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

UAV autonomous navigation with thermal infrared camera

Natália G. Silva[1], Neusa M. F. Oliveira[2], Elcio H. Shiguemori[3], Haroldo F. C. Velho[4], José R. G. Braga[5], Marielcio G. Lacerda[3]

[1]*Departamento de Pesquisa e Desenvolvimento, Instituto de Aplicações Operacionais Av. Brg. Faria Lima, nº 1941, 12227-000, São Paulo, Brasil nataliagalvao.t17@gmail.com*
[2]*Instituto Tecnológico de Aeronáutica, Departamento de Ciência e Tecnologia Aeroespacial Av. Brg. Faria Lima, nº 1941, 12227-000, São Paulo, Brasil neusa@ita.br*
[3]*Instituto de Estudos Avançados, Departamento de Ciência e Tecnologia Aeroespacial Av. Brg. Faria Lima, nº 1941, 12227-000, São Paulo, Brasil elcio@ieav.cta.br, marielciomgl@fab.mil.br*
[4]*Laboratório de Computação e Matemática Aplicada, Instituto Nacional de Pesquisas Espaciais Av. dos Astronautas, 1.758, 12227-010, São Paulo, Brasil haroldo.camposvelho@inpe.br*
[5]*Coordenação de Observação da Terra, Instituto Nacional de Pesquisas Espaciais Av. dos Astronautas, 1.758, 12227-010, São Paulo, Brasil jgarciabraga@gmail.com*

Abstract Image processing approach has been employed as an alternative to the global navigation satellite system (GNSS) signal for the unmanned aerial vehicle (UAV) autonomous navigation. The UAV image is caught during the fly. A supervised artificial neural network is configured to be an image edge extractor. Two images are employed to the UAV positioning: a satellite georeferenced image, and image from the UAV. The image processing procedure consists to extract edges from the images, and to apply a correlation operator on the two segmented former images. The maximum correlation coordinate will be to indicate the UAV position. A low cost thermal infrared camera of type FLIR (Forward Looking Infrared) Duo is coupled to two UAVs (DJI Phantom 3 Standard, DJI Matrice 600 Pro) for two experiments, during daytime and nighttime. The algorithm is also implemented in Raspberry Pi and its processing time was evaluated. The results have shown good results for the UAV autonomous navigation using the low cost thermal infrared camera.

Keywords: UAV autonomous navigation, thermal infrared camera, image edge extraction.

1 Introduction

In several applications, the use of manned aircraft has been expensive, in terms of development costs, maintenance, and training of pilots [1]. In this sense, the use of Unmanned Aerial Vehicles (UAVs) has received considerable attention in recent decades. Some characteristics motivate its use, mainly, low cost and ability to use in environments that are difficult for man to access [2, 3].

A limitation for the insertion of UAVs in the airspace is related to the quality and reliability of the information about its location [2]. In most systems, this information is provided by the Inertial Navigation System (INS), which can present an accumulative error, especially low-cost INS [4]. These errors are generally corrected by the Global Navigation Satellite System (GNSS), for example, the North American Global Positioning System (GPS) [5].

The accuracy of GNSS systems, however, can be compromised by several factors: number of visible satellites and environmental-climatical conditions, for example; on equatorial/tropical zones, the scintillation due to ionospheric bubbles and the ion spring effect can cause interference on the GNSS signal, and interference by jamming, adding noise to the signal, or spoofing, falsifying the signal [5–7].

The mitigation of GPS signal failures can be achieved by fusioning data from several sensors [1]. UAVs, however, have restrictions on payload, energy consumption, and processing time. Therefore, the number of embedded sensors is limited. An alternative is to use image-based navigation. [1].

There are several approaches to visual navigation [8, 9], among them, template matching. which is a technique that consists of matching the image captured by the UAV to a region of a georeferenced image, already stored in the embedded system. For this, different techniques can be used, for example, the maximum correlation [7, 10].

