UAVs Applied to the Counting and Monitoring of Animals

Pablo Chamoso, William Raveane, Victor Parra, and Angélica González

Computer and Automation Department, University of Salamanca, Spain
{chamoso,wr,parra,angelica}@usal.es

Abstract. The advantages of intelligent approaches such as the conjunction of artificial vision and the use of Unmanned Aerial Vehicles (UAVs) have been recently emerging. This paper presents a focused on obtaining scans of large areas of livestock system. Counting and monitoring of animal species can be performed with video recordings taken from UAVs. Moreover the system keeps track of the number of animals detected by analyzing the images taken with the UAVs cameras. Several tests have been performed to evaluate this system and preliminary results and the conclusions are presented in this paper.

Keywords: Unmanned Aerial Vehicle, Convolutional Neural Networks, livestock detection.

1   Introduction

Possibilities of applying UAVs in the professional world have increased in recent years thanks to advances in technology, especially aerial imaging. Aerial imaging provides detailed images and quick monitoring of large areas, making it a very efficient way to solve the problem proposed in this study.

The problem presented in this study is related to cynegetic activities. So far, the different methods that are carried out for conducting censuses to optimize hunting and harvesting in farms and private reserves consists of conducting pathways, which rely on feeders and hunting results.

The case study presented in this paper adapts the developed system as a methodology for conducting population censuses of farm animals. A farm with cows was used for this case study. Area sweeping techniques in combination with visual detection, recognition techniques and UAV system control are used and will be explained throughout this article.

An important part of the system is its ability to visually identify targets on the ground. As with any computer vision application, this is a difficult problem to solve due to the high dimensionality of the input data. The numerical representation of a photograph will vary greatly with any subtle change in illumination, camera conditions, or even object appearance, among other factors. Shadows cast by the target objects only complicate matters as they usually impede effective segregation from the terrain background. These conditions can be overwhelming for most traditional image

processing techniques, such as Chan-Vese segmentation [17], which has excelled in many other image separation tasks, but fails with the type of data involved in this work.

The Convolutional Neural Network (CNN) [14][37] was introduced as a general solution to the image recognition problem of variable inputs. A CNN consists of a multi-layer artificial neural network with a built-in feature extraction process and translational tolerance of the input image space. Therefore, this type of network is capable of accurately identifying images of a target object among cluttered background noise. This is achieved by learning the distinctive features that characterize the class the object belongs to, regardless of the relative position at which it appears in the input image sample.

The article is structured as follow: the next section describes the general background of livestock detection, UAVs and CNN. Section 3 is a system overview in which the system components and the visual recognition techniques used are described in detail. Finally, the last section shows the results and conclusions obtained.

2   Background

There are three areas that converge in this study: (i) livestock species detection and accounting, which is the main objective of the study; (ii) multirotor systems, such as hardware tools used as part of the counting process, in this case the UAV; (iii) image processing techniques used for carrying out the analysis of the images obtained by the UAV, in this case, Convolutional Neural Network. This section describes the current state of the art of each technique. The next section explains how they have been used in this study.

2.1   Livestock Detection

Species accounting is of great interest in activities such as hunting, biology and agriculture. There have been several studies on cynegetic use [23][3] that take into account the advantages [26] and disadvantages [1] with regard to the environment, the economy and tourism.

The difficulty of this task arises from the diversity of soils, species-specific characteristics, and spatial aggregation of animals in the field [15]. In this sense, several authors have made various proposals to solve this problem based on statistical and biological methods [45][25][7].

Thanks to advances in computer science, the task of counting animals can be performed automatically (as opposed to a manual count), as has been addressed successfully in several studies [24][2][38].

2.2   Multirotor System

One of the most popular technological advances in recent years is the multirotor or multicopter, a type of UAV capable, among other characteristics, of aerial filming as

described in [39]. The use of this system to obtain an aerial video with sufficient quality is a tool that, in combination with image analysis to detect animals, can save considerable time and money in the process of counting animals.

To carry out the case study transmission of large data from multirotors, it is necessary to have a powerful computer that can perform real time processing. The transmission of large amounts of data requires at least one channel with high bandwidth. Most multirotor systems that exist today use analogue channels to transmit flight orders to the multirotor and send the captured video [47][20][46][6][41].

