XLI CILAMCE
XLI Ibero-Latin American Congress on Computational Methods in Engineering
16-19 November, 2020 FOZ DO IGUAÇU PARANÁ - BRAZIL

A Bag of Visual Words-Based Approach to Identify Scenarios Suspected of Being Breeding Sites of the *Aedes aegypti* Mosquito from Aerial Images Acquired by UAVs

Gustavo A. Lima[1], Daniel T. Bravo[1], Sidnei A. Araujo[1]

[1] *Programa de Pós-graduação em Informática e Gestão do Conhecimento (PPGI), Universidade Nove de Julho – UNINOVE, Rua Vergueiro, 235/249 – Liberdade, São Paulo/SP, Brasil*
gustavoaraujo59@hotmail.com, danieltbravo@uni9.pro.br, saraujo@uni9.pro.br

Abstract. In addition to the programs of the Brazilian Ministry of Health to prevent and combat the *Aedes aegypt* mosquito, the use of unmanned aerial vehicles – UAVs (known as drones) has proved to be an important alternative to assist the work of health surveillance teams. However, the analysis of aerial images acquired by such equipment are usually manual, which can be time-consuming for health workers. Thus, it becomes important the proposition of computational approaches able to recognize and interpret patterns in such images. This work proposes a computer vision approach to identify scenarios that represent potential breeding sites of the *Aedes aegypti* mosquito from aerial images acquired by UAVs. The proposed approach employs the Bag of Visual Words – BoVW technique combined with the Support Vector Machine – SVM classifier (BoVW + SVM), taking into account two descriptors based on the SIFT – Scale Invariant Feature Transform algorithm (Transformed color SIFT and RGB-SIFT), and the descriptors Color Level Co-occurrence Matrices – CLCM and Color Histograms. For conducting the experiments we compose a database of images, acquired in urban regions of the city of São Paulo, which contemplate real and simulated scenarios suspected of being breeding sites of the mosquito (gutters and roofs with accumulation of objects, open-air garbage containing old tires, old tires, pet bottles, plastic and paper packaging and other open containers that can accumulate water). The results obtained in the experiments with BoVW+SVM, in terms of hit rate, were: RGB-SIFT (91,37%), Transformed color SIFT (85,34%), CLCM (64,65%) and Color Histograms (86,20%).

Keywords: Drone, Pattern Recognition, Computer Vision, Bag of Visual Words, Support Vector Machine.

1 Introduction

The epidemics of dengue, zika, chikungunya and urban yellow fever, caused by the *Aedes aegypti* mosquito, have been causing great concern to health authorities, not only in Brazil, but worldwide. According to the Brazilian Ministry of Health, in 2019 (until August 24), 1,439,471 dengue cases were recorded across the country, with a 599.5% increase compared to the same period from 2018 (205,791). The incidence rate, which considers the proportion of cases per inhabitant, is 690.4 cases per 100 thousand inhabitants (MS [1]).

The Brazilian Ministry of Health, in order to combat the *Aedes aegypti* mosquito, launches advertising campaigns annually that involve state and municipal managers, and the population. In 2019, the federal government spent about R$ 22 million on advertising campaigns of this nature.

The fight against the *Aedes aegypti* mosquito has demanded countless other efforts, since the programs for developing information campaigns and mobilizing people do not always achieve the desired effect. In this sense, several studies have been carried out to speed up the search for breeding sites of the mosquito, mainly in places of difficult access for health surveillance agents and residents, with the use of unmanned aerial vehicles (UAVs) for the acquisition of aerial images in regions with a higher incidence of diseases (Diniz & Medeiros [2]).

UAVs have been widely used in recent years in tasks that require the acquisition of aerial images, which are analyzed manually or automatically. Among the various applications are: remote sensing (Aguirre-Gómez *et al.*

[3], Albuquerque *et al*. [4]), health (Capolupo *et al*. [5]; Passos *et al.* [6]), precision agriculture (Alves, Ferreira and Custódio. [7]) and combating the mosquito breeding sites (Agrawal et al [8]; Mehra *et al*. [9]; Diniz & Medeiros *et al*. [2] and Bravo. [10]).

Although there are works in the literature proposing the use of UAVs to identify objects and scenarios suspected of being possible mosquito breeding sites, in most of them manual image analysis is carried out, as in Diniz and Medeiros *et al*. [2].

