IJSRM
INTERNATIONAL JOURNAL OF SCIENCE AND RESEARCH METHODOLOGY
An Official Publication of Human Journals

Human Journals
Research Article
December 2020 Vol.:17, Issue:2

Construction of a Computer Vision System for Tuna Meat Classification: A Case Study of Sensory Analysis Based on Color Parameters

IJSRM
INTERNATIONAL JOURNAL OF SCIENCE AND RESEARCH METHODOLOGY
An Official Publication of Human Journals

Erika Carlos Medeiros*[1], Leandro Maciel Almeida[2], José Gilson de Almeida Teixeira Filho[3]

[1,2,3]*Av. Prof. Moraes Rego, 1235 – Cidade Universitária, Recife, PE – Brasil – CEP: 50670-901. Federal University of Pernambuco, Center of Informatics, Brazil*

Submitted: 12 November 2020
Revised: 02 December 2020
Accepted: 22 December 2020

HUMAN JOURNALS

www.ijsrm.humanjournals.com

Keywords: Tuna Meat; Freshness; Computer Vision System; Color Space

ABSTRACT

This article shows, through successive refinements, supported by the Design Science Research (DSR) methodology, constructing a Computer Vision System (CVS). The construction of CVS made it possible to confirm that samples previously labeled by specialists belong to the same level of freshness classification. The dataset consists of 46 samples, and the CVS was able to analyze automatically 26 samples through histograms, Red, Green, Blue (RGB), and Hue, Saturation, Value (HSV) color space parameters. The analyses show that it is possible to build a system to automate tuna meat's freshness classification based on colorimetric parameters' sensory analysis.

INTRODUCTION

The initial step of an artificial intelligence project, using statistical techniques and Machine Learning (ML), goes through a good definition of the problem to be solved. Once the problem has been defined, the dataset's choice becomes the critical point for developing solutions that cause positive impacts in the context in which they are applied [1].

The scientific community, corporations, and industries have made available real and free datasets to stimulate artificial intelligence solutions for increasingly diverse and urgent problems. Artificial intelligence researchers believe that solutions based on artificial intelligence in business or industrial processes, initially solved by humans, can be more efficient and faster [1].

The researchers tried to consider various free datasets as alternatives in developing solutions; however, there is no specific dataset for this research. One of the significant challenges for the data scientist is converting a business problem to a data science problem. In this context, datasets are not always available. In this sense, it is essential and appropriate to build your dataset.

This article accounts for the researchers' experience in building a dataset of images of tuna meat samples to classify them for freshness through sensory analysis based on their color. It should be noted that there is no dataset for tuna meat samples available publicly and intended for science. The work also presents extractor software built to automatically extract color parameters from the samples' images to analyze freshness levels.

Experts use the tuna classification in the world to identify better quality and turn them into Japanese cuisine. Expert classifiers use sensory analysis through sight, touch, and smell to identify distinct attributes. There are significant differences between the amounts paid for tuna at different rates, which moves a specific and specialized market worldwide [2].

The dataset consists of 46 images of tuna meat samples, extracted from the anatomical region below the side fin, with a cutting instrument called sashibo. The fish used is of the species Bigeye Tuna (or Thunnus Obesus), fished in the Western Tropical Atlantic Ocean, the coast of the state of Rio Grande do Norte Northeast region of Brazil. The researchers monitor sample collection and capture your images in a fishing company in Recife, State of Pernambuco, in

Brazil's Northeast region. The fish, transported from Rio Grande do Norte to Pernambuco, were in refrigerated trucks frozen in ice, at a temperature of -2^{0} C.

The dataset aims to analyze the samples' color patterns to determine the fish's freshness rating. The researchers analyzed such parameters by extracting parameters from the Red, Green, and Blue (RGB) color space and converting it to the Hue, Saturation, and Value (HSV) color space parameters and histogram analysis.

This article, in addition to this introduction, consists of four more sections. Section 2 describes the most important tuna species in Brazil and how tuna is classified worldwide. Section 3 reports on the steps taken to build a dataset of tuna meat samples for analysis and classification. Section 4 presents the process used for the automated extraction of color space parameters. Finally, section 5 presents the conclusions, as well as limitations and future work.