The proposed system presents an alternative to these systems by taking advantage of technological advances. It tries to take advantage of powerful Wi-Fi antennas currently available in the market, to unify the process of sending information through a single digital channel, including flight commands, telemetry and video in high resolution. Currently there are some systems that utilize UAV communication using sockets [4], but their use is restricted to UAV type aircraft.

This form of communication makes it possible to control the multirotor as well as to obtain information about the status of the sensors it can carry, such as the Global Positioning System (GPS) or altimeter. Of course, it can also send video images taken by its camera. Furthermore, the use of a computer instead of a radio station as a multirotor control element allows a processing capacity that the station does not have, which makes the computer able to control the multirotor intelligently without a human pilot.

2.3 Convolutional Neural Network

The layer architecture for the CNN used will vary according to the application. However, the count of output neurons in the final layer will always be associated with two output classes: one for the object of interest, and the other for background noise. Hence, every time the network is executed over an image patch, it will output two values, each of which will be interpreted as the confidence level with which the network believes the corresponding class correctly describes the analyzed sample.

The CNN is trained with data manually collected from previous UAV test flights, and later artificially augmented. The augmentation consists of increasing the number of training data available by applying a series of transformations to translate, rotate and scale each manually collected example. Such a process can yield up to 40 artificial samples for every original image patch, thus giving the network a lot more data to train with. The training data is separated into two sets for each of the classes, target and background.

The network is finally trained with the prepared data sets through the stochastic gradient descent method for back-propagation [5], which offers an optimal route to minimizing the classification error for this type of highly mutable training data

3   System Overview

The architecture of the multirotor for obtaining images consists of three main parts: hardware, software and communication protocol.

The hardware part refers to the system running the software and is present at both ends of the communication. At one end is the UAV, a multirotor with 6 engines and 6 arms bearing a Single Board Computer (SBC) with a 700MHz processor and 512MB memory capable of running a Linux distribution. The connectivity of the SBC allows connecting a camera via Ethernet and a Wi-Fi antenna Universal Serial Bus (USB); it also gets the information that Speed Electronic Controllers (ESCs) need to control the multirotor motors. At the other end of the communication is an access point whose potency depends on the distance to be achieved, and can cover distances of up to 3 kilometers. This access point will connect the SBC UAV control computer with a high performance laptop with a USB gamepad connected for manual control.

Regarding the software, SBC runs a piece of software that mainly reads, processes and delivers telemetry. It is responsible for carrying out the calculations for the behavior and stability of the multirotor. For its part, the computer runs a piece of software specifically developed (in Java) to control the multirotor both manually and autonomously through waypoints. The user can see all the information from the sensors that the multirotor sends in real time, such as information relating to power consumption, intensity and quality of the Wi-Fi signal, image and video (Fig. **3**). This information will generate log files in XML format that will serve to make a more accurate and detailed analysis of the captured images, as it will provide additional useful information such as height and GPS.

Finally, the communication protocol between the two parts is based on a multi-agent platform called PANGEA (Platform for Automatic coNstruction of orGanizations of intElligent Agents) [48] based on the IRC protocol. Many previous studies [9][13][31][30] support the use of multi-agent systems [12][34][33][10] combined with techniques such as neural networks as an optimal solution to similar problems but in other fields of research [32][22][27][36][28][40][29][21][11][8][35]. With the use of this platform, described in [48], each of the parties involved in the communication will behave as an agent following the proposed PANGEA scheme. In this way, the system is ready for a possible extension where one computer can control different multirotors at the same time and multirotors can communicate with each other to achieve common goals more efficiently [16][44][42][43][19][18].

3.1   Runtime Operation

At runtime, object recognition is carried out by analyzing individual frames received from an on-board camera. A sliding window approach is taken to sequentially analyze small, adjacent, and overlapping image patches positioned in a grid pattern over the frame. Each of these image patches is evaluated by the trained network, and the output values are recorded.

This process is repeated at three scales. The middle scale is chosen to approximate the relative size at which the target objects are expected to be seen – a value which can be calculated trigonometrically based on the current flight altitude of the UAV. The two remaining scales are set at 85% and 115% of the middle scale window size. These additional scales provide supplementary information which later helps to boost the readings obtained at each grid coordinate.

