Applied Artificial Intelligence
An International Journal

ISSN: 0883-9514 (Print) 1087-6545 (Online) Journal homepage: www.tandfonline.com/journals/uaai20

A Proposal of an Animal Detection System Using Machine Learning

William H. S. Antônio, Matheus Da Silva, Rodrigo S. Miani & Jefferson R. Souza

To cite this article: William H. S. Antônio, Matheus Da Silva, Rodrigo S. Miani & Jefferson R. Souza (2019) A Proposal of an Animal Detection System Using Machine Learning, Applied Artificial Intelligence, 33:13, 1093-1106, DOI: 10.1080/08839514.2019.1673993

To link to this article: https://doi.org/10.1080/08839514.2019.1673993

Published online: 08 Oct 2019.

APPLIED ARTIFICIAL INTELLIGENCE
2019, VOL. 33, NO. 13, 1093–1106
https://doi.org/10.1080/08839514.2019.1673993

A Proposal of an Animal Detection System Using Machine Learning

William H. S. Antônio, Matheus Da Silva, Rodrigo S. Miani, and Jefferson R. Souza

Computer Science, Federal University of Uberlândia, Minas Gerais, Brazil

ABSTRACT
One of the current challenges is to reduce collisions between vehicles and animals on roads, such accidents resulting in environmental imbalance and large expenditures in public coffers. This paper presents the components of a simple animal detection system and also a methodology for animals detection in images provided by cameras installed on the roads. This methodology allows the features extraction of regions of the image and the use of Machine Learning (ML) techniques to classify the areas into two classes: animal and non-animal. Two ML techniques were compared using synthetic images, traversing the pixels of the image using five distinctive approaches. Results show that the KNN learning model is more reliable than Random Forest to identify animals on roads accurately.

Introduction

In Brazil, about 61% of freight and 95% of passengers are transported through highways (Pesquisa CNT de rodovias 2015: relatrio final 2015). It is estimated that 475 million animals are killed every year on Brazilian roads, 90% of which are small vertebrates, 9% are medium-sized vertebrates, and 1% are large vertebrates; which can put the passengers of the vehicle at risk (CBEE 2017).

A consequence of the road network growth is the increase in the number of traffic accidents. Such accidents are considered as the ninth most significant cause of death in the world, being able to move to the seventh largest in 2030 (AMBEV 2014). The importance of a system that detecting animals in movement in the threat of highway is precise to act in the reduction of these increasing rates of accidents that every year increases. Brazil has an extensive network of roads because it is a country of continental proportions, the animal's detection will help researchers to identify which animal species are common in areas. Considering the human life losses, and threats to the conservation of biodiversity caused by trampling; there is a need for urgent measures to minimize this effect.

CONTACT Jefferson R. Souza jrsouza@ufu.br Faculty of Computer Science, Federal University of Uberlândia, Minas Gerais, Brazil

Color versions of one or more of the figures in the article can be found online at www.tandfonline.com/uaai.

We propose an architecture and methodology for animals detection using images; methodology has three steps: feature extraction, training, and testing sets and supervised learning. First, using five approaches, features are extracted from a synthetic image dataset. These images are hand-labeled and divided into training and testing sets. Finally, ML techniques are trained using training set and evaluated in test set.

The main contributions of this research are:

(1) An architecture for animal detection on roads that also warns the drivers;
(2) Five block-based approaches were developed to extract the features of animals and non-animals on roads;
(3) An animals detection model was selected based on the performance-wise evaluation of ML techniques.

The remainder of this paper is organized as: Section 2 is related works; Section 3 shows an overview of the proposed animal detection system. Section 4 presents the methodology for animal detection; results are shown in Section 5. Section 6 draws conclusions and suggests for future work.

Related Work

Researchers on wildlife collisions and consequently roadkills are essential tools for biodiversity conservation (Carvalho, Custódio, and Oswaldo 2015). Proposing wildlife warning signs to warn drivers about the potential or actual presence of wild animals is a strategy to reduce wildlife-vehicle collisions. According to (Huijser et al. 2015) such warning signs can be categorized as i) standard wildlife warning signs; ii) enhanced wildlife warning signs; iii) temporal wildlife warning signs, or iv) animal detection systems. Usual wildlife warning signs are typically manufactured in the same style as other traffic warning signs while enhanced wildlife warning signs and temporal wildlife warning signs tend to be larger than standard signs and have lights attached to them.