Only a few works, such as Agrawal *et al.* [8], Mehra *et al*. [9] and Bravo [10] consider the automatic analysis of the images. However, the approaches proposed in the work of Agrawal *et al*. [8] and Mehra *et al*. [9] are incomplete with respect to the spatial location of objects and scenarios in the images.

Bravo [10] proposed approaches for automatic detection and location of objects and scenarios suspected of being breeding sites of the *Aedes aegypti* mosquito in aerial images acquired by UAVs. The author explores an approach based on Bag of Visual Words (BOVW) which, although it is fast and robust, the best success rate, in terms of the mAP-50 metric, was 64.53%, which represents a very lower result compared to the results obtained in related applications in the literature. In this sense, the author suggests in his work some ways to improve this approach. He also suggests the development of a computer vision system (SVC) that can be embedded on UAV. This work explores a computer vision approach to identify scenarios that represent potential breeding sites of the Aedes aegypti mosquito from aerial images acquired by UAVs. The proposed approach employs the Bag of Visual Words – BoVW technique combined with the Support Vector Machine – SVM classifier (BoVW+SVM), taking into account two descriptors based on the SIFT – Scale Invariant Feature Transform algorithm (RGB-SIFT and Transformed color SIFT), besides the descriptors Color Level Co-occurrence Matrices – CLCM and Color Histograms.

2 Theoretical Background

2.1 Aedes aegypti: Behavior and favorable places for its proliferation

The *Aedes aegypti* mosquito has daytime habits. In residences it can be found preferably in shaded and dark places. The male mosquito feeds on plant saps, while the female needs human blood for the maturation of eggs, which are deposited separately on the internal walls of objects, close to water surfaces, a place that offers them better survival conditions (MS. [1]).

Regarding the flight capacity, it is known that the mosquito has the possibility of access to heights, for example, reaching water tanks, gutters and terraces of urban buildings. However, its flight potential would not reach a four-story building. Despite this, it can reach higher heights if it is housed in elevators, packaging in general, toys, toolboxes and a multitude of other objects that can take it to the roof of any building.

According to the Ministry of Health, there are several locations (scenarios) that represent potential breeding grounds for the *Aedes aegypti* mosquito, such as water tanks, water accumulated on the slab, garbage in plastic bags in open bins, tires, empty bottles, dishes under pots of plants or some other object that can accumulate rainwater, as the examples illustrated in Figure 1.

(a) (b) (c)

Figure 1: Examples of possible breeding sites of the *Aedes aegypti* mosquito.

Therefore, it is important to be careful with all the places that can accumulate water, since the eggs are resistant to desiccation and can survive in the environment for up to 450 days, being just a small amount of water as a small puddle enough for the larvae to hatch. If the container containing the water is well treated and with the recommended amount of chlorine, the mosquito does not develop, due to the chlorine functioning as a repellent.

In this work suspected scenarios are characterized by the existence of inorganic garbage located outdoors, comprising objects that can accumulate water such as old tires, pet bottles, plastic and paper packaging, among others.

2.2 Computer Vision

Computer vision (CV) can be defined as a subarea of image processing that studies the development of methods and techniques that enable a computer system to interpret images. In other words, a computer vision system (CVS) aims to provide a machine with the capacity to describe a scene contained in a digital image (Gonzalez & Woods. [11]).

In addition, an efficient CVS must be able to extract a set of features that accurately describe a scene and **is** small enough to reduce processing time and, at the same time, efficient to enable the construction of practical applications such as vision systems for industrial robots, autonomous vehicles, detection of events in surveillance systems, automated reading of license plates of vehicles, industrial inspection and recognition of biometric patterns, among others (Zhang & Lu. [12]).

In recent years, the interest in the field of computer vision has been growing and the advances achieved have enabled the development of practical applications in various segments of the industry such as automotive, pharmaceutical and cosmetics, ceramic, metal-mechanics, textiles, electronics and the agribusiness (Rodríguez-Pulido *et al*. [13]).

2.3 Image descriptors

A descriptor has the objective of returning important features (attributes) of an image through a set of values. This set is called a "feature vector" and is used by the classification algorithms to classify the objects contained in the image in their respective classes (Silva *et al*. [14]). The feature recognition in images is important to access visual information at the level of objects and scene types.