For every window analyzed, the output values of the CNN are processed with the softmax activation function described in Equation 1, where y_i is the resulting value of

the network for output neuron i. This transformation results in a probability-like value $P(i)$, which for any given class i estimates the likelihood that the analyzed window belongs to it.

$$P(i) \equiv \sigma(y, i) = \frac{e^{y_i}}{\sum_{n=0}^{N} e^{y_n}} \tag{1}$$

Equation 2 defines L_{xys} as the likelihood value that the analyzed window at grid coordinate $(x;y)$ and scale s belongs to the target class. Combining these values over the entire grid for every x, y and s will lead to a discrete 2D probability distribution over the original input frame that will indicate the positions at which a target object has been detected.

Boosting the values at each coordinate through a simple nearest-neighbor clustering algorithm, where the likelihood value for each point in the grid is increased by strong adjacent readings, can further enhance this distribution but left unchanged by a lack thereof. This process is detailed in Equation 3.

$$L_{xys} \equiv P(target|x, y, s) \tag{2}$$

$$B_{xy} = \sum_{i=x-1}^{x+1} \sum_{j=y-1}^{y+1} \sum_{s=0}^{2} \begin{cases} L_{ijs} & if \ L_{ijs} \geq 1/2 \\ 0 & if \ L_{ijs} < 1/2 \end{cases} \tag{3}$$

The boosted value B_{xy} can then passed through a threshold, determined ad hoc for the particular application, to finally produce a quantized representation of the probability distribution.

An additional benefit of boosting involves the elimination of most false positives, occurrences from which neural networks are never exempt, but which, fortunately, often tend to appear without any neighboring support, thus disappearing after applying this process. Likewise, the use of this mechanism will make the system more tolerant towards false negatives, as these will not have such a significant impact on the final result, due to the network often finding reinforcing values at adjacent positions.

3.2   Visual Recognition

For this application, the network utilized follows an architecture given by 64x64-18C7-MP4-96C5-MP2-4800L-2, wherein there are two convolutional feature extraction stages and one hidden linear layer. The training data is prepared as previously described, with the *target* class consisting of individual animal samples. A small subset of this data set can be seen in Fig. 1.

Fig. 1. A subset of the training data used for the CNN, where the dataset is divided into two classes: target (top row) and background (bottom row)

Fig. 2 depicts the runtime implementation of the system, where the multiple stages of detection can be seen. The final quantized likelihood distribution can be used to easily count the cattle currently in view.
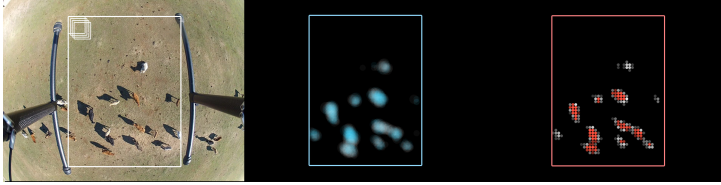

Fig. 2. A sample video frame (left), analyzed with the CNN over a grid pattern producing a probability distribution Lxys plotted over the entire frame (middle), values which can then be boosted and quantized as B_{xy} to better distinguish individual targets (right)

4   Results and Conclusions

This study was developed to test software that controls multirotors and allows visualization of all information captured from the air in real time. It also has a tool that allows for autonomous routes based on waypoints. The multirotor is equipped with two cameras: a flight chamber which provides guidance for the pilot who is controlling the UAV, and a second high resolution camera placed perpendicular to the images obtained from counting the species. Further, the CNN used as a counting technique produced good results, as shown below.
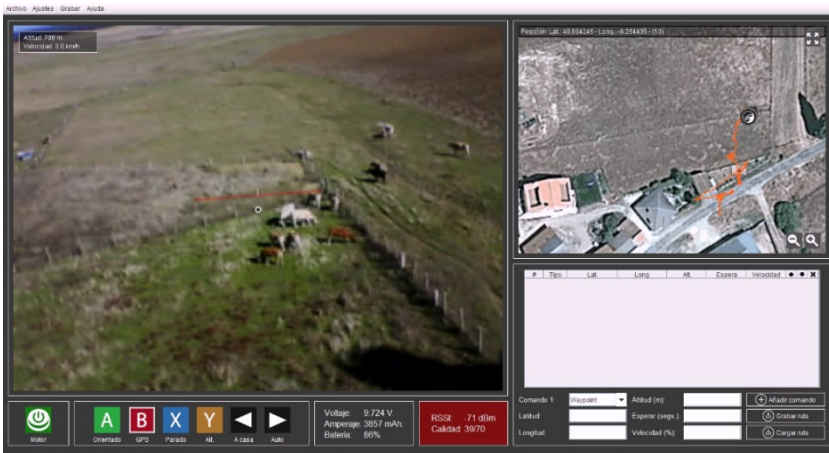