Pattern matching can use information from the inertial system to reduce the search on the georeferenced

image. In addition, information from embedded sensors such as a compass and altimeter can be used to assist in correcting scale and rotation of images obtained in flight [11, 12].

Most studies focused on autonomous aerial navigation use imaging sensors in the visible spectrum, considering ambient conditions with good light [8]. However, there is a demand for low-light night navigation. In this context, thermal infrared imaging is a viable complementary system for navigation at night [13, 14].

In this work, template matching techniques were employed. Thermal images obtained by a low-cost thermal sensor embedded in UAVs were used. A georeferenced image in the visible spectrum was also considered as a reference. The algorithm was implemented in a Raspberry Pi CPU hardware, in order to evaluate the capabilities of real time processing, for application in autonomous navigation.

2 Methodology

2.1 Thermal infrared sensor

A sensor system is designed for converting and storing an image scenario into energy values [14]. A specific type of sensor is formed by arrays of photodetectors or thermal detectors, they are called Focal Plane Array (FPA). These matrix arrangements of photodetectors are able to convert the energy reaching into electrical signal values for processing [15]. After this conversion process, the image is discretized into digital numbers.

The thermal sensor employed was a low-cost camera, imaging in visible and thermal bands. The specifications provided by the manufacturer were summarized in Table 1.

Table 1. Camera Technical Data

Thermal Imager	Uncooled VO_x Microbolometer
Spectral Band	$7.5 - 13.5\ \mu m$
Sensor Resolution (pixels)	160×120
Lens Options	$57^o \times 44^o$
Thermal Frame Rates	$7,5\ Hz\ (NTSC)$; $8,3\ Hz\ (PAL)$

2.2 Automatic image registration

A technique for estimating UAV position by images is automatic registration Conte and Doherty [6], Braga et al. [7], Goltz et al. [10], Braga [12], Silva [13], identifying the overlap of images captured in different moments, points of view and also with different sensors [12].

In the image register, illustrated in Fig. 1, a correlation window goes from the top to the bottom calculating the correlation between the image windows: the reference image and the image captured in flight. At the end, the entire image is scrolled. The position estimation is given by the correlation maximum. [13].

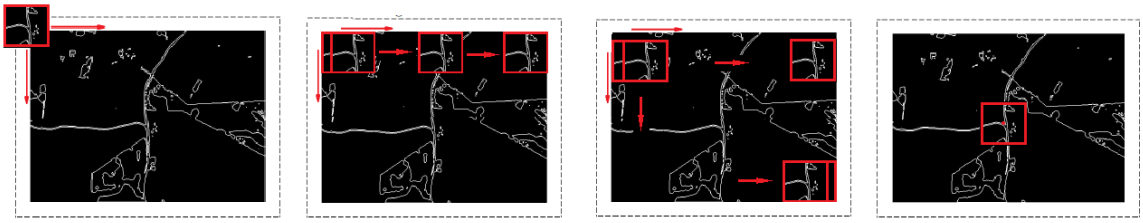

Figure 1. Correlation window of the georeferenced image and the computation of its cross correlation with the aerial image of the UAV. On the right, the final position estimate is represented by the central pixel of the window that indicates the highest correlation. Fonte: Image adapted from Braga [12].

The cross correlation between the UAV image I, with M $\times$ N *pixels*, and the georeferenced image T, with P $\times$ Q *pixels*, is given by [12]:

$$C(k,l) = \sum_{m=0}^{M-1} \sum_{n}^{N-1} \big(I(m,n) \times \bar{T}(m-k, n-l)\big), \tag{1}$$

where $C\ (k,l)$ is the matrix correlation with k and l indices, where $k \in [-(P-1), (M-1)]$ and $l \in [-(Q-1), (N-1)]$; and $\bar{T}$ is the complex conjugate T.

The highest value in the cross-correlation matrix corresponds to the central pixel of the image captured in flight, estimating the UAV position. Since the image of the overflown region is georeferenced,the geographical coordinates of the point of greatest cross-correlation and location of the UAV are known. Thus, the position estimation process by automatic image registration is completed [12].

