Semiautonomous pipeline inspection using infrared thermography and unmanned aerial vehicles

R. Usamentiaga, *Senior, IEEE*

***Abstract*— Pipeline inspection is crucial to ensure the safe and efficient operation of pipelines, as well as to prevent potential hazards and disruptions in the transportation of materials. To address this challenge, this work proposes an innovative semiautonomous inspection method that uses unmanned aerial vehicles and infrared thermography to efficiently inspect pipelines. The approach combines the benefits of automated flight and high-speed data acquisition with advanced data processing techniques, enabling the detection of hidden defects and abnormal temperature patterns in pipelines. The proposed procedure presents a comprehensive operational approach that covers the entire inspection process, including flight configuration, data acquisition, and processing. The proposed method for data processing is automated, utilizing vision-based detection and tracking techniques and leveraging the power of deep learning algorithms to ensure robust analysis and inspection. A novel active learning procedure is also proposed, further improving the efficiency and effectiveness of the pipeline inspection process. Extensive tests demonstrate the effectiveness of the proposed procedure in industrial applications. The proposed thermographic system enables the detection and localization of insulation defects and product leaks in pipes, which are critical for maintaining pipeline integrity.**

***Index Terms*— Pipeline inspection, Unmanned aerial vehicles, Infrared thermography, Detection and tracking, Active learning**

I. Introduction

Pipeline inspection helps to ensure the safety and integrity of pipelines, which are critical infrastructure for the transportation of a wide range of materials including liquids, gases, and slurries [1]. These materials can include raw materials, such as oil, natural gas, and ore, as well as finished products, such as gasoline, diesel fuel, and metal products. Pipelines are often the most cost-effective and efficient means of transporting these materials over long distances, as they are able to transport large volumes of material quickly and reliably [2]. This is especially important in the case of oil and natural gas, which are often transported over great distances to reach their destination. In addition to transportation, pipelines can also be used to connect various stages of the production process, which helps to streamline production and make it more efficient.

Several types of pipeline inspection methods are commonly used to assess the condition and integrity of pipelines [3]. The most common methods are based on visual inspection, non-destructive testing (NDT) and in-line inspection (ILI). Visual inspection methods observe the pipeline looking for visible signs of damage or deterioration. This can be done by trained personnel on foot or using aerial platforms such as drones. NDT methods inspect the condition of a pipeline without causing damage or disruption. Examples include ultrasonic testing, radiographic testing and magnetic particle testing. On the other hand, ILI methods involve the use of specialized tools or devices, such as smart PIGs (Pipeline Integrity Gauges). These devices are inserted into the pipeline and propelled along its length to collect data on its condition, such as wall thickness, corrosion, and deformities.

Drones, also known as unmanned aerial vehicles (UAVs), are experiencing growing utilization for visual inspections across various sectors such as construction, manufacturing, and transportation [4]. Novel applications are also becoming popular, such as remote sensing for wildlife and environmental monitoring [5]. In the context of pipeline inspection, drones can be equipped with cameras or other types of sensors to capture images or data about the condition of the pipeline. Visual inspection using drones provides several advantages, as they can access hard-to-reach or hazardous areas and provide more cost-effective solutions than traditional inspection methods. Moreover, they can cover large areas quickly while collecting large amounts of data, making them an efficient option for inspections that would otherwise be time-consuming. The acquired data can be used not only for the generation of detailed reports but also for the identification of trends over time. Overall, visual inspection using drones provides a safe, cost-effective, and efficient means of inspecting pipelines and other structures [6].

Visual inspections using drones are generally carried out by trained personnel that manually operate all aspects of flight, including the control of the camera or other sensors. This way, the position of the drone is adjusted to specific areas of interest. An example of this approach can be found in [7], where a drone-based bridge inspection protocol is established, proposing a five-stage inspection methodology to achieve optimum data collection. In [8] an analysis is presented about the utilization of drone technology in the

This work has been partially funded by the project PID2021-124383OB-I00 of the Spanish National Plan for Research, Development and Innovation.

R. U. is with the Department of Computer Science and Engineering, University of Oviedo, 33204 Gijón, Asturias, Spain (e-mail: rusamentiaga@uniovi.es).

construction industry, with a focus on health and safety, efficiency, and cost of the site inspection process. A major challenge of all these approaches is scale, as only limited areas can be fully inspected. A complete review of the history of unmanned aerial vehicles, along with a comprehensive review of studies focused on UAV-based NDI of industrial and commercial facilities can be found in [9]. In recent research, the application of this technology has extended to include active thermographic inspection [10]. This work focuses on the effectiveness and challenges associated with utilizing active infrared thermography on an unmanned aerial vehicle (UAV) platform. The main objective is to evaluate the performance of small, low-powered excitation sources suitable for active thermography and their ability to detect subsurface defects on composite materials. Other practical issues are analyzed in [11].

