

A Supervised Approach to Electric Tower Detection and Classification for Power Line Inspection

Carlos Sampedro, Carol Martinez, Aneesh Chauhan, and Pascual Campoy

Abstract— Inspection of power line infrastructures must be periodically conducted by electric companies in order to ensure reliable electric power distribution. Research efforts are focused on automating the power line inspection process by looking for strategies that satisfy the different requirements of the inspection: simultaneously detect transmission towers, check for defects, and analyze security distances. Following this direction, this paper proposes a supervised learning approach for solving the tower detection and classification problem, where HOG (Histograms of Oriented Gradients) features are used to train two MLP (multi-layer perceptron) neural networks. The first classifier is used for background-foreground segmentation, and the second multi-class MLP is used for classifying within 4 different types of electric towers. A thorough evaluation of the tower detection and classification approach has been carried out on image data from real inspections tasks with different types of towers and backgrounds. In the different evaluations, highly encouraging results were obtained. This shows that a learning-based approach is a promising technique for power line inspection.

I. INTRODUCTION

Electric power companies invest significantly on the inspection and preemptive maintenance of the power line infrastructure. The most common strategy is to perform aerial inspection of the power line corridor, at regular intervals. The traditional (and the most common) approach to inspection uses a manned helicopter, equipped with multiple sensors - e.g. differential global positioning system (DGPS), coupled with inertial measurement unit (IMU); light and radar sensor, Lidar; visual, infrared and ultra-violet cameras etc. - mounted on gyroscope stabilized gimbals, and an expert crew, for recording and documenting the relevant data captured from these sensors. This data, which is recorded over thousands of kilometers, is later manually examined to detect potential faults and damage on different power line components (such as, cables, towers, insulators etc.). This process is not only extremely time consuming, but also very expensive and prone to human error. Moreover, the manned flights, which are carried out very close to the live power cables, are highly dangerous to the crew. With these problems in mind, power industry is actively seeking solutions to automate different aspects of the power line inspection.

In the last two decades, multiple complementary research directions have been investigated for automating the task of

Computer Vision Group (www.vision4uav.com), Centro de Automática y Robótica CAR UPM-CSIC, Universidad Politécnica de Madrid, José Gutiérrez Abascal 2, 28006 Madrid, Spain. Email: {carlos.sampedro, pas-cual.campoy}@upm.es.

This work was supported by the Spanish Ministry of Industry under the National R&D Program INNPACTO IPT-2012-0491-120000 led by Gas Natural Unón Fenosa, in which INTA, Pryisma, and Diagnóstica participate.

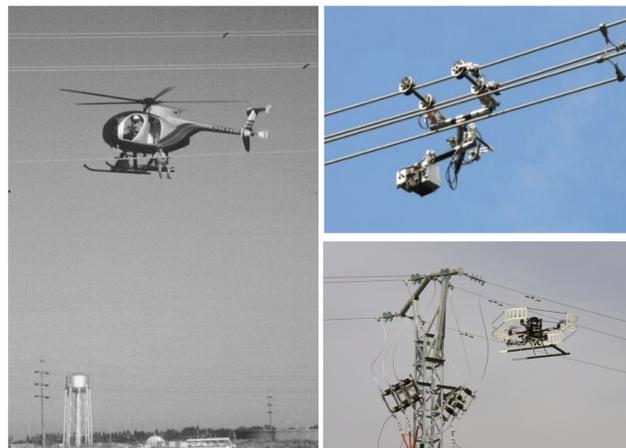


Fig. 1. Three main types of aerial inspection platforms: Manned helicopter [8]; Expliner, a rolling on wire (ROW) robot [5]; and Unmanned Aerial Vehicles (UAVs) [9].

visual inspection (see Fig. 1). One key direction has been on developing unmanned robotic vehicles for autonomously inspecting the power line corridor [1]. Two prominent lines of research have emerged:

- Unmanned Aerial Vehicles (UAV) [2], [3], [4]; and
- Rolling on Wire (ROW) robots, also known as climbing robots [5], [6].

More recently, some authors have also proposed a hybrid climbing-flying robot which combines the advantages of UAVs and ROW robots into a single platform [7].

In addition to the type of robotic platform, another key research direction has been on applying existing or suitably modified computer vision algorithms for automating the inspection process.

Power line inspection is usually divided into two steps: data collection and fault identification. In autonomous data collection, computer vision approaches have been investigated primarily on UAVs and remotely piloted vehicles, for visual control [3], tracking of power lines [10] or electric towers [11], [12], and obstacle detection [13]. The second step involves the detection of faults in the power line infrastructure [14]. This is usually carried out at a base station once the data has been collected. Automated inspection, in this step, is directed towards detection and localization of electric wires [4], towers [15], insulators [14], conductors, wire-clamps etc., in the captured video data.