Classification Of Tuna Meat

Tuna is a great oceanic migrant present in oceans worldwide. The annual value of the deal is nearly $12 billion. Fish can weigh more than 600Kg and have a large amount of myoglobin in their muscles, which gives them the characteristic of red meat. The Thunnus Obesus, or Bigeye Tuna, is the second tuna species most captured by the Brazilian longline fleet. This article analyzes these species. Bigeye Tuna is fish with red meat accentuated when fresh and with a remarkable flavor and can accumulate a fair amount of fat just below the coat, but less fat than other species [2].

Japan imports these species at a high price, and the local market consumes lower quality fish [3]. When it comes to tuna classification, meat quality is considered the most important parameter [4]. Sample analysis and type require a well-lit environment, emphasizing that the color temperature and light intensity can influence the result. Natural light for sample analysis is inappropriate since the power and the tone of the sunlight vary during the day and in different climatic conditions, suggesting using artificial lights with cold temperatures above 5500K (bluish-white) [5].

Experts questioned by the researchers claim to feel fatigued and, therefore, difficulties adapting the vision to different illuminations when they perform classifying tuna activity on the same day, justifying the need for artificial classifiers not subject to this condition.

Tuna classification is a world standard through the sensory analysis of a sample of meat extracted from the anatomical region below the lateral fin, using a cutting tool called sashibo. This technique prevents the fish from having to be cut. Visual and odor analysis is required to classify the freshness of samples [2]. In the context of this research, only visual analysis using color was considered, with fish containing meat with an intense red and open color, more valued than those with a paler red color, sometimes with a light pink or yellow hue. According to the color aspect, tuna meat samples' classification receives labels that indicate quality and freshness standards. These codes may change from country to country. In the context of this research, a decreasing classification of five tags was used: # 1, # 2 +, # 2, # 2-, # 3 (from very intense red to little intense red). This labeling's choice is justified because specialists in Pernambuco and Rio Grande Norte use it and the suggested labeling (unless labeled # 2-) in [5].

In the next section, we present the steps to build the tuna meat samples dataset extracted with sashibo.

MATERIALS AND METHODS

This section shows the steps that led to the construction of a dataset of samples of tuna meat of the Bigeye Tuna species, collected between November 10, 2020, and December 15, 2020, in the fish industry located in Recife, Pernambuco, Brazil. The steps are considered construction cycles, and the methodology that supports them is Design Science Research (DSR).

The DSR must have six steps, the sequential steps of the method, which is the sequence followed in this research [6, 7]. We present the DSR steps in Table 1.

Table No. 1: DSR Steps

Sequence	Step
I	Identify the problem and its motivation.
II	Define the objectives for a solution after evaluating the available alternatives.
III	Design and develop artifacts for the chosen solution alternative.
IV	Demonstrate the use of the artifact through its practical experimentation.
V	Measure the results of practical experimentation to return to step III or go straight to step.
VI	Communicate the solution at the end of the process.

Briefly, after identifying the problem to be solved and choosing the solution to be implemented, it is suggested that solution artifacts be built and evaluated, being refined if necessary, to improve the final solution. Considering the artifact as satisfactory communicates the final solution. In this research, the problem identified is: how to automate the freshness classification of tuna meat samples extracted through sashibo?. Among the solution alternatives studied in the literature, we understood that the construction of a Computer Vision System (CVS) to capture and analyze the samples' images shows itself as the most effective, current, and innovative alternative. The next sections present the CVS built for the solution, and the refinements applied for its improvement.

Capture System Construction

This section presents the steps for building a CVS to capture and analyze tuna meat samples' images, extracted using sashibo.

The dataset consists of 46 samples collected and labeled by a specialist and photographed by the researchers. The collection, made in four visits, took place in a dirty area of the fish industry, located in Recife, in Pernambuco, Brazil.