Autonomous and semiautonomous systems are recently becoming a possible alternative approach for visual inspections using drones. This type of solution can increase inspection speed and also improve cost efficiency. A solution for inspecting straight oil and gas pipelines is presented in [12], where an autonomous procedure is proposed combining image processing and a controller that ensures the position and orientation of the drone are aligned with the pipeline. Pipeline inspection is also studied in [13] to monitor oil and gas pipelines to detect leakages and cracks. In [14] an autonomous approach is proposed for power line inspection, establishing a perspective model and including a novel pan/tilt monocular-based navigation scheme. Despite the promising research work about autonomous and semiautonomous inspection systems, there are still some shortcomings and challenges that need to be solved, including reliability, safety and data management issues.

This work proposes a semiautonomous solution for pipeline inspection combining infrared thermography and unmanned aerial vehicles. This work follows a passive approach, i.e., no external heat sources are applied [15]. To address the challenge of scale in pipeline inspection, this work proposes a semi-autonomous approach where drones fly fully autonomously without any human interactions. This is achieved by using previously configured flight plans. The pre-programmed flight follows a path where the pipeline is located. During the flight, the drone acquires infrared video at high-speed rates, enabling the drone to fly very quickly and cover long distances. In the proposed procedure, while the battery is swapped, the storage card with the recording is also substituted, and a new pre-programmed flight is configured with a different section of the pipeline system. During the new flight, the previous recording is processed through automated workflows to robustly turn the raw data into quantitative analytics, resulting in simplified, decision-ready information, backed by data. Data processing involves the use of vision-based detection and tracking methods, with deep learning serving as the foundation for robust data analysis and inspection. The proposed work addresses important challenges found in automated visual inspection, such as navigation inaccuracies, variability caused by factors such as different weather conditions, processing speed, image quality, data management and object recognition and tracking to provide measurements about individual pipes, including a novel active learning methodology. The proposed approach is based on infrared images that can be used to detect hidden defects that may not be visible to the naked eye and also are sensitive to temperature variations, allowing for the detection of hot spots and abnormal temperature patterns. Moreover, it can be used to inspect pipelines during the night and to identify temperature trends over time, enabling the comparison of temperature readings in different sections of the installation. This work conducts extensive tests in an industrial facility as it is a sector that heavily relies on pipelines for the transportation of materials, demonstrating the effectiveness of the proposed approach.

In this work, the utilization of infrared technology is paramount, as it plays a crucial role in detecting and locating defects in pipe insulation and identifying leaks in the transported product. Although the proposed approach does not specifically address particular defects, it offers a comprehensive characterization of each pipe's temperature using various temperature intensity features, including mean, median, standard deviation, mode, maximum, minimum, quartiles, moments, skewness, kurtosis, and peak height. These measurements enable the easy identification of abnormal temperature patterns. Furthermore, this approach facilitates the creation of historical records, which can be utilized for predictive maintenance purposes. The incorporation of infrared thermography significantly contributes to verifying the integrity of the pipeline, ensuring that the entire transport system remains fully sealed. This is of utmost importance as leaks not only result in substantial energy and economic losses but also pose a threat to the environment.

This work presents important contributions that collectively advance the field of pipeline inspection and offer innovative solutions to the challenges faced in this domain. the proposed method tackles important challenges that are often overlooked in the literature, providing a novel approach to overcome them. Specifically, it focuses on effectively managing large datasets, high-speed data acquisition, and GPS unreadability, making a significant contribution to the field of pipeline inspection. Furthermore, traditional inspection methods face limitations in inspecting pipelines that span great distances. To address this challenge, the proposed approach introduces a semiautonomous inspection method using unmanned aerial vehicles and infrared thermography. The proposed procedure offers a comprehensive approach that covers the entire inspection process, including flight configuration, data acquisition, and processing. A novel active learning procedure is introduced, further enhancing the efficiency and effectiveness of the pipeline inspection process. This procedure optimizes data processing, reducing the time and effort required for analyzing large amounts of data. The proposed work also tackles various challenges encountered in automated visual inspection, such as navigation inaccuracies, variability due to weather conditions, processing speed, image quality, data management, and object recognition and tracking. By effectively addressing these challenges, the proposed method provides accurate measurements regarding individual pipes. A novel tracking procedure is proposed that combines semantic information about the pipeline

