COMPILER
Vol. 12, No. 2, November 2023, pp. 83~92
ISSN: 2252-2403 (Print), ISSN: 2549-2403 (Online)
Aaccredited Fourth Grade by Kemenristekdikti, Decree No:200/M/KPT/2020
DOI: 10.28989/compiler.v12i2.1760

Arc Circularitie Naive Bayes for Occupational Safety Helmet Detection

Taufiq Rizaldi[1,*], Hermawan Arief P[2], Gilang Rizqi Palmasugra[3]
[1,2,3]Department of Information Technology, Politeknik Negeri Jember.

Article Info

Article history:

Received August 3, 2023
Accepted August 23, 2023
Published November 30, 2023

Keywords:

Arc Circularitie
Helmet
Image Processing
Naïve Bayes
Occupational Safety and Health

ABSTRACT

Occupational Safety and Health (OHS) is an effort to guarantee and protect the safety and health of every worker through efforts to prevent work accidents and work-related diseases. Safety Helmet is one of the components that must exist and be used in accordance with Occupational Safety and Health standards. Detection of safety helmets usage is one of the efforts to support these activities. The application of Arc Circularity Naive Bayes is used to detect whether an object meets the ratio of a circle by utilizing RGB and HSV image filtering and classification using Naïve Bayes. That method is used to detect whether a worker uses a safety helmet or not, it also detects helm color. The average value of accuracy is 50.8, precision is 58.3%, recall is 66.0%, and f1-score is 59.5% which are calculated using the Confusion Matrix.

Corresponding Author:

Taufiq Rizaldi,
Department of Information Technology,
Politeknik Negeri Jember,
Jl. Mastrip PO BOX 164, Jember - Jawa Timur- Indonesia.
Email: *taufiq_r@polije.ac.id

1. INTRODUCTION

The Ministry of Manpower in the fourth quarter of 2021 noted that there were 7,928 work accident cases that occurred, with a total of 9,224 workers [1]. They were 2,097 workers experienced type A accidents (hits generally indicate contact or contact with sharp objects or hard objects) and 485 workers experienced type B accidents (hit or generally due to falling, sliding, flying, etc.). Workers who work in an environment with the potential for work accidents are required to implement Occupational Health and Safety (K3) operational standards where one of the components is the use of safety helmets. A safety helmet that functions to protect the head from objects that fall directly on the head and can absorb shock or blows is called a project helmet. Safety helmets have a standard shape and are differentiated into several colors [2]: red, blue, and yellow. Supervision of safety helmets usage is one of the efforts to reduce the number of accidents in the working world. The large number of workers who still heed the use of work safety helmets causes many work accidents to occur [1] so that technological assistance is needed for monitoring measures.

Object detection using images [2] [3] [4] [5] and video [6] has become one of the focuses of research that has been carried out in the field of computer vision. Using object detection technology for workplace safety is one of the solutions offered to reduce the number of work accidents. Several methods are used in the object detection process such as Haar Cascade [5], Support Vector Machine (SVM) [2], Component Principal Analyst (CPA) [4], Naïve Bayes [6] [7] and YOLO [3][12]. Naïve Bayes only requires a small amount of training data compared to other methods that require quite a lot of data to achieve a high level of accuracy [6], this makes Naïve Bayes widely used to optimize the object detection process.

Arc Circularities is a measurement method to determine whether the marked object meets the circle ratio or rate. The utilize values range from >=1.0 for perfect circles to <1.0 for semi-circles and other shapes to determine whether the object is included in the circle ratio or rate [9]. The focus of this study is to apply Arc Circularities which are used to detect the edges of an object whether it enters the circle area after a color filtering process with RGB and HSV so that a circularity ratio is obtained, the filtering results are then processed with Naïve Bayes to get a label whether the object, in this case, is a worker wearing a helmet or not, so that it can be early detection of the use of OHS equipment, in this case, a helmet to reduce the risk of work accidents. The focus of this paper is to apply and connect computations from Arc Circularities and Naïve Bayes to produce prototypes for detecting the use of safety helmets and classifying them according to type.