The researchers' first visit to the industry took place on November 10, 2020. At the time, the researchers captured images of eight samples of tuna. The fish had been landed three days ago and transported to Recife, coming from Natal, Rio Grande do Norte in a refrigerated truck, frozen at -2^0 C. Each fish had its temperature measured using a digital skewer thermometer when removed from the car.

The samples, labeled by specialists, were photographed individually in the industry's dirty area, under yellow incandescent artificial light (present in the environment). There was no direct illumination in the sample to capture the image — the equipment used to photograph the pieces of meat is on the iPhone XR smartphone rear camera. The camera has 12 megapixels and was used in automatic mode with *Focus Pixels*, without flash, without zoom, and without applying filters. The researchers stored the captured samples on the smartphone's camera roll, and there was no standardization of the distance and angle between the camera and the sample in the capture. The researchers placed the samples in cards made of white craft paper. The fish sample extraction and the capture of its image took a maximum time of three minutes.

The images were downloaded from the smartphone, *via* USB cable, in jpeg format, on a MacBook Pro 2017 notebook, i5 2.3GHz, 32GB of RAM. For the initial analysis of the images, the researchers used Photoshop software.

Researchers' impressions of the capture system used in the first visit are shown in Table 2.

Table No. 2: Impressions of the Capture System Used in the First Visit

Impressions
The quality of the images was considered satisfactory for generating the histogram by Photoshop.
The uniform white paper in the background facilitated its removal, leaving in the image only the region of interest (ROI), the meat sample.
Shadows were formed due to the lack of direct illumination in the sample, forcing its manual removal.
The fact that researchers using the câmera without a support or tripod caused the image capture to shake.
The lack of standardization and inclination between the camera and the samples made it difficult to define the ROI.

From the impressions, the researchers implemented a refinement for the next visit: purchase of studio with top opening, positioning the camera, and LED lighting, for standardizing the capture of images, as shown in Figure 1.

The studio for capturing tuna meat samples was acquired commercially and had 20 cm^3. LED lights with adjustable temperature from 3200K to 6500K located on the top and top opening for positioning the camera.

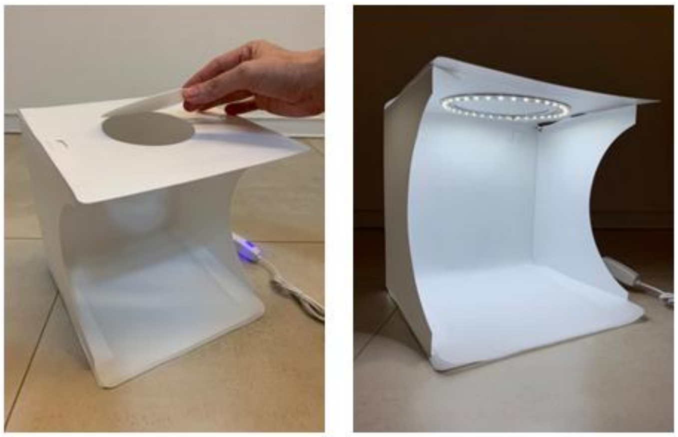

Figure No. 1: Studio For Capturing Images From Tuna Meat Samples

The researchers carried out the second visit on December 10, 2020. On this occasion, they photographed seven samples using the same camera and settings as the first visit. The fish had been landed a day ago, on the coast of Rio Grande do Norte, and transported to the industry in a refrigerated truck, frozen at -2^0 C. The fish removed from the car had their temperature measured with a digital skewer thermometer.

The meat samples were extracted with sashibo and labeled by a specialist. The researchers wrote down the label and used styrofoam to transport the samples to another environment. They used to carry out the capture work, a room provided by the industry, and set up the studio shown in Figure 1. The camera was positioned in the top opening, with the studio's LED lights' temperature being set to 6500K (cold temperature and more bluish nuance). In this capture, maintaining a standard distance of 20 cm, the samples (individually) were removed from the styrofoam and transported to the uniform white paper card and then photographed. The images were later downloaded from the smartphone *via* USB cable in jpeg format, on the same notebook used on the previous visit and analyzed using software coded by the researchers to automate the background removal, leaving only the region of interest the image (code available at https://bityl.co/4si1). Since it was extracted from the fish's body, the time taken to capture a sample lasted an average of three minutes.