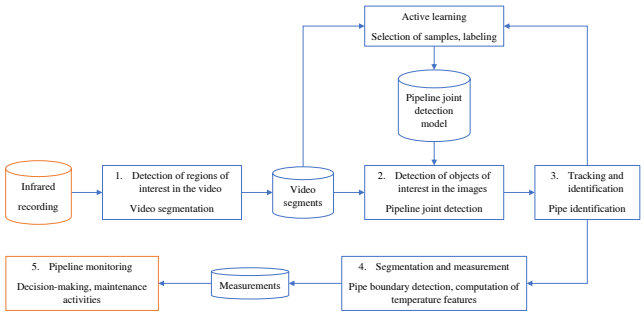


Fig. 1. Flowchart of the proposed method.

configuration. By incorporating geometric information about the pipeline, the tracking approach significantly enhances the robustness of the procedure. Consequently, false positive detections have a negligible influence on the identification of correct pipeline joints.

II. Proposed Approach

In the proposed approach, a drone flies along the pipeline route while acquiring infrared video at high-speed rates. When the battery runs out, both the battery and the storage card are swapped, and a new flight plan is configured for a different section of the pipeline. Data processing is performed while the new flight is executed. Therefore, at any given moment both the recording and processing tasks are executed at the same time for different sections of the pipeline. This improves the efficiency of the inspection procedure, providing a near real-time response.

The methodology proposed is intended to swiftly examine the recorded data, thereby affording the opportunity for data analysis prior to the drone's return for a fresh battery replacement. This capability grants the operator the ability to repeat the examination of a specific section of the pipeline if deemed necessary.

The proposed solution for the data processing task is outlined in Figure 1. When combined with mission planning, it offers a complete operational strategy that encompasses the entire inspection process. A comprehensive description of the components involved in the proposed solution is provided in the following sections.

A. Mission planning

The flight plan is created using a series of waypoints, or specific locations, along the pipeline route. This ensures that the drone flies over and inspects the entire pipeline. Waypoints are connected using line segments. Thus, each linear section of a pipeline is defined using two waypoints. These waypoints mark the beginning and end of the linear section of the pipeline that the drone will inspect. Using two waypoints per linear section of the pipeline is an efficient and effective way to ensure that the drone thoroughly inspects the entire pipeline while minimizing the configuration needed. Radiometric infrared video recording is activated from the first waypoint to the last. Moreover, each frame in the infrared video is tagged with a GPS coordinate using a GPS module that is integrated with the camera.

Drone batteries typically do not last very long, only around 30 minutes. This is a key consideration when planning a mission. For a long pipeline, multiple flights are needed. This way, the drone is configured to inspect only a particular section at a time. After each flight, the battery of the drone is swapped for a fresh one before continuing the mission. The number of flights depends on the specific characteristics of the pipeline, the flight conditions and the capabilities of the drone. It is also important to consider that high flying speeds result in motion blur in the acquired video, leading to inaccurate measurements. As a result, the configuration of the drone's velocity must be based on the rate at which images can be acquired.

The altitude of the drone must be carefully considered when conducting temperature measurements using infrared thermography as it influences the spatial resolution of the camera. The necessary sensor resolution and lens are determined by the size of the pipeline and the distance between the camera and the pipeline. An erroneous spatial resolution may result in either an overestimation or underestimation of temperature if the pipeline temperature differs from the background since the measurement will be an average of both the pipeline and the background.

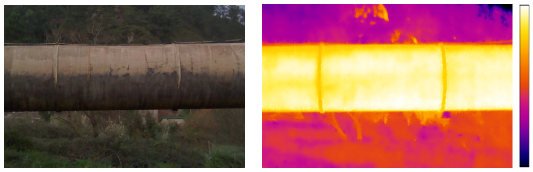
B. Detection of regions of interest in the video

A drone flight results in a radiometric infrared video recorded along the pipeline path, which corresponds to a section of the entire pipeline. The processing of the videos starts by dividing the video into a series or segments.