In the recent years, research community has primarily focused on power line detection. For this task, a common preprocessing strategy is to detect line segment candidates in

an image, and find the segments which form part of the power lines [4], [12]. Line detection is primarily useful for detecting faults such as sagging and proximity to the vegetation or man-made structures.

Some researchers have also focused on the detection and segmentation of electric towers in the images [11], [12], [16], [15]. Similar to the power line detection approaches, preprocessing stage for tower detection generally involves locating the line segments in the image. Several authors apply either further filters or predefined rules/heuristics to remove the noisy line segments; and then apply different segmentation approaches to extract the complete tower from the image: e.g. a template matching approach is used in [11]; graph-cut based segmentation is used in [16]; a rule-based, as well as watershed segmentation is used in [12]. On the other hand in [15], instead of lines, corners were considered the key identifying features of a tower. They used a modified corner detector to detect and track the tower tops.

Although different approaches to tower detection and segmentation have reported promising results, most of the results have been reported on just one type of tower, and several simplifying assumptions were made (mostly with respect to the color, shape and appearance of the tower). However, the electric towers are extremely diverse in shape, appearance and size, as well as they differ in color and the material used for construction (wood, ceramic, steel etc.). Fig. 2 displays 4 common types of towers. Most the state of the art results cannot be generalized to different tower types.

To achieve the goal of complete autonomy, researchers must aim towards developing more general approaches which are able to detect more than one type of tower. Our paper is an effort in this direction. In this paper, a supervised learning approach for solving the tower detection and classification problem is proposed.

Two main contributions are presented in this paper. First, we consider tower detection as a supervised learning problem. To our knowledge, supervised learning for electric tower detection has not been previously reported in literature. We propose a solution using a 2-class neural network classifier for tower-background classification. Moreover, we also raise the problem of classification of different types of electric towers, which has not been addressed previously in literature. We approach this problem using a 4-class neural network for classifying 4 types of electric towers, which is our second key contribution.

A complete solution is proposed for combining tower detection and classification, by integrating the tower type classifier with the tower detection workflow. A sliding-window approach [17] is used to first locate the tower in a given image. In this approach, the tower-background classifier is applied to the subregions/windows inside the image to detect the presence of the tower in that region. Once the tower has been located, the tower type classifier is used to identify the type of tower.

In contrast to most of the recent approaches (see [15] for an exception), which make assumptions regarding the



Fig. 2. Power line inspection. This is a complex problem for computer vision: varying lighting conditions, background changes, power line infrastructure is very complex and heterogeneous (different shapes and sizes), among other problems.

global shape and appearance of the tower, local features are explored in this paper. In particular, a state of the art feature descriptor, Histograms of Oriented Gradients (HOG)[18], is used to train the mentioned MLP (multi-layer perceptron) neural networks for tower detection and classification.

The rest of the paper is organized as follows: Section II states the problem addressed in this paper and describes several challenges which need to be addressed. Our approach to tower detection and classification is presented in Section III. The results are reported and discussed in Section IV; and the final section concludes the paper, as well as points towards future research directions.

II. PROBLEM STATEMENT

For many years, ground patrols and also helicopter patrols have been in charge of the inspection of power line infrastructures. Currently, different projects are looking for automating either the acquisition process or the analysis process, or both, with the main objective of being able to detect and diagnose different faults of the power line infrastructure by using new sensors or by using new inspection platforms (e.g. ROW robots [5], [6]; UAVs [2], [10], [4]).

In all these new possible approaches, computer vision plays an important role for automatically moving the camera in order to maintain the electric tower inside the field of view of the camera, and for identifying and categorizing the different faults in the power line infrastructure.

Nonetheless, although computer vision is a key technique for automating the power line inspection process, it is in fact a very challenging task for this purpose. Power line infrastructures are heterogeneous and complex, for example, as can be seen in Fig. 2, electric towers come in a wide variety of shapes and sizes, and the location of their components also varies depending on the type of tower (e.g. the position of the insulators changes).

Background changes is another problem that the visual system has to deal with. As can be seen in Fig. 2, depending on the terrain, different visual features can be used to segment the electric tower from the background, or to segment the wires. However, because of the high variability of the terrain and the variety of electric towers, it is difficult to find a unique feature (e.g. the color of the towers is not unique) that can work in all the possible scenarios. Illumination changes also play an important role. For example, Fig. 2 shows that in some of the images the contrast between the lines and the background is low and not sufficient to segment the wires.

Another important factor that must be taken into consideration when automating the power line inspection, is the quality of the images, which changes depending on the kind of inspection that is conducted and on the vehicle used for inspection. As can be seen in Fig. 3, when an intensive inspection is conducted (Fig. 3, images on the right) details are perceived much better, and therefore it would be more feasible for a computer vision algorithm to detect faults on those images. Nevertheless, this kind of inspection requires the helicopter to go slow and also to stop in every tower, which implies a considerable increase of the inspection price. In general, for accurate inspections, the quality of the images should be good, but this is currently difficult to ensure, especially at low prices.

