Automatic recognition of emergency vehicles in images from video recorders

Julian Balcerek, Paweł Pawłowski, Krzysztof Dolny, Julia Skotnicka
Division of Signal Processing and Electronic Systems, Institute of Automatic Control and Robotics,
Poznan University of Technology, Poland
{julian.balcerek, pawel.pawlowski}@put.poznan.pl,
{krzysztof.dolny, julia.skotnicka}@student.put.poznan.pl

***Abstract* — In this article, a system for automatic recognition, including detection and classification, of emergency vehicles based on images from video recorders is proposed. The system utilizes YOLO v8 artificial neural network. Recognition accuracy tests were conducted on a specially prepared database of frames from real recordings. It includes marked and unmarked police vehicles (which are very difficult to distinguish and typically not included in automatic recognition systems), fire, ambulance, military vehicles and other reference images. The experiments concerned the recognition of vehicle types and the state of emergency lighting. The highest value of total (for all classes) F1 measure is over 86%. In selected classes the F1 reaches 91%, while precision is 94% and sensitivity is 89%. They can be considered as satisfactory results, what indicates the possibility of using the presented system in practice.**

Keywords — recognition of emergency vehicles; vision system; artificial neural networks; video recorders; image recognition; image processing

I. Introduction

In today's crowded road traffic, it is important to give way to a vehicle that can save human life or property, as quickly and collision-free as possible. Unfortunately, for example, in the United States in year 2022, the majority of fatalities in accidents involving emergency vehicles occurred in multi-vehicle accidents and half of the fatalities involved occupants of other vehicles [1]. Emergency vehicle recognition is therefore an important issue from the point of view of safety and ensuring appropriate vehicle traffic flow.

Emergency vehicles can be recognized by the video monitoring systems [2, 3]. This may improve traffic control, e.g. by appropriately adjusting traffic lights [4]. However, information from surveillance cameras may only come from selected, permanent point locations and may not cover certain areas. Furthermore, recognition can also be performed directly from the vehicles participating in the road traffic [5]. It can then provide precise and individualized information about the emergency vehicle, intended for a given driver or systems of a non-privileged vehicle. Recognition can be performed based on combination with the detection of audio signals [5–8] or even only on siren sound [9].

In recognition based on image analysis, a distinction can be made between approaches based on classification only or on classification and detection. The classification can differentiate various classes of objects: two classes, i.e. ambulances and other vehicles [10], four classes, i.e. ambulances, fire brigade vehicles, police vehicles and other vehicles [11, 12]. There may also be more complex system based on various sets of classes, e.g. including the city guard and army, occurring in Polish conditions. Detection of vehicle dimensions and emergency lighting state on emergency vehicles were also implemented [13].

In the case of detection-based recognition, ambulances only [14] or parallel to various classes of non-privileged vehicles [15] can be detected. Various emergency vehicles can be also detected as one class [16, 17].

Video recorders used by drivers can be a source of information for recognition of selected objects in road traffic. Such research was conducted on image databases collected in East Asian [18] or European [19] conditions. Video recorders can be accessory devices, or a function of users' smartphones, and can also be used in vehicles that are not factory equipped with cameras. They can also be used to recognize emergency vehicles.

This work focuses on the full emergency vehicle recognition process, i.e. taking into account not only classification but also detection. Various sets of emergency vehicle classes are detected. Additionally, it is possible to detect the state of emergency lighting. Real frames from videos recorded in Polish conditions from video recorders were used for testing.

II. Database of images of emergency vehicles

In order to use artificial neural networks (ANN) to recognize emergency vehicles, it was necessary to complement and appropriately divide the images depicting the objects to be recognized. It was also necessary to prepare appropriate annotations.

The vehicle database used for training and testing ANN models was made from frames taken by video recorders. The films, from which the image frames were obtained, contained recordings from various models of car video recorders with various parameters. A maximum of three frames showing a given vehicle were collected from a single recording. Care was taken to ensure that individual frames had different environments and the arrangement of vehicles on the road so that they were not too similar.