This work employs a robust detection and tracking method to identify the pipe number present in any image, effectively addressing the challenge of linking an image to a specific pipe without relying on GPS coordinates associated with the image. The approach starts by identifying regions of interest in the video, which correspond to the path between waypoints in the flight plan. Thus, the division of the video into regions of interest effectively breaks down the video into a series of segments, each corresponding to a specific part of the pipeline. This allows for more accurate and efficient tracking of the pipe number throughout the entire video. The use of detection and tracking in combination with the division of the video into regions of interest, greatly improves the overall performance and accuracy of the pipe number identification process, without the need for GPS coordinates.

The proposed approach for the detection of regions of interest in the video is as follows:

1) The GPS coordinates of the waypoints in the flight plan are transformed into a projected coordinate system (Web Mercator projection), where the units are in meters. This process uses the EPSG (European Petroleum Survey Group) codes. The result is a sequence of 2D line segments that represents the model, where the points correspond to waypoints in the flight plan.
2) The GPS coordinates of all frames in the video are also transformed into the same projected coordinate system. The result is a sequence of 2D points that represents the data, i.e., the location of the drone in the projected coordinate system. The position of these points is affected by noise introduced by GPS inaccuracies.



(a) (b)

Fig. 2. Pipeline with two joints recognizable. (a) Image in the visible spectrum. (b) Image in the infrared spectrum (temperature goes from 0 to 10ºC).

3) The Iterative Closest Point (ICP) algorithm [16] is used to align the data with the model. The algorithm finds the best alignment between the data and the model by iteratively minimizing the distance between corresponding points in the data and the model.
4) The closest points in the data to the points in the model correspond to the waypoint positions. The video is divided into a series of segments according to these points, where each segment corresponds to a specific part of the pipeline.
5) To ensure precise detection and tracking of the pipes, every segment undergoes thorough processing. The proposed approach for analyzing these video segments is extensively described in the subsequent sections. Depending on the configuration, specific segments that are deemed irrelevant to the inspection procedure can be directly ignored.

The proposed approach allows for more precise and efficient tracking of the pipe number throughout the video by breaking it down into segments that correspond to specific parts of the pipeline, and by using the ICP algorithm to detect these segments in the video. The approach used in this work limits errors in pipe identification to specific segments, preventing them from propagating to other parts of the pipeline. Furthermore, this method allows for parallel processing as different segments of the pipeline can be processed simultaneously.

C. Detection of objects of interest in the images

In this work, the considered objects of interest in the images are not the pipes, but the pipeline joints, as they have distinct shapes and features that make them easier to detect and recognize compared to the pipes themselves, particularly in the infrared spectrum. A pipeline joint is a section of a pipeline that is designed to connect two separate pieces of pipe. Thus by detecting pipeline joints, the position of an individual pipe can also be determined and located, making it an effective approach for object detection.

Figure 2 shows a pipeline in the visible and infrared spectrum where two welded joints can be appreciated. As can be seen, the joints can be better discriminated in the infrared spectrum. The proposed approach in this work leverages the distinct shapes and characteristics of the joints to streamline measurements and analysis of individual pipes.

The proposed method for pipeline joint detection in infrared images utilizes deep learning techniques, which have been proven to provide superior accuracy and efficiency compared to traditional object detection methods. Studies such as [17] have demonstrated that deep learning can overcome the limitations of traditional methods and achieve high performance at a reduced computational cost. Additionally, research has shown that deep learning-based object detection is particularly effective in handling low-contrast infrared images [18]. In this work, a state-of-the-art object detection model is trained on a dataset of infrared images of pipeline joints. The trained model is then applied to all frames of the infrared video, providing the location of pipeline joints in the images.

D. Tracking and identification

The proposed approach avoids relying on GPS coordinates to identify pipes in an image due to the unreliable nature of GPS readings. The GPS readings on the drone often produce inaccurate positions, with errors sometimes exceeding 5 meters. Additionally, the acquisition frequency of GPS readings is lower than that of the camera, leading to multiple images being associated with the same GPS coordinate even when the drone is in motion. These issues become more pronounced when the drone is moving at high speeds.