2. RESEARCH METHOD

The stages in this study began with the process of collecting training data obtained from the Kaggle dataset. The data obtained was an image of a helmet that already had a label according to the classification class in the metadata of each image. Classification class data in the form of safety helmet image data in blue, white, yellow, green, red and so on. Image data as primary data used for training data totaling 5000 data where in each data there was an image of a helmet, a human head and a human image, each image had a bounding box annotation and was stored in the PASCAL VOC format. Secondary data of 30 data used were obtained from image capture using a camera.

The data obtained was processed so that there were two types of images, namely the original image and the post-processed image. The classification process flow is shown in Figure 1. At the stage of identifying the correctness of the entered image, the system started by recording the image through the camera which then was processed using a model that had been trained and labeled. Then the system performed classification and comparison of the image input and the previous model input, if the system predicted that the result was the correct output then the system issued a positive result and vice versa if the system predicted a negative result, then the system issued a negative result. The output results were used as parameters for decision making by the system if there was a violation or not. In this study the cases used were blue and yellow helmets.

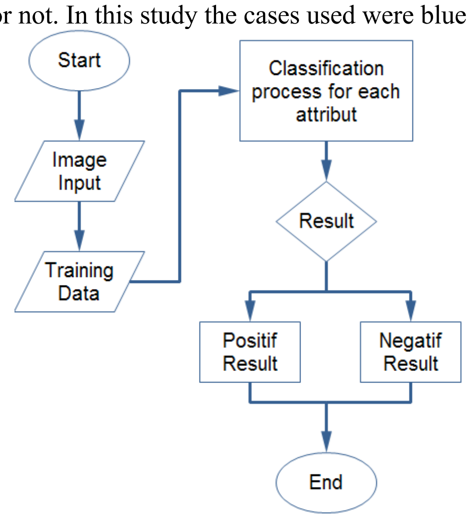

Figure 1. Image Classification Flow

The first process was the process of taking and labeling data. In the data labeling process, the image features were extracted in the form of color classification using the RGB and HSV color filtering methods. After making some adjustments to the average RGB and HSV values, three values were taken which were the reference for each of its features, namely the mean value, standard deviation, and amount of data. The results obtained after the color adjustment process and obtaining the circularity ratio [8] are shown in Table 1.

Table 1. Training Models

RGB Average			Circularity Rate	Average				Label
Red	Green	Blue		V0	V1	V2	V3	
49.9315339	148.5253456	63.89005925	0.622304955	3.89005925	48.5253456	49.9315339	0	Kuning
22.7667494	7.8473945	46.2158089	0.78539816	6.21588089	07.8473945	22.7667494	0	Kuning
11.5775	0.425	54.42	0.650136982	4.42	00.425	111.5775	0	Kuning
69.14550265	57.94021164	23.57037037	0.83355573 2	23.57037037	57.94021164	69.14550265	0	Kuning
97.30926158	71.53629537	14.37234043	0.55623815 9	14.37234043	71.53629537	97.30926158	0	Kuning
39.63653754	63.41654711	110.6958393	0.69813170 1	110.6958393	63.41654711	39.63653754	0	Biru
44.27243108	67.10350877	126.0533835	0.49013620 3	126.0533835	67.10350877	44.27243108	0	Biru
63.17220339	90.58779661	161.6061017	0.85905052 2	161.6061017	90.58779661	63.17220339	0	Biru
23.39046653	57.93610548	94.76470588	0.75398223 7	94.76470588	57.93610548	23.39046653	0	Biru
22.34576023	65.79824561	111.0270468	0.75089511 2	111.0270468	65.79824561	22.34576023	0	Biru

Information:
RGB average : Represents the average value of the Red, Green, and Blue color space.
Circularity Rate : Is the ratio of the shape of the detected image contour to the circle shape.
V0, V1, V2, V3 average: Is the standard deviation value of the Red, Green, Blue, Alpha color space.