The database included, among others: several video recordings of emergency vehicles prepared by the authors, where iPhone 11 and Samsung Galaxy S10 smartphones were used to acquire video sequences. Illustrative frames, with many details removed for anonymization, are shown in Fig. 1.

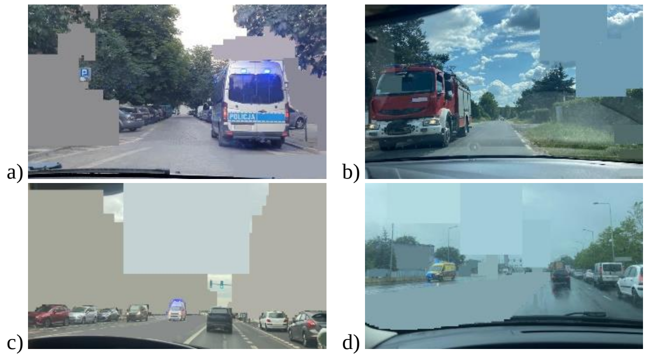

Figure 1. Sample frames from authors' video recordings of emergency vehicles in Poland: (a) police, (b) fire brigade, (c) ambulance in white painting, and (d) ambulance in yellow painting

The images collected in the database included emergency vehicles in Poland, which had characteristic markings and paintings. The military vehicle photo database also included American vehicles recorded on Polish roads.

The resolution of the collected images varied due to the use of various devices. Some of the photos in the database had resolutions of up several megapixels. The minimum resolution below which images were not saved was 100×100 pixels. The collected photos had different lighting levels because they were taken at different times of the day and at night. This often resulted in additional noise in the image. In some cases it was not easy, even for a human, to quickly and clearly determine the type of vehicle in the images (see Fig. 1c and 1d).

The prepared database contained of approximately 1.2 thousand images. It included images of marked and unmarked police vehicles, ambulances, fire brigade and army vehicles. We assumed that there could be more than one emergency vehicle in one image. Therefore, the number of objects was greater than the number of images in the database.

Seven sets of emergency vehicle classes were distinguished from the collected images, taking into account various types of recognition experiments. The detailed information about assigned image sets is presented in Table I.

The first set assumes the simplest division and it was called the basic set. It was designed to conduct the largest number of tests. It includes three classes of emergency vehicles: police, fire brigade, and ambulance.

The second set includes an additional class of emergency vehicles, i.e. military vehicles.

The third set includes an additional class of unmarked police vehicles, which are distinguishable from civilian vehicles only by their blue emergency lighting. Vehicles of this type do not have any characteristic markings or painting.

The fourth set was introduced to recognize only unmarked police vehicles. This made it possible to check to what extent it is possible to automatically distinguish relatively similar vehicles, such as unmarked police vehicles and civilian passenger cars.

The fifth set includes seven various classes of vehicles. It takes into account the state (on or off) of the blue emergency lighting. Military vehicles were not included in this set because we could not collect enough photos of military vehicles with their emergency lights on. Information about the state of the emergency lighting is important from a practical point of view. It could act as a potential warning to the driver to give free way to an approaching emergency vehicle. This would give the driver more time to react.

The sixth set focuses on the recognition of the emitted blue emergency light without distinguishing the type of emergency vehicle. From the driver's point of view, it is not necessarily important which type of vehicle he has to give way to. This division could be used in practical solutions used in autonomous vehicles.

The last, seventh set was similar to the previous one, but it also included unmarked police vehicles.