This work proposes an identification and tracking method based on the tracking-by-detection paradigm, which has become the leading paradigm in multi-object tracking. In the proposed approach, the object detector is applied to each video frame. Then, the tracking method associates these detections to tracks, in this case, pipeline joints. The underlying detector may produce false positive and missed detections, thus the tracking method needs to fill detection gaps and ignore false positives. The robust association of detections and pipeline joints is carried out using a motion model, predictions about the future positions of the objects, and semantic information about the pipeline. Motion prediction is estimated using Kalman filtering in the image space and frame-by-frame data association is based on the Hungarian method considering the bounding box overlap. This work does not use appearance information for tracking, unlike in [19], as all the objects being considered have the same appearance. Additionally, using such methods significantly increases the computational time required. In the proposed tracking approach, the state of the pipeline joint being tracked in the scene is modeled using a vector that contains the bounding box coordinates of the object in the image and the velocity. The prediction of the future state of the object uses the standard Kalman filter with a constant velocity motion model. The object detector is applied to each video frame and the resulting detections are compared with the predicted state of the tracked objects. The matching procedure can result in three potential outcomes:

1) A newly detected object matches a tracked object (match). In this case, the state of the tracked object is updated with the position of the newly detected object.
2) A newly detected object does not match any tracked object. A new track hypothesis is initiated in a tentative state given the position of the object is aligned with the semantic information about the pipeline. A tentative tracked object is considered a valid pipeline joint only if it is both consistently detected over multiple consecutive

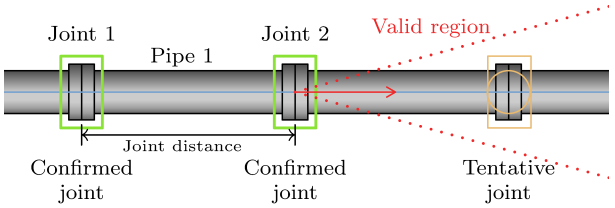


Fig. 3. Semantic information about the pipeline during tracking.

frames and aligned with the semantic information about the pipeline. Once a tentative tracked object is confirmed as a valid pipeline joint, it is assigned a pipeline joint identification number.

3) If no detected object matches a tracked object, it is marked as missed in that frame. If a tracked object goes beyond a certain number of consecutive missed detections, it is deemed to have exited the scene and removed from the list of tracked objects.

The matching procedure is solved based on the bounding box coordinates of the newly detected objects and the tracked objects. A cost matrix is calculated using the intersection-over-union (IOU) distance between them. The optimal assignment is determined using the Hungarian algorithm, considering an IOU threshold to reject insufficient overlapping.

The key contribution of the proposed tracking method is the use of semantic information. Incorporating the geometric information about the pipeline in the tracking approach significantly enhances the robustness of the procedure, with false positive detections having a negligible influence on the identification of the correct pipeline joints. The method uses the prior gathered information to confirm or reject the tentative pipeline joint, rather than relying solely on the initial detection.

When determining if a tentative pipeline joint can be confirmed, various features of the pipeline are taken into account. Figure 3 shows an example where two pipeline joints are confirmed and the line passing through both coordinates is calculated. Since the drone flight is not perfectly aligned with the pipeline, and sudden motion corrections may occur during the flight navigation, an uncertainty factor is taken into account when calculating the valid region for new pipeline joint detections. Any detections outside this region are considered false positives and ignored, as they are unlikely to be actual pipeline joints.

When tracking and assigning identification numbers to pipeline joints, a single mistake has a cascading effect on the accuracy of the identification of the remaining objects. This is because pipeline joints are identified sequentially, and one misidentified pipeline joint leads to a chain reaction of incorrect identifications for the next joints. Therefore, the tracking procedure is crucial for a robust and accurate identification system. The incorporated semantic information in the procedure minimizes the effect of false positive detections and the potential for a cascading effect of inaccuracies on pipeline joint identification.

E. Active learning

In this work, a methodology is proposed to address two of the most challenging issues when creating a dataset: (i) the cost of labeling, and (ii) the diversity of the samples. The methodology combines concepts from active learning, incremental learning and continual learning [20]. The goal is to improve the performance of the detection model over time, designing a methodology where the model can continuously learn new information without forgetting old information, the stability-plasticity dilemma, at the minimum possible cost.

In the proposed methodology, the annotation process is simplified using assisted labeling. A model is first trained on a small amount of labeled data, and then it is used to generate initial labels for the remaining data. The human annotator then reviews the generated labels and corrects any errors. This process is repeated multiple times as new data becomes available, with the detection model becoming more accurate with each iteration as it incorporates feedback from the human annotator. This approach speeds up the annotation process, increases the accuracy and consistency of the labels and reduces the overall cost. The process is further simplified with the usage of an interactive image annotation tool.