Researchers' impressions of the capture system used in the second visit are shown in Table 3.

Table No. 3: Impressions of the Capture System Used in the Second Visit

Impressions
The quality of the images was considered satisfactory for generating the histogram by Photoshop.
The extractor software performed automatic background extraction. The researchers could see that the extractor could not define the ROI of the images collected on the first visit, having different behavior in the second visit's photos.
The uniform white paper in the background facilitated its removal, leaving only the meat sample's ROI in the image.
There were no shadows of the samples in the images, thanks to direct lighting in them.
The researchers captured the samples' images without the tremor seen on the first visit, thanks to the camera's support at the studio's top.
It was standardized lighting, the distance between camera and samples, and inclination of the camera about the surface on which the samples were.
The capture of images in the studio, to the detriment of an uncontrolled environment, facilitated the automation of background extraction, and consequently, the ROI.
The handling of meat samples, being transported manually, from the styrofoam to the paper card, caused many samples to break.

The researchers implemented the refinements, shown in Table 4, for the next visit.

Table No. 4: Implementations for the Next Visit

Implementations
The researchers reduced the distance between the camera and the samples to 12.5 cm.
The researchers created white cards and placed them in styrofoam to accommodate the samples coming from the sashibo. The cards were detached and taken to the studio.

The researchers, on December 11, made the third visit to the fish industry. In this visit, they photographed nineteen samples using the studio and the improvements implemented since the second visit. The fish had been landed a day ago, on the coast of Rio Grande do Norte, and transported to the industry in a refrigerated truck, frozen at -2^0 C. The fish were removed from the car and had their temperatures measured with the same thermometer used in previous visits.

The meat samples were extracted with sashibo and labeled by a specialist. The researchers noted the label with a pilot pen on a uniform white paper card placed in styrofoam to be transported to the same room used on the second visit. The researchers maintained a distance of 12.5 cm between the camera and the samples, and the lighting settings used in the previous visit, to photograph the samples. The cards (each containing just one sample) were detached from the styrofoam and taken to the studio, preventing the samples' handling. The average time taken to capture the samples, taking into account their extraction from the fish's body, was 3 minutes. Following the standard, the images were downloaded offline in jpeg format on the work MacBook Pro and analyzed with extractor software.

On the third visit, there were the same impressions on the part of the researchers observed on the second visit, plus the perception of greater clarity of the samples, given the shorter distance between them and the camera, and the certainty that without manipulation of the samples made these would not break. The researchers understood that the capture system built based on refinements was quite efficient on this visit. However, researchers decided to adapt a white LED lighting to illuminate the samples' background.

This last refinement was implemented on the fourth visit, held on November 15, 2020. A directional light fixture with LED diffused lights were purchased in local stores to maintain the same color temperature studio lighting (6500K, white, cold).

On the fourth visit, 12 samples were extracted with sashibo and labeled by the same specialist. The researchers used the same capture studio used on the third visit, plus the last refinement (lighting in the sample background). They noticed a significant improvement in the images' quality, seen with the naked eye, concerning the images captured on the third visit. However, when analyzing the pictures using the extractor software, it wasn't easy to define the region of interest of the image and extract the background in an automated way. The researchers went back to use lighting in the studio's background, being considered its configuration used in the third visit suitable for capturing tuna meat samples and can be easily reproduced for other contexts. Figure 2 shows images of samples collected during the four trips.