TABLE I. THE NUMBER OF IMAGES AND EMERGENCY VEHICLE OBJECTS IN PARTICULAR CLASSES FOR SELECTED DATABASE DIVISIONS

<table>
<tr><th>set of classes</th><th>number of classes</th><th>police</th><th>fire brigade</th><th>ambu-lance</th><th>un-marked police</th><th>military vehicle</th><th>refer-ence images</th><th>total</th></tr>
<tr><td>1</td><td>3</td><td rowspan="3">288/296</td><td rowspan="3">246/262</td><td rowspan="3">237/243</td><td rowspan="2">–</td><td>–</td><td rowspan="7">30 (test only)/0</td><td>801/801</td></tr>
<tr><td>2</td><td>4</td><td>197/282</td><td>998/1083</td></tr>
<tr><td>3</td><td>4</td><td rowspan="3">164/164</td><td rowspan="5">–</td><td>965/965</td></tr>
<tr><td>4</td><td>1</td><td>–</td><td>–</td><td>–</td><td>194/164</td></tr>
<tr><td>5</td><td>7</td><td>153/157 (on signals) 135/139 (without signal)</td><td>131/135 (on signals) 115/127 (without signal)</td><td>126/126 (on signals) 111/117 (without signal)</td><td>965/965</td></tr>
<tr><td>6</td><td>2</td><td colspan="3">410/418 (on signals) 361/383 (without signal)</td><td>–</td><td>801/801</td></tr>
<tr><td>7</td><td>2</td><td colspan="4">574/582 (on signals) 361/383 (without signal)</td><td>965/965</td></tr>
</table>

Taking into account the use of ANNs, in each set the images were divided into subsets. Thus, in each class, in addition to the reference images, training, validation and test sets were prepared, with images divided in the following proportions: 70%, 15% and 15%, respectively.

For testing purposes only, the so-called reference images were prepared. These were frames from video recorders, which did not contain emergency vehicles, but civilian vehicles. The number of these images was 30 and they were the same for each set. The relatively small number of reference images was due to the fact that these images only appeared in the test sets. This made it possible to check how often civilian vehicles were incorrectly identified as emergency vehicles. It was assumed that including them in the training or validation set would not add significant impact to the training.

An appropriate annotation has been prepared for all images in the database. It provided information about where a specific type of emergency vehicle was located in the image. It was necessary for proper training, validation and testing.

The CVAT tool [20] was used to prepare the annotations. It included functionalities that significantly accelerated the process of marking a large number of objects.

Some of the images showed vehicles obscured to varying degrees by other objects. In order for a given vehicle to be marked, it had to meet certain criteria: the vehicle bounding box had to be at least 50×50 pixels in size, at least one side of a vehicle had to be visible (not covered by other objects) in 50% or more. Non-emergency vehicles were not marked.

Finally, the prepared database of images from video recorders, containing annotations relating to emergency vehicles, allowed for training, validation and tests of software for automatic recognition of this type of vehicles.

III. Artificial neural networks architecture

A dedicated software in Python programming language was prepared for automatic recognition of emergency vehicles using ANNs [21]. This programming language is often used in this field. The main development environments were PyCharm [22] and Jupyter Notebook [23]. The process of training and testing the operation of given neural network models was possible thanks to the Ultralytics library [24].

Ultralytics, including the YOLOv8 project, is a tool specializing in real-time object detection. According to available information about performed tests, the provided models achieve a frame rate that allows processing even several dozen images per second. It is focused on providing interactive tools for analyzing object detection results [24].

The YOLOv8s (small) and YOLOv8l (large) models belong to the YOLO architecture. They are often used in real-time object detection and classification due to its combination of speed and processing accuracy. Small and large are two models of different sizes, differing, among others, in: number of parameters and speed of operation [24]. We decided to compare these two models and thus examine the impact of the level of model complexity on the obtained results.

Before first training, the number of epochs was set to 186 (depending on the number of images in the training and validation set), batch size to 10, the implementation of the Adam algorithm was selected as the optimizer, and the image resolution was set to 320×320 pixels [24]. There were ten thousand training steps per training. Subsequent training and validation processes, depending on test specifications, were conducted in a similar way. Model training was conducted using the Intel Core i7-9750H processor with Windows 10 operating system and 16 GB of RAM.