Conversely, if a faster inspection (non-intensive inspection) is conducted (see Fig. 3, image on the left), the quality of the images will degrade (blurred images) and only external problems could be analyzed (e.g. the structure of the tower). This is also a problem that could be found when exploring a UAV-based approach. With UAVs, constant vibrations and payload restrictions make the acquisition of high quality images a very difficult task, and therefore, making the process of detecting faults in those images extremely difficult.

Other problems such as constant viewpoint changes (e.g. especially when cameras are manually moved) and scale changes of the electric tower and its components add additional complexity to the idea of applying computer vision to solve this problem, in which, depending on the adopted strategy, could require a system that automatically defines which is the best frame to be used for detecting faults.

Currently, there is no solution that satisfies the different requirements of automated power line inspection: simultaneously detect electric towers, detect and analyze faults, and also analyze security distances to the power line infrastructure. Power line inspection is still an open area of research, where in terms of cost-benefits, it is important for electric companies not only to have a system that can deal with the different requirements of the inspections, but also to have a system that can do it at high speeds.

In this paper, we explore the electric tower detection and classification problem applying a machine learning approach, using low quality images. We believe this is a key step to develop more complex tasks such as fault detection and analysis.

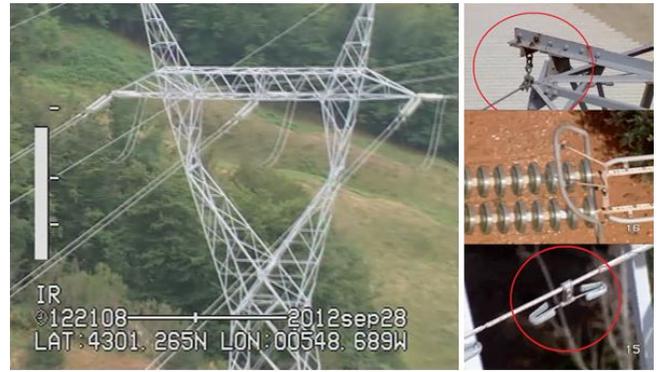


Fig. 3. Visual inspection. The image quality changes depending on the kind of inspection. Intensive inspections allow high quality images (images on the right), whereas a fast inspection (non-intensive) make it difficult to conduct a detailed analysis of the components of the electric tower (image on the left).

III. TOWER DETECTION AND CLASSIFICATION STRATEGY

The objective of the proposed strategy is to determine the position of the electric tower and the type of tower, in single images. Due to the difficulty of the task (e.g. wide variety of backgrounds), a learning-based approach is used. The strategy is based on two stages. In the first stage (tower detection stage), a neural network classifier is trained for tower-background classification, and in the second stage (tower classification stage), a 4-class neural network classifier is trained for identifying the type of tower. In both stages, HOG features [18] are used to train two MLP neural networks. Once the two MLP classifiers have been trained, they are applied for tower detection and classification for power line inspection. In the following paragraphs the system architecture is described.

A. System Architecture

The proposed strategy for power line inspection is based on the interaction between a tower detection and a tower classification stage as shown in Fig. 4. As input the system receives a color image; and the output of the system, if it finds a tower, corresponds to the position of the tower and the type of tower contained in the image.

Fig. 4(a) describes the workflow of the tower detection stage. In order to apply the trained tower detection classifier to the input image, the color image is converted into grayscale, and a sliding-window approach is used to scan the image. As shown in Fig. 4(a), a small window SW of a predefined size is slid over the image. In our strategy two different window sizes are used ($SW_1:160 \times 290$ pixels, and $SW_2:130 \times 260$ pixels). The size of these windows has been defined based on the average size of the tower images used for training the classifiers. Each window SW , provided by the sliding window algorithm is resized to 64×128 pixels, and then from this image HOG features are extracted. The resulting HOG feature vector (of size 3780) is passed as input to the MLP classifier trained for tower detection, where

the window **SW** will be classified as Tower or Background, using the following criteria:

$$\text{Class} = \begin{cases} 1 & , \text{ if } (a_1 \geq 0.98 \ \& \ a_2 \leq 0.001) \\ 2 & , \text{ otherwise} \end{cases}$$

where a_1 and a_2 are the activation values of the output layer neurons for Class 1 (Tower) and Class 2 (Background), respectively.

The position in the image of all windows **SW** that have been classified as Class 1 (Tower) are then saved (see Fig. 4(a), red boxes, image on the right). Finally, when the sliding window algorithm has finished scanning the image, the result from the detection stage is obtained as the bounding box that covers all windows **SW** that were saved. This ROI (region of interest) shown in Fig. 4(a) (green box, image on the right), corresponds to the final result from the detection stage.