In the considered approach, data labeling and model training is repeated over time as new and relevant data becomes available. However, a recording for just a few minutes generates thousands of images. Thus, it is important to actively select the most informative examples to be labeled and used for training. The two most common approaches in active learning are uncertainty sampling and diversity sampling. Uncertainty sampling involves identifying instances where the model is uncertain or most likely to make an error. Diversity sampling, on the other hand, involves identifying instances that represent the diversity of the unannotated data. Both uncertainty sampling and diversity sampling are difficult to implement, particularly in the context of object detection. One of the main challenges in implementing uncertainty sampling for object detection is determining the model's level of uncertainty, particularly for missed detections. Since the model does not detect the object, it does not have a confidence score for the object, making it difficult to determine the model's level of uncertainty for that object. Additionally, in object detection, there may be multiple objects in an image, and determining which objects the model is most uncertain about can be challenging. Diversity sampling can also be difficult to implement in object detection, as it involves identifying instances that represent the diversity of the unannotated data. This can be challenging, as object detection datasets often contain a large number of images with similar objects in similar poses and contexts. Additionally, identifying the diversity of objects in an image can be difficult, as some objects may be small or occluded, making them difficult to detect.

F. Segmentation and measurement

The combined detection and tracking procedure is used to determine the position of individual pipes in the images. However, this only indicates a region where the pipe is located as the pipeline joints are identified with an approximated

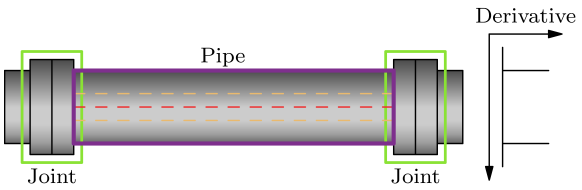


Fig. 4. Segmentation of pipe using profiles and derivative.

bounding box. To accurately measure the temperature of each individual pipe, a more precise segmentation procedure is required. This work proposes an image processing method to segment each pipe in order to determine its temperature and other relevant variables. This approach provides a precise analysis of the pipes, enabling accurate measurements and reliable results.

The segmentation process is triggered when the left and right joints of the pipe are identified in the image. These detections provide an approximate position of the pipe, and the line connecting their centers represents the center profile of the pipe. The procedure is shown in Figure 4, where the red dashed line denotes the center profile. To determine the exact region of the pipe in the image, the method considers the center profile and its derivatives in relation to parallel profiles (orange lines above and below the center profile). When these profiles reach the end of the pipe, there is a pronounced peak in the derivative, which precisely indicates the position of the pipe boundaries in the image.

The result of the segmentation procedure is a set of regions or segments, each of which corresponds to a different pipe in the image. With these segments, various measurements can be performed. However, before conducting these measurements, it is necessary to convert the measured radiation signal in the infrared images into temperature, taking into account all relevant parameters. This procedure needs to consider that radiation measured by the infrared sensor comes from different sources, not only the pipe in the scene. As such, a rigorous temperature compensation procedure is required [21].

Converting radiometric data into temperature is crucial for several reasons, even though it may not be necessary for the segmentation or identification of joints. Converting radiometric data into temperature allows for quantitative analysis, which provides precise and objective measurements. It also enables comparisons with reference values. These reference values can be the temperature of a similar component under the same conditions or the maximum allowable temperature for the component. By comparing the measured temperatures to these references, it becomes possible to determine if the component's temperature readings exceed critical thresholds or deviate from expected values, indicating potential issues or abnormalities. Further more, temperature conversion allows for standardized reporting and analysis of thermal data and enhances the diagnostic capability of infrared thermography. The details about the procedure to convert radiometric data into temperature can be found in [22].

Temperature measurement is implemented only for the pixels within the segmented regions, as it is computationally intensive. Afterward, various intensity features that characterize the temperature within the pipes are calculated. These features include the mean, median, standard deviation, mode, maximum, minimum, quartiles, moments, skewness, kurtosis, and peak height. Additionally, measurements are also performed considering different positions in the pipe to compare temperature differences, such as between the center and the edges.

III. Results and Discussion

The proposed approach is evaluated with data acquired from a laboratory prototype and real data acquired in a large pipeline system in an industrial facility. The laboratory prototype was used to evaluate the performance of the semiautonomous algorithm under controlled conditions, while the real pipeline system provided an opportunity to evaluate the method in a realistic industrial setting. The results of this study provide valuable insights into the potential of this method for the reliable inspection of pipelines in industrial environments.