The main tool used to analyze training results was Tensorboard [25]. It was actually a set of tools used to visualize training and validation results for the YOLOv8 library. The Tensorboard provided information on various training parameters such as steps per second, classification loss, location loss and total loss. The Matplotlib library [26], implemented and used in Python, was used to visualize the data. It was a library dedicated to drawing charts.

An example of the total loss function for training the YOLOv8s model, for the first training set, is shown in Figure 2. It can be seen that there was no stabilization phenomenon even after 1000 training steps.

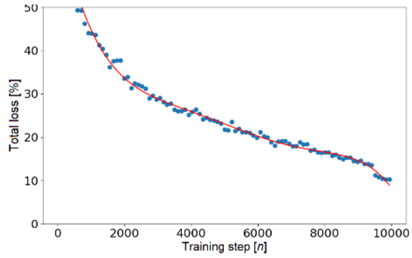

Figure 2. Total loss function for training the YOLOv8s model, for the first training set

IV. Tests and experimental results

The effectiveness of the prepared software was tested on the collected database of images. Selected recognition performance metrics, based on a typical confusion matrix were used. A true positive detection occurred when the model correctly identified an object that actually existed in the image. A false positive detection occurred when an object of a given class was incorrectly identified in a location where it was not located. A true negative detection occurred when the model failed to identify an object that was not present in the image. A false negative detection occurred when the model failed to identify an object in the image.

To describe the results we used precision, sensitivity and F1 measure. Precision provides information about how many objects for which a given class was identified actually belonged to that class. Sensitivity, provides information on how many objects, out of all objects actually belonging to a given class, were classified well.

The obtained recognition precision values for given classes, taking into account detection, could be lower than if they were calculated for classification step only, because they also took into account situations when the model detected a given object in a place where there was no emergency vehicle. The overall precision values were calculated without taking into account the case when the model did not detect any object that was in the image.

From the same reason, the obtained recognition sensitivity values for given classes, taking into account detection, could be lower than for classification step only. Here also, the overall sensitivity values were calculated without taking into account the case when the model detected a given object in an area where there was no emergency vehicle.

The F1 measure is the harmonic mean of precision and sensitivity, defined by the formula [27]

$$\text{F1 measure} = 2 \cdot \frac{\text{precision} \cdot \text{sensitivity}}{\text{precision} + \text{sensitivity}} \cdot 100\ [\%]\,. \qquad (1)$$

In the recognition tests of the basic emergency vehicles (division 1), ANN models YOLOv8s and YOLOv8l were compared in terms of effectiveness, processing speed. Additionally we tested various image resolutions, in order to select the network with the best metric values.

Both models, for images with a resolution of 320×320 pixels, achieved a similar F1 result of approximately 85% and maintained the balance between precision and sensitivity (see Table II).

TABLE II. EMERGENCY VEHICLE RECOGNITION RESULTS FOR PARTICULAR CLASSES FOR DIVISION 1 (S – YOLOV8S, L – YOLOV8L)

measure [in %]	police	fire brigade	ambu-lance	total
precision	78.72 (s) 91.89 (l)	83.72 (s) 80.49 (l)	88.24 (s) 94.29 (l)	83.06 (s) 88.50 (l)
sensitivity	80.43 (s) 70.83 (l)	97.30 (s) 89.19 (l)	83.33 (s) 89.19 (l)	86.55 (s) 81.97 (l)
F1	79.57 (s) 80.00 (l)	90.00 (s) 84.62 (l)	85.71 (s) 91.67 (l)	84.77 (s) 85.11 (l)

The speed of image processing was compared in tests for two trained models for the same image resolution. The YOLOv8s model achieved 3 times higher average number of images processed per second than the YOLOv8l model (18.18 and 5.21 images per second, respectively).

The YOLOv8s model was tested for various image size values. The resulting measures are presented in Table III. For a resolution of 320×320 pixels, the model obtained the highest F1 percentage result and a satisfactory number of frames processed per second.