After the labeling process on the training data, the mean, average, and amount of data obtained from filtering RGB and HSV images was carried out. There were 35 data with yellow class labels and there were also 35 data with blue class labels which then these data were searched for the value of the descriptor [7] [8]. The formula for determining the mean and standard deviation is as follows:

$$\text{Mean } (\bar{x}) = \frac{\sum x}{N} \quad (1)$$

Where:
$\sum x$: Sum of all data values
N : Total number of frequencies.

$$\text{Standard Deviation (S)} = \sqrt{\sum \frac{(X_i - \bar{x})^2}{n-1}} \quad (2)$$

Where:
X_i : population standard deviation
$\bar{x}$: population size
n : each value from the population

So that the results are obtained for the yellow class features as in table 2 and the features for the blue class as in table 3, where count is the number of detected contours. Then after doing one of the sample classes and their features, a manual calculation will be carried out to find out the results of the classification using the Gaussian Naïve Bayes formula.

Table 2. Yellow Class Features

	Red	Green	Blue	V0	V1	V2
Mean	118.93114682657142	109.77081956428572	40.01462730094287	40.01462730094287	109.77081956428572	118.93114682657142
Std. Dev.	25.809366913522595	28.129554032106466	22.188323803602366	22.188323803602366	28.129554032106466	25.809366913522595
Count	35	35	35	35	35	35

Table 3. Blue Class Features

	Red	Green	Blue	V0	V1	V2
Mean	40.27771189362856	69.10946251171431	102.71613330257144	40.01462730094287	109.77081956428572	118.93114682657142
Std. Dev.	27.02430067498438	27.687033941465796	33.977380493092404	33.977380493092404	27.687033941465796	27.02430067498438
Count	35	35	35	35	35	35

Based on table 2 and 3, class features are obtained for each helmet color class which later was used for comparison during the classification process with Gaussian Naïve Bayes [6] so that it can be determined whether the detected object is a helmet and belongs to the yellow or blue color class.

3. RESULTS AND ANALYSIS

Based on the calculation of the descriptor value as in Table 2 and Table 3, it can be calculated from the input image taken. The detected object was calculated for its Circularity Rate value, then feature extraction was carried out and after adjusting the HSV values, the value was obtained as shown in table 4. Circularity Rate calculations use the formula from Arc Circularities [9], namely:

$$C = 4\pi \times \frac{\text{area}}{\text{perimeter}^2} \quad (3)$$

Where
C: Circularity shape
Area: Area or content in a geometric shape
Perimeter: Fill the circle closest to the outermost area of the circle
In the case example there is an input image to be detected and classified

Table 4. Input Image Value

Avg Red	Avg Green	Avg Blue	V0	V1	V2	Circularity Rate
123.120454545454 55	110.751948051948 05	51.0954545454545 44	51.095454545454 544	110.75194805194 805	123.12045454545 455	0.81985456

Table. 4 displays some of the extracted features, namely the average value and intensity of the RGB color channel and the circularity rate feature which calculates the ratio of the image shape above to the circle shape. Then the calculation and comparison of the results of the feature extraction values were done to get the results which were shown in Table 5 and Table 6 with the following formula:

$$P(x_i|y)=\frac{1}{\sqrt{2\pi\sigma_y^2}}\exp\left(-\frac{(x_i-u_y)^2}{2\pi\sigma_y^2}\right) \quad (4)$$

Where

Xi : The data hypothesis X to i is a data with a specific class

y : Data with unknown class

σ : Standard deviation of a class

μ : The mean or average of a class

Table 5. Gaussian Naive Bayes Yellow Class Group

	R	G	B	V0	V1	V2
Input	123.12	110.751	51.09	51.09	110.751	123.12
Mean	118.0193	108.6309	39.31	39.31	108.6309	118.0193
St. Dev.	25.6189	27.72	22.123	22.123	27.72	25.6189
Exponent	0.98	0.997	0.8678	0.8678	0.997	0.98
Calculation (Xn)	0.0152	0.0143	0.0156	0.0156	0.0143	0.0152
Circularity Rate						0.81985456
Final Result						4.2585