The drone model used in the industrial facility was the DJI Matrice 300 RTK equipped with a Zenmuse XT2 camera, which is a dual-sensor design with a thermal and a visual camera. Some key features of the FLIR infrared camera include a resolution of 640 $\times$ 512 and a frame rate of 30 Hz. The manufacturer reports a sensitivity lower than 50 mK. The long-wave infrared camera operates in the range of 7.5-13.5 µm, using an uncooled microbolometer detector.

Data processing is performed using a computer with an Intel Core i9-11900K CPU running with 8 cores at 3.5 GHz and 128 GiB of RAM. The computer is also equipped with a GeForce RTX 3090 GPU with 24 GiB of RAM. All experiments are run under Linux 5.4.0.

A. Detection of regions of interest in the video

Figure 5 shows the coordinates of a drone during the inspection of a pipeline in an industrial facility. The planned flight path is depicted by waypoints (shown as large red points in the figure). The drone follows this path from left to right and bottom to top, but deviations and inaccuracies in the drone's position can be observed. Additionally, the drone's speed changes as it moves from one waypoint to the next, slowing down at turns, which results in varying densities of positions along the path in different areas. Despite these issues, the proposed approach for detecting regions of interest in the video effectively aligns the GPS coordinates measured by the drone with the flight plan. The video can then be divided into segments based on the positions of the frames closest to the waypoints. Each segment corresponds to a specific part of the pipeline and can be analyzed independently.

B. Active learning

The active learning approach proposed in this work was implemented for one year. During this time, the size of the dataset increased steadily. The final dataset consists of 15201 images, with 34983 pipeline joint instances. For almost 30 iterations, more varied data and different scenarios were introduced. New relevant data was automatically determined

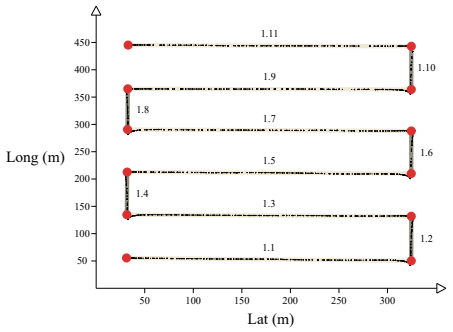


Fig. 5. Detection of regions of interest during the inspection of a pipeline

TABLE I
OBJECT DETECTION PERFORMANCE FOR DIFFERENT MODEL ARCHITECTURES.

Model	Par. (M)	P	R	F_1	AP_{50}	$AP_{50:95}$
YOLOv5n	1.77	0.993	0.991	0.992	0.995	0.869
YOLOv5s	7.23	0.992	0.992	0.992	0.995	0.880
YOLOv5m	20.85	0.995	0.990	0.992	0.995	0.886
YOLOv5l	46.11	0.993	0.991	0.992	0.995	0.882
YOLOv5x	86.17	0.993	0.991	0.992	0.995	0.886
YOLOv8n	3.01	0.993	0.989	0.991	0.995	0.881
YOLOv8s	11.14	0.992	0.992	0.992	0.995	0.886
YOLOv8m	25.86	0.993	0.988	0.990	0.995	0.867
YOLOv8l	43.63	0.992	0.990	0.991	0.995	0.885
YOLOv8x	68.15	0.993	0.990	0.991	0.995	0.887

based on the performance of the tracking procedure. When inspections included new types of pipeline joints and different weather conditions, the performance degraded. In some cases, the object detector was able to generalize, detecting new type of pipeline joints slightly different than the samples in the dataset or successfully processing data acquired at different times of the day. However, in other cases, the detection model did not perform well, with missed detections for new pipeline joints and false positives, particularly in the case of reflections depending on the position of the sun. In all these scenarios, which were easily determined based on the results of the tracking procedure, new samples were actively selected to extend and diversify the dataset with a wider variety of samples.

During the active learning approach, in each iteration, a group of new samples is actively selected as the most informative examples. The samples are automatically labeled using the current detection model, and both the images and the labels are uploaded to the interactive image annotation tool CVAT (Computer Vision Annotation Tool). No human intervention is required for any of these tasks. A human annotator is then required to review and correct any errors in the labels generated by the model. This process is faster and easier than performing full annotation, as a high percentage of the samples are already correctly labeled. For example, a group of 500 images could take a human annotator 10 minutes to review and correct.

The process of updating the object detector begins by retraining it with the new dataset. Once the retraining is complete, the model's performance is evaluated using a test set that is separate from the training set. This process was repeated multiple times with updated versions of the dataset. The evaluation results of this iterative approach showed that the model can learn new information without forgetting the old information. This agile approach emphasizes iterative, incremental development and frequent delivery of an updated and improved object detector model.

