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 features are used to train two MLP (multi-layer perceptron) neural networks. The first classifier is used for background-foreground separation, and the second multi-class MLP is used for classifying 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 that were conducted 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 power line inspection.

In the last two decades, multiple complementary research directions have been investigated for automating the task of visual inspection. One key direction has been on developing

Fig. 1. Three main types of aerial inspection platforms: Manned helicopter [17]; Expliner, a rolling on wire (ROW) robot [11]; and Unmanned Aerial Vehicles (UAVs).

unmanned robotic vehicles for autonomously inspecting the power line corridor (see Figure 1) [1], [2], [3]. Two prominent lines of research have emerged:

- Unmanned Aerial Vehicles (UAV) [4], [5], [6], [7], [8], [9]; and
- Rolling on Wire (ROW) robots, also known as climbing robots [10], [11], [12], [13], [14], [15].

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 [16].

The primary objective is to design and develop techniques that allow these unmanned mobile platforms to perform completely autonomous navigation and inspection of the power line corridors. The degree of autonomy, however, ranges from one robot to another, and complete autonomy is yet to be achieved.

The robots (both UAVs and ROWs), currently in use by the industry, require human intervention for navigation, as well as data acquisition, and complete autonomy of the mobile platform remains a very active area of research. It has been emphasized in recent literature that a marked improvement over the state of the art is necessary at multiple fronts, such as, visual servoing, obstacle detection and avoidance, robust control algorithms, battery recharging etc. [1], [2].

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 segmented

into two steps - data collection and fault identification - where, the field of computer vision has contributed to help automate both steps.

Autonomous data collection will require mobile robots which can navigate and acquire relevant sensor data, with minimal human supervision. In this step, computer vision approaches have been investigated primarily on UAVs and remotely piloted vehicles, for visual control [4], [5], tracking of power lines [5], [6], tracking of electric towers [18], [19], and obstacle detection [9].

The second step involves the detection of faults in the power line infrastructure. 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, towers, insulators, conductors, wire-clamps etc., in the captured video data. Using the visual information, the state of the art has focused on power line and electric tower detection. The detection of power lines and towers can then be used to find various defects/faults in the power line infrastructure.

Several works have been aimed at detecting faults - in insulators, conductors, tower joints, structure discoloration etc. - using data from other sensors than a video camera (although see [20] for an exception). This paper presents a computer-vision based solution, applied to tower detection and classification. Therefore, the current discussion is limited to the works which use information captured from 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 [7], [8], [19], [21], [22]. 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 detection and segmentation of electric towers in the images [23], [18], [19], [24], [25]. Similar to the power line detection approaches, preprocessing stage for tower detection generally involves locating the line segments in the image. This is unsurprising, since linearity is a common and a very useful feature for identifying most likely regions in the image with high likelihood of being a power line or a tower candidates.

For locating and segmenting the tower from the set of most likely candidates, several authors apply either further filters or predefined rules/heuristics to remove the noisy line segments [18], [19], [24], [25]. The remaining line segments are considered as belonging to the tower. Some authors then apply different segmentation approaches to extract the complete tower from the image: e.g. a template matching approach is used in [18]; graph-cut [26] based segmentation is used in [24]; a rule-based, as well as watershed segmentation [27] is used in [19]. Golightly and Jones [23] presented a different approach from the state of the art where, instead of lines, the corners were considered the key identifying features of a tower. They used a modified corner detector [28] to detect and track the tower tops.
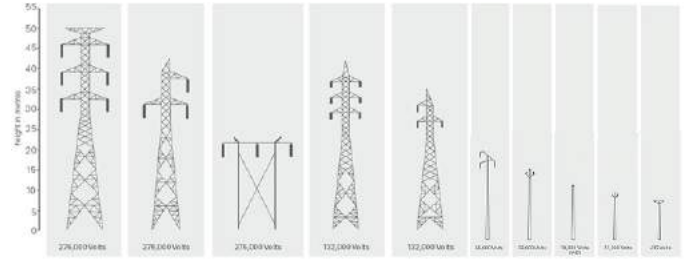

Fig. 2. Exemplars of a few types of electrical towers [29]