Fig. 3. Ground Control Station software during a flight

The CNN can be trained surprisingly fast, reaching a plateau for the training criterion (MSE on classification) after only a few SGD epochs. A summary of the CNN training results is given in Table 1 Confusion matrix of the visual classification CNN trained with the augmented sample data over 10 SGD epochs..

Table 1. Confusion matrix of the visual classification CNN trained with the augmented sample data over 10 SGD epochs

Class	Training Data Classification 10,816 Samples			Testing Data Classification 2,704 Samples		
	Cattle	Background	Accuracy	Cattle	Background	Accuracy
Castle	**5,202**	206	96.2%	**1,241**	111	91.8%
Background	109	**5,299**	98.0%	10	**1,342**	99.3%
Global			**97.1%**			**95.5%**

We can obtain an animal count for every analyzed frame by running a connected component-labeling algorithm over the boosted and thresholded B_{xy} values in that image. The results of this process applied on evenly spaced frames from a 3 minute video sequence are reported in Table 2.

Table 2. Results of the cattle counting application over 70 sample frames analyzed. They are divided according to the counting result, which makes it evident that most errors occur in overcrowded frames, when there is an average of 10 or more visible targets.

Result	Frames analyzed	Targets total	Targets per frame	Targets counted	Δ Count total	Δ Count per frame	Δ Count per target
Over-count	2	22	11.00	24	+2	+1.00	0.091
Correct	57	309	5.42	309	0	0.00	0.000
Undercount	11	112	10.18	99	-13	-1.18	0.116
Total	70	443	6.32	432	-11	-0.16	**0.025**

Finally, it should be noted that the use of PANGEA as the platform for communication between the computer and the multirotor opens new research lines since PANGEA provides the capacity and mechanisms for several multirotors to communicate simultaneously. This means that it will be possible to explore areas in an autonomous and collaborative way [18], allowing the task of flying over the area to be distributed among the various required multirotors, thus reducing the flight time or increasing the area covered.

References

[1] San Miguel, A., Ochoay, J., Pérez Carral, C.: Wildlife Management in Mediterranean Forest Ecosystems. Problems and current situation of Shrublands. Montes (35), 33–36 (1994)

[2] B. T. a. C. J.: Real-time Face Detection and Tracking of Animals. In: de 8th Seminar of Neural Network Applications in Electrical Engineering, Belgrade (2006)

[3] Buenestado, B.V.: Aprovechamiento y gestión de la caza en España. Una reflexión a propósito de los cercados cinegéticos. Actas del VI Coloquio de Geografía Rural, 257–272 (1991)

[4] Bodin, W.K., Redman, J.J., Thorson, D.C.: U.S. Patent No. 7,286,913. U.S. Patent and Trademark Office, Washington, DC (2007)

[5] Bottou, L.: Stochastic learning. In: Bousquet, O., von Luxburg, U., Rätsch, G. (eds.) Machine Learning 2003. LNCS (LNAI), vol. 3176, pp. 146–168. Springer, Heidelberg (2004)

[6] Buss, H., Busker, I.: Mikrokopter (2008)

[7] C. S. a. R. T. J. McKinlay: Integrating Count Effort by Seasonally Correcting Animal Population. CCAMLR Science 17, 213–227 (2010)

[8] Pinzón, C.I., Bajo, J., De Paz, J.F., Corchado, J.M.: S-MAS: An adaptive hierarchical distributed multi-agent architecture for blocking malicious SOAP messages within Web Services environments. Expert Systems with Applications 38(5), 5486–5499

[9] Tapia, D.I., Abraham, A., Corchado, J.M., Alonso, R.S.: Agents and ambient intelligence: case studies. Journal of Ambient Intelligence and Humanized Computing 1(2), 85–93 (2010)