In order to consider sensors in different regions of the electromagnetic spectrum, the images are processed, which allows segmentation by edge detection. This allows you to correlate the border images [13].

2.3 Extraction of edges by artificial neural network

The automatic image registration allows match images from different sensors, with adequate accuracy in position estimation, and low influence of variations of ambient lighting Silva [13]. In addition, working only with the edges of the images can reduce the amount of data processed [12].

Conte and Doherty [6], Braga et al. [7], Goltz et al. [10], Braga [12], Silva et al. [16] successfully use Artificial Neural Networks (ANNs) as filters for edge detection. According to Li et al. [17], when compared with traditional edge extractors, ANNs provide greater accuracy.

The input vector for the ANN extracting edges can be a vector of nine elements, which corresponds to the elements of a window (matrix) 3×3 pixels. This window for computing the image correlation, between a reference (georeference image) and the image obtained by the UAV, scrolls through a binary image pixel to pixel. In the process, the binarization threshold is determined by the Otsu method [12].

The Multilayer Perceptron Neural Network (MLP), trained to recognize edge and non-edge patterns, has been published in previous papers Braga et al. [7] e Silva et al. [16]. MLP has shown to be efficient as an extractor of binary edges. Its architecture has been determined as an optimization problem, solved by using the Multiple Particle Collision Algorithm (MPCA) meta-heuristic for automatic network configuration [16]. The parameters of training and neural network architecture: constant at the moment; learning rate; number of hidden layers; and number of neurons in each hidden layer; obtained with the MPCA are presented in Table 2.

Table 2. Neural Network Architecture

Input Layer	9 Neurons
Hidden Layers	18 Neurons
Output Layer	2 Neurons
Activation Function	Hyperbolic Tangent
Learning Rate	0.73
Momentun Rate	0.85
Maximum Number of Epochs	15000
Maximum Acceptable Error	10^{-8}

2.4 Estimation of UAV position

The UAV position estimation process performs the treatment of georeferenced and thermal images, for extracting edges of images from different sensors (Figure 2). The processing of the georeferenced image in the visible spectrum embraces of changing to gray levels; binarization, responsible for employing a value of threshold for an image with values of zero and one; filtering, to remove discretizing noise; edge extraction by ANN; selection of correlation, reducing the search region; and the cross correlation, to estimate position.

The processing of the aerial image in the thermal infrared spectrum involves scaling, in order to match the scales to the thermal and georeferenced images; conversion to gray levels, binarization, filtering and extraction of