Although different approaches to tower detection and segmentation have reported promising results, these results must be viewed from the perspective of the assumptions authors make. Most of the results have been reported on just one type of tower. Since the tower type is fixed, several simplifying assumptions are 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 (see Figure 2), as well as they differ in color and the material used for construction (wood, ceramic, steel etc.). Therefore, most the state of the art results cannot be generalized to several 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 literature, tower detection has been seen as a filtering pipeline. Line segment detectors are applied to detect lines in images. To capture the tower from these set of line segments, further refinement steps (heuristics or more filters) are applied. This strategy, as previously mentioned, is not generic and not useful for more than one type of tower.

In this paper, 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. This is one of the key contributions of this paper. Moreover, we also raise the problem of classification of different types of electric towers. The problem of tower classification has not been addressed previously in literature. We approach this problem using a 4-class neural network for classifying 4 types of electric towers. This is our second key contribution. Finally, a complete solution is implemented which allows simultaneous tower detection and classification.

In contrast to most of the recent approaches (see [23] for an exception), which make assumptions regarding the global shape and appearance of the tower, local features are explored. In particular, a state of the art feature descriptor, histogram of oriented gradients (HOG)[30], is used. Note that the corner features used in [23] can be considered as local features. The choice of learning algorithm is a multi-layer perceptron (MLP) neural network. A labeled dataset (collected from several real manned aerial inspections) of cropped images of towers (of 4 different types) and

background, is used to train the MLP. Figure 6(a) shows the sample images from this dataset. To apply the trained classifier to real images, a sliding-window approach ([31], [32]) is used to first locate the tower in a given image. In this approach, the classifier is applied to the subregions/windows inside the image to detect the presence of the tower in that region.

In addition to tower detection, knowing the type of a tower can greatly facilitate the localization of regions more prone to faults or damage (e.g. insulators, tower tops, power line connection, conductors etc.). Tower classification, therefore, can be a valuable step in automatizing the inspection process. A 4-class MLP was trained on the previously mentioned dataset of cropped tower images. Finally, a complete solution is proposed for combining tower detection and classification, by integrating the tower type classifier with the tower detection workflow.

A thorough evaluation of the tower detection and classification approach was carried out. In the independent evaluation of these approaches, highly encouraging results were obtained. A small drop in performance is observed when the complete detection-classification workflow is evaluated. The reduced performance resulted from the errors carried forward from tower detection from the detection stage to the tower classification stage.

The rest of the paper is organized as follows: Section II states the problem addressed in this paper concretely 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 insfraestructures. 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 defects of the power line infrastructure by using new sensors or by using new inspection platforms (e.g. robots [10], [3], [15]; UAVs [4], [6], [7]).

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 defects and failures of 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 technique. Power line infrastructures are heterogeneous and complex, for example, as can be seen in Fig. 3, 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. 3, depending

Fig. 3. Power line inspection. This is a complex problem for computer vision: varying lighting conditions, background changes, position of components of the tower depends on the kind of tower (insulators' position), and power line infrastructures are very complex an heterogeneous, among other problems.

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, in Fig. 3, it can be seen 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. The image quality changes depending on the kind of inspection that is conducted and on the vehicle used for inspection. As can be seen in Fig. 4, when an exhaustive inspection is conducted (images on the right) details are perceived mush better, and therefore it would be more feasible for a computer vision algorithm to detect defects on those images. Nevertheless, this kind of inspection requires the helicopter to go slow and also to stop in every tower, something that with manned helicopters implies a considerable increase of the inspection price. In general, for accurate inspections, the quality of the images should be good but, this on the other hand, is currently difficult to ensure, especially at low prices.

Conversely, if a faster inspection is conducted (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.

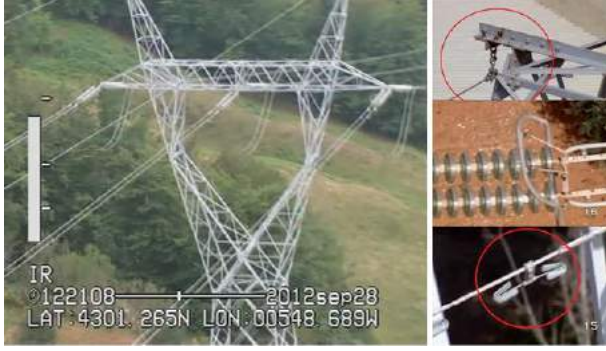Other problems such as constant viewpoint changes (e.g.