The idea behind animal detection systems is to use computer vision techniques from video cameras images to detect animals that approach the road and warn the drivers using some sign, for instance, an intelligent sign that lights up. Huijser and McGowen (Huijser, Muir, and Davis 2003) identified 27 locations where animal detection systems are or have been in place in North America and Europe and 15 planned locations. Most of the systems target animals such as Deer and Moose covering distances from 50 m to 5000 m. All of the surveyed systems use infrared sensors (break-the-beam) to detect animals. Such systems can help drivers to avoid 81.53% of wildlife collisions. A similar system (warning system with infrared sensors) (Grace, Smith, and Noss 2017) was deployed in Florida-USA due to an increased number of panther mortalities. The authors conclude that the proposed method is more beneficial during the tourist season since drivers in such season generally drove faster than those in the off-season.

Li et al. (2011) propose a hierarchical wild animal detection and notification system based on wireless sensor network using detection nodes, relay nodes, and sink nodes. Detection nodes report the information to the relay node. The relay node controls the detection nodes and sinks node cumulate the notifications received by relay nodes and warn the drivers by lightening the warning signal. The system was not evaluated.

Sharma and Shah (2017) proposed an animal detection system in the context of Indian roads. The authors used HOG and boosted cascade classifier for animal detection. The techniques were developed in OpenCV. The authors get a recall to 80.4%, specificity close to 83.5% and accuracy close to 82.5%.

Matuska et al. (2014) developed an automatic system to monitor the movement of animals. To create a classification model they choose two classifier methods: Bags of Visual Keypoint and Support Vector Machine. The training set into five classes: fox, deer, wolf, brown bear, and wild boar; each class has 60 images with animal and testing set has 50 images. The highest classification score 94% was achieved by combination SURF, SIFT, and FlannBased. The authors also propose an animal detection system using cameras and intelligent road signs (Matuska et al. 2016). Their solution is similar to ours; however, they do not provide details about the computing units that compose that system and how they will communicate with each other.

Forslund and Bjarkefur (2014) proposed a night vision animal detection. From images captured, was used a Region of Interest to classify animal using an efficient sliding windows approach. The data have been labeled and utilized by learning algorithms designed to extract discriminative information efficiently. The detection of animals is up to 200 m away from the car.

Zhou (2014) explores the performance of different image features and classification algorithms in animal detection application. Haar-like, Histogram of Oriented Gradients (HOG) and Local Binary Patterns (LBP) features were used for extraction features. To classify data was used the AdaBoost, SVM and HOG SVM classifiers. The results indicate that HOG+AdaBoost detector has the most energy-saving features.

Overview of an Animal Detection System

A typical environment for animal detection system is depicted in Figure 1. Our approach is focused on developing an animal detection system that would be able to detect animals in the road and warns drivers. This system was thought to be efficient without wasting too many resources. The system is composed of: i) camera modules with presence sensors and computing units, ii) microwave presence sensors ”break-the-beam,” iii) a low-cost server

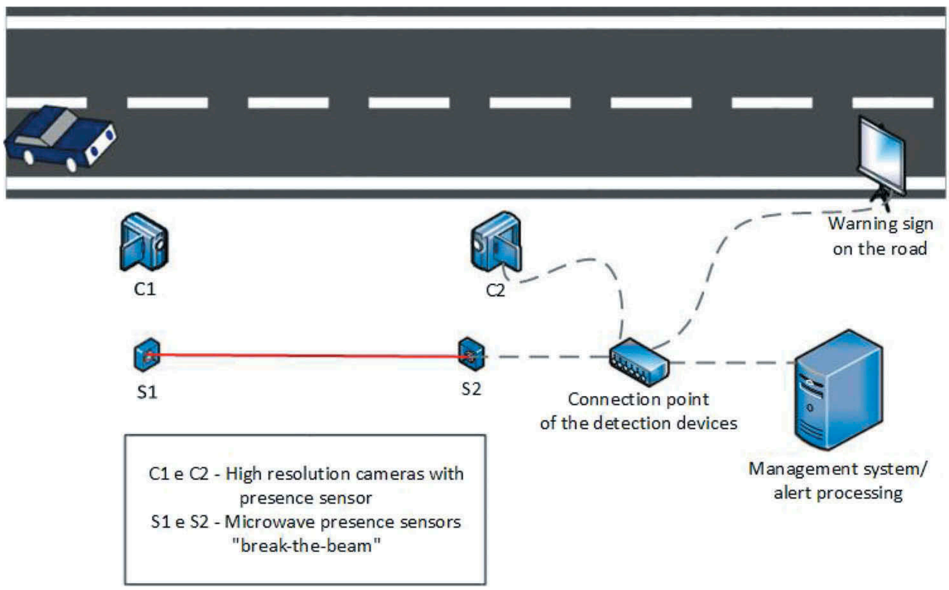


Figure 1. Proposed architecture.

with the management system and alert processing, iv) a warning signal on the road and v) a network equipment (a switch) that makes the interconnection between all of these devices.