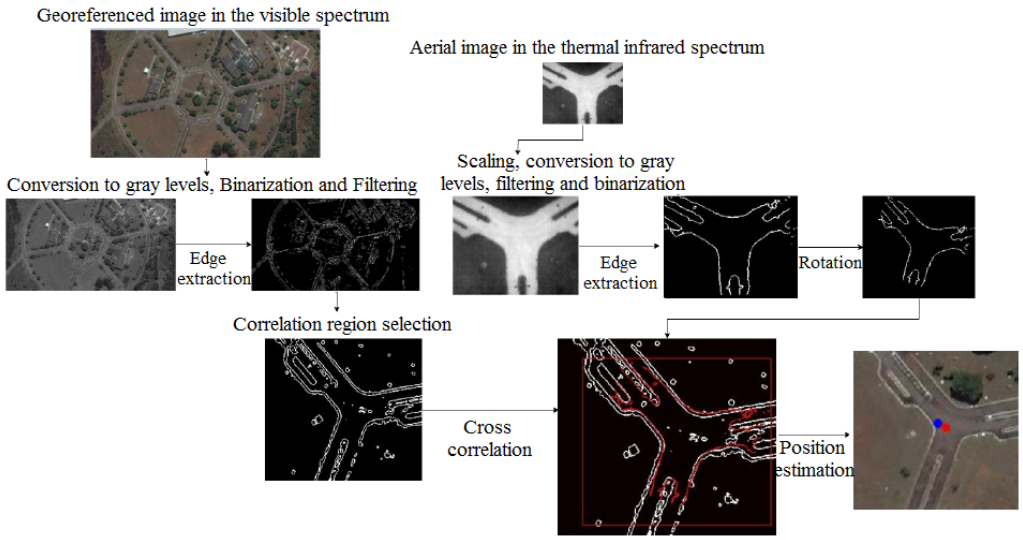

Figure 2. Position estimation algorithm

edges by ANN as well as in the georeferenced image; rotation adjustment to leave both images pointing north; and the cross-correlation with the visible image.

Silva [13] verified this position estimation algorithm for a FLIR thermal camera in the spectral range of 3.4 to $4.8\ \mu m$, resolution of *pixels* 640×512, and generation rate *frames* of $25\ Hz$. In this work, the used FLIR camera operates in a different spectral range (7.5 to $13.5\ \mu m$), lower resolution in *pixels* (160×120), and lower generation rate *frames* ($8.3\ Hz$). Silva [13] did not consider the implementation on embedded hardware.

2.5 Embedded system processing

As an embedded system, a Raspberry Pi 3, Model B, was considered. The compatibility with the library *Open Source Computer Vision* (OpenCV) and its documentation motivated the choice.

The algorithm was implemented using the $C++$ language. The funcions grayscale conversion, filtering, binarization, template matching, rotation and resizing were implemented using the OpenCV library.

3 Results

3.1 First Experiment

In the first experiment, images were collected during the day using the FLIR thermal camera integrated in the UAV Phantom 3. The flight was programmed considering the cover, using the Mission Planner application. Flights were performed in a parking at an average altitude of $100\ m$, at a time close to $7:30$ a.m., with average speed of $3\ m/s$. For the flights, the data provided by the INS and the GPS, accessed via flight log were considered as correct UAV positioning. A georeferenced image of the parking obtained by Google Earth Pro was considered.

A number of 120 thermal aerial images were considered as a database for the analysis, with location given according to information from the GPS integrated to the Phantom 3. This location information was adopted as a correct measure of the position of the UAV. The latitude, longitude, and altitude, as well as the yaw, pitch, and roll angles adopted as the true position in the estimation algorithm were reported in the UAV flight log.

The performance metric used was the Error Good Matching (EGM), a position estimate with error, calculated by Euclidean distance, less than five meters. The value of five meters is adopted, as it corresponds to half the maximum error of the embedded GNSS systems [6].

The Gaussian filter and edge extraction by ANN were used in the automatic registration process Silva [13] and Braga et al. [7]. The estimation error was obtained in relation to each of the 120 points of evaluation of the trajectory, given according to Fig. 3.

In Fig. 3, the dashed red line indicates the reference equal to 5 meters for errors. Points below this red line are considered EGM. 55 estimates were considered EGM. This corresponds to approximately $46\ \%$ of the analyzed points of the trajectory.