Fig. 4. Visual inspection. The image quality changes depending on the kind of inspection. Exhaustive inspection allow high quality images (images on the right), whereas a fast inspection make it difficult to conduct a detailed analysis of the components of the electric tower (image on the left).

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 defects.

Currently, there is not a complete solutoion that satisfies the different requirements of automated power line inspection: simultaneously detect electric towers, check for defects, and also analyze security distances. Therefore, in terms of cost-benefits, it is important for energy companies to solve this problem and try to find a system that can deal with the different requirements of the inspections at high speed.

In this paper we explore the electric tower detection and classification problem applying a machine learning approach, using low quality images. Therefore, the system will help in reducing the maintenance cost of the electric system by being able to cope with one of the problems of increasing the vehicle speed (reducing the quality of the images).

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 (Histogram of Oriented Gradients) features [30] 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 for tower detection and classification 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. 5. 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. 5(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 a grayscale image, and a sliding-window approach is used to scan the image. As shown in Fig. 5(a), a small window W of a predefined size is slid over the image. In our strategy two different window sizes are used (W$_1$:160 × 290 pixels, and W$_2$:130 × 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 W, 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 W will be classified as Tower or Background, using the following criteria:

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

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 the windows W that have been classified as Class 1 (Tower) are then saved (see Fig. 5(a), red boxes, image on the right). Finally, when the sliding algorithm has finished scanning the entire image, the result from the detection stage is obtained as the bounding box that covers all windows W that were saved. This ROI (region of interest) shown in Fig. 5(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. 5(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. 5(b)).

B. HOG descriptor

Histograms of oriented gradients (HOG) are used in this paper as features to describe the shape of the electric towers, and its use 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 [30].

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 $[-1, 0, 1]^{\text{T}}$ masks

(a) Tower detection stage.

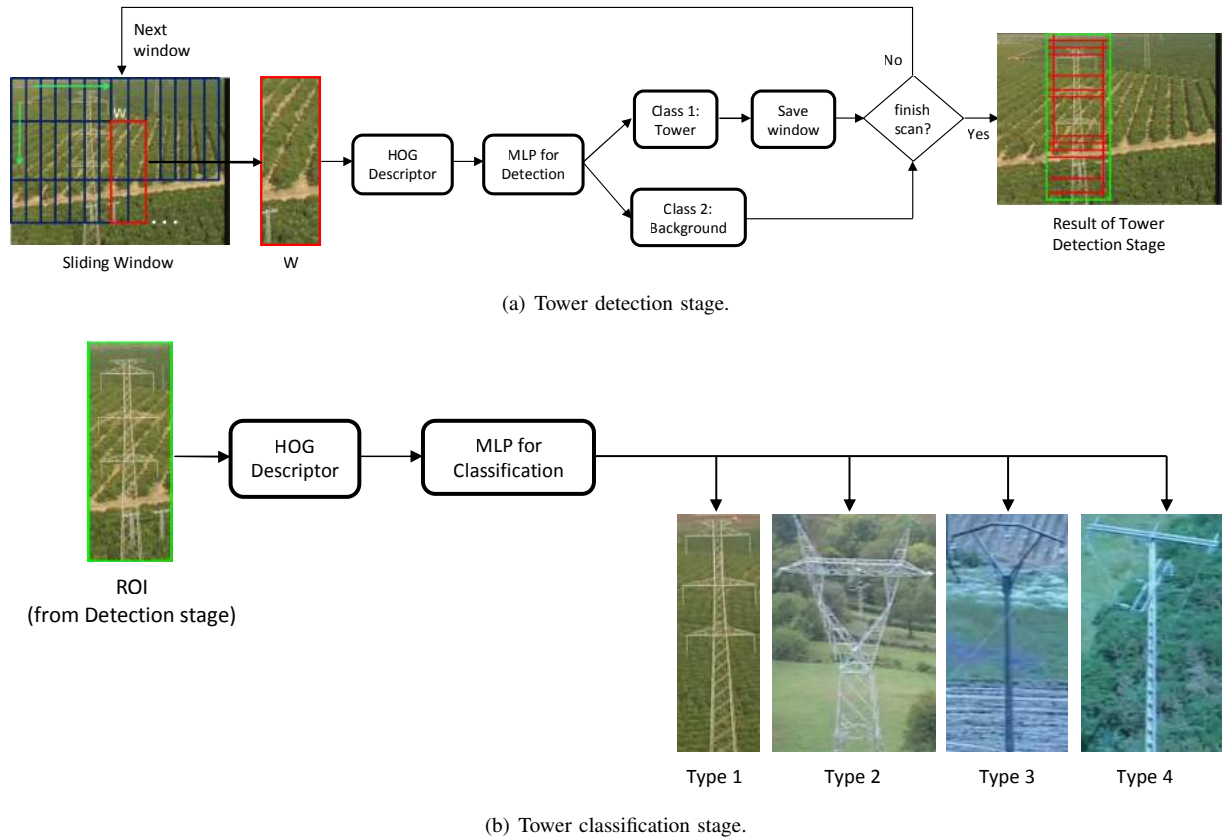(b) Tower classification stage.