The camera, together with the microwave sensors ”break-the-beam,” captures the images and passes them to their computing unit which, in turn, initiates the detection algorithms. Soon after, we use the machine learning algorithms to classify the images of moving objects and, if an animal with a risk of danger to the vehicles on the road is detected, a message will be sent to the server of the management system and alert processing. The server will process the message and send an alert to the warning signal installed on the road.

Detecting Animals Using Machine Learning

This section describes the method to identify an animal on roads by using images. Figure 2 shows the methodology in three sub-sections.

Five approaches (Figure 3) were developed block-based proposed in (Souza et al. 2015), in which features are extracted from square areas and feed the model. These five different approaches have been created to traverse the image from end to end so that the size and shape of the classification block and context will vary to investigate which approach manages to extract the characteristics of the images better. The result is used for the classification of the square regions (animal or non-animal). We applied a pair of blocks of different sizes: a smaller non-overlap, known as classification block (ClaB), whose area is classified by the

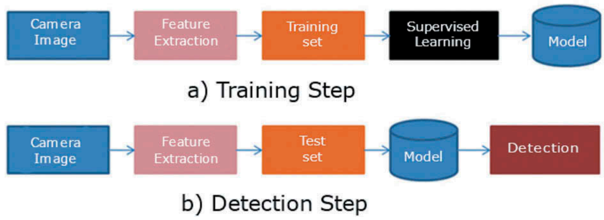


Figure 2. Methodology. Data flow is represented by arrows.

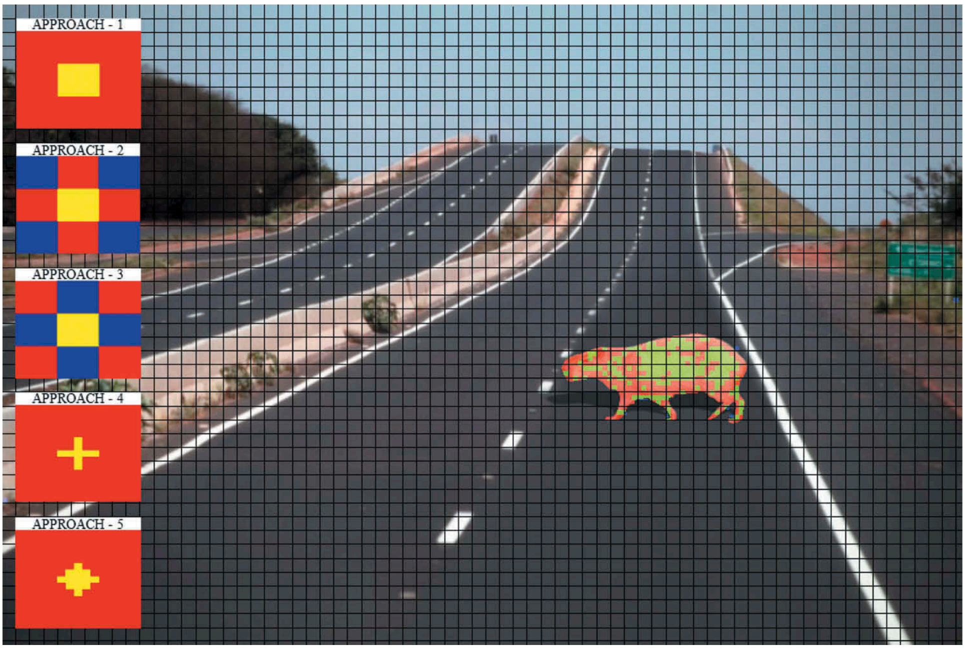


Figure 3. Five distinctive approaches are proposed.

model, and a larger contextual block (ConB), whose features are expected to yield contextual cues. The same features are extracted in both blocks and concatenated as input to a model.

Approach 1 is based on the standard block of squares areas: a smaller that is the ClaB (yellow color) and a larger that is the ConB (red color). Approach 2 was developed on a cross, a smaller square area in the ClaB, and a larger on a cross-format to produce the contextual clues. Approach 3 was developed on the ConB square area edges to extract the contextual cues, and ClaB remains unchanged. The three initial approaches had no change in the ClaB, only in the ConB format. Approach 4 was developed on a cross in the ClaB and the ConB in a square area to extract context. Approach 5 was developed on a target in the ClaB, and a square area ConB. For all, the contextual blocks are larger, and the classification block is in the center. ClaB is limited to 5×5

pixels, and ConB is limited to a 15 × 15 pixels area to find the better animals classification results. We implemented the animal's detection methodology presented using the Python 2.7 programming language.