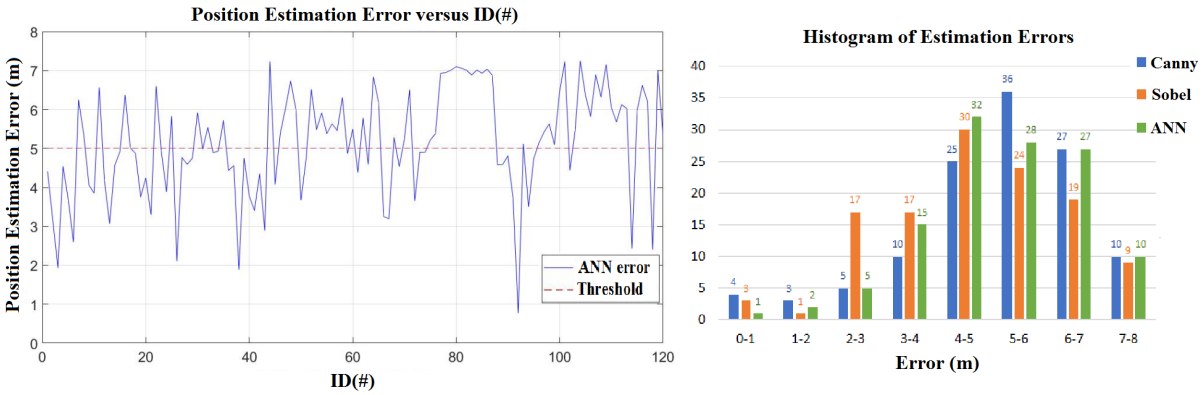

Figure 3. On the left, position estimation error obtained by the automatic image registration algorithm considering the 120 evaluation points.On the right, number of position estimates in each error range, using the automatic image registration algorithm configured with the RNA, Sobel and Canny extractors.

It should be noted that there were no position estimation errors above the maximum error of the GPS system ($10\ m$). The maximum error presented was $7.234\ m$, which has the same order of magnitude as the expected error when using a global positioning system for navigation. This indicates that the estimation algorithm is a candidate in air navigation by thermal camera. However, it is necessary to evaluate the processing time of this algorithm in an embedded system.

In another evaluation, the Sobel and Canny edge detection performance were evaluated. Gaussian filtration was applied to eliminate noise in Sobel extraction process. The Canny filter is already based on Gaussian filter. In Fig. 3, the histogram of the distribution of errors for each edge detectors is also shown. For example, in ANN detection, 32 position estimates resulted in errors between 4 and 5 meters.

By analyzing Fig. 3, the distribution of the estimation errors was close for the three extractors. Cconsidering the Error Good Matching metric and the 120 position estimation points, obtained with automatic image registration, adopting extractors ANN, Sobel and Canny, were obtained values of 55, 68, and 47 were verified by ANN, Sobel, and Canny edge extractors, respectively.

3.2 Second experiment

In the second experiment, images obtained at night from the FLIR thermal camera on board the UAV Matrice were considered. The reference positioning information was provided by a GPS Real Time Kinematic (GPS RTK). This test aimed at evaluating the ability to execute night flights by computer vision. The UAV captured the images at $80.5\ m$ altitude, at a time close to 07:30 pm, with an average speed of the drone of $3\ m/s$.

208 thermal images were obtained, the locations were measured by a GPS RTK. The location in latitude, longitude, and altitude, were provided by the RTK system. The bow direction and pitch, raw, and yaw angles were provided by the INS.

The georeferenced image used in this experiment was obtained using Google Earth Pro. It is an image in the visible spectrum captured during the day, with resolution of $0.2649\ m/pixel$, and $1920 \times 1080\ pixels$ of dimension.

The estimation errors for the 208 evaluation points of the night trajectory, using the Gaussian filter, are presented in Fig. 4. The dashed red line delimits the region of points classified as EGM ($5\ m$).

208 thermal images were used to test the position estimation algorithm with MLP edge detection and Gaussian filtering. 123 estimates were considered EGM, $59.13\ \%$ from trajectory points. Also in Fig. 4, no errors exceeded the maximum GPS system error ($10\ m$). The maximum error obtained was $7.539\ m$. Thus, considering the experiments, the estimation algorithm could be applied for autonomous aerial navigation by thermal infrared imaging. Therefore, it is necessary to evaluate the processing time of this algorithm in an embedded system.

In another analysis, the performance of the estimation algorithm, adopting Sobel and Canny, were evaluated. In Fig. 4, there is a histogram of the distribution of errors for each of the edge detectors. The distribution of estimation errors was similar for the three extractors.

The EGMs obtained using the ANN, Sobel, and Canny edge extractors were 123, 109, and 112, respectively. Thus, the ANN edge detector performed better in position estimation, with $59.13\ \%$ of the assessment points.