Fig. 5. System architecture. The proposed strategy for power line inspection is based on the interaction between a tower detection Fig. 5(a) and a tower classification stage Fig. 5(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 (Histogram of Oriented Gradients) features are used.

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 $0 - 180\,°$ ("unsigned" gradient). Then, as it is mentioned in [30], 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. 5. Both neural networks use a sigmoid

activation function and the algorithm used for training these networks is the *Resilient Backpropagation* algorithm [33]. One of the advantage of this algorithm is its low computational cost [33], 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 evaluation of MLPs and the complete system. The methodology to train and evaluate the two classifiers is also presented. After performance evaluation of the respective MLPs independently, the complete system (the complete tower detection and classification pipeline) is assessed. Although the results obtained are very encouraging, several immediate directions for further improvements are highlighted.

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

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. Figure 6 can give the reader an idea of the resized labeled images of different types of towers and background.

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 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 (Figure 6(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.
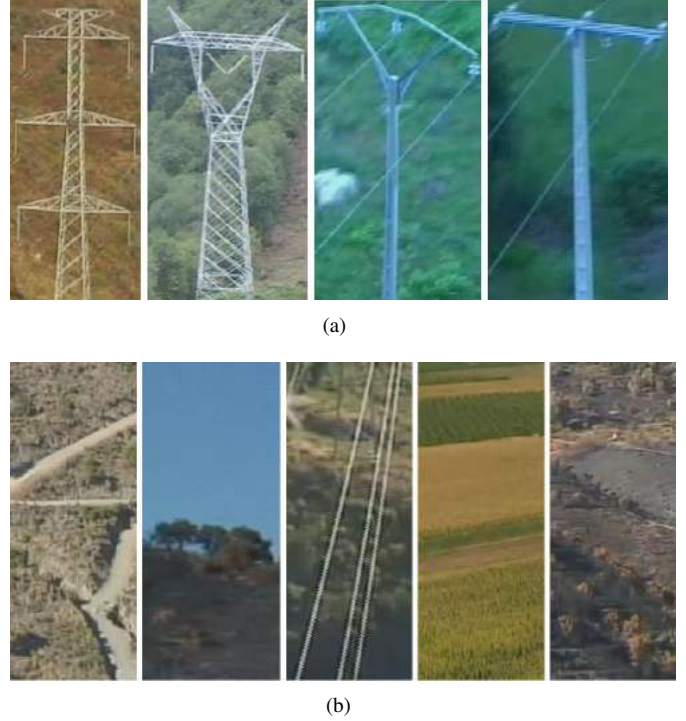

Fig. 6. Examples of cropped images of: (a) 4 cropped tower images and (b) 5 background images. These and similar images are used for training and evaluation of the MLPs for tower detection and tower-type classification.