The result of the tower detection stage is used as input to the tower classification stage, as described in Fig. 4(b). This ROI is resized to 64×128 pixels, and then, HOG features are extracted. The resulting HOG feature vector (of size 3780) is passed as input to a 4-class MLP trained for tower classification, which will be in charge of defining to which class the ROI belongs to: Type 1, Type 2, Type3, or Type 4 (see Fig. 4(b)).

B. HOG descriptor

Histograms of Oriented Gradients (HOG) are used in this paper as features to describe the shape of electric towers, and its application for power line inspection is explored. The general idea of the use of the HOG descriptor is that the local appearance and shape of an object can often be described by the distribution of intensity gradients or edge directions, as it is mentioned in [18].

The first stage of the algorithm consists in calculating the gradient along two directions in order to obtain the magnitude and direction of the gradient at every pixel. This is conducted applying the 1-D $[-1,0,1]$ and $[-1,0,1]^T$ masks to the 64×128 resized image. Then, the image is divided into small regions of 8×8 pixels size, called “cells”. For each cell, a local 1-D histogram of gradients is calculated over all the pixels in the cell. This histogram consists in 9 orientation bins, evenly spaced over $0 - 180^\circ$ (“unsigned” gradient). Then, as it is mentioned in [18], in order to reduce aliasing, votes are interpolated bilinearly between the neighboring bin centres, and the gradient magnitudes of the pixels in the cell are used to vote into the histogram.

The next step of the algorithm consists in normalizing the oriented histograms in order to get invariance to illumination changes and foreground-background contrast. This is conducted using blocks of 2×2 cells. The blocks are overlapped 50% so that each cell histogram contributes with several components to the final feature vector, each of them normalized with respect to a different block of cells. The final HOG feature vector is obtained by collecting all the values from the normalized blocks. With this procedure, a HOG

descriptor of size 3780 is obtained, which will be used for tower detection and classification for power line inspection.

C. MLP classifiers

Two feed-forward backpropagation neural networks are used for the tower detection and the tower classification stages shown in Fig. 4. Both neural networks use a sigmoid activation function and the algorithm used for training these networks is the *Resilient Backpropagation* algorithm [19]. One of the advantage of this algorithm is its low computational cost, which allows to quickly train and evaluate different neural network configurations.

The configuration of the neural network used in the tower detection stage is a 3-layers MLP with 10 neurons on the hidden layer and 2 neurons on the output layer, and for the tower classification stage, a 3-layers MLP with 40 neurons on the hidden layer and 4 neurons on the output layer, is used.

IV. EXPERIMENTS AND RESULTS

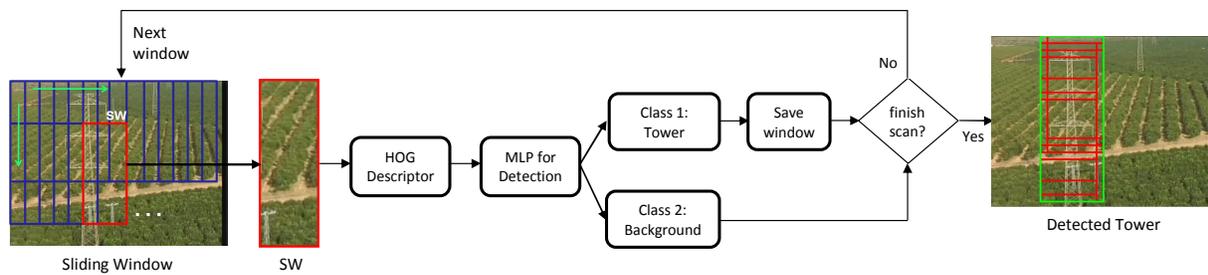
This section begins by describing the data used for training and evaluating the MLPs and the complete system (the complete tower detection and classification pipeline). The methodology to train and evaluate the two classifiers is also presented. After an independent performance evaluation of the MLPs, the evaluation of the complete system is assessed.