Table 6. Gaussian Naive Bayes Blue Class Group

	R	G	B	V0	V1	V2
Input	123.12	110.751	51.09	51.09	110.751	123.12
Mean	40.27	69.109	102.716	102.716	69.109	40.27
St. Dev.	27.02	27.68	39.977	39.977	27.68	27.02
Exponent	0.0090	0.3225	0.4343	0.4343	0.3225	0.0090
Calculation (Xn)	0.00013	0.0046	0.0043	0.0043	0.0046	0.00013
Circularity Rate						0.81985456
Final Result						1.85904

Based on the final result of the Gaussian Naïve Bayes calculation for the yellow class group, which is 4.2585 and the circularity rate value is 0.81985456, while for the blue class, the manual calculation value is 1.85904 and the circular rate value is 0.8198456, so that comparisons are made between the yellow class group and the blue class group. It was obtained the calculation with the highest calculation result in the yellow class. The appearance of the computation process in application is shown as in Figure 2.

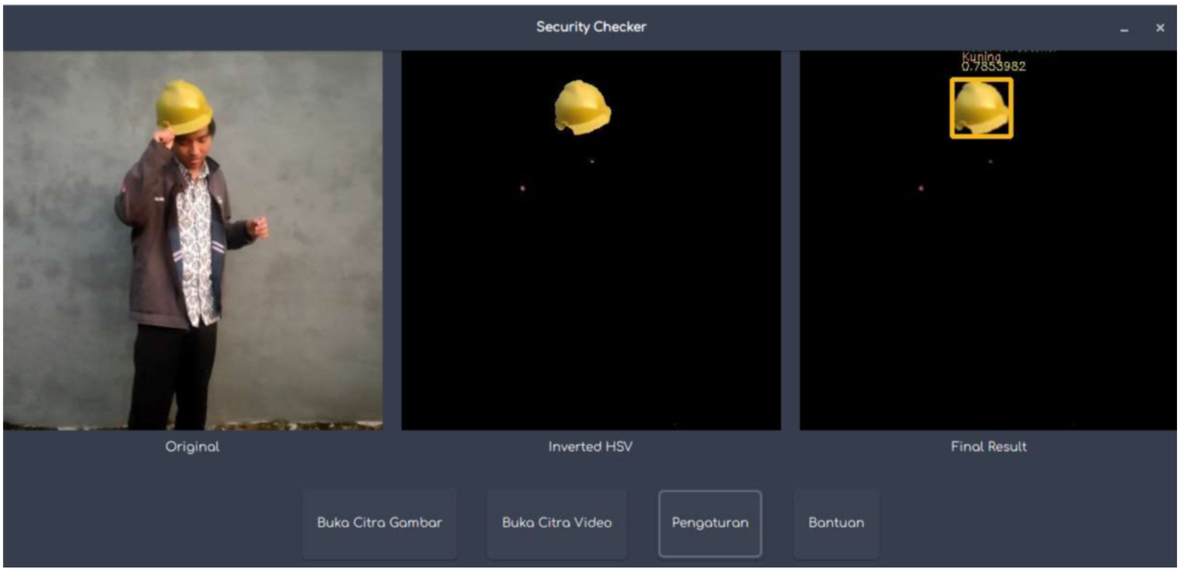

Figure 2. Application View

In Figure 2, the detected safety helmet image that get from a camera and put into the application was marked with a border color and a color label that matches the classification results along with the detected circular rate.

3.1. Evaluation

At the evaluation stage, fifteen (15) data were taken for image test data using yellow and blue safety helmets each and tested three (3) times each using different places and lighting according to the time taken. Each test will be assessed by determining TP (True Positive), TN (True Negative), FP (False Positive) and FN (False Negative) where:

- TP (True Positive) : If the prediction result is positive and it is verified that the result is indeed positive.
- TN (True Negative) : If the prediction result is negative and it is verified that the result is indeed negative.
- FP (False Positive) : If the prediction result is positive and it is verified that the result should be negative.
- FN (False Negative) : If the prediction result is negative and it is verified that the result should be positive.

