Tiny Machine Learning for Classifying Specialty Coffees

Isabela V. de Carvalho Motta, Felipe A. P. de Figueiredo, and Samuel B. Mafra

***Abstract*— The consumption of specialty coffee has increased around the world. Specialty coffee is free from impurities and defects. Specialty coffee is produced in smaller quantities, and its production process is hard and expensive. Traditional coffee beans have defects that affect the flavor of the coffee. It is essential to select the type of coffee to guarantee the best cost and quality. When a human manually classifies the type of coffee, there may be interference due to the human condition. This process has the disadvantage of being subjective. Few studies have used machine learning methods to predict specialty coffee classification by analyzing images. This article proposes a framework for classifying specialty coffees by applying Tiny Machine Learning techniques. We developed a model that can help accurately analyze and classify the coffee process and inform the quality of coffee without human interference. We achieved 100% accuracy, and our model can be used effectively in the coffee industry.**

***Keywords*— IoT, TinyML, Machine Learning, Computer Vision, Precision Agriculture, Coffee.**

TABLE I: Arabica Coffee Classification

Groups	Subgroups	Aroma and Flavor
Specialty Coffee	Strictly soft	Extremely smooth, and sweet.
Specialty Coffee	Soft	Pleasant, smooth and sweet.
Specialty Coffee	Just Soft	Slightly sweet, and smooth, without astringency.
Specialty Coffee	Hard	Astringent, and harsh, without strange flavors.
Traditional Coffee	"Riada"	Light typical iodorfomic flavor.
Traditional Coffee	"Rio"	Typical and sharp iodoform flavor.
Traditional Coffee	"Rio Zona"	Very strong aroma and flavor of iodoform or phenol.

I. Introduction

The importance of coffee classification is linked to the fact that coffee is one of the most popular and appreciated drinks worldwide, with an industry that generates billions of dollars annually. According to the Brazilian Coffee Industry Association (ABIC), it is estimated that coffee industry sales in 2022 reached 100 billion dollars, an increase of 54.6% compared to 2021. Also, according to ABIC, Brazilian coffee exports increased by 4% between 2021 and 2022. More than 34,000 bags of specialty Arabica coffee were exported in 2022 from Brazil [1], [2]. According to the Brazilian Ministry of Agriculture and Livestock, the price of Arabica coffee increased by 88% compared to 2020 [3].

Coffee classification plays a fundamental role in determining quality and market value and identifying the sensory characteristics of the drink [4]. Over the years, coffee classification has evolved, influenced by factors such as the growing demand for specialty coffees, the search for more objective evaluation methods, and the need for standardization of international trade.

Coffee classification is based on visual and physical criteria, such as the beans' size, shape, and color [5]. The industry has developed different classification systems and methods to meet its needs, establish quality parameters, and facilitate communication between the partners.

In Brazil, one of the main coffee-producing countries, the classification system known as sieving was implemented [6]. In this system, the grains are separated using sieves of different sizes, and a visual analysis is performed to identify possible defects.

Defects and impurities in coffee beans reduce their quality. Defective beans affect the coffee flavor, devaluing it and impacting the producer's profitability. Several factors can cause these defects, such as problems in production and storage and physiological and genetic changes. Identifying defects makes it possible to improve management and prevent the occurrence of these defects that depreciate the coffee.

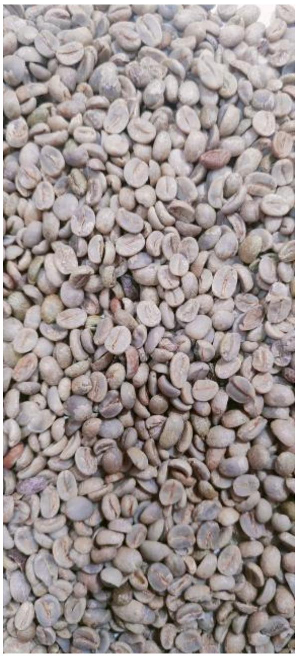
According to Normative Instruction Number 8 of the Brazilian Ministry of Agriculture and Livestock, the criteria for classifying the type of coffee, specialty or traditional, are the species, the shape of the bean, the granulometry, the aroma and flavor, the drink, the color, and the quality [7]. Table I indicates the classification criteria according to the aroma and flavor of the coffee.