[10] Tapia, D.I., De Paz, J.F., Rodríguez, S., Bajo, J., Corchado, J.M.: Multi-agent system for security control on industrial environments. International Transactions on System Science and Applications Journal 4(3), 222–226 (2008)

[11] Tapia, D.I., Alonso, R.S., De Paz, J.F., Corchado, J.M.: Introducing a distributed architecture for heterogeneous wireless sensor networks. In: Omatu, S., Rocha, M.P., Bravo, J., Fernández, F., Corchado, E., Bustillo, A., Corchado, J.M. (eds.) IWANN 2009, Part II. LNCS, vol. 5518, pp. 116–123. Springer, Heidelberg (2009)

[12] Tapia, D.I., Rodríguez, S., Bajo, J., Corchado, J.M.: FUSION@, a SOA-based multi-agent architecture. In: International Symposium on Distributed Computing and Artificial Intelligence 2008 (DCAI 2008), pp. 99–107 (2008)

[13] Fdez-Riverola, F., Corchado, J.M.: CBR based system for forecasting red tides. Knowledge-Based Systems 16(5), 321–328 (2003)

[14] Fukushima, K.: Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. Biological Cybernetics 36(4), 193–202 (1980), doi:10.1007/BF00344251

[15] Sileshi, G.: The excess-zero problem in soil animal count data and choice of appropriate models for statistical inference. Pedobiologia 52(1), 1–17 (2008)

[16] Garijo, F., Gómes-Sanz, J.J., Pavón, J., Massonet, P.: Multi-agent system organization: An engineering perspective. In: Pre-Proceeding of the 10th European Workshop on Modeling Autonomous Agents in a Multi-Agent World (MAAMAW 2001) (2001)

[17] Getreuer, P.: Chan-Vese Segmentation. Image Processing on Line 2012 (2012), doi:10.5201/ipol.2012.g-cv

[18] Gómez, J., Patricio, M.A., García, J., Molina, J.M.: Communication in distributed tracking systems: an ontology-based approach to improve cooperation. Expert Systems 28(4), 288–305 (2011)

[19] Griol, D., García-Herrero, J., Molina, J.M.: Combining heterogeneous inputs for the development of adaptive and multimodal interaction systems. Advances in Distributed Computing and Artificial Intelligence Journal 6, 37–53 (2013) ISSN 2255-2863

[20] Haehnel, H.: Remote controlled flying robot platform. In: Third International Conference on Digital Information Management, ICDIM 2008, pp. 920–921. IEEE (November 2008)

[21] Bajo, J., De Paz, J.F., Rodríguez, S., González, A.: Multi-agent system to monitor oceanic environments. Integrated Computer-Aided Engineering 17(2), 131–144 (2010)

[22] Bajo, J., Corchado, J.M.: Evaluation and monitoring of the air-sea interaction using a CBR-Agents approach. In: Muñoz-Ávila, H., Ricci, F. (eds.) ICCBR 2005. LNCS (LNAI), vol. 3620, pp. 50–62. Springer, Heidelberg (2005)

[23] J. C. G. R. R. A. B. a. J. M. V. M. Ángel Farfán: Game harvest characterisation of the mammals in Andalusia. Galemys 16(1), 41–59 (2004)

[24] J. D. R. a. K. R. G. Felix A. Wichmann: Animal detection in natural scenes: Critical features revisited. Journal of Vision 10(4) (April 2010)

[25] Anderson, J.I.J.M.: Tropical Soil Biology and Fertility. A Handbook of Methods. de CAB International, Wallingford (1993)

[26] Gallego, J.I.R.: Caza y turismo cinegético como instrumentos. Anales de Geografía de la Universidad Complutense 30(2) (Octubre 2010)

[27] Fraile, J.A., Bajo, J., Corchado, J.M., Abraham, A.: Applying wearable solutions in dependent environments. IEEE Transactions on Information Technology in Biomedicine 14(6), 1459–1467 (2011)

[28] De Paz, J.F., Rodríguez, S., Bajo, J., Corchado, J.M.: Case-based reasoning as a decision support system for cancer diagnosis: A case study. International Journal of Hybrid Intelligent Systems 6(2), 97–110 (2009)