A. Experimental Set-up

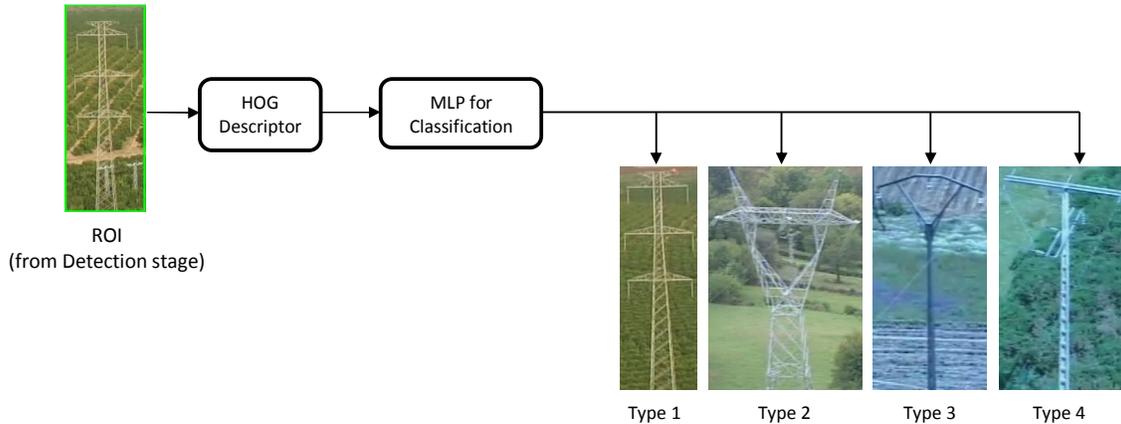
Currently there are no publicly available datasets of power line inspection. Proprietary aerial inspection data was made available by an electric power company. The data consists of 11 videos captured during multiple manned aerial inspections. 6 of these videos primarily contain inspections of towers supporting high voltage lines (Type 1 and Type 2 towers) and the other 5 videos contain inspections of towers for medium voltage lines (Type 3 and Type 4). The inspections were non-intensive, therefore the video quality is relatively poor. The resolution of the frames is also low: for Type 1 and Type 2 towers, the average frame size is 550×480 , and for Type 3 and Type 4, the average frame size is 720×576 .

From these videos, a dataset of cropped images was created where each of those images was either labeled as Background or as Tower, indicating, in the latter case, also the type of the tower. To collect this data, two software tools were created:

- Data acquisition tool: Given all the frames of a video, this tool allows a human user to traverse through each frame sequentially or randomly. From any chosen image, the user can select a rectangular region, which can contain a tower or part of the background. Finally, for a selected region, the tool allows to provide the label, for example, if the region containing the tower is selected, user can provide the type of the tower, otherwise label the region as Background.
- Label correction tool: Labeling process is a time consuming and tiring process. It is possible that some



(a) Tower detection stage.



(b) Tower classification stage.

Fig. 4. System architecture. The proposed strategy for power line inspection is based on the interaction between a tower detection Fig. 4(a) and a tower classification stage Fig. 4(b). In the first stage (tower detection stage), a neural network classifier is trained for tower-background classification, and in the second stage (tower classification stage), a 4-class neural network classifier is trained for identifying the type of tower. In both stages, HOG (Histograms of Oriented Gradients) features are used.

labeling errors can occur. This tool allows the user to see the cropped images and the associated labels, and correct them in case there is a labeling mistake.

The data acquisition tool was used to collect and label 3200 image regions (1600 regions containing tower and 1600 containing background) from 11 videos. For 1600 tower images, 400 image regions of each type were labeled. Later, the label correction tool was applied to remove any labeling mistakes. Finally, all the image regions were resized to the size of 64×128 . Fig. 5 can give the reader an idea of the labeled images of different types of towers and background.

The experiments have been carried out using the Matlab Neural Network Toolbox, and the HOG descriptor implementation developed in [20].

B. Training and Evaluation Methodology

In order to train and evaluate the MLP for detection, 3200 images have been divided into 3 sets: training, cross validation, and test set. 1200 images of each class (tower and background) have been used for training, while 200 images of each class are used for the cross validation and 200 of each class for the test set. The images belonging to the tower class have to be equally distributed according to each type, such that 300 images of each type of tower are used for training, and 50 images of each type are used for validation, and 50 for test.

For training and evaluating the MLP for classifying tower types, 1600 images of electric towers (Fig. 5(a)) have been divided into training, cross validation and test set. From these images, 300 of each tower type (Type 1 to Type 4) have been used for training, 200 (50 images per tower type) for cross-validation and another 200 for testing.

C. Results and Discussion

Table I shows the confusion matrix obtained on testing the MLP used in the detection stage. A total test error of 3.25% is attained. A false positive rate of 2.5% was achieved, which means that only 5 of the 200 background test images were incorrectly classified as tower. On the other hand, we obtain a false negative rate of 4%, which indicates that 8 tower images, out of 200 used for testing, were predicted as background. These results suggest that, although overall performance of the classifier is good, tower images get predicted as background more often than background images as tower.