This study considers the criteria of bean shape, particle defects, impurities, size, and color to classify coffee into Specialty and Traditional categories.

Specialty coffees go through a more laborious production process, require special care from planting to roasting, and are made of pure, unmixed, high-quality coffee beans. On the other hand, traditional coffee has inferior and defective beans. This mixture reduces the quality of the coffee as it changes its flavor and reduces its cost. The difference between specialty and traditional coffee can be seen in Figures 1a and 1b.

Artificial intelligence (AI) has the potential to revolutionize the coffee industry, offering efficient solutions to various challenges. AI has been widely researched and has proven important and efficient in people's daily lives. In the coffee

Isabela V. de Carvalho Motta, Felipe A. P. de Figueiredo, and Samuel B. Mafra are with the Instituto Nacional de Telecomunicações, Santa Rita do Sapucaí-MG, e-mails: isabela.motta@dtel.inatel.br; felipe.figueiredo@inatel.br; samuelbmafra@inatel.br.



(a) Specialty Coffee

(b) Traditional Coffee

Fig. 1: Examples of coffee classes.

industry, AI can help farmers in many aspects [8], from optimizing production to improving quality control.

One of the uses of AI that has helped in several areas is the classification of images [9]. Artificial Neural Networks (ANNs) are based on the structure of the biological brain [10]. ANNs work through several layers and connections that process information throughout it, extracting features from the input data. Each layer refines the representation as the data propagates through the network, enabling the neural network to learn complex hierarchical features and relationships, ultimately leading to meaningful predictions. ANNs exhibit characteristics that have attracted the attention of researchers, such as the ability to map input features to classes, a task known as classification.

Image classification is done through ANNs that are trained to find patterns. This pattern-finding process requires several images to be labeled into classes so that ANNs can learn through examples.

The main contribution of this work is the development of a specialty coffee classifier using the Edge Impulse platform [11], which can be embedded in microcontrollers and smartphones using Tiny Machine Learning (TinyML) techniques [12][1]. This classifier aims to leverage AI, particularly ANNs, for image classification in the context of coffee beans. By applying ML, the research addresses the important task of distinguishing between specialty and traditional coffee based on various criteria such as bean shape, particle defects, impurities, size, and color. This approach holds significance in the coffee industry, contributing to quality assessment and market value determination. The use of Edge Impulse and TinyML suggests a focus on deploying efficient and lightweight models on resource-constrained devices for practical applications in coffee classification.

[1]https://studio.edgeimpulse.com/public/315018/latest

II. Related Works

Few studies in the literature classify specialty coffee beans. The authors of [13] classify green coffee beans in different regions of Brazil as specialty or traditional. Coffee samples were collected from different regions of Brazil, and a multispectral camera captured images of the beans at different wavelengths. Four machine learning models were tested, and the Support Vector Machine (SVM) model performed the best, with an accuracy of 96%.

Most research aimed at classifying grain quality is focused on classifying defects. The authors of [14] developed a system for detecting defects in coffee beans. Firstly, the system classifies the coffee beans into sour, black, broken, moldy, shell, insect-damaged, and good beans. The system works like a mechanical sieve sprayed into an outlet bin, and the bean is identified as defective. The screening accuracy reached more than 90%.

In [15], the authors used the slim Convolutional Neural network (CNN) approach to identify coffee quality by classifying defective beans. The study used a dataset with 5435 images to identify normal beans, male beans, broken, insect-damaged, black, shell, and sour beans. The Slim CNN generates a model with few parameters, and the initiative is suitable for edge computing. Their results reached 92% of accuracy.

The study described in [16] presented research to help farmers classify coffee beans based on their color. The ANN was created to identify four color groups: whitish, cane green, green, and blueish-green. The system used a Naive-Bayes classifier and achieved 100% accuracy.