In general, pattern recognition systems in images and video use machine learning based on the features extracted from the images that are used in training to distinguish categories of objects and scenes. However, there can be great variations in the conditions of viewing from one image to another, complicating the description process and, consequently, the task of recognizing objects and scenes. In addition, changes in lighting can greatly affect the performance of the recognition algorithm if the descriptors used are not robust to these changes. In this work, the four descriptors were used to extract features presented below, being the first two invariant to changes in lighting:

RGB-SIFT: is based on the Scale Invariant Feature Transform (SIFT) descriptor. RGB-SIFT extracts from each color channel of the RGB system a vector of features with size of 128 bytes (Koen *et al*. [15]). This descriptor has been used in several works in the literature, producing good results.

Transformed Color Sift (TC-SIFT): is invariant to scale, rotation and small changes in lighting. Just like RGB-SIFT, it produces a descriptor for the three RGB channels containing 384 features (Koen *et al*. [15]).

Color-Level Co-Occurrence Matrices (CLCM): is a powerful method for statistical analysis of images that considers color and texture information (Haralick *et al*. [16]). The primary method, called Gray Level Co-occurrence Matrices (GLCM), applied in the description of textures, takes in account the probability of two values of gray intensity being involved by a certain spatial relationship, defined by an angle (a) and a distance (d). From this matrix of probabilities, different statistical measures can be calculated in order to characterize the texture present in the image. Haralick *et al*. [16], for example, proposed 14 statistical measures, calculated from the co-occurrence matrices. The 6 most relevant measures, which were employed in this work, are: second angular momentum, entropy, contrast, variance, correlation, and homogeneity. The main difference between GLCMs (for gray level images) and CLCM (for color images) is that in the first method the matrices are calculated for a single channel, while in the second the matrices are calculated based on the relationship among the color components of the image. Thus, adopting a={0, 45, 90, 135} and d=1, an RGB image generates 27 CLCM matrices (4 for each individual channel and 15 for the 3 combinations of the color channels, that is, 12+15).

Color Histogram (CH): it considers the distribution of gray levels in the image are invariant to rotation and translation**.** It can be represented by a graph indicating the number (or percentage) of pixels in the image for each level of gray (Pedrini; Schwartz, [17]). In the case of color images, each color channel is quantized in a certain number of bins and the color count in each bin is used in the calculation of the histogram. In other words, for an RGB image, for example, the histogram of each one of the three-color channels is computed respecting the number of bins. In this work, for example, we considered 128 bins generating 3 histograms that compose the vector of 384 characteristics.

2.4 Bag of Visual Words (BOVW)

The BoVW technique, also known as Bag of Visual Features (BoVF), is used to robustly combine the features extracted from an image into a single feature vector. It is a very popular technique due to its simplicity, as it is based on the unordered representation of local descriptors applied to an image and are, therefore, conceptually and computationally simpler than many alternative methods (Bravo, [10]).

Figure 2 shows the steps of the BoVW technique, which are: (i) extracting the features of the images; (ii) generation of a dictionary of visual words through a grouping algorithm (for example, k-means); (iii) occurrence (frequency) count of each visual word contained in the image for the creation of visual word histograms that describe the images (Agrawal *et al*. [8]; Mehra *et al*. [9]).

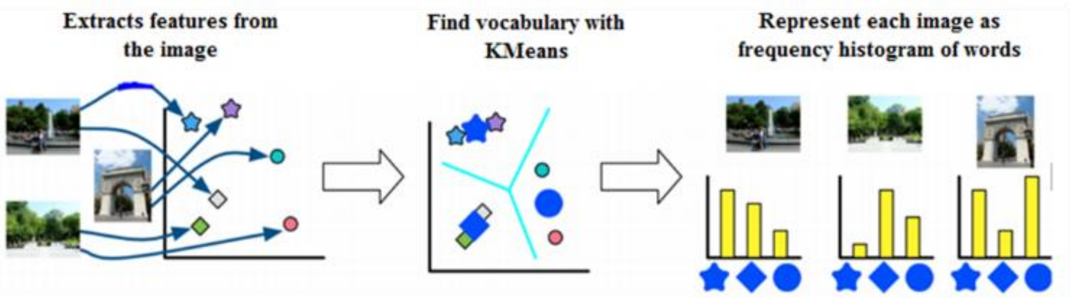


Figure 2: Steps of BoVW technique. Source: Mehra *et al*. [9].