The errors in the detection stage will have significant influence on the complete system, since the prediction errors get carried forward to the tower classification stage. More specifically, the regions detected as Tower, which were actually Background, will always lead to prediction errors in the tower-type classification stage. That is, from the perspective of the complete system, it is more favorable to have less



(a) Tower training examples



(b) Background training examples

Fig. 5. Examples of cropped images of: (a) 4 tower images and (b) 5 background images. These and similar images are used for training and evaluation of the MLPs for tower detection and tower-type classification.

false positives than false negatives in the detection stage. Since the false positives in the evaluation of the detection stage, are relatively low, we believe this MLP configuration is suitable for being applied to the complete tower detection-classification pipeline.

TABLE I

CONFUSION MATRIX OBTAINED FOR THE TOWER DETECTION TEST SET

Predicted class \ Actual class	Tower	Background
Tower (%)	96	2.5
Background (%)	4	97.5

Table II presents the confusion matrix corresponding to the tower classification MLP tested with the test set of tower images. In this results it can be seen that towers of Type 1 and 2 are the most likely to be well classified, obtaining a classification accuracy of 98% and 96% respectively, while towers Type 3 and 4 are the hardest one in the classification task, obtaining a classification accuracy of 94% and 92% respectively. It is interesting to see that most of the false positives obtained for Type 3 correspond to tower Type 4 and vice versa. These obtained classification results seem to be reasonable due to the fact that Types 3 and 4 correspond to medium-voltage towers (Fig. 5(a), the two images on the right), which are mainly identified by their vertical pole. In contrast, towers Type 1 and 2 correspond to high-voltage towers (Fig. 5(a), the two images on the left), which have a more complex structure, and therefore a more complex HOG pattern, very different from the one of the other tower types.

TABLE II

CONFUSION MATRIX OBTAINED FOR THE TOWER CLASSIFICATION TEST SET

Predicted class \ Actual class	Type 1	Type 2	Type 3	Type 4
Type 1 (%)	98	4	2	2
Type 2 (%)	0	96	0	0
Type 3 (%)	0	0	94	6
Type 4 (%)	2	0	4	92

D. Evaluation of the complete system

Finally, the complete system is evaluated. Therefore, given an input image, the tower detection and classification stages are applied together, as explained in section III-A, obtaining as output a final ROI (result of de detection stage) with the associated tower type (result of the classification stage). For the purpose of evaluating the complete system, 120 completely new images (not cropped, like the labeled ones used for training the neural networks) were collected from the video data, 60 containing no tower, and 60 (15 per tower type) containing an electric tower. These images were not used in the training and evaluation of the MLPs.

Table III presents the confusion matrix of the tower detection stage when the complete system was tested. In this table, it can be seen that when there is an electric tower in the image, it is detected in 92% of the cases. That is, only in 5 images, of the 60 images containing a tower, the detection stage has missed it.

On the other hand, the false positive rate (background region detected as a tower) is 25%. Although, this result appear contrary to the one presented in Table I, it is important to take into account that Table III shows the results after scanning the complete image with a sliding window approach. Thus, Table III should be analyzed from the perspective of the number of sliding windows per image, which in this test can range anywhere between 550 and 900. Making a cautious assumption that most of the windows in an image only contain background regions, then the real false positive rate is much lower.

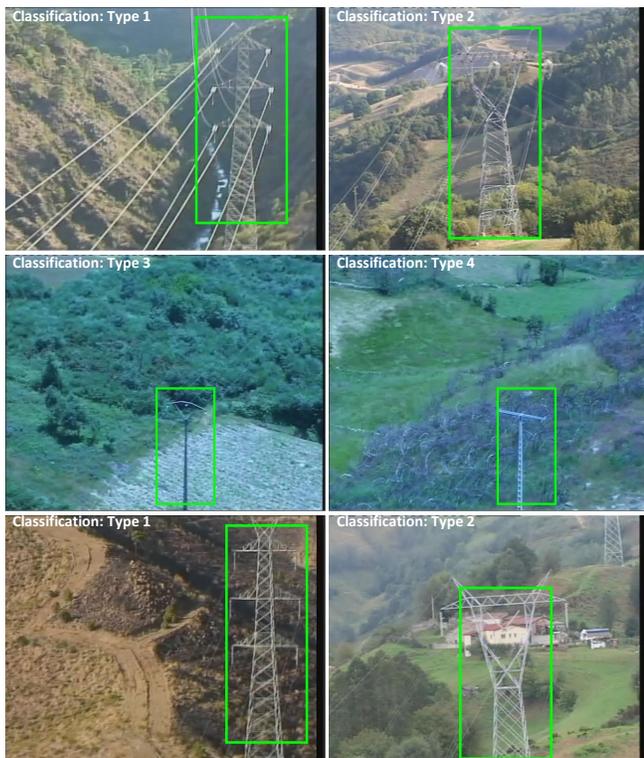
TABLE III

CONFUSION MATRIX OF THE TOWER DETECTION STAGE OF THE COMPLETE SYSTEM.