III. Background

A. Edge Machine Learning

Edge Machine Learning (Edge ML) marks a significant advancement in the deployment of machine learning models by shifting computation from centralized servers to local devices at the network's edge. This approach minimizes latency, enhances privacy, and conserves bandwidth by processing data directly on smartphones, IoT devices, or edge servers. Edge ML empowers these devices to make intelligent and real-time decisions.

B. Tiny Machine Learning

TinyML is an emerging area of artificial intelligence where machine learning algorithms are implemented on resource-constrained devices, such as smartphones and microcontrollers, including Internet of Things (IoT) devices, to perform automated tasks [17].

This emerging discipline focuses on implementing and executing machine learning models on devices with limited power, memory, and processing capabilities. By bringing the power of machine learning to the edge, TinyML facilitates real-time decision-making and inference directly on small devices. This movement can potentially revolutionize diverse domains, from healthcare and agriculture to smart homes and wearable technology.

C. Convolutional Neural Network

Most research on tinyML focuses on the use and deployment of CNNs. A CNN is a specialized deep-learning architecture designed for processing and analyzing visual data, including images. Inspired by the visual processing in the human brain, CNNs use convolutional layers to learn hierarchical representations of features from input data automatically. These networks are particularly effective in tasks like image classification, object detection, and image recognition. The advantages and motivations for building a classification model based on tinyML techniques are local device intelligence, distributed computation, energy consumption, and flexibility [18].

IV. Material and Methods

The project is in the experimental development stage. Its main objective is to test the proposed solution in the laboratory and validate the idea systematically and practically.

A. Edge Impulse

Edge Impulse is a Software-as-Service (SaaS) platform created in 2019 to facilitate the development of machine learning projects using edge computing. It enables users to create, train, and deploy machine learning models directly on edge devices. With a user-friendly interface, Edge Impulse supports various sensor data types, making it suitable for various applications, including predictive maintenance, anomaly detection, and gesture recognition. The platform provides tools for data collection, model training, and seamless integration, making it accessible for developers looking to implement efficient and real-time machine learning solutions at the edge. After training a model on the Edge Impulse platform, users can load a compressed version of the model on a microcontroller or a smartphone.

B. Roboflow

Roboflow is a platform designed to streamline and enhance the process of training computer vision models [19]. It simplifies the management and preprocessing of image datasets for machine learning, offering tools for data augmentation, annotation, and version control. With Roboflow, users can efficiently prepare and optimize their image data before training models, saving time and improving model performance. The platform supports various computer vision frameworks, making it accessible for developers and data scientists working on various projects, from object detection to image classification.

C. Data acquisition

The data acquisition phase involved the acquisition of coffee bean images from the field using a high-resolution digital camera and with the help of a coffee specialist who helped identify the types of coffee. In total, 68 Arabica coffee bean images were collected. The coffee beans come from the Santa Rita do Sapucaí, Minas Gerais, Brazil.

D. Labeling

Image labeling or annotation involves associating meaningful labels or categories with images in a dataset. This process is crucial for training supervised machine learning models, particularly in computer vision tasks. Each image is assigned a specific label corresponding to the object, scene, or concept depicted in the image. All collected images were annotated to create “gourmet” and “traditional” classes using the Roboflow platform.

E. Data Augmentation

Data augmentation is a common technique that applies several transformations to the original training images to expand the dataset artificially [20]. These transformations can be rotations, zooms, flips, random shifts, and changes in brightness. The goal is to create additional training examples of variations of the original images, helping the model generalize better to different conditions and improve its robustness.

Data augmentation helps prevent overfitting and improve the performance of machine learning models by presenting them with a broader range of examples. By increasing the variability in the training data, the model becomes more adaptable and is better prepared to handle real-world scenarios with different orientations, lighting conditions, and perspectives.

Data augmentation techniques were applied to the collected images to increase the number of images synthetically. The data augmentation techniques used in this work include flipping, cropping, and shearing. The flip was done horizontally. In the case of crop, a 20% maximum zoom was used. Shear added variability to perspective to help the model be more resilient to camera and subject pitch and yaw. Saturation was not used because this feature changes the color of the images. After applying data augmentation techniques to the 54-image training dataset, it became a 162-image dataset.