TABLE III. COMPARISON OF PERFORMANCE METRICS OBTAINED FOR THE YOLOV8S MODEL FOR PARTICULAR IMAGE RESOLUTIONS

image resolution [in pixels]	measures [in %]			
	precision	sensitivity	F1	fps
96×96	80.56	73.11	76.65	57.25
192×192	79.83	77.87	78.84	30.60
320×320	83.06	86.55	84.77	18.18
640×640	84.03	83.33	83.68	7.52

The confusion matrix corresponding to this test is presented in Table IV. The "other/lack" class in the confusion matrix referred to the case where the model did not detect an object that was in the image (column), or to the case where it detected a given object in an area where no emergency vehicle was located (row).

TABLE IV. CONFUSION MATRIX OBTAINED FOR THE YOLOV8S MODEL FOR IMAGE RESOLUTION OF 320×320 PIXELS

		predicted classes			
		police	fire brigade	ambu-lance	other/ lack
real classes	police	37	0	2	7
	fire brigade	0	36	0	1
	ambulance	0	0	30	6
	other / lack	10	7	2	–

In subsequent tests, only one ANN model was used, i.e. YOLOv8s for images with a resolution of 320×320 pixels. The results of emergency vehicle recognition for all seven sets of data are presented in Table V.

TABLE V. EMERGENCY VEHICLE RECOGNITION RESULTS (S – YOLOV8S, L – YOLOV8L)

set of classes	measure [in %]		
	precision	sensitivity	F1
1	83.06 (s) 88.50 (l)	86.55 (s) 81.97 (l)	84.77 (s) 85.11 (l)
2	87.80 (s)	85.21 (s)	86.49 (s)
3	79.33 (s)	80.95 (s)	80.13 (s)
4	50.00 (s)	68.00 (s)	57.63 (s)
5	60.42 (s)	58.00 (s)	59.19 (s)
6	59.09 (s)	61.90 (s)	60.46 (s)
7	70.00 (s)	65.77 (s)	67.82 (s)

When one vehicle class was added, i.e. military vehicles, the value of the F1 measure even increased slightly compared to the first division, to 86.5% (see Table V, set 2). This result indicates that military vehicles are detectable and distinguishable and do not negatively affect the prediction effectiveness of the artificial neural network model. It was noted that misidentification may occur when the car has a dark body color and large dimensions, as is the case with typical military vehicles.

In turn, adding the class of unmarked police vehicles reduces the resulting F1 metric values to 80.1% (see Table V, set 3). Detection errors after adding vehicles of this class were most likely due to the fact that these vehicles were similar to civilian ones.

These observations were confirmed by the results for the division, where only unmarked police vehicles were recognized, and the value of the F1 metric dropped to 57.6% (see Table V, set 4).

In the next case, vehicles were divided not only by type, but also by whether they have emergency lighting turned on, and unmarked police vehicles were also included. The achieved results of the F1 measure of 59.2% (see Table V, set 5) were much worse than in the case when the ANN did not check the state of the emergency lighting, but assigned emergency vehicle type (set 3).

The next set only takes into account the division into whether the emergency light was turned on or off for police, fire brigade and ambulance vehicles. The obtained result (60.5%, see Table V, set 6) was lower than for vehicle class recognition (set 1), but slightly higher than in the case of detailed simultaneous recognition of class and the state of the emergency light (set 5).

Recognition of the state of the emergency lighting only, taking into account unmarked police vehicles, resulted in an F1 value of 67.8% (see Table V, set 7). This means that it was more accurate than with the simultaneous division into vehicle classes (set 5). The better result than for sixth set may be due to the fact of an addition of unmarked police vehicles, a larger number of images were used during learning while the number of classes remained unchanged.