The visual word dictionary is responsible for determining the features and patterns that represent the structure of an image. It is worth remembering that a universal dictionary cannot be generalized, that is, valid for all types of applications. Different domains need specific dictionaries, since the representation can change. In this case, the challenge is how to determine the ideal dictionary for each type of application according to the type of representation used to describe the features.

2.5 Support Vector Machine (SVM)

The SMV, introduced by Cortes and Vapnik [18], is a supervised learning method used mainly for classification and regression (Suykens & Vandewalle. [19]). The basic concept of SVMs comprises the construction of a hyperplane as a decision surface so that the margin of separation between classes is maximum. The objective of training through SVMs is to obtain hyperplanes that divide the sample in such a way that the generalization limits are optimized, with the points located close to these limits called support vectors (Bravo. [10]).

An example of data classification is illustrated in Figure 3, in which the method seeks to determine the parameters of a line (in other words, an optimal linear separator) in order to maximize the distance of the support vectors, that is, the best support vectors that maximizes the margin. However, there are situations in which this is not possible, and in this case, SVM is allowed to classify some data incorrectly, to a certain degree.

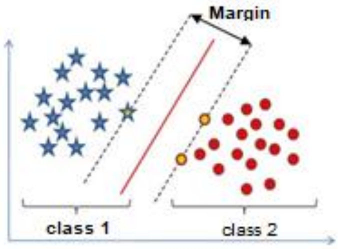
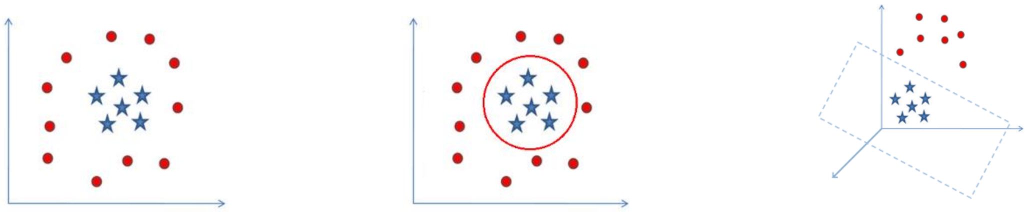


Figure 3: Exemple of classification using SVM.

There are many cases in which it is not possible to divide satisfactorily the training data by a hyperplane. An example is shown in Figure 4, in which the use of a curved boundary would be more appropriate in the class separation. In Figure 4a, it can be observed a case in which the data are arranged in a non-linear way. In order to separate them into two distinct classes, a curved boundary (red circle in Figure 4b) was defined, which subsequently allowed for linear division through the definition of a feature space (Figure 4c). SVMs deal with non-linear problems by mapping the training set from its original space, referred as input, to a new larger space, called the feature space.



(a) (b) (c)

Figure 1: (a) non-linear data set; (b) non-linear boundary in the input space; (c) linear boundary in the feature space.

SVMs have different kernels that are used to solve nonlinear space problems. The most used are: Polynomial (which manipulates a polynomial function whose degree can be defined during training); Sigmoidal (allows SVM to behave similarly to the MLP network); and Gaussian (SVM behaves like an RBF network). SVMs also have some regularization parameters, the main ones being described below:

C parameter. indicates to the SVM optimization how much you want to avoid misclassifying each training example. Large values of C indicates that the optimization will choose a smaller-margin hyperplane if that hyperplane does a better job in terms of classifying correctly the instances of training.

***Gamma* parameter**. Gamma parameter defines how far the influence of a single training example (represented by a pint in the feature space) reaches. A low gamma value means that all points have a greater reach and, inversely, a high gamma value means that all points have a close reach. If gamma has a very high value, the decision limit will depend only on the points that are very close to the class separation line, which effectively results in ignoring some of the examples that are very distant from the decision limit. This occurs because the closest points gain more weight and result in a wavy curve.