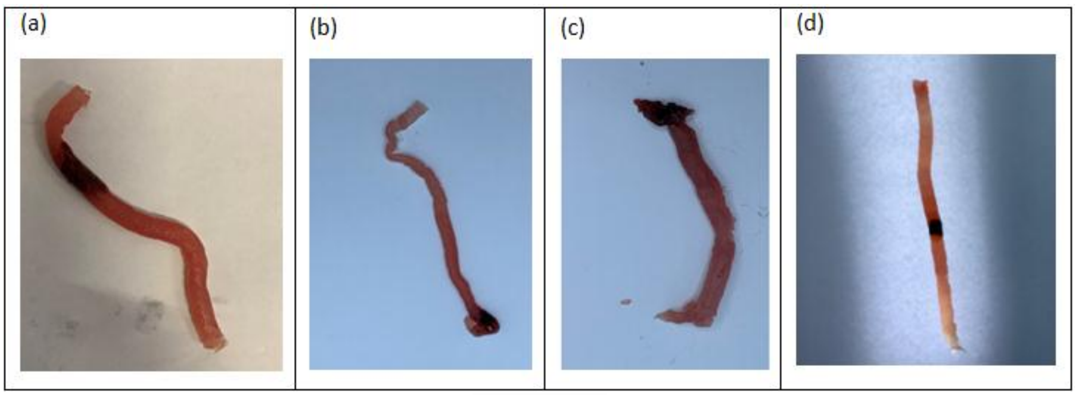

Figure No. 2: Samples collected on four visits. (a) Without standard lighting. (b) With illumination on the top and a distance of 20 cm from the camera. (c) With lighting on the top and a length of 12.5 cm from the câmera. (d) With lighting on the top and background and with space 12.5 cm from the câmera

In summary, the researchers noticed that building the dataset has been motivated by identifying a problem. For the solution, following the steps of the DSR, the first capture artifact (first visit) was built, consisting only of a camera. After analyzing the images, The researchers proceeded to refinement by deriving a refined artifact, consisting of a camera and a photographic studio, enabling the standardization of lighting, the distance between the sample and camera. After analyzing the captured images, they proposed a new refinement in the CSV, giving rise to artifact three used on the third visit. The improvement consists of artifact two, with a base placed between the studio's surface and the top to decrease the capture distance. Finally, the researchers proposed artifact four from the previous one by placing lighting on the studio's surface to illuminate its background. Artifact three was deemed adequate for capturing the samples, declining from artifact four, with the capture system reaching its final implementation.

Color Space Parameter Extraction

The extractor software construction was performed using the Python language to automate the images' ROI (the meat sample). Codes have been added for automatic extraction of parameters that can be used in an automatic classifier based on Machine Learning algorithms. The histogram and parameters of two color spaces: RGB and HSV, were extracted from the images.

In the RGB color space, a color is represented by a triple (r, g, b) that encodes the amount of red (r), green (g), and blue (b) present in color. The HSV color space is an alternative representation of the RGB color space, formed by the parameters hue (or Matiz), saturation, and value. The hue parameter checks the type of color, covering all colors in the spectrum. The saturation parameter represents the proportion of the amount of color about the average gray color—the less gray in the color composition, the more saturated. Finally, the value parameter defines the color's brightness, with a lower value, making it less bright and a higher value, making it more colorful. No model describes all the aspects of colors; it is essential to note that different colored spaces specify other characteristics of colors [8].

Using the Python OpenCV library, it was possible to extract the RGB parameters from the samples' images in order b, g, r (depending on the library implementation details). The parameters b, g, and r were then converted into the HSV color space parameters, using the same library. The HSV space's choice was to express the colors intuitively, such as brightness and hue, facilitating the comparison between values extracted from different samples.

RESULTS AND DISCUSSION

This section presents the results of the entire sample set analysis, exemplifying the process by studying four samples, shown in Figure 3, through the respective histograms and parameters of the HSV space, captured between the second and third visits. The researchers discarded from the dataset the eight images collected in the uncontrolled environment of the first visit, as the extractor software did not automatically locate the ROI. The same happened with sixteen pictures on the fourth visit since the background lighting hindered the region's automatic location of interest. The dataset included, in the end, 26 images of tuna meat samples.

Figure 3 shows four samples belonging to two different freshness classifications, labeled by the specialist, at the time of extraction and capture. The left part of Figure 3 shows the ROI. You can see the respective histogram on the right side of the image and the HSV color space parameters at the bottom.