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 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 II presents the confusion matrix corresponding to the 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

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

Actual class Predicted 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

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 seems to be reasonable due to the fact that Types 3 and 4 correspond to medium-voltage towers (Figure 6(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 (Figure 6(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.

D. Evaluation of the complete system

Actual class Predicted class	Tower	Background
Tower (%)	91.67	26
Background (%)	8.33	74

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

Finally, the complete system is evaluated using the trained MLPs for tower detection and classification. 110 completely new images (not cropped, like the labeled ones used for training the neural networks) were collected from the video data, 50 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.

In Table III 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 26%. These results appear contrary to the ones in Table I. The results reported in Table III are shown with respect to the complete images, which is an important criterion. The same results can also be viewed from the perspective of the number of sliding windows per image. The MLP for detection makes prediction over each sliding window. Two different window sizes are explored for detection (see Section III-B), and since images from different videos are of varying size, the number of classifications per image can range anywhere between 240 and 480. Making a cautious assumption that most of the windows in an image only contain background regions, the real false negative rate is much lower.

All the regions predicted as containing a tower, in the tower detection stage, get passed as input to the MLP for classifying the type of the tower. The results are reported in Table IV. The false negatives in the detection stage belong to the towers of Types 2 (1 incorrect detection), 3 (2 incorrect detections) and 4 (2 incorrect detections). Note that, in Table IV, the results do not show the false positives (background detected as a tower) of the detection stage.

The results obtained in the classification stage 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 do not get confused with Types 3 and 4. Type 4 towers get predicted correctly in 87% of the cases.

Figure 7 shows a few tower detections and classifications obtained during the evaluation of the complete system. [1] As shown in Figure 7(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 Figures 7(b) and 7(c).

A few promising directions can be explored to achieve better discrimination. At the feature-level, in addition to the HOG features, simpler feature spaces can be simultaneously explored. On the other hand, at the classification level ensemble-based classifiers can be explored [34]. In this case, multiple complementary features can then be exploited to train several classifiers where eventual decision is derived from their ensemble. Moreover, it is also important to note that, although good results in tower classification and detection were obtained, the dataset itself was relatively small. More labeled data is expected to further improve, both the detection and the classification stages.

---

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

V. Conclusions and future directions

Power line infrastructures are heterogeneous and complex, making automatic power line inspection a difficult problem. Therefore, there is a current considerable interest in this area of research. To achieve the goal of autonomous inspection, research efforts must aim towards developing general approaches that satisfy the several requirements: e.g. simultaneous detection of power lines and electric towers, check for defects in several power line components, analyze 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 defects analysis, especially when the main source of information comes from poor quality images.

To our knowledge, the problem of tower detection and classification in video sequences has not been addressed as a machine learning problem. The key novelty of this paper is the investigation of a supervised learning framework for providing a complete solution for simultaneous detection and classification of different types of electric towers.

In particular, the learning paradigm investigated is based on two feed-forward back-propagation multi-layer perceptron (MLP) neural networks. The first MLP is a 2-class classifier which is trained for tower-background segmentation. A sliding window approach is applied, where, given an image, each window is classified individually to predict whether the region inside an image is Background or a Tower. Several windows, where towers are detected, are combined to provide a single region of interest (ROI), which is the final prediction of the complete detection process.

The second MLP is a 4-class classifier which is trained for identifying 4 different types of electric towers. The problem of tower classification, from video sequences, has not been explored in the literature previously, and is another key novelty of the paper. The ROI detected as a tower by the first MLP is taken as input by the second MLP and classified as belonging to one of the 4 types of towers explored in this paper.

A thorough evaluation of the tower detection and classification approach has been carried out using image data from real visual inspection tasks. In the independent evaluation 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 viewpoints etc. This shows that a learning-based approach is a promising direction for power line inspection.

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, such features may not be an ideal representation for all types of towers. It was discussed that perhaps simpler features may be useful for towers with a simple structure (Types 3 and 4). This problem was more visible during the evaluation of the complete system, especially with higher number of mis-classifications for the Type 3 tower. Therefore, immediate future work is lined towards exploring other feature spaces. Another promising direction is to use classifier ensembles where multiple classifiers are trained on different features. This can enhance the performance of the detection as well as the classification stages. Finally, visual tracking is also anticipated to significantly enhance the results from tower detection.