The results for testing for helmets with a blue color class are shown in table 7, where it was tested on images in which there were safety helmets.

Table 7. Blue Color Class Test

No	Time	Class	Objek		Circular Rate	TP/FP/FN/TN
			Real	Detected		
1	09:10 WIB	Blue	1	1	0.8399	1/0/0/0
2			1	2	0.8614	1/0/0/0
					0.8982	0/1/0/0
3			1	1	0.3907	0/0/1/0
4			1	2	-	0/1/0/0
					0.7853	1/0/0/0
5			1	2	0.6981	1/0/0/0
					0.6819	0/1/0/0
6			1	1	0.8081	1/0/0/0
7	09:14 WIB		1	1	0.7228	1/0/0/0
8			1	1	0.6838	1/0/0/0
9			1	1	0.8594	1/0/0/0
10			1	1	0.7099	1/0/0/0
11			1	1	0.7784	1/0/0/0
12			1	2	0.8284	1/0/0/0
					0.4944	0/0/1/0
13	17:25 WIB		1	1	0.7853	1/0/0/0
14			1	1	0.7853	1/0/0/0
15			1	2	0.7853	1/0/0/0
					0.5643	0/1/0/0

From the test results for blue safety helmets there are cases where there are cases of FN and FP due to the Circular Rate of an object whose value is not up to 0.5 so it is considered not a safety helmet (example in Figure 3) and there is a BLOB (Binary Large Object) so considered as a helmet object which in reality is not a safety helmet (example in Figure 4).

Figure 3. Example of detection with a Circular Rate value of less than 0.5 (1).

Figure 3 is an example when the system detects 1 blob contour that appears as a safety helmet but it is detected as "Helmet Not Detected". There is an imprecise detection because the contour does not meet the minimum circle ratio of 0.5 so that this detection is considered as FN.

Figure 4. Example of detection with non-helmet object detection (1).

Figure 4 is an example when the system detects the presence of 2 blob contours but there are contours that are considered by the system as contours that enter the blue color class and display "Helmets Detected", so the results of this detection are considered as FP because the contours detected are in the form of inaccurate blobs said to be a safety helmet. The results for the test for helmets with the yellow class are shown in table 8 with almost the same conditions as the testing time for the blue class.

Table 8. Yellow Color Class Test

No	Class	Objek		Circular Rate	TP/FP/FN/TN
		Real	Detection		
1	Yellow	1	1	0.7539	1/0/0/0
2		1	1	0.8198	1/0/0/0
3		1	1	0.7539	1/0/0/0
4		1	1	0.7853	1/0/0/0
5		1	2	0.6981	0/1/0/0
				0.3846	0/0/1/0
6		1	1	0.714	1/0/0/0

7	1	1	0.785	1/0/0/0
8	1	2	0.7539	1/0/0/0
			0.7234	0/1/0/0
9	1	3	0.6981	1/0/0/0
			0.7539	0/1/0/0
			0.6162	0/1/0/0
10	1	2	0.7853	1/0/0/0
			0.7853	0/1/0/0
11	1	1	0.8198	1/0/0/0
12	0	1	0.4554	0/0/1/0
13	1	1	0.7841	1/0/0/0
14	1	1	0.6099	1/0/0/0
15	1	1	0.6558	1/0/0/0

In the yellow class test, there are several which included in FN and FP because it occurs as in the blue class test (Examples of Figures 5 and 6). Another factor that causes the FN and FP values to occur is the system predicts several other objects obtained as safety helmet objects. The detected object is not a type of safety helmet, but the object in this case reflects sunlight so that it makes the object appear yellow and the system also provides the results of the color classification of several objects as blue objects so that they are included in the FN and FP classes.

Figure 5. Example of detection with non-helmet object detection (2).

In figure 5 the system detects 3 different blob contours, but there are 2 contours detected as "Helmets Detected" so this contour is considered as False Positive. It is because this contour computationally fulfills the circularity ratio and the classification of the color class is correct but the physical image of the resulting predicted value is not correct.