Feature Extraction

To obtain the features, four color spaces were chosen: RGB, LAB, HSV, GRAYSCALE, and ENTROPY (an image segmentation algorithm). With these space colors, we process all the possible combinations without repetition between them; and we have reached the five features.

The RGB color space (Color Feature Extraction 2000) is the most common one used for images on a laptop because the computer display is using the combination of the primary colors (red, green, blue). Each pixel in the screen is composed of three points stimulated by red, green and blue. However, RGB space is not suitable for dealing with shadow and darkness. We prefer to transform the image data of RGB to other uniform space.

LAB (Schwiegerling 2004) is a second attempt at providing a perceptually uniform color space. In this color space, the distance between two points also approximately tells how different the colors are in luminance, chroma, and hue. Another way to characterize a color is concerning the HSV model (Gonzalez, Woods, and Edd 2004). The hue (H) of a color refers to which pure color it resembles. The saturation (S) of color describes how white the color is. The value (V) of a color, called its lightness, describes how dark color is.

Grayscale (Gonzalez, Woods, and Edd 2004) is a range of shades of gray without apparent color. The darkest possible shade is black, which is the total absence of transmitted light. The lightest possible shade is white, the total transmission of light at all visible wavelength. Intermediate shades of gray are represented by equal brightness levels of the three primary colors (red, green and blue) for transmitted light, or equal amounts of the three primary pigments (cyan, magenta, and yellow) for reflected light.

The entropy of features (De and Sil 2012) is calculated by forming a feature matrix (M) with the number of rows corresponds to the number of training images and the number of columns representing the dimension of each feature. The information in each feature is obtained using Shannon entropy (Gonzalez, Woods, and Edd 2004) as:

$$E_j = -\left(p_j \quad log \quad p_j\right)$$

where p_j represents the probability of occurrence of j^{th} feature in a training image. To compute the entropy of each feature, feature matrix M is considered a representation of an image and each element of M denotes pixel value of image. The M matrix has been scanned from left-hand top corner pixel to right-hand bottom corner pixel, each pixel is considered asm reference pixel that corresponds to a specific in an image.

Training and Testing Set

To perform the construction of a sample database, the images were manually classified and only features obtained from the pairs of blocks in which each classification block belongs to the same class (using 5×5 classification block, all of the 25 pixels should be equally classified). Then, a division occurs in the initial dataset, where specific amounts of images are separated for training and tests. Thus verifying that using the training set, it is possible to classify a new image input. This demonstrates the significant difference between the training set and tests. The samples are obtained and normalized with the following formula:

$$f_{norm} = \frac{f - m(\mathbf{f})}{std(\mathbf{f})}$$

where f_{norm} was the normalize value, f is the feature values, $\mathbf{f}$ is the vector containing all training values of a certain feature and m and std return the mean and standard deviation of a vector, respectively. The testing set is normalized by the same parameters m and std calculated in the training set.

Supervised Learning

We have selected two supervised ML algorithms based on their precise generalization performances: K-Nearest Neighbors (KNN) (Shalev-Shwartz and Ben-David 2014) and Random Forest (RF) (Shalev-Shwartz and Ben-David 2014).

KNN is instance-based learning, where the function is only approximated locally, and all computation is deferred until classification. KNN is the simplest of all algorithms, which classifies a test sample according to the class of its K nearest training samples (Figure 4). K nearest is defined according to a distance metric (our case, Euclidean distance).

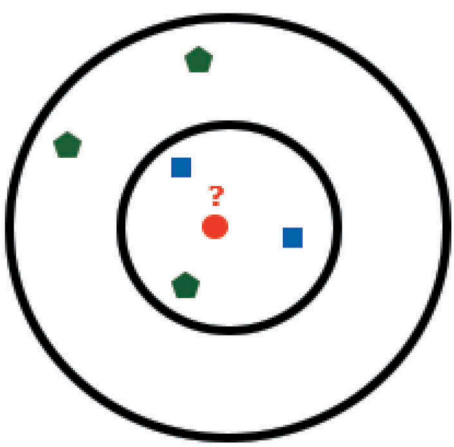


Figure 4. KNN with square and pentagon classes. Red circle is new query classified square with $K = 3$ and pentagon $K = 5$.

