2wg.1

[3] https://www.fao.org/sustainable-development-goals/indicators/1231/en/

[4] https://www.epa.gov/recycle

[5] https://sdgs.un.org/goals/goal12

[6] https://ec.europa.eu/eurostat/en/web/products-eurostat-news/-/ddn-20220925-2

[7] https://www.ilsole24ore.com/art/o-spreco-cibo-cala-italia-12percento-ma-pattumiera-finiscono-9-miliardi-l-anno-AELeR2fC?refresh_ce=1

[8] https://www.agshift.com/hydra

[9] Azhar, A. R., Salehuddin, M. Z. Mohd, Saiful Bakhtiar, M. F. B., Kamaruddin, M. S. Y. B. (2013). Customer satisfaction with hypermarket fresh food's characteristics. Faculty of Hotel & Tourism Management, Universiti Teknologi MARA, Malaysia.

[10] Vukadinovic, Danijela & Anderson, David. (2022). X-ray Baggage Screening and Artificial Intelligence (AI). 10.2760/46363.