F. Datasets

The dataset was divided into two sets (i.e., training and test), with the images being resized to 96 x 96 pixels. The two datasets have 162 and 14 images, respectively. Note that each image has various beans, meaning our task is not a detection problem but a binary classification problem on two classes (i.e., traditional and specialty coffee beans).

G. Model Design and Training

The Edge Impulse platform created and trained an ANN model that identifies specialty and traditional coffee beans through images. The project, called an impulse in Edge Impulse’s context, was created as an image classification. The previously created database was uploaded, and classification was defined as the task being tackled.

Henceforth, we call this default model the first model. MobileNetV2 [21] was used with 16 neurons in a final layer, 0.1 dropout. The input layer has 76,800 features. This model was configured with a learning rate of 0.0005, 20 training cycles, and a batch size of 32 images. This model achieved an accuracy of 100% on both the training and validation sets.

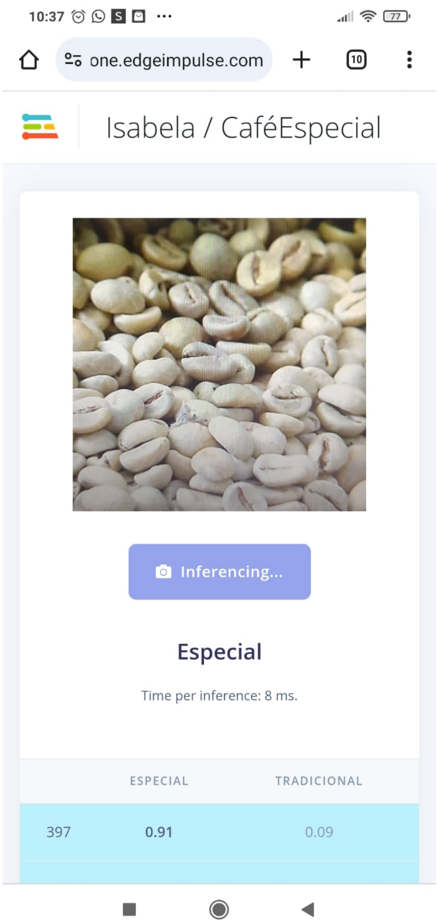


Fig. 2: Live Classification Mode

The second model tested using the edge impulse tool was MobileNetV1, with no final dense layer and 0.1 dropout. The input layer has 76,800 features. This model was configured with a learning rate of 0.0005, 20 training cycles, and a batch size of 32 images. This model achieved an accuracy of 78.8% on the validation set and 64.29% on the training set.

The third model tested using the edge impulse tool was EfficientNet. The input layer has 76,800 features. This model was configured with a learning rate of 0.0005, 20 training cycles, and a batch size of 32 images. It achieved an accuracy of 84.8% on the validation set and 92.86% on the training set.

H. Live Classification

The live classification mode is another Edge Impulse feature that allows testing the created model in the real world using smartphones or a microcontroller. In this experiment, we use a smartphone to evaluate the model. Figure 2 below shows the smartphone screen classifying the type of coffee in real-time.

I. Deployment

After finding the optimum model, achieving the best results, and validating it through the live classification mode, it is time to deploy it. All signal processing, configuration, and learning blocks are quantized to be deployed into resource-constrained devices. Quantization reduces the precision of all blocks' variables and constants from float32 to int8, so the blocks take up less memory (RAM and flash), require less computation, and consume less power. In general, the impact on accuracy is minimal.

Once quantized, the model can be deployed to any device as a C++ library. Edge impulse, though the TensorFlow Lite library [22], packages all signal processing, configuration, and learning blocks, i.e., the model itself, into a single package. This package can be included in applications to run the model locally. This package makes the model run without an internet connection, minimizes latency, and runs with minimal power consumption.