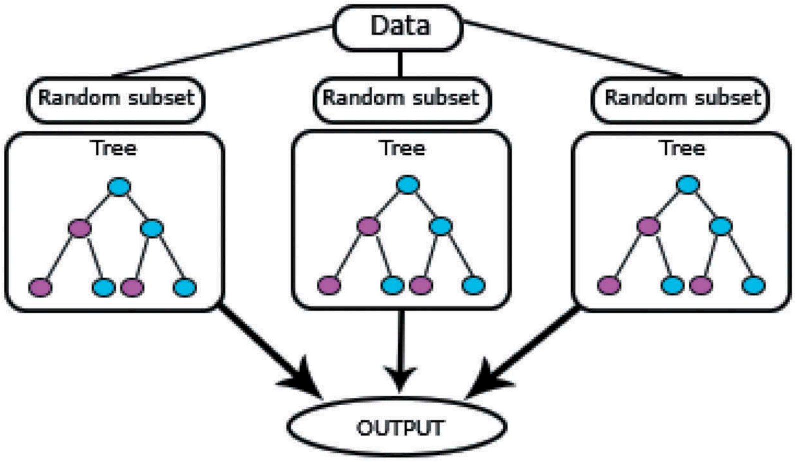


Figure 5. RF: Subsets of the trees are created and the result is a combination of the trees outputs.

RF generates randomness for model when the trees grow. Instead of finding the best feature of splitting a node, it looks for best among a random subset of resources. This procedure creates a wide diversity, which in most cases results in a better model. Figure 5 shows the RF for multi-class classification.

Results

This section describes the implementation of the Client/Server application and the results of the animal's detection system using the two ML algorithms.

Communication

The communication between the detection system and the warning signal to drivers needs a simple structure. Detections made by the cameras and presence sensors are performed and sent to a processing system which, in turn, does all the necessary processing and sends the message "Be careful! Animal on the Road" signal to the sign installed on the road. Since our goal is to provide a cheap and straightforward implementation, we decided to use a Raspberry Pi as processing system. This device will host a Client/Server application to communicate with cameras and sensors.

For the development of the Client/Server application of this work, Java 8 technology and Netbeans IDE (version 8.2) were used on a 64-bit Windows 10 Home machine. The hardware configurations of this machine are as follows: 2.50 GHz Intel Core i7 processor (3rd generation), 6 GB RAM and 1 TB HD. Our idea is to assess the communication between the camera modules and the processing system that will send the warning messages to the sign. Next, the application was installed on a Raspberry Pi 3 B+.

We developed a TCP client/server implementation to manage the communication between the camera module (client) and the processing system (Raspberry Pi – server). Socket implementations were used to perform the interprocess communications of the client and the server. The server waits for requests coming from the network and, based on these requests, generates some processing. After processing, the server may return some information to the requesting client application. We use multiple threads implementations at the server for the server to receive requests from one or more clients and treat them concurrently without having to queue them, allowing a faster delivery and reception of the messages coming from the animal detection systems.

The following steps describe the communication between the processing system (Raspberry Pi) and the client (camera module):

(1) Server (Raspberry Pi) starts and waits for client requests;
(2) Client (camera module itself or a computing unit) sends a message to the server informing that an animal was found near the road;
(3) Server returns an ”OK” message so that the client knows that the message was received with success by the server;
(4) Server sends a message to the sign, so it displays “Be careful! Animal on the Road” warning;
(5) The thread is closed by the server, and the connection is closed too.

The server (Raspberry Pi) starts at TCP port 4242 and waits for new client requests (camera module itself or a computing unit). The client then sends a connection request to the server as ”localhost” on the same port the server is listening (4242). The connection is established, and the client sends a message: ”Animal on the Road.” The server receives this message and sends an ”OK” back, signaling the successful treatment of this message and closes the connection with this client. Note that only the server is not terminated and it is still waiting for new requests on port 4242.

Animal Detection

We implemented the described system and tested the performance using the synthetic images with an image editing software to evaluate the capability and performance of the methodology. Synthetic images were created from selected images of animal species that occur in the region, all of which were produced at a resolution of 984×656 pixels.

Training and Model Selection

The dataset consists of 20 images, which we separated five images for training and one image for the test. The features are extracted from all the blocks, being classified according to a classifier and result is a classified image. The assessment is performed using the Ground Truth (GT), to differentiate which pixels are animals and what is not an animal.

The set of training data and tests were applied for each combination of color spaces and each of the approaches. Two ML algorithms were employed, with all five best combinations of color spaces and five contextual approaches. The first experiments focused on finding a model that obtained the highest F-measure result. In all of them, a 15×15 contextual block and a 5×5 classification block were used. The KNN tests the k-neighbors were set equal to 3, RF number of trees in the forest was set equal to 100. The predictive performance of KNN and RF is shown in Tables 1, 2, 3 and 4.