Predicted class \ Actual class	Tower	Background
Tower (%)	91.67	25
Background (%)	8.33	75

Table IV presents the confusion matrix of the tower classification stage when the complete system was evaluated. Note that, in Table IV, the results do not show the false positives (background detected as a tower) of the detection stage.

This results presented in Table IV are very promising. The towers with complex structure, Types 1 and 2, lead to 93% and 87% accuracy. Due to the complexity of the structure, as captured by the HOG features, these two types of towers



(a)



(b)



(c)

Fig. 6. Detection and classification results obtained during evaluation of the complete system: (a) Examples of correct detection and classification for different types of towers; (b) Examples of incorrect detection; and (c) Examples of incorrect classification.

do not get confused with Types 3 and 4. Type 4 towers get predicted correctly in 87% of the cases. It has to be mentioned that results presented in Table IV (classification stage) are highly dependent on the final ROI obtained from the detection stage, which depends on the sliding window algorithm (currently based in two window sizes). That is,

TABLE IV

CONFUSION MATRIX OF THE TOWER CLASSIFICATION STAGE OF THE COMPLETE SYSTEM.

Predicted class \ Actual class	Type 1	Type 2	Type 3	Type 4
	Type 1 (%)	93.33	6.66	20
Type 2 (%)	6.67	86.67	6.67	0
Type 3 (%)	0	0	60	0
Type 4 (%)	0	0	13.33	86.67

when accurate tower detections are achieved, classification results of the complete system could lead to similar results such as those presented in Table II.

Fig. 6 shows a few tower detection and classification results obtained during the evaluation of the complete system.¹

As shown in Fig. 6(a) good results are obtained in highly cluttered backgrounds, with varying illumination, color, texture, and for the different types of tower that we have considered. In the figure, it can be seen that the towers are properly detected even with a very complex background with vertical structures in the terrain and even with houses or other parts of electric towers in the scene.

Several poor cases were also observed in the detection, as well as, in the classification stages, as shown in Fig(s). 6(b) and 6(c) respectively. However, it is important to note that the results were achieved with a relatively small dataset. More labeled data is expected to further improve, both the detection and the classification stages.

V. CONCLUSIONS AND FUTURE DIRECTIONS

Power line infrastructures are heterogeneous and complex, making automatic power line inspection a difficult problem. To achieve the goal of autonomous inspection, research efforts must aim towards developing general approaches that satisfy several requirements: e.g. simultaneous detection of power lines and electric towers, fault detection in several power line components, analysis of security distances, among others. The current paper is an effort in this direction, with emphasis on electric tower detection and classification in aerial inspection data. We believe this is a key stage to be able to develop more complex tasks such as fault analysis.

A learning paradigm, based on two feed-forward back-propagation MLP neural networks, has been investigated in this paper for solving the tower detection and classification problem during power line inspection. The first MLP has been trained for tower-background segmentation, and a second MLP has been trained for identifying 4 different types of electric towers. Both MLPs were trained using HOG features. To our knowledge, the problem of tower detection and classification in video sequences has not been addressed as a machine learning problem, which are the key novelties of this paper.

A thorough evaluation of the tower detection and classification approach has been carried out using image data from

¹A video demonstration of the reported results has been made available at: <http://youtu.be/iZmuOOXB4ps>

real visual inspection tasks. In the independent evaluations of the MLPs, highly encouraging results were obtained. Tower detection was shown to be very robust in several challenging environments with cluttered background, varying illumination, different tower shapes and viewpoints, etc. This shows that a learning-based approach is a promising direction for power line inspection, which can be generalized to work in multiple environments, and with multiple tower types and power line components, if the appropriate data for training the neural networks is available.

One of the main reasons for the good performance is due to the use of local shape and appearance features, HOG, for image region representation. However, in addition to the HOG features, simpler feature spaces can be simultaneously explored, especially for towers with a simple structure (e.g. medium voltage towers).

Therefore, immediate future work is lined towards exploring other feature spaces to achieve better discrimination. Another promising direction is to use ensemble classifiers where multiple classifiers are trained on different features. This can enhance the performance of the detection as well as the classification stages. Visual tracking is also anticipated to significantly enhance the results from tower detection. Finally, future direction is also focused on extending the system for automatic fault detection and analysis by fusing information from different sensor (e.g. infrared cameras and Lidar)

VI. ACKNOWLEDGEMENTS

This work has been supported by the Spanish Ministry of Industry under the National R&D Program INNPACTO IPT-2012-0491-120000, the European funding FEDER and the Spanish Ministry of Science under grant MICYT DPI2010-20751-C02-01. The authors would like to thank the Spanish companies Gas Natural Unión Fenosa and Prysma for the aerial inspection data supplied within mentioned R&D project.