Examples of emergency vehicle recognition results are shown in Fig. 3. Many details were removed for anonymization.

a) b) c) d) e) f)

Figure 3. Sample recognition results: (a) correct recognition of a fire brigade vehicle for set 2, (b) correct recognition of a police vehicle at the signal for set 5, (c) correct recognition of vehicle on the emergency signal for set 6, (d) correct recognition of a police vehicle for set 2, (e) incorrect recognition of an ambulance as a military vehicle for the set 2, and (f) incorrect recognition of an unmarked police vehicle for the set 4

In the case of Fig. 3e, the misrecognition may have resulted from the fact that the vehicle did not have emergency lighting turned on, there was reflection in the glass, and the colors were not contrasting enough. Detection errors in other cases could result from poor quality recordings, small size of vehicles or similar dimensions or colors. The high similarity between unmarked police vehicles and non-privileged vehicles, especially with the center brake light on (see Fig. 3f), may have contributed to misidentification. In the case of incorrect recognition of the turned off emergency lighting, the error could result from the fact that the blue background was not sufficiently contrasting with the lights. In an extreme case, a red truck was incorrectly recognized as a fire brigade vehicle. Due to the reflective colors and illumination by the lights of other vehicles, recognizing emergency vehicles was not a problem even at night. Several cases occurred where a car was incorrectly detected as an emergency vehicle occurred, but situations where an image area other than the vehicle was incorrectly assigned as an emergency vehicle were extremely rare.

In Table VI comparison of various approaches to emergency vehicles recognition is presented.

TABLE VI. COMPARISON OF CHOSEN APPROACHES TO THE DETECTION AND CLASSIFICATION OF EMERGENCY VEHICLES

work	dataset; CNN model	type of experiment	resulting effectiveness
[10]	photos featuring emergency trucks; VGG-16	classification, 2 classes: ambulances, various kinds of vehicles	overall classification accuracy: 98%
[11]	images gathered from Kaggle and Google images; VGG-16	classification, 4 classes: ambulance, firetruck, police, standard car	overall accuracy, precision, recall and F1-score: 86%
[12]	customized dataset collected from different sources; DenseNet201	classification, 4 classes: ambulance, fire brigade, police, non-emergency vehicles	test accuracy: imbalanced data 96.7%, balanced data 98.6%
[13]	customized dataset of emergency vehicles in Poland collected using Google Images; 7 various models, e.g. EfficientNetV2B0	classification, 5 various sets of 6 main classes: ambulance, police, fire brigade, city guard, army, non-privileged, distinction: dimensions, emergency lighting state	best set test accuracy results: 99.1%
[14]	custom dataset including annotated images and videos of traffic scenes containing ambulances; YOLOv5	detection, annotated images and videos of traffic scenes containing ambulances	best model precision 93.6%, recall 84.7%, accuracy 93.2%
[15]	6 classes extracted from the COCO data set, and 3 classes externally added to the training data set from local images; YOLOv5	detection: ambulance and 8 other classes: person, car, van, truck, bus, bicycle, threewheeler, motorcycle	ambulance accuracy and detection precision: 98%
[16]	obtained from Kaggle; 6 various models, e.g. MobileNet	detection, 2 classes: emergency, nonemergency vehicles	best model configuration accuracy: 92%
[17]	no detailed information found; YOLOv: 5, 7, 8	detection, 2 classes: normal vehicles, emergency vehicles	best neural network model precision 90.2%, recall 92.2%
This paper	available frames from video recorders recorded in Polish conditions; YOLOv8	detection: 7 various sets of 5 main classes of emergency vehicles: police, fire brigade, ambulance, unmarked police, military vehicle, distinction: emergency lighting state	best set total F1 score result: 86.5%

In favor of the proposed system, compared to similar solutions listed in the table above, four facts should be noted. First, the images that the ANNs analyzed were frames from video recorders that needed to be searched for and extracted, before collected in the dataset. Secondly, it was not only the classification but prior the detection of emergency vehicles. Third, there were five main classes of emergency vehicles, including unmarked police vehicles. Fourth, there were seven different sets of classes, with the largest number of classes being also seven, including recognition of the emergency lighting state.

V. CONCLUDING REMARKS

The article presents automatic recognition of emergency vehicles in recordings from video recorders. Such types of recordings varies in quality. Thus, the highest value of the F1 measure is over 86%. In selected classes the F1 reaches 91%, while precision is 94% and sensitivity is 89%. They can be considered as satisfactory results, what indicates the possibility of using the presented system in practice.