Figure 6. Example of detection with a Circular Rate value of less than 0.5 (1).

In figure 6, the system detects 1 blob contour, but the system predicts this contour object as "Helmet Not Detected" so the final result of this object is "False Negative". The reason for this result is that the object cannot meet the minimum circularity ratio specified in this case of 0.5. In terms of physical image, this object can be classified as a safety helmet object in terms of shape and color class, but the system predicts the opposite so that the prediction results can be said to be inaccurate.

The next test is in the form of detecting objects other than safety helmets with the aim of seeing whether the system can detect and classify the detected objects as safety helmets or other objects. The types of objects tested are blue and yellow objects that resemble safety helmets, namely hats, beanie caps and other accessories as shown in Table 9.

Table 9. Testing of Non-Helmet Objects

No	Class	Objek		Circular Rate	TP/FP/FN/TN
		Real	Detection		
1	Blue	1	1	0.698	0/1/0/0
2		1	2	0.698	0/1/0/0
				0.724	0/1/0/0
3		1	1	0.588	0/1/0/0
4		1	1	0.683	0/1/0/0
5		1	1	0.733	0/1/0/0
6	Yello	1	1	0.268	0/1/0/0
7		1	1	0.753	0/1/0/0
8		1	2	0.845	0/2/0/0
				0.753	
9		1	1	0.572	0/1/0/0
10		1	3	0.7853	0/3/0/0
				0.7853	
				0.7853	

Based on the results of trials conducted on non-safety helmet objects, 8 objects were detected for the yellow class, and 6 objects were detected for the blue class. Several objects were detected from the 14 objects detected, there were false positive values from both the yellow and blue color classes. This false positive result is caused by the results of an inaccurate classification of the detected object, the system detects the presence of a circularity shape and the correct color classification but the detected object is not a safety helmet. Thus, there are errors or inaccuracies in the classification and prediction of the system due to the lack of information received for processing. An example of a system error when it detects a non-helmet object is shown in Figure 7.

Figure 7. Example of Testing Non-Helmet Objects.

The system detects 2 different blob contours, the contour detection results are marked by a yellow and blue border. The prediction results from these two contour objects produce False Positive values, because the first contour in terms of circularity ratio and color classification is correct but the prediction value can be said to be inaccurate and it is the second contour because the contour object which is not clearly described is predicted by the system as object with the result "Helmet Detected". Based on the results of tests carried out on models trained using the Gaussian Naïve Bayes Classifier method as a classification and Confusion Matrix, the values for accuracy, precision, recall, and f1-score for each location & type of test performed can be summarized in table 10.

Table 10. Test Results

Test result	Accuracy	50.8
	Precision	58.3
	Recall	66.0
	F1-Score	59.5

From the total of all tests, the results obtained from the Confusion Matrix which serves to determine the location of the misclassification of the Gaussian Naïve Bayes Classifier as shown in table 11.

Table 11. Confusion Matrix Data

	Predicted Incorrect	Predicted Correct	Total
Actual False	0	23	23
Actual True	4	27	31
Total	4	50	54

There were 4 data that were predicted as incorrect, and 50 data that were predicted as correct. However, in the 50 data that were considered correct, there were errors of 23 incorrectly predicted data. This data should have an incorrect class, but the model predicts it as a true class.

4. CONCLUSION

The test results on 35 image samples consisting of 15 yellow image samples, 15 blue image samples and 10 non-safety helmet image samples obtained from the Naïve Bayes Classifier Algorithm to produce a classification model with average values of accuracy are 50.8, precision 58.3%, recall 66.0%, and f1-score 59.5% which were calculated using the Confusion Matrix. The use of Arc Circularitie and Naive Bayes has an accuracy rate based on the Confusion Matrix of 54%, The addition of Adaptive Thresholding is needed so that the level of accuracy in the detection process can produce a better level of accuracy.