Table 1 refers the KNN using three approaches in the ConB. Normal contextual block offered best F-measure in the test set.

Table 2 refers the KNN using three approaches in the ClaB. Target classification block offered best F-measure in test set.

Table 3 refers the RF using three approaches in the ConB. Normal contextual block offered best F-measure in test set. Table 4 refers to RF using three approaches in the ClaB. Cross classification block offered best F-measure in test set.

The Best Approaches

We proposed five different approaches to traverse the image. After performing all the experiments, the two best approaches that obtained a more

Table 1. F-measure of the ConB approaches using KNN.

KNN	Contextual block		
Color spaces	Normal	Cross	Edges
GRAY + ENTROPY	**0.6113**	0.5525	0.5823
LAB + RGB + HSV + ENTROPY	0.5092	0.5112	0.4747
LAB + RGB + ENTROPY	0.5055	0.5169	0.4587
GRAY + LAB + RGB + ENTROPY	0.5011	0.5120	0.4548
GRAY + LAB + RGB + HSV + ENTROPY	0.5003	0.5037	0.4623

Table 2. F-measure of the ClaB approaches using KNN.

KNN	Classification block		
Color spaces	Normal	Cross	Target
GRAY + ENTROPY	0.6133	0.6219	**0.6243**
LAB + RGB + HSV + ENTROPY	0.5092	0.5278	0.5199
LAB + RGB + ENTROPY	0.5055	0.5215	0.5144
GRAY + LAB + RGB + ENTROPY	0.5011	0.5196	0.5196
GRAY + LAB + RGB + HSV + ENTROPY	0.5003	0.5190	0.5162

Table 3. F-measure of the ConB approaches using RF.

RF	Contextual block		
Color spaces	Normal	Cross	Edges
GRAY + ENTROPY	**0.5601**	0.5204	0.4702
LAB + RGB + HSV + ENTROPY	0.1555	0.2232	0.1299
LAB + RGB + ENTROPY	0.4013	0.3809	0.1670
GRAY + LAB + RGB + ENTROPY	0.3189	0.2902	0.2110
GRAY + LAB + RGB + HSV + ENTROPY	0.1707	0.2086	0.1165

Table 4. F-measure of the ClaB approaches using RF.

RF	Classification block		
Color spaces	Normal	Cross	Target
GRAY + ENTROPY	0.5601	**0.5891**	0.5845
LAB + RGB + HSV + ENTROPY	0.1555	0.1999	0.1693
LAB + RGB + ENTROPY	0.4013	0.3274	0.3612
GRAY + LAB + RGB + ENTROPY	0.3189	0.2345	0.4852
GRAY + LAB + RGB + HSV + ENTROPY	0.1707	0.1756	0.1430

considerable F-measure were represented in Figure 6. The experimental results were better when we elaborated new types of approaches in the ClaB. Approach 5 obtained the best F-measure of 0.6243 and the second position we had the Approach 4 with an F-measure of 0.6219. Figure 7 shows the set of synthetic images that were used to train the supervised ML algorithm and to perform the tests to obtain the F-measure metric.

Approach 5 with a target format in the classification block was efficient. After the experiments with each of the supervised algorithms, KNN showed the best F-measure.