The test results show that it is possible not only to recognize types, but also the state (on/off) of the emergency lighting. Recognizing unmarked police vehicles, due to their similarity to civilian vehicles, is an even more difficult task, but not impossible. Moreover, the developed vision system was able to detect vehicles even in bad weather conditions and at night.

In further research, the currently collected image database could be supplemented with additional frames from car video recorders. Trained models of artificial neural networks used in the implemented mobile application could be tested during long drives.

The presented system, as an element of the driver warning system, can contribute to faster driver reaction and thus minimize the number of accidents on the road. It can also shorten the time it takes for emergency vehicles to arrive at the scene of the report and incident.

The research was financed by a research subsidy 0211/SBAD/0224.

REFERENCES

[1] National Safety Council, Injury Facts, Motor Vehicle, Road Users, Emergency Vehicles, https://injuryfacts.nsc.org/motor-vehicle/road-users/emergency-vehicles/ (access: 09.06.2024).

[2] Roy S., Rahman M. S., "Emergency Vehicle Detection on Heavy Traffic Road from CCTV Footage Using Deep Convolutional Neural Network", International Conference on Electrical, Computer and Communication Engineering (ECCE), 7-9 February, 2019, doi: 10.1109/ECACE.2019.8679295.

[3] Razalli H., Ramli R., Alkawaz M. H., "Emergency Vehicle Recognition and Classification Method Using HSV Color Segmentation", 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA), Langkawi, Malaysia, 2020, pp. 284–289, doi: 10.1109/CSPA48992.2020.9068695.

[4] Thakare N., Morey A., Gajbhiye K., Bhalerao S., Nehare P., Meshram A., "Advanced Traffic Clearance System for Emergency Vehicles", International Research Journal on Advanced Engineering Hub (IRJAEH), vol. 02, issue: 05, May 2024, pp. 1174–1180, e ISSN: 2584-2137, doi: 10.47392/IRJAEH.2024.0161.

[5] Garg A., Gupta A. K., Shrivastava D., Didwania Y., Bora P. J., "Emergency Vehicle Detection by Autonomous Vehicle", International Journal of Engineering Research & Technology (IJERT), vol. 8, issue 05, May 2019, pp. 190–194, ISSN (Online): 2278-0181.

[6] Zohaib M., Asim M., ELAffendi M., "Enhancing Emergency Vehicle Detection: A Deep Learning Approach with Multimodal Fusion", Mathematics, 12, 1514, 2024, doi: 10.3390/math12101514.

[7] Tran V. -T., Tsai W. -H., "Audio-Vision Emergency Vehicle Detection", IEEE Sensors Journal, vol. 21, no. 24, 15 Dec. 2021, pp. 27905-27917, doi: 10.1109/JSEN.2021.3127893.

[8] Raj V. S., Sai J. V. M., Yogesh N. A. L., Preetha S. B. K. and R L., "Smart Traffic Control for Emergency Vehicles Prioritization using Video and Audio Processing", 6th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 2022, pp. 1588-1593, doi: 10.1109/ICICCS53718.2022.9788119.

[9] Sathruhan S., Herath O. K., Sivakumar T., Thibbotuwawa A., "Emergency Vehicle Detection using Vehicle Sound Classification: A Deep Learning Approach", 6th SLAAI International Conference on Artificial Intelligence (SLAAI-ICAI), Colombo, Sri Lanka, 2022, pp. 1-6, doi: 10.1109/SLAAI-ICAI56923.2022.10002605.

[10] Sanjai P. C., Sam J. J., Iyengar S., Kalimuthu K., "Ambulance Detection in Traffic Signals Using Image Processing", 5th International Conference on Inventive Research in Computing Applications (ICIRCA), Coimbatore, India, 2023, pp. 927-931, doi: 10.1109/ICIRCA57980.2023.10220909.