TABLE II: Confusion Matrix of the MobileNetV2 Model

	Specialty	Traditional	Uncertain
Specialty	100%	0%	0%
Traditional	0%	100%	0%
F1 Score	1.00	1.00	

TABLE III: Confusion Matrix of the MobileNetV1 Model

	Specialty	Traditional	Uncertain
Specialty	88.9%	0%	11.1%
Traditional	40%	20%	40%
F1 Score	0.84	0.33	

V. Results and Discussions

MobileNetV2 presented 100% accuracy on the validation set, and the confusion matrix of the training set is shown in Table II. This model correctly classifies all images in the test set. As can be seen, it presents optimum values on the validation set, exhibiting no confusion when classifying the images.

The second model used MobileNetV1 and presented 78.8% accuracy on the validation set, and the confusion matrix of the training set is shown in Table III. We can see that this model had problems classifying specialty coffee beans. About 40% of Specialty images were incorrectly classified as Traditional.

Table IV shows the confusion matrix of the third model tested with EfficientNet.

The comparison between the three models is presented in Table V. In our first model, the loss was 0.03. In the second model, the loss was 0.46. In the third model, the loss was 0.26. Therefore, the first model, i.e., the MobileNetV2, outperforms better.

Table VI presents information on the quantized and unquantized model's latency, size, and accuracy. As can be easily seen, the quantized version of the model is faster, smaller in terms of memory consumption, and presents the same accuracy as its unquantized version.

TABLE IV: Confusion Matrix of the EfficientNet Model

	Specialty	Traditional	Uncertain
Specialty	88.9%	0%	11.1%
Traditional	0%	100%	0%
F1 Score	0.94	1.00	

TABLE V: Performance comparison between the models

Model	Accuracy		Loss
	Validation	Train	
MobileNetV2	100%	100%	0.03
MobileNetV1	78.8%	64.29%	0.46
EfficientNet	84.8%	92.86%	0.26

TABLE VI: Comparison of quantized and unquantized models

	Quantized (int8)	Unquantized (float32)
Latency	189 ms	292 ms
RAM	721.7k	2.4M
Flash	585 KB	1.6 MB
Accuracy	–	92.86

VI. Conclusion

Coffee classification is, undoubtedly, one of the most essential aspects of the coffee business. Coffee is important in the global economy and has a huge global market size. Therefore, correctly classifying coffee is a very important task and must be taken seriously.

This work proposed a novel, cheap, and light framework for classifying coffee beans into specialty and traditional. This framework can assist farmers in detecting the coffee type and deciding the best time to sell their production.

The methodology presented consisted of evaluating the problem of classification of specialty coffees. The proposed solution was developed on the Edge Impulse platform and tested on smartphones. The MobileNetV2 model reached a value of 100% accuracy with a small memory footprint and latency, confirming that TinyML is an ideal approach to image classification in a simple, safe, and low-cost way., producing excellent results.

We believe that the proposed methodology and model are crucial to reducing the subjectivity in coffee classification carried out by humans.

However, improving the found results further by increasing the dataset should be sought. In the future, this model can be expanded with more data from coffee beans to identify multi-classes: Soft, Just Soft, Hard, "Riada", "Rio", and "Rio Zona". Additionally, future studies should compare other pre-trained deep learning models. Furthermore, a future research direction should be developing a cheap application that can run on smartphones without an internet connection so that farmers can use it in rural areas, for example.

VII. Acknowledgment

This work was carried out with the help of "ML Comércio de Café". This work was partially funded by CNPq (Grant Nos. 403612/2020-9, 311470/2021-1, and 403827/2021-3), by Minas Gerais Research Foundation (FAPEMIG) (Grant Nos. APQ-00810-21 and APQ-03162-24) and by the project XGM-AFCCT-2024-2-5-1 supported by xGMobile – EMBRAPII-Inatel Competence Center on 5G and 6G Networks, with financial resources from the PPI IoT/Manufatura 4.0 from MCTI grant number 052/2023, signed with EMBRAPII.