VI. Acknowledgements

References

[1] J. Katrasnik, F. Pernus, and B. Likar, "A survey of mobile robots for distribution power line inspection," *Power Delivery, IEEE Transactions on*, vol. 25, no. 1, pp. 485–493, 2010.

[2] 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.

[3] K. Toussaint, N. Pouliot, and S. Montambault, "Transmission line maintenance robots capable of crossing obstacles: State-of-the-art review and challenges ahead," *Journal of Field Robotics*, vol. 26, no. 5, pp. 477–499, 2009.

[4] P. Campoy, P. J. Garcia, 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.

[5] 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.

[6] 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.

[7] 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.

[8] B. Wang, X. Chen, Q. Wang, L. Liu, H. Zhang, and B. Li, "Power line inspection with a flying robot," in *Applied Robotics for the Power Industry (CARPI), 2010 1st International Conference on*. IEEE, 2010, pp. 1–6.

[9] 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.

[10] M. Bähringer, J. Berchtold, M. Bächel, C. Dold, M. Bätikofer, and M. Feuerstein, "Inspection of high voltage power lines - a new approach." Proc. of The 12th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines (CLAWAR), 2009.

[11] 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.

[12] S. Montambault and N. Pouliot, "The hq linerover: contributing to innovation in transmission line maintenance," in *Transmission and Distribution Construction, Operation and Live-Line Maintenance, 2003. 2003 IEEE ESMO. 2003 IEEE 10th International Conference on*. IEEE, 2003, pp. 33–40.

[13] 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.

[14] J. Sawada, K. Kusumoto, Y. Maikawa, T. Munakata, and Y. Ishikawa, "A mobile robot for inspection of power transmission lines," *Power Delivery, IEEE Transactions on*, vol. 6, no. 1, pp. 309–315, 1991.

[15] J. Zhao, R. Guo, L. Cao, and F. Zhang, "Improvement of linerover: A mobile robot for de-icing of transmission lines," in *Applied Robotics for the Power Industry (CARPI), 2010 1st International Conference on*. IEEE, 2010, pp. 1–4.

[16] 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.

[17] T. M. Shoemaker and J. E. Mack, *The lineman's and cableman's handbook*. McGraw-Hill, 2007, ch. 42, Figure 42.2.

[18] 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.

[19] 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.

[20] D. Jones, I. Golightly, J. Roberts, and K. Usher, "Modeling and control of a robotic power line inspection vehicle," in *Computer Aided Control System Design, 2006 IEEE International Conference on Control Applications, 2006 IEEE International Symposium on Intelligent Control, 2006 IEEE*. IEEE, 2006, pp. 632–637.

[21] B. Song and X. Li, "Power line detection from optical images," *Neurocomputing*, 2013.

[22] T. W. Yang, H. Yin, Q. Q. Ruan, J. D. Han, J. T. Qi, Q. Yong, Z. T. Wang, and Z. Q. Sun, "Overhead power line detection from uav video images," in *Mechatronics and Machine Vision in Practice (M2VIP), 2012 19th International Conference*. IEEE, 2012, pp. 74–79.

[23] 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.

[24] 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.

[25] J. Tilawat, N. Theera-Umpon, and S. Auephanwiriyakul, "Automatic detection of electricity pylons in aerial video sequences," in *Electronics and Information Engineering (ICEIE), 2010 International Conference On*, vol. 1. IEEE, 2010, pp. 342–346.

[26] Z. Wu and R. Leahy, "An optimal graph theoretic approach to data clustering: Theory and its application to image segmentation," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 15, no. 11, pp. 1101–1113, 1993.

[27] L. Vincent and P. Soille, "Watersheds in digital spaces: an efficient algorithm based on immersion simulations," *IEEE transactions on pattern analysis and machine intelligence*, vol. 13, no. 6, pp. 583–598, 1991.

[28] J. Cooper, S. Venkatesh, and L. Kitchen, "Early jump-out corner detectors," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 15, no. 8, pp. 823–828, 1993.