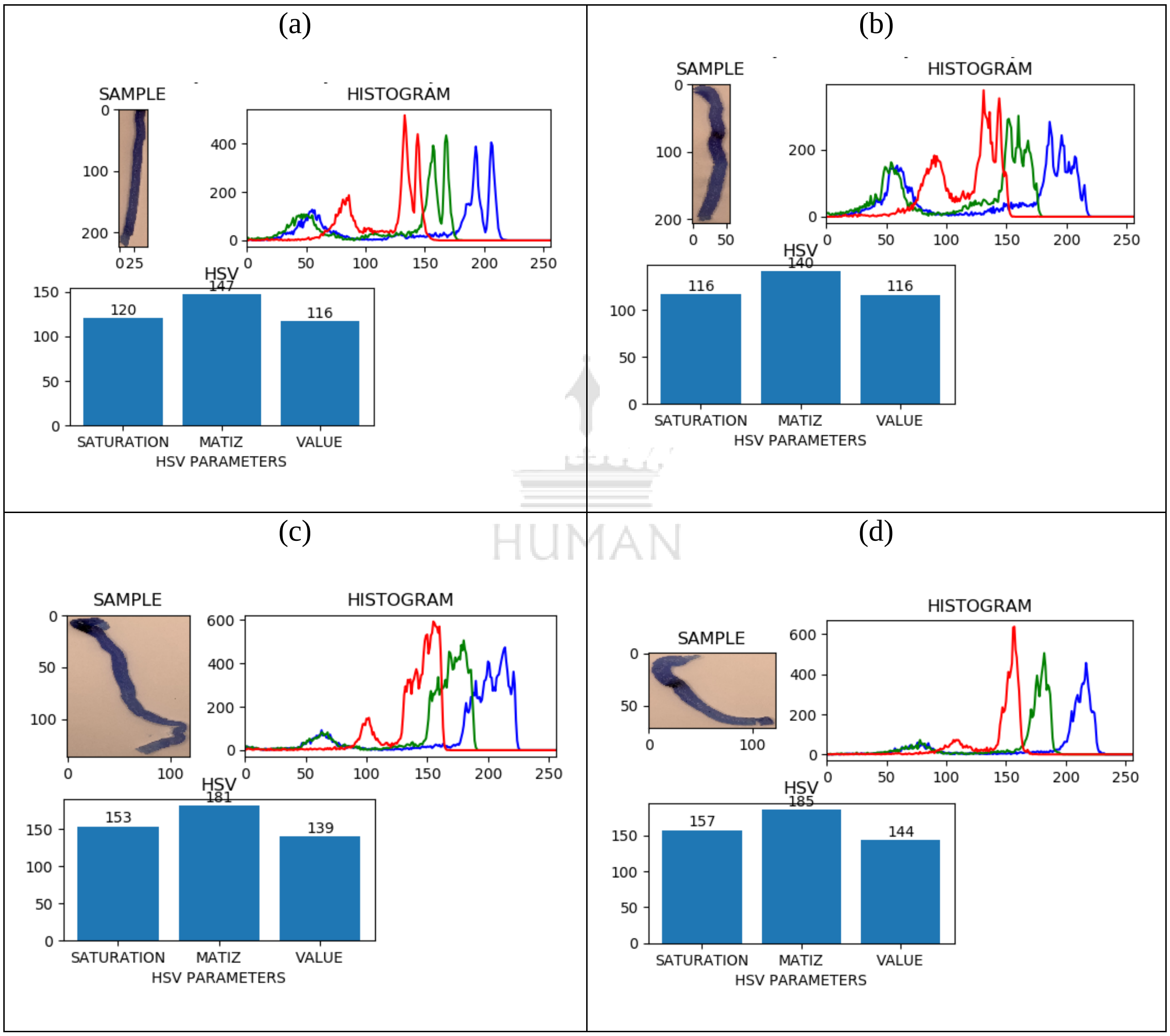

Figure No. 3: Histograms and Parameters of The HSV Space of Four Different Samples