[29] De Paz, J.F., Rodríguez, S., Bajo, J., Corchado, J.M.: Mathematical model for dynamic case-based planning. International Journal of Computer Mathematics 86(10-11), 1719–1730 (2009)

[30] Corchado Rodríguez, J.M.: Redes Neuronales Artificiales: un enfoque práctico. Servicio de Publicacións da Universidade de Vigo, Vigo (2000)

[31] Corchado, J.M., Lees, B.: Adaptation of cases for case based forecasting with neural network support. In: Soft Computing in Case Based Reasoning, pp. 293–319 (2001)

[32] Corchado, J.M., Fyfe, C.: Unsupervised neural method for temperature forecasting. Artificial Intelligence in Engineering 13(4), 351–357 (1999)

[33] Corchado, J.M., Aiken, J., Rees, N.: Artificial intelligence models for oceanographic forecasting. Plymouth Marine Laboratory (2001)

[34] Corchado, J.M., Aiken, J.: Hybrid artificial intelligence methods in oceanographic forecast models. IEEE Transactions on Systems, Man, and Cybernetics, Part C: Applications and Reviews 32(4), 307–313 (2002)

[35] Corchado, J.M., Bajo, J., De Paz, J.F., Rodríguez, S.: An execution time neural-CBR guidance assistant. Neurocomputing 72(13), 2743–2753 (2009)

[36] Corchado, J.M., De Paz, J.F., Rodríguez, S., Bajo, J.: Model of experts for decision support in the diagnosis of leukemia patients. Artificial Intelligence in Medicine 46(3), 179–200 (2009)

[37] LeCun, Y., Bottou, L., Bengio, Y., Haffner, P.: Gradient-based learning applied to document recognition. Proceedings of the IEEE 86(11), 2278–2324 (1998), doi:10.1109/5.726791

[38] Parihk, M., Pately, M., Bhat, D.: Animal Detection Using Template Matching Algorithm. International Journal of Research in Modern Engineering and Emerging Technology 1(3) (2013)

[39] Mahony, R., Kumar, V., Corke, P.: Multirotor aerial vehicles: Modeling, estimation, and control of quadrotor (2012)

[40] Borrajo, M.L., Baruque, B., Corchado, E., Bajo, J., Corchado, J.M.: Hybrid neural intelligent system to predict business failure in small-to-medium-size enterprises. International Journal of Neural Systems 21(4), 277–296 (2011)

[41] Ramli, H., Kuntjoro, W., Makhtar, A.K.: Advanced Autonomous Multirotor Response System. Applied Mechanics and Materials 393, 299–304 (2013)

[42]   Rodriguez, S., Julián, V., Bajo, J., Carrascosa, C., Botti, V., Corchado, J.M.: Agent-based virtual organization architecture. Engineering Applications of Artificial Intelligence 24(5), 895–910

[43]   Rodríguez, S., Pérez-Lancho, B., De Paz, J.F., Bajo, J., Corchado, J.M.: Ovamah: Multiagent-based adaptive virtual organizations. In: 12th International Conference on Information Fusion, FUSION 2009, pp. 990–997 (2009)

[44]   Rodríguez, S., de Paz, Y., Bajo, J., Corchado, J.M.: Social-based planning model for multiagent systems. Expert Systems with Applications 38(10), 13005–13023 (2011)

[45]   Macrofauna, S., Lavelle, P., Senapati, B., Barros, E.: Trees, Crops and Soil Fertility: Concepts and Research Methods, pp. 303–323 (2003)

[46]   Svanfeldt, M.: Design of the hardware platform for the flight control system in an unmanned aerial vehicle (Doctoral dissertation. Linköping) (2010)

[47]   Tretyakov, V., Surmann, H.: Hardware architecture of a four-rotor UAV for USAR/WSAR scenarios. In: Workshop Proceedings of SIMPAR 2008-International Conference on Simulation, Modeling and Programming for Autonomous Robots (2008)

[48]   Zato, C., Villarrubia, G., Sánchez, A., Bajo, J., Corchado, J.M.: PANGEA: A New Platform for Developing Virtual Organizations of Agents. International Journal of Artificial Intelligence TM 11(A13), 93–102 (2013)