3 Proposed approach and composed database of images

3.1 Proposed approach to detect suspicious scenarios

The working of the approach proposed in this work based on BoVW is illustrated in Figure 5, and is described by the following steps:

Step 1. Extraction of features from subimages (windows) extracted from the training images. We consider the following descriptors: RGB-SIFT, TC-SIFT, CH and CLCM. The descriptors RGB-SIFT and TC-SIFT extract feature vectors of 128 bytes from the images. From the CLCM matrices, we compute 6 Haralick descriptors: second angular moment, entropy,

contrast, variance, correlation and homogeneity. The CH generates feature vectors of 384 elements (128 bins for each color channel of the RGB space).
Step 2. Creation of the visual words dictionary from the extracted features using the k-means algorithm. The dictionary size (number of clusters adopted in k-means) was empirically defined as 110. Each cluster represents a visual word.
Step 3. Representation of each subimage from the dictionary through histograms of visual words;
Step 4. Training the SVM classifier using the histograms of visual words;
Step 5. Classification of the windows extracted from each image belonging to the set of tests using the trained SVM.

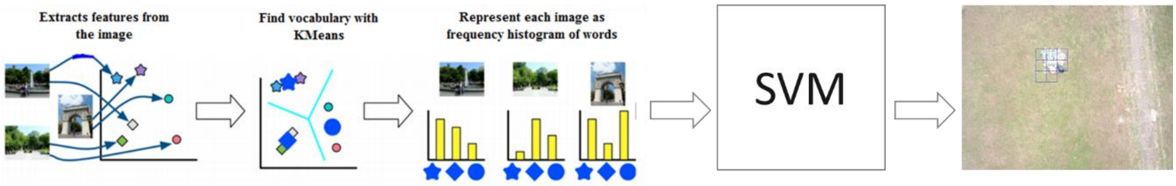


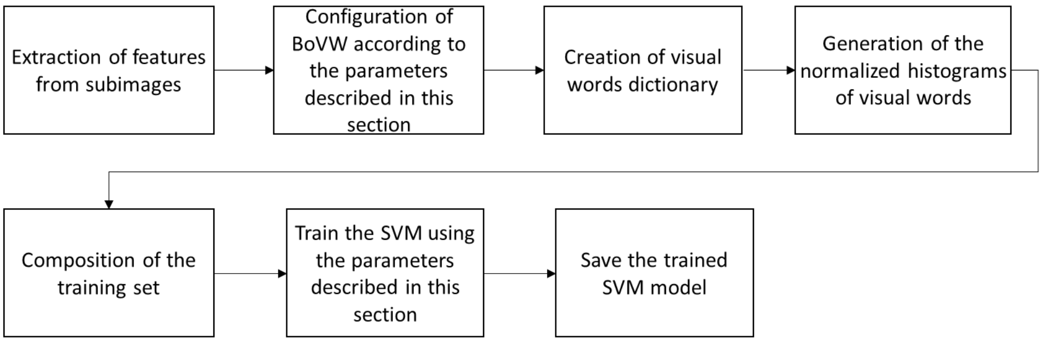
Figure 5: Steps of the employed procedure using BoVW technique. Source: Adapted from MEHRA *et al.*[9]

The SVM was trained with 1,386 sub-images of 200×200 pixels, manually extracted from 85 images from the DS1 and DS2 datasets (described below), which were separated for the training and validation steps. For the evaluation of BoVW+SVM, 15 other images from DS1 and DS2 were employed. The training instances (feature vectors extracted from the sub-images) were separated into two classes:

- class_0: non-suspect scenarios
- class_1: suspect scenarios

The accuracy of the SVM was obtained comparing the images classified automatically and those classified manually.

The following parameters were configured for SVM training: learning rate of 0.001, C = 1.0, Kernel = rbf, gamma = 0.001. The diagrams in Figure 6 detail the functioning of the training SVM (a) and evaluation of BoVW+SVM approach (b).



(a)

(b)

Figure 6: Diagrams of the operation of the proposed BoVW+SVM. (a) training and (b) evaluation.

As can be seen, the extraction of features and classification using BoVW+SVM are based on the sliding window strategy, in which each sub-image of 200×200 pixels extracted from an input image is classified.

3.2 Database of images

For conducting the experiments, we compose a database of images acquired in areas located in the city of São Paulo – Brazil. It contemplates two datasets, named as DS1 and DS2 which are detailed below.