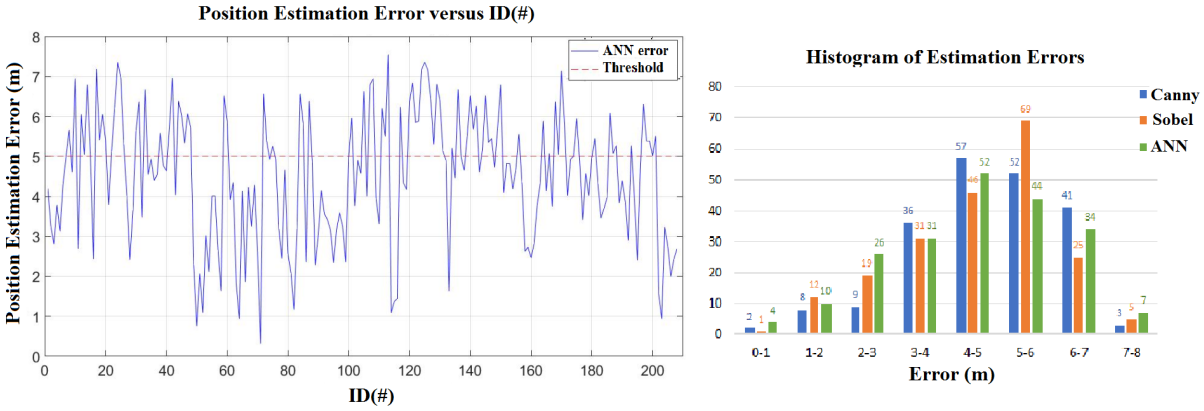

Figure 4. Position estimation error obtained by the automatic image registration algorithm at 208 points (left). Position estimates in each error range, using the automatic image registration algorithm configured with RNA, Sobel and Canny (right).

3.3 Embedded Processing

The average of CPU-time for the position estimation, executed on the Raspberry Pi 3, are presented in cref TemposPi. These times correspond to the processing of a thermal image, with 160×120 pixels. In order to save time and processing, the image of the edges of the georeferenced image is previously stored. In Table 3, average processing time of the 208 thermal images captured at night are shown.

From Table 3, the greater computational effort was verified by computing the correlation. It is observed that the processing time on the Raspberry Pi 3 has not been a limiting factor for computer vision navigation using edge extraction by MLP neural network and automatic registration. A limitation would be the time interval between the capture of two subsequent thermal images. Therefore, the process to estimate the UAV positioning takes less than 1 second.

Table 3. Average times of execution of the steps of scale correction, rotation adjustment, filtering, clipping and correlation in Raspberry Pi 3.

Algorithm Step	Execution Time (seconds)
Scale Correction	0.00128
Rotation Adjustment	0.00680
Filter Application	0.00179
Edge Extraction (LUT)	0.00857
Cropping the Georeferenced Image	0.00032
Correlation between Images	0.03976
Total time	0.06845

4 Final Remarks

The results of this study indicate that self-configuring by the MPCA for an ANN for edge detection is a competitive technique, allowing to be used in autonomous night navigation by thermal imaging. The UAV position was estimated in the same order of magnitude as the conventional GPS error. Therefore, the described methodology for autonomous navigation by thermal images can be applied as an alternative solution in case of failure or unavailability of the GNSS signal.

Acknowledgements. This study is by the Institute for Advance Studies (IEAv), a division from the Department of Aeronautics Science and Technology (DCTA). Author HFCV would also like to thank the National Council for Scientific and Technological Development (CNPq, Brazil) for the research grants (CNPq: 312924/2017-8).

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work, and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors, or has the permission of the owners to be included here.