[11] Hakim bin Che Mansor M. A., Mohamad Kamal N. A., bin Baharom M. H., Adib bin Zainol M., "Emergency Vehicle Type Classification using Convolutional Neural Network", IEEE International Conference on Automatic Control & Intelligent Systems (I2CACIS), Shah Alam, Malaysia, 2021, pp. 126-129, doi: 10.1109/I2CACIS52118.2021.9495899.

[12] Manguri K. H., Mohammed A. A., "Emergency vehicles classification for traffic signal system using optimized transfer DenseNet201 model", Indonesian Journal of Electrical Engineering and Computer Science, vol. 32, no. 2, November 2023, pp. 1058–1069, ISSN: 2502–4752, doi: 10.11591/ijeecs.v32.i2.

[13] Balcerek J., Pawłowski P., Blajer G., Filipkiewicz J., Koćwin K., "Automatic recognition of emergency vehicles" (in Polish), Przegląd Elektrotechniczny, R. 99, no. 10, 2023, pp. 274–277, doi: 10.15199/48.2023.10.57.

[14] Patel A., Degadwala S., Vyas D., "Enhancing Traffic Management with YOLOv5-Based Ambulance Tracking System", IEEE Canadian Conference on Electrical and Computer Engineering (CCECE), Regina, SK, Canada, 2023, pp. 528-532, doi: 10.1109/CCECE58730.2023.10288751.

[15] Munasinghe K. D. S. A., Waththegedara T. D., Wickramasinghe I. R., Herath H. M. O. K., Logeeshan V., "Smart Traffic Light Control System Based on Traffic Density and Emergency Vehicle Detection", Moratuwa Engineering Research Conference (MERCon), Moratuwa, Sri Lanka, 2022, pp. 1-6, doi: 10.1109/MERCon55799.2022.9906184.

[16] S M., Shetty K., Kodipalli A., Rao T., and S K., "Emergency vs Non-Emergency Vehicle Classification: Enhancing Intelligent Traffic Management Systems", International Conference on Network, Multimedia and Information Technology (NMITCON), Bengaluru, India, 2023, pp. 1-6, doi: 10.1109/NMITCON58196.2023.10276208.

[17] Srikanteswara R., H S. D., Hegde A., Hegde P. T., "Emergency Vehicle Recognition Via Automated Smart Traffic Manager", International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT), Bengaluru, India, 2023, pp. 1-6, doi: 10.1109/EASCT59475.2023.10393727.

[18] Yao Y., Xu M., Wang Y., Crandall D. J., Atkins E. M., "Unsupervised Traffic Accident Detection in First-Person Videos", IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), Macau, China, 2019, pp. 273-280, doi: 10.1109/IROS40897.2019.8967556.

[19] Balcerek J., Pawłowski P., Trzciński B., "Vision system for automatic recognition of animals on images from car video recorders" (in Polish), Przegląd Elektrotechniczny - 2023, R. 99, no 10, pp. 278-281, doi: 10.15199/48.2023.10.57.

[20] CVAT.ai Corporation, CVAT, Open Data Annotation Platform, https://www.cvat.ai/ (access: 13.11.2023).

[21] Python Software Foundation, "Python", https://www.python.org (access: 29.12.2023).

[22] JetBrains s.r.o., "PyCharm", https://www.jetbrains.com/pycharm/ (access: 29.12.2023).

[23] Project Jupyter, "Jupyter", https://jupyter.org (access: 29.12.2023).

[24] Ultralytics Inc., Ultralytics YOLOv8 Docs, https://docs.ultralytics.com/ (access: 10.11.2023).

[25] Google Brain Team, TensorBoard, https://www.tensorflow.org/tensorboard?hl=pl (access: 29.12.2023).

[26] The Matplotlib development team, "Matplotlib: Visualization with Python", https://matplotlib.org (access: 18.04.2023).

[27] Sasaki Y., "The truth of the F-measure", Teach Tutor Mater, 26.10.2007, pp. 1–5.