REFERENCES

- [1] A. Pagnano, M. Höpf, and R. Teti, "A roadmap for automated power line inspection. maintenance and repair," *Procedia CIRP*, vol. 12, pp. 234–239, 2013.
- [2] P. Campoy, P. J. García, A. Barrientos, J. del Cerro, I. Aguirre, A. Roa, R. Garcia, and J. M. Muñoz, "An stereoscopic vision system guiding an autonomous helicopter for overhead power cable inspection," in *Robot Vision*. Springer, 2001, pp. 115–124.
- [3] I. Golightly and D. Jones, "Visual control of an unmanned aerial vehicle for power line inspection," in *Advanced Robotics, 2005. ICAR'05. Proceedings., 12th International Conference on*. IEEE, 2005, pp. 288–295.
- [4] Z. Li, Y. Liu, R. Walker, R. Hayward, and J. Zhang, "Towards automatic power line detection for a uav surveillance system using pulse coupled neural filter and an improved hough transform," *Machine Vision and Applications*, vol. 21, no. 5, pp. 677–686, 2010.
- [5] P. Debenest, M. Guarnieri, K. Takita, E. F. Fukushima, S. Hirose, K. Tamura, A. Kimura, H. Kubokawa, N. Iwama, and F. Shiga, "Expliner-robot for inspection of transmission lines," in *Robotics and Automation, 2008. ICRA 2008. IEEE International Conference on*. IEEE, 2008, pp. 3978–3984.
- [6] S. Peungsungwal, B. Pungsiri, K. Chamnongthai, and M. Okuda, "Autonomous robot for a power transmission line inspection," in *Circuits and Systems, 2001. ISCAS 2001. The 2001 IEEE International Symposium on*, vol. 3. IEEE, 2001, pp. 121–124.
- [7] J. Katrasnik, F. Pernus, and B. Likar, "New robot for power line inspection," in *Robotics, Automation and Mechatronics, 2008 IEEE Conference on*. IEEE, 2008, pp. 1195–1200.
- [8] T. M. Shoemaker and J. E. Mack, *The lineman's and cableman's handbook*. McGraw-Hill, 2007, ch. 42, Figure 42.2.
- [9] Computer Vision Group, "Universidad Politécnica de Madrid." <http://www.vision4uav.com/>.
- [10] D. Jones, "Power line inspection-a uav concept," in *Autonomous Systems, 2005. The IEE Forum on (Ref. No. 2005/11271)*. IET, 2005, pp. 8–pp.
- [11] C. Whitworth, A. Duller, D. Jones, and G. Earp, "Aerial video inspection of overhead power lines," *Power Engineering Journal*, vol. 15, no. 1, pp. 25–32, 2001.
- [12] C. Sun, R. Jones, H. Talbot, X. Wu, K. Cheong, R. Beare, M. Buckley, and M. Berman, "Measuring the distance of vegetation from powerlines using stereo vision," *ISPRS journal of photogrammetry and remote sensing*, vol. 60, no. 4, pp. 269–283, 2006.
- [13] M. Williams, D. Jones, and G. Earp, "Obstacle avoidance during aerial inspection of power lines," *Aircraft Engineering and Aerospace Technology*, vol. 73, no. 5, pp. 472–479, 2001.
- [14] M. Oberweger, A. Wendel, and H. Bischof, "Visual recognition and fault detection for power line insulators." 19th Computer Vision Winter Workshop, 2014.
- [15] I. Golightly and D. Jones, "Corner detection and matching for visual tracking during power line inspection," *Image and Vision Computing*, vol. 21, no. 9, pp. 827–840, 2003.
- [16] W. Cheng and Z. Song, "Power pole detection based on graph cut," in *Image and Signal Processing, 2008. CISP'08. Congress on*, vol. 3. IEEE, 2008, pp. 720–724.
- [17] H. Harzallah, F. Jurie, and C. Schmid, "Combining efficient object localization and image classification," in *Computer Vision, 2009 IEEE 12th International Conference on*. IEEE, 2009, pp. 237–244.
- [18] N. Dalal and B. Triggs, "Histograms of oriented gradients for human detection," in *Computer Vision and Pattern Recognition, 2005. CVPR 2005. IEEE Computer Society Conference on*, vol. 1. IEEE, 2005, pp. 886–893.
- [19] M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: The rprop algorithm," in *Neural Networks, 1993., IEEE International Conference on*. IEEE, 1993, pp. 586–591.
- [20] C. McCormick, "HOG descriptor in matlab (version 1.2) [computer program]." Available at <http://chrisjmccormick.wordpress.com/2013/05/09/hog-descriptor-in-matlab/> (Accessed 20 November 2013), 2012.
- [21] L. Rokach, "Ensemble-based classifiers," *Artificial Intelligence Review*, vol. 33, no. 1-2, pp. 1–39, 2010.