REFERENCES

[1] Y. Adiratna, S. Astono, M. Fertiaz, S. C. A. O. Sugistria, H. Prayitno, R. I. Khair, A. Brando e B. A. Putri, Profil Keselamatan Dan Kesehatan Kerja Nasional Indonesia Tahun 2022, Jakarta: Kementerian Ketenagakerjaan Republik Indonesia, 2022.

[2] A. K. Wijaya e S. Devella, "Pengenalan Penggunaan Helm Proyek Berstandar Pada Citra Foto Berdasarkan SIFT Dengan SVM," *Jurnal Ilmiah Teknologi dan Komputer,* vol. 3, nº 2, 2022.

[3] Z. Fangbo, Z. Huailin e N. Zhen, "Safety Helmet Detection Based on YOLOv5," em *International Conference on Power Electronics, Computer Applications (ICPECA)*, Shenyang, China, 2021.

[4] S. Firdiyan e H. Hasti, "Deteksi Helm Keselamatan Kerja Barbasis Android Menggunakan Metode PCA (Component Principal Analys)," *Jurnal Dinamika Informatika,* vol. 10, nº 2, p. 84–90, 2021.

[5] S. E. Setyati e H. Armanto, "Deteksi Alat Pelindung Kepala (Helm) Menggunakan Metode Haar Cascade Classifier," *Journal of Informatic Unisla,* vol. 4, nº 1, 2019.

[6] N. J. Peksi, B. Yuwono e M. Y. Florestiyanto, "Classification of Anemia with Digital Images of Nails and Palms using the Naive Bayes Method," *Telematika: Jurnal Informatika dan Teknologi Informasi,* vol. 18, nº 1, pp. 118-130, 2021.

[7] M. M. H. Santoso e R. Muliono, "Analysis Naïve Bayes In Classifying Fruit by Utilizing Hog Feature Extraction," *Journal of Informatics and Telecommunication Engineering,* vol. 4, nº 1, 2020.

[8] R. Kosasih e A. Fahrurozi, "Pengklasifikasian Bunga Dengan Menggunakan Metode ISOMAP Dan Naive Bayes Classifier," *Jurnal Ilmiah Informatika dan Komputer,* vol. 22, nº 3, 2017.

[9] A. S. Talaulikar, S. Sanathanan e C. N. Modi, "An Enhanced Approach for Detecting Helmet on Motorcyclists Using Image Processing and Machine Learning Techniques," em *International Conference on Advanced Computing and Communication Technologies*, 2018.

[10] M. M. Saritas e A. Yasar, "Performance Analysis of ANN and Naive Bayes Classification Algorithm for Data Classification," *International Journal of Intelligent Systems and Applications in Engineering,* vol. 7, nº 2, pp. 88-91, 2019.

[11] R. R. V. e. Silva, K. R. T. Aires e R. d. M. S. Veras, "Detection of Helmets on Motorcyclists," *Multimed Tools Appl,* vol. 77, 2018.

[12] M. Hatami, T. Tukino, F. Nurapriani, W. Widiyawati e W. Andriani, “DETEKSI HELMET DAN VEST KESELAMATAN SECARA REALTIME MENGGUNAKAN METODE YOLO BERBASIS WEB FLASK,” *Jurnal Pendidikan Sains dan Teknologi,* vol. 10, nº 1, 2023.

[13] S. D. , “ENERAPAN ALGORITMA NAÏVE BAYES UNTUK MENGETAHUI KUALITASPRODUKSI HELMET HONDA PADA PT TERANG PARTS INDONESIA,” *Journal of Research and Publication Innovation,* vol. 1, nº 3, 2023.

[14] G. L. e T. Shilei, “Improved YOLOv5 Network Model and Application in Safety Helmet Detection,” em *IEEE International Conference on Intelligence and Safety for Robotics (ISR)*, Tokoname, Japan, 2021.

[15] R. C, X. H, Z. Z e Z. W, “Multi-Scale Safety Helmet Detection Based on SAS-YOLOv3-Tiny,” *Applied Sciences,* vol. 11, nº 8, 2021.