The DS1 is composed by 119 RGB images of simulated scenarios acquired in an area of the University of São Paulo (USP). Such images, with resolution of 4000×3000 pixels, were acquired using a DJI Phantom 4 UAV equipped with a DJI 20 MP RGB camera. On flights, three distances above the ground were considered: 7, 10 and 13 m. These distances provide values of GSD ranging from 0.30 to 0.56 cm/px.

The DS2 is composed by 111 images with resolution of 3000×2250 pixels contemplating real scenarios. They were acquired in the districts of Guaianases and Ferraz de Vasconcelos, using the same DJI Phantom 4 UAV mentioned above, equipped with a GoPro HERO4 Silver camera. The flying altitude ranged from 3 to 5 m above the ground, providing GSD values below 0.30 cm/px.

4 Results

To evaluate the performance of BoVW+SVM in identifying suspicious scenarios, 15 images were used in the experiments and the results obtained for the 4 descriptors (RGB-SIFT, TC-SIFT, CH and CLCM), in terms of hit rate, are shown in Table 1. It is important to mention that the accuracy obtained in the training validation considering the 4 descriptors were: RGB-SIFT (95%), TC-SIFT (93%), CH (91%) and CLCM (88%).

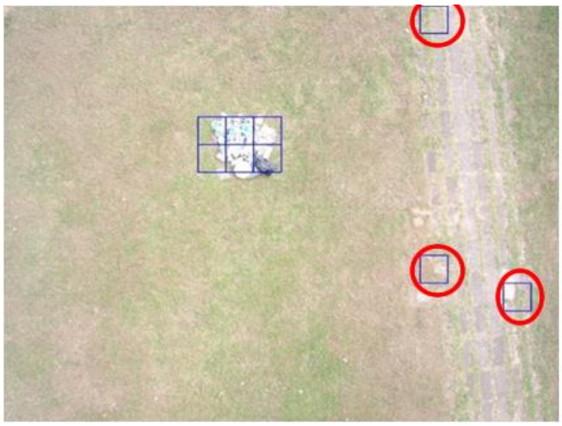
Table 1: Results of the detections made by BoVW+SVM on the 15 images used in the testing step.

Image	**Number of suspicious scenarios existing in the image**	**Correct detections (true positives - TP)**			
		RGB-SIFT	**TC-SIFT**	**HC**	**CLCM**
Img_01	6	6	6	6	5
Img_02	6	6	5	6	6
Img_03	8	8	6	6	4
Img_04	12	12	12	11	9
Img_05	9	7	8	8	5
Img_06	4	4	4	3	4
Img_07	4	4	3	4	2
Img_08	6	5	3	3	3
Img_09	6	6	6	5	4
Img_10	12	10	8	12	8
Img_11	8	7	6	7	5
Img_12	12	9	10	9	7
Img_13	7	7	7	7	4
Img_14	9	9	9	8	6
Img_15	7	6	6	5	3
Total	116	106	99	100	75
Hit rate (%)		91.38	85.34	86.21	64.66

From Table 1 one can see that the best results were achieved with the use of the RGB-SIFT descriptor. Adopting such descriptor, BoVW+SVM was able to correctly detect 106 out of 116 suspicious scenarios present in the 15 images analyzed, generating a hit rate of 91.38%. However, the total number of cases of false positives – FP (63) indicates the need for improvements. It can be observed that the images acquired at a higher distance from the ground tend to generate a greater number of FP and/or false negatives (FN), due to the fact that the scenarios become smaller, making it difficult to extract some important features from them, decreasing the generalization capacity of the classifier.

The figures 7 to 10 illustrate examples of images analyzed by BoVW+SVM considering the four descriptors, in which there are cases of TP, FP and FN.

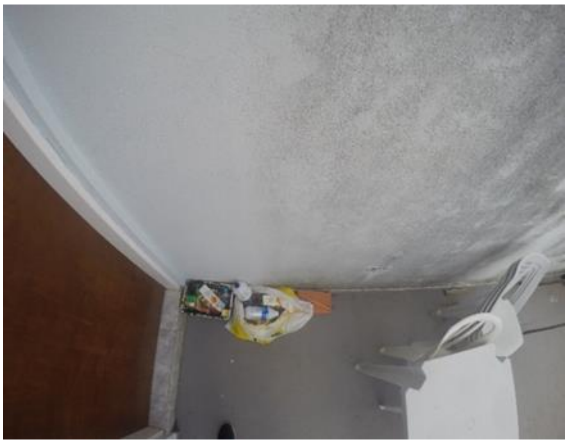
(a) input image



(b) output image

Figure 7: Result of BoVW+SVM considering the descriptor RGB-SIFT

In Figure 7 it is possible to notice that although all windows belonging to the scenario were classified correctly, there were 3 FP cases, highlighted by red circles in Figure 7b. In Figure 8, which shows a result achieved using the TC-SIFT descriptor, a case of FN is presented, highlighted with a white circle, where a window that was part of the scenario was not detected.