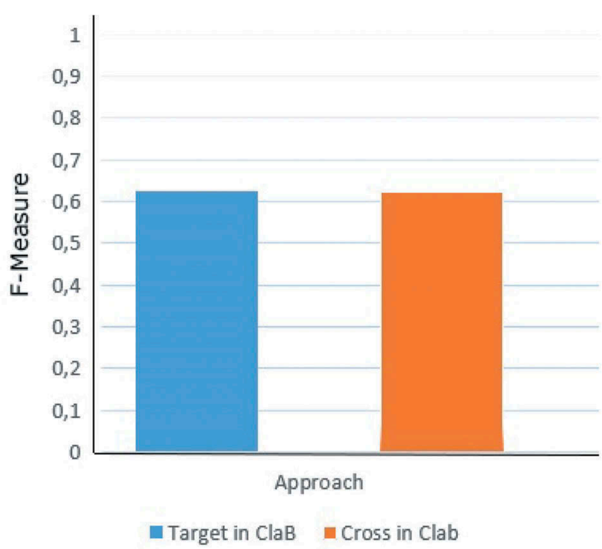
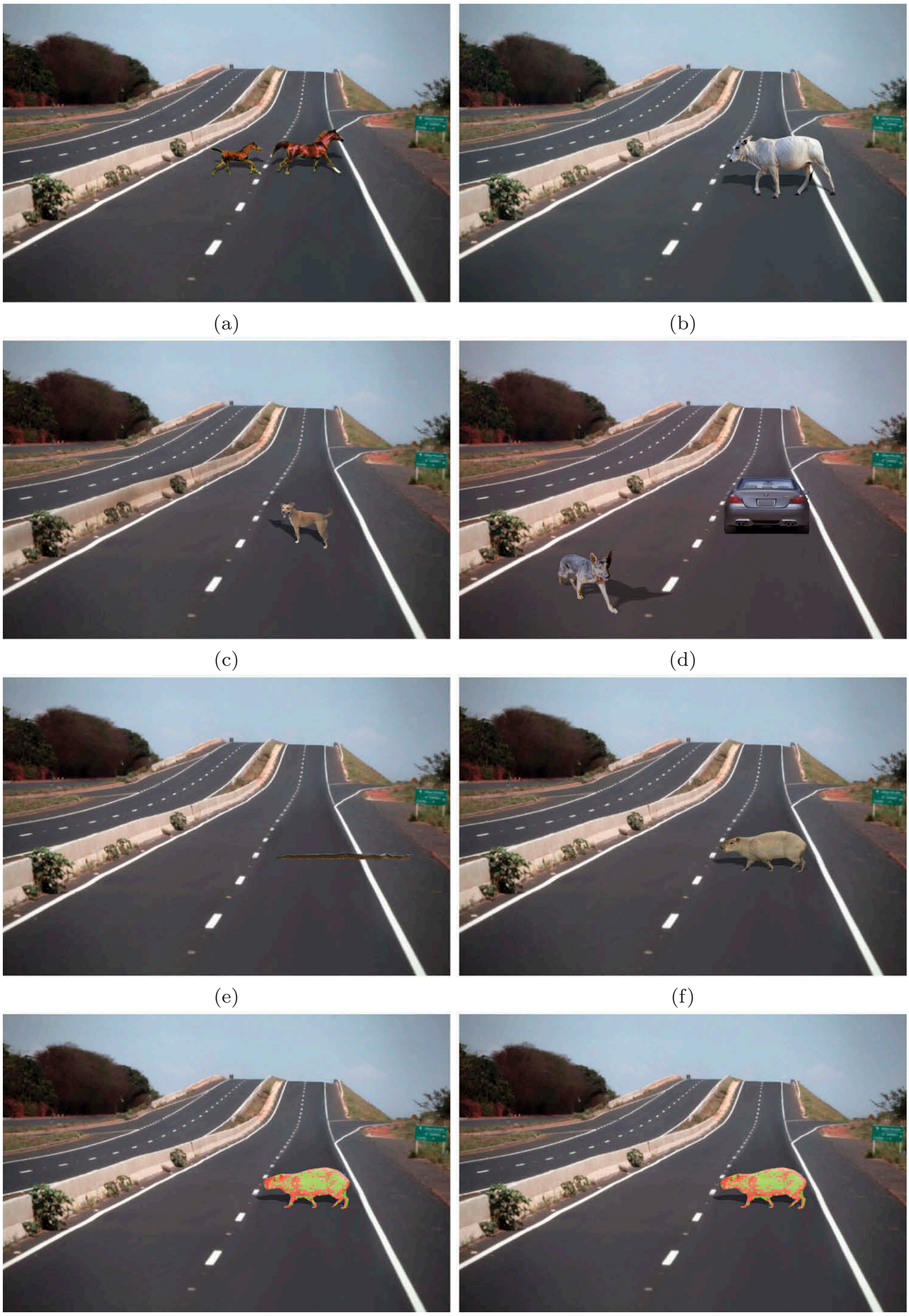


Figure 6. Comparison of the two best approaches.



(a) (b) (c) (d) (e) (f) (g) (h)

Figure 7. (a) Horse, (b) Bull, (c) Dog 1, (d) Dog 2, and (e) Snake (training data). (f) Capybara (test data). (g) First best result from Table 2 on test set. (h) Second best from Table 2 on test set. The pixels painted in green shows the ML algorithm classified as animal and hit. The pixels painted in blue show algorithm classified as animal but missed. Pixels painted in red shows the demarcated area as animal in the GT.

Conclusions

This paper proposes an architecture to warn the drivers on roads and also a system to detect the presence of animals on roads; capable of learning features of each pixel to differentiate between what is an animal and non-animal using synthetic images. KNN and RF were used to process the features, and the classification results were compared to identify which type of approach and ML was best according to the F-measure. The purpose is the detection of animals on Brazilian roads; however, the same technique can be readily applied to another goal that has any visually observable condition.