References

[1] de C. Motta, I. V., Huy-Hieu Pham and Felipe A. P. de Figueiredo. 2024 "Machine Learning Techniques for Coffee Classification: A Comprehensive Review of Scientific Research" Preprints. https://doi.org/10.20944/preprints202406.1462.v1

[2] Associação Brasileira da Indústria de Café. Available: https://estatisticas.abic.com.br/ [Accessed on November 20, 2023].

[3] Ministério da Agricultura e Pecuária. Available: https://www.gov.br/agricultura/pt-br [Accessed on November 20, 2023].

[4] DE OLIVEIRA, Emanuelle Morais et al. A computer vision system for coffee beans classification based on computational intelligence techniques. Journal of Food Engineering, v. 171, p. 22-27, 2016.

[5] FARIDAH, Faridah; PARIKESIT, Gea OF; FERDIANSJAH, Ferdiansjah. Coffee bean grade determination based on image parameter. TELKOMNIKA (Telecommunication Computing Electronics and Control), v. 9, n. 3, p. 547-554, 2011.

[6] SANTOS, Wagna Piler Carvalho dos et al. Evaluation of sample preparation (grinding and sieving) of bivalves, coffee, and cowpea beans for multi-element analysis. Microchemical Journal, v. 89, n. 2, p. 123-130, 2008.

[7] Embrapa. Available: http://www.sapc.embrapa.br/ [Accessed on November 20, 2023].

[8] DHANYA, V. G. et al. Deep learning-based computer vision approaches for smart agricultural applications. Artificial Intelligence in Agriculture, 2022.

[9] COPPIN, Ben. Inteligência artificial. Grupo Gen-LTC, 2015.

[10] NORVIG, Peter; RUSSELL, Stuart. Inteligência artificial. Rio de Janeiro: Grupo GEN, 2013.

[11] Edge impulse. Available: <https://edgeimpulse.com/>. [Accessed on November 20, 2023].

[12] WARDEN, Pete; SITUNAYAKE, Daniel. Tinyml: Machine learning with TensorFlow lite on Arduino and ultra-low-power microcontrollers. O'Reilly Media, 2019.

[13] GOMES, Winston Pinheiro Claro et al. Application of multispectral imaging combined with machine learning models to discriminate special and traditional green coffee. Computers and Electronics in Agriculture, v. 198, p. 107097, 2022.

[14] SHAO, Bing et al. Deep Learning-based Coffee Beans Quality Screening. In: 2022 IEEE International Conference on e-Business Engineering (ICEBE). IEEE, 2022. p. 271-275.

[15] WANG, Yan-Feng; CHENG, Chen-Chiou; TSAI, Jenn-Kai. Implementation of Green Coffee Bean Quality Classification Using Slim-CNN in Edge Computing. In: 2022 IEEE 5th International Conference on Knowledge Innovation and Invention (ICKII). IEEE, 2022. p. 133-135.

[16] DE OLIVEIRA, Emanuelle Morais et al. A computer vision system for coffee beans classification based on computational intelligence techniques. Journal of Food engineering, v. 171, p. 22-27, 2016.

[17] DISABATO, Simone; ROVERI, Manuel. Tiny machine learning for concept drift. IEEE Transactions on Neural Networks and Learning Systems, 2022.

[18] REN, Haoyu; ANICIC, Darko; RUNKLER, Thomas A. The synergy of complex event processing and tiny machine learning in industrial IoT. In: Proceedings of the 15th ACM International Conference on Distributed and Event-based Systems. 2021. p. 126-135.

[19] Roboflow. Avalaible: https://roboflow.com/ [Acessed November, 20, 2023].

[20] SHORTEN, Connor; KHOSHGOFTAAR, Taghi M. A survey on image data augmentation for deep learning. Journal of Big Data, v. 6, n. 1, p. 1-48, 2019

[21] DONG, Ke et al. MobileNetV2 model for image classification. In: 2020 2nd International Conference on Information Technology and Computer Application (ITCA). IEEE, 2020. p. 476-480.

[22] DAVID, Robert et al. Tensorflow lite micro: Embedded machine learning for tinyML systems. Proceedings of Machine Learning and Systems, v. 3, p. 800-811, 2021.