References

[1] Kanellakis, C. & Nikolakopoulos, G., 2017. Survey on computer vision for uavs: Current developments and trends. *Journal of Intelligent & Robotic Systems*, vol. 87, n. 1, pp. 141–168.
[2] Valavanis, K. P., 2008. *Advances in unmanned aerial vehicles: state of the art and the road to autonomy*, volume 33. Springer Science & Business Media.
[3] Maes, W. H., Huete, A. R., & Steppe, K., 2017. Optimizing the processing of uav-based thermal imagery. *Remote Sensing*, vol. 9, n. 5, pp. 476.
[4] Conte, G. & Doherty, P., 2009. Vision-based unmanned aerial vehicle navigation using geo-referenced information. *EURASIP Journal on Advances in Signal Processing*, vol. 2009, pp. 10.
[5] Infrastructure, T., 2001. Vulnerability assessment of the transportation infrastructure relying on the global positioning system. Technical report, Technical Report, Center, John A. Volpe National Transportation Systems.
[6] Conte, G. & Doherty, P., 2008. An integrated uav navigation system based on aerial image matching. In *Aerospace Conference, 2008 IEEE*, pp. 1–10. IEEE.
[7] Braga, J. R., Velho, H. F., Conte, G., Doherty, P., & Shiguemori, É. H., 2016. An image matching system for autonomous uav navigation based on neural network. In *Control, Automation, Robotics and Vision (ICARCV), 2016 14th International Conference on*, pp. 1–6. IEEE.
[8] Lenskiy, A. A. & Lee, J.-S., 2010. Terrain images segmentation in infra-red spectrum for autonomous robot navigation. In *Strategic Technology (IFOST), 2010 International Forum on*, pp. 33–38. IEEE.
[9] Rathinam, S., Almeida, P., Kim, Z., Jackson, S., Tinka, A., Grossman, W., & Sengupta, R., 2007. Autonomous searching and tracking of a river using an uav. In *American Control Conference, 2007. ACC'07*, pp. 359–364. IEEE.
[10] Goltz, G., Shiguemori, E., & Velho, H., 2011. Position estimation of uav by image processing with neural networks. In *X Congresso Brasileiro de Inteligncia Computacional-CBIC*, pp. 1–6.
[11] Kuroswiski, A. R., 2017. Navegação Aérea Autônoma Baseada em Visão Computacional. 107f. Trabalho de Conclusão de Curso. (Graduação em Engenharia Eletrônica) - Instituto Tecnológico de Aeronáutica, São José dos Campos.
[12] Braga, J. R. G., 2018. *Navegação Autônoma de VANT por Imagens LIDAR*. PhD thesis, Instituto Nacional de Pesquisas Espaciais.
[13] Silva, W. D., 2016. Navegação autônoma de vant em período noturno com imagens infravermelho termal. Master's thesis, Dissertação de Mestrado do Curso de Pós-Graduação em Computação Aplicada - Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos.
[14] Oliveira, L. T. D., 2017. Avaliação do uso de sensor termal a bordo de vant através de análises radiométricas, espectrais, espaciais e posicionais. Master's thesis, Dissertação de Mestrado do Curso de Pós-Graduação em Sensoriamento Remoto - Instituto Nacional de Pesquisas Espaciais (INPE), São José dos Campos.
[15] Silva, R. E. D., 2018. Desempenho em alcance e caracterização de sistemas eletro-ópticos termais. Master's thesis, 105f. Dissertação de Mestrado em Ciências no Programa de Pós-Graduação em Ciências e Tecnologias Espaciais, Área de Física e Matemática Aplicadas - Instituto Tecnológico de Aeronáutica, São José dos Campos.
[16] Silva, W. D., Shiguemori, E. H., Vijaykumar, N. L., & Campos Velho, H. F. D., 2015. Estimation of uav position with use of thermal infrared images. In *Sensing Technology (ICST), 2015 9th International Conference on*, pp. 828–833. IEEE.
[17] Li, W., Wang, C., Wang, Q., & Chen, G., 2008. An edge detection method based on optimized bp neural network. In *Information Science and Engineering, 2008. ISISE'08. International Symposium on*, volume 2, pp. 40–44. IEEE.