We believe that the contributions of this paper are a crucial step for the long-term preservation of species and the decrease in vehicle accidents, causing a reduction of expenses to the public coffers and avoiding deaths of drivers and passengers.

As future work, we will install cameras on roads to get a set of real images. We will also use other techniques as the deep learning to check which method is best for our case in several conditions. Finally, we will develop an algorithm that will make the difference between these detected animals and this application works efficiently in the real-time. This information generated can help in a better understanding of the behaviors of each species of animal that lives in the fauna and facilitating the work of professionals or other areas as biologists.

Acknowledgments

This work was supported in part by the Federal University of Uberlândia and in part by the FAPEMIG and CNPq under Grant 400699/2016-8 (Brazilian Agencies).

References

AMBEV. 2014. *Retrato da Segurança Viária no Brasil*, Observatório Nacional de Segurança Viária.

Carvalho, C. F., A. E. I. Custódio, and M. Oswaldo Junior. 2015. Wild vertebrates roadkill aggregations on the br-050 highway, state of Minas Gerais, Brazil. *Bioscience Journal* 31 (3). doi:10.14393/BJ-v31n3a2015-27468.

CBEE. 2017. Atropelmetro. [Online]. cbee.ufla.br/portal/index.php.

De, I., and J. Sil. 2012. Entropy based fuzzy classification of images on quality assessment. *Journal of King Saud University - Computer and Information Sciences* 24:165–73. doi:10.1016/j.jksuci.2012.05.001.

Forslund, D., and J. Bjrkefur. 2014. Night vision animal detection. IEEE Intelligent Vehicles Symposium Proceedings, 737–42.

Gonzalez, R. C., R. E. Woods, and S. L. Eddins. 2004. Digital Image processing using MATLAB, Pearson Prentice Hall.

Grace, M. K., D. J. Smith, and R. F. Noss. 2017. Reducing the threat of wildlife-vehicle collisions during peak tourism periods using a roadside animal detection system. *Accident Analysis and Prevention* 109:5561. doi:10.1016/j.aap.2017.10.003.

Huijser, M. P., M. Christa, O. Mattias, and S. Martin. 2015. Wildlife warning signs and animal detection systems aimed at reducing wildlife vehicle collisions, 198–212. John Wiley & Sons.

Huijser, M. P., J. Muir, and U. C. Davis. 2003. Overview of animal detection and animal warning systems in North America and Europe. International Conference on Ecology and Transportation.

Li, Y., F. Ben, Z. Wang, C. Li, and R. Venkatesan. 2011. Hierarchical Wild Animal Detection and Notification System based on Wireless Sensor Networks — Design and Implementation, Natural Sciences and Engineering Research Council.

Matuska, S., R. Hudec, M. Benco, P. Kamencay, and M. Zachariasova. 2014. A novel system for automatic detection and classification of animal. 76-80, ELEKTRO.

Matuska, S., R. Hudec, P. Kamencay, and T. Trnovszky. 2016. A video camera road sign system of the early warning from collision with the wild animals. *Civil and Environmental Engineering* 12:42–46. doi:10.1515/cee-2016-0006.

Pesquisa CNT de rodovias 2015: relatrio final. 2015. [Online]. pesquisarodoviascms.cnt.org.br/Relatorio%20Geral/PESQUISA_CNT2015_BAIXA.pdf.

Schwiegerling, J. 2004. *Field guide to visual and ophthalmic optics*. International Conference on Optical Instruments and Technology.

Shalev-Shwartz, S., and S. Ben-David. 2014. *Understanding machine learning: From theory to algorithms*. NY, USA: Cambridge Uni. Press.

Sharma, S. U., and D. J. Shah. 2017. A practical animal detection and collision avoidance system using computer vision technique. *IEEE Access* 5:347358. doi:10.1109/ACCESS.2016.2642981.

Souza, J. R., C. C. T. Mendes, V. C. Guizilini, K. C. T. Vivaldini, A. Colturato, F. T. Ramos, and D. F. Wolf. 2015. Automatic Detection of Ceratocystis Wilt in Eucalyptus Crops from Aerial Images, IEEE International Conference on Robotics and Automation (ICRA), 3443-3448, 2015.

Wu, J.; Lim, J.; Kankanhalli, M. S.; Hong, D. Color Feature Extraction. 2000. 4967. Boston, MA: Springer US.

Zhou, D. 2014. *Real-time animal detection system for intelligent vehicles*. Thesis to the Faculty of Graduate and Postdoctoral Studies, University of Ottawa.