[29] G. of Southern Australia. Identifying powerlines. [Online]. Available: http://goo.gl/2ZHwaO

[30] 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.

[31] V. Ferrari, L. Fevrier, F. Jurie, and C. Schmid, "Groups of adjacent contour segments for object detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, vol. 30, no. 1, pp. 36–51, 2008.

[32] 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.

[33] 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.

[34] L. Rokach, "Ensemble-based classifiers," *Artificial Intelligence Review*, vol. 33, no. 1-2, pp. 1–39, 2010.
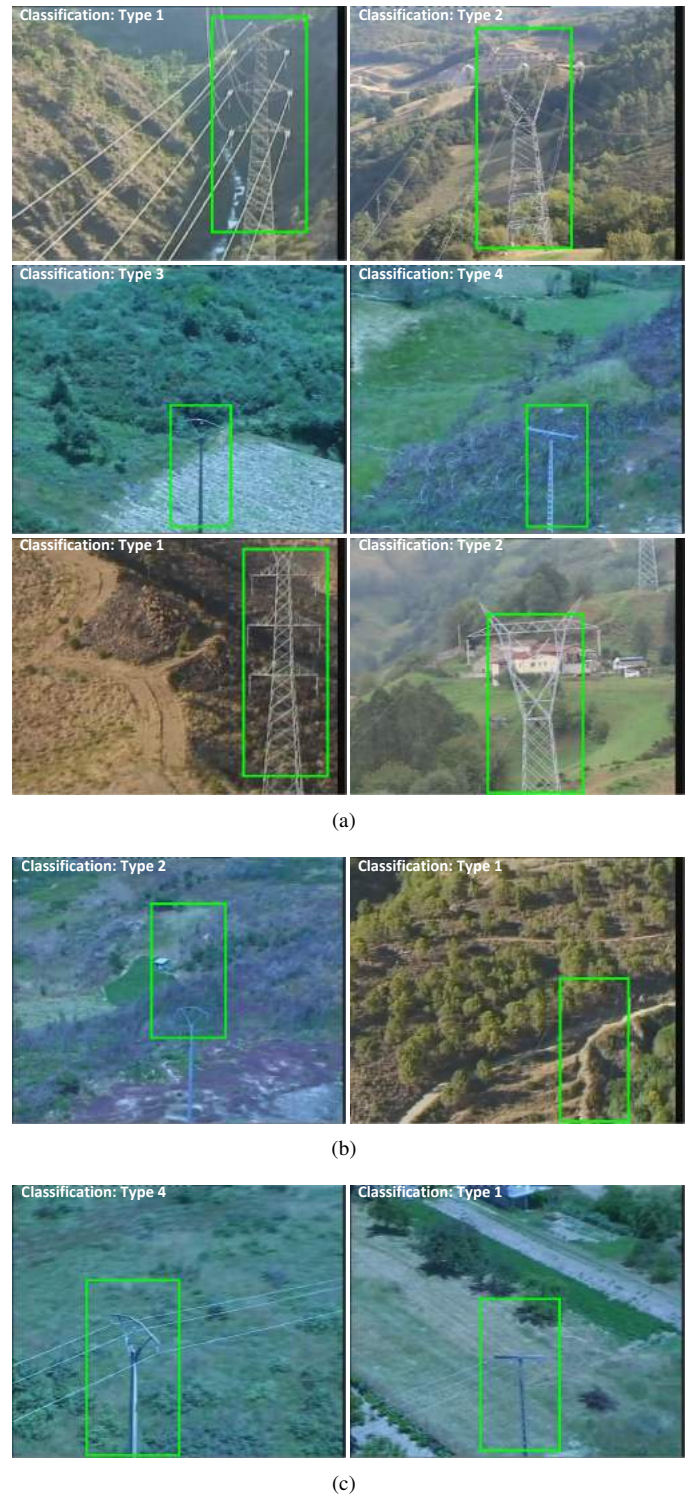
Fig. 7. Detection and classification results obtained during evaluation of the complete system: (a) Examples of correct detection and classification for different types of towers; (b) Examples of incorrect detection; and (c) Examples of incorrect classification.