Samples (a) and (b) in Figure 3 were labeled by a specialist with freshness classification 2, while the same specialist labeled samples (c) and (d) with freshness classification 3. The researchers realized that the saturation, hue, and value of Figure 3 (a) and (b), suggesting that the samples belong to the same classification. The same is valid for Figure 3 (c) and (d). On the other hand, there is a distance between the values of the HSV color space parameters of (a) and (b) concerning the samples (c) and (d) of Figure 3, suggesting that the former have freshness distinct. After analyzing the HSV parameters of the 26 images of pieces of meat, it was possible to perceive the approximation of the h (hue), s (saturation), and v (brightness) values for samples classified at the same freshness level by the specialist. It was also possible to notice their distance when the samples had different classifications. The researchers identified four clusters due to the proximity of h, s, and v values, suggesting four freshness classification levels. The researchers concluded that it is possible to use the HSV parameters and histogram points to automate tuna meat by analyzing their samples' images. The capture studio and the extractor compose a powerful CVS for analyzing tuna meat.

CONCLUSIONS

This article detailed the collection of 46 samples of tuna meat, extracted from the anatomical region, below the side fin, with sashibo. It was possible to perceive, through successive refinements, the construction of a CVS, constant in the image capture system, and extractor code of color parameters. The analysis of the images through extractor software allowed us to conclude that the automation of the ROI and the extraction of color parameters are effortless when we capture images in a controlled environment.

From 46 initial samples, only 26 were possible to analyze automatically using the proposed CVS. It was possible to confirm, through analysis of parameters of the HSV color space, the segmentation into four distinct freshness classifications, made initially by the specialist.

We understand that this work has limitations, the low number of samples in the dataset, and the different conditions in which were collected. For continuing this work, we suggest to correct the limitations and, above all, to analyze the samples through parameters of other color spaces, such as the HSI color space. We also suggest using the k-means and different algorithms for grouping samples and confirming the specialist's label.

REFERENCES

1. Raschka S, Mirjalili V. Python Machine Learning. 3rd Ed. Packt Publishing Ltd; 2017
2. Nóbrega C, Mendes P, Mendes E. Factors That Determine The Quality Of Bigeye Tuna, Caught In The Western Tropical Atlantic Ocean. Arquivo Brasileiro De Medicina Veterinária E Zootecnia. 2014;66(3):949-958
3. Forget F. Iniciativa Da FAO Ajuda Na Redução Da Sobrepesca Do Atum [Internet]. ONU News. 2020 [Cited 14 December 2020]. Available From: Https://News.Un.Org/Pt/Story/2020/01/1702372
4. Blanc M. Grading Of Tunas For The Sashimi Market. SPC Fisheries Newsletter; 2002
5. Digregorio R. Tuna Grading And Evaluation: The Complete Tuna Buyer's Handbook. 2nd Ed. EUA: Urner Barry Publisher; 2017
6. Peffers K, Tuunanen T, Rothenberger M, Chatterjee S. A Design Science Research Methodology For Information Systems Research. Journal Of Management Information Systems [Internet]. 2007 [Cited 13 December 2020];24(3):45-77. Available From: Https://Www.Tandfonline.Com/Doi/Abs/10.2753/MIS0742-1222240302
7. Wieringa R. Design Science As Nested Problem Solving. Proceedings Of The 4th International Conference On Design Science Research In Information Systems And Technology - DESRIST '09 [Internet]. 2009 [Cited 13 December 2020];:1-12. Available From: Https://Dl.Acm.Org/Doi/Abs/10.1145/1555619.1555630
8. Hardeberg J. Acquisition And Reproduction Of Color Images. [Parkland, Fla.]: Universal Publ; 2001.

	***Erika Carlos Medeiros – Corresponding Author*** *Federal University of Pernambuco, Center of Informatics, Brazil* *Av. Prof. Moraes Rego, 1235 – Cidade Universitária, Recife, PE – Brasil – CEP: 50670-901*
	***Leandro Maciel Almeida*** *Federal University of Pernambuco, Center of Informatics, Brazil* *Av. Prof. Moraes Rego, 1235 – Cidade Universitária, Recife, PE – Brasil – CEP: 50670-901*
	***José Gilson de Almeida Teixeira Filho*** *Federal University of Pernambuco, Center of Informatics, Brazil* *Av. Prof. Moraes Rego, 1235 – Cidade Universitária, Recife, PE – Brasil – CEP: 50670-901*