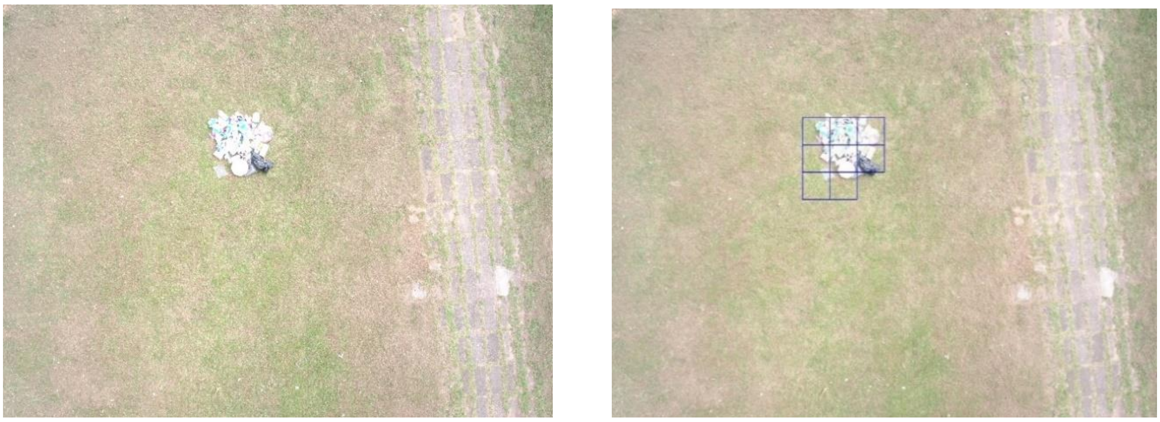


(a) input image

(b) output image

Figure 8: Result of BoVW+SVM considering the descriptor TC-SIFT

Figure 9 shows a case in which all windows belonging to a scenario were classified correctly using the HC descriptor. Finally, Figure 10 shows an example of the classification result considering the CLCM descriptor, which presented the highest number of FP.



(a) input image (b) output image

Figure 9: Result of BoVW+SVM considering the descriptor HC

(a) input image (b) output image

Figure 10: Result of BoVW+SVM considering the descriptor CLCM

Based on the accuracy obtained in the training validation, we expected that the hit rate of TC-SIFT was higher than the hit rate produced by CH, but this was not what happened in the evaluation of BoVW+SVM. In addition, we also expected better results for the first three descriptors. These problems can be related both to the generalization capacity of SVM and to the parameterization of BoVW, for example regarding the number of clusters for each descriptor. Another negative aspect is related to the number of cases of FP (RGB-SIFT=63, TC-SIFT= 91, CH=85 and CLCM=582). Based on these observations, new experiments will be carried out in future works, in order to mitigate and overcome such disadvantages.

5 Conclusions and future works

In this work we proposed an approach based on BoVW+SVM to detect scenarios representing potential breeding sites of the Aedes aegypti mosquito from aerial images acquired by UAVs. For extracting features from images, we consider the descriptors RGB-SIFT, TC-SIFT, CH and CLCM. The obtained results, in terms of hit rate (RGB-SIFT=91.38%, TC-SIFT=85.34, CH=86,21% and CLCM=64,66%), indicate that the proposed approach is a promising alternative to solve the investigated problem. However, the number of FP cases (RGB-SIFT=63, TC-SIFT=91, CH=85 and CLCM=582) demonstrate that improvements need to be made in the approach aiming to increase its effectiveness. Mitigating the improvements to be made, as well as conducting new experiments considering combinations of the employed descriptors, and a larger number of images in the evaluation procedure are in our plans for future works.

Acknowledgments

This work was supported by the FAPESP–São Paulo Research Foundation (Proc. 2019/05748-0), and by the CNPq–Brazilian National Research Council (research scholarship granted to S. A. Araújo, Proc. 313765/2019-7).

References

[1]M. Saúde, "https://www.saude.gov.br/noticias/agencia-saude/45257-ministerio-da-saude-alertapara-aumento-de-149-dos-casos-de-dengue-no-pais", Ministerio da saúde, 2019.
[2] M. T.Diniz & J. B. Medeiros, "Mapeamento de criadouros de reprodução de aedes aegypti na cidade de Caicó/RN com o auxílio de veículo aéreo não tripulado". *Revista GeoNordeste*, 2018.
[3] R. Aguirre-Gómez, O. Salmerón-García, G. Gómez-Rodríguez & A. Peralta-Higuera, "Use of unmanned aerial vehicles and remote sensors in urban lakes studies in Mexico". *International Journal of Remote Sensing*, 2016.
[4] R. W. Albuquerque, M. O. Costa, M. E. Ferreira, L. A. Jorge, L. H. Sarracini & E. O. Rosa., "Uso do índice MPRI na avaliação de processos de Restauração Florestal (RF) utilizando sensor RGB a bordo de VANT quadricóptero". *Simpósio Brasileiro de Sensoriamento Remoto (SBSR)*
[5] A. Capolupo, S. Pindozzi,C. Okello & L. Boccia. "Indirect field technology for detecting areas object of illegal spills harmful to human health". *applications of drones, photogrammetry and hydrological models*.
[6] W. L. Passos, T. M. Dias, H. M. Alves Junior, B. D. Barros, G. M. Araujo, A. A. Lima & S. Lima Netto, "Acerca da Detecção Automática de Criadouros do Mosquito Aedes aegypti". *Anais do XXXVI Simpósio Brasileiro de Telecomunicações e Processamento de Sinais (SBrT)*.
[7] M. O. Alves, R. V. Ferreira & V. B. Custodio, "Interpretação de imagens de drone e do sensor OLI/ Landsat 8 para identificação de pragas e doenças na cana-de-açúcar". *Anais do XVIII Simpósio Brasileiro de Sensoriamento Remoto (SBSR)*.
[8] A. Agrawal, U. Chaudhuri, Chaudhuri & G. Seetharaman, "Detection of potential mosquito breeding sites based on community sourced geotagged images". *Spie. 2014*
[9] M. Mehra, A. Bagri, X. Jiang & J. Ortiz., "Image analysis for identifying mosquito breeding grounds". *IEEE International Conference on Communication and Networking*, 2016
[10] D. T. Bravo. Identificação automática de possíveis criadouros do mosquito Aedes aegypti a partir de imagens aéreas adquiridas por VANTs. *Tese submetida ao curso Doutorado em Informática e Gestão do Conhecimento, como requisito para título de doutor em Informática e Gestão do Conhecimento*, 2019.
[11] R. C . Gonzalez, R. E. Wood, Digital Image Processing. Massachusetts. *São Paulo: Edgard Blucher*, 2002.
[12] D. Zhang; Dengsheng & G. Lu., "Review of shape representation and description techniques". *Pattern Recognition*, Vol. 37, n.1, pp. 1-19, 2004.
[13] F. Rodriguez-Pulido, "Analysis of food appearance properties by computer vision applying ellipsoids to colour data". *Computers and Electronics in Agriculture*, v. 99, pp. 108–115, 2013.
[14] R. Silva, K. Aires, T. Santos, K. Abdalla, & Veras, R., "Segmentação classificação e detecção de motociclistas sem capacete". *XI Simpósio Brasileiro de Automação Inteligente (SBAI)*.
[15] Koen, T. Gevers & Snoek, C., "Evaluating Color Descriptors for Object and Scene Recognition", *IEEE* 2010.
[16] R. M. Haralick & K. Shanmugam, I. Dinstein, Textural Features for Imagem Classification, 1973.
[17] H. Pedrini & W. R. Schwartz. Análise de imagens digitais: princípios, algoritmos e aplicações, 2008.
[18] C. Cortes, & Vapnik, "Support vector machine. Machine learning", 20(3), 273-297, 1995.
[19] J. A. Suykens & J. Vandewalle., "Least squares support vector machine classifiers. Neural processing letters", 9(3), 293-300, 1999.