C. Detection of objects of interest in the images

Table I shows the performance of object detection for different model architectures and the final version of the dataset.

The performance metrics presented in the table encompass precision, which measures the percentage of correct detections out of the total number of detections; recall, which calculates the percentage of correct detections among the total number of elements; and F_1, the harmonic mean of precision and recall. For object detection tasks, the output of object detectors consists of bounding boxes that can partially intersect with the real objects. Therefore, determining a correct detection relies on the chosen intersection threshold. Currently, the most common approach is to evaluate the Intersection over Union (IoU) metric for each detected object. Average precision (AP) represents the area under the precision-recall curve, which is generated by varying the confidence threshold. This metric provides a single scalar value that summarizes the system's performance across different thresholds. Typically, AP is calculated for a specific IoU threshold, with 0.5 being the most commonly used value, referred to as AP_{50}. However, in certain cases, AP is averaged for multiple IoU thresholds. For instance, AP@50:5:95 denotes the average AP across IoU thresholds ranging from 50% (0.5) to 95% (0.95) in increments of 5% (0.05). These performance metrics, which describe the effectiveness of object detectors, are extensively discussed in [17].

Multiple models with a varying number of parameters and computational requirements, ranging from the smaller YOLOv5n (1.8 million) to the larger YOLOv8x (more than 250 million) are considered. The number of parameters is a measure of the complexity of the model, with more parameters generally leading to better accuracy but also higher computational cost.

The results presented in Table I are obtained using a consistent configuration for the training hyperparameters. The initial learning rate was set to 0.01, the solver used was SGD (Stochastic Gradient Descent), the momentum was set to 0.9, the batch size was 16 and a one-cycle learning rate scheduler was employed with 1000 epochs. To prevent overfitting during training, an early stopping approach was used. This approach involved monitoring the performance of the model with a validation set during training. If there was no improvement on the validation set for a specified number of iterations, the training process was terminated. During training, data augmentation techniques such as clipping, rotation, flipping, adjusting hue, saturation, exposure, and altering aspect ratio are applied to artificially increase the picture data. Additionally, mosaic data enhancement is also utilized. These techniques are also used to prevent overfitting by introducing variation to the training

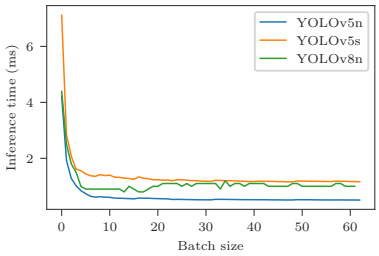


Fig. 6. Inference time versus batch size for various model architectures.

data. The complexity disparity between the models is reflected in their training time. Simpler models take approximately 10 hours to converge, while the convergence of more complex models requires over 50 hours.

The results in Table I indicate that there is no substantial increase in performance when utilizing a more complex model for the considered dataset. Variations in $AP_{50:95}$ can be observed for more complex models, which represent a minimal performance improvement when taking into account the added complexity of the model. Moreover, no major differences could be found for YOLOv5 and YOLOv8 architectures, except for their $AP_{50:95}$ scores. The metric $AP_{50:95}$ is a specific variation of the AP metric that calculates the average precision (area under the precision-recall curve) at different intersection over union (IoU) thresholds, specifically at 50% and 95%. It is used to evaluate the model's performance in detecting objects of varying sizes, as well as its ability to accurately assign bounding boxes to objects. It is a more comprehensive metric than using a single IoU threshold, such as the AP_{50}, because it measures the model's performance across a range of different object sizes. When two models have the same AP_{50} but different $AP_{50:95}$, it means that they have similar performance in detecting objects with an IoU threshold of 50%, but they perform differently at different IoU thresholds. The model with the higher $AP_{50:95}$ is likely to have a better overall performance, as it is able to detect objects of different sizes more accurately. Objects of different sizes can also refer to objects observed at different distances, for example images acquired when a drone is at different altitudes. $AP_{50:95}$ measures the performance of the model at different IoU thresholds, which also corresponds to the model's performance at detecting objects at different scales and distances.

The impact of model complexity can be observed in Figure 6, which shows the computational cost for the three smallest model architectures. This figure shows the inference time when considering different batch sizes, which refer to the number of samples passed through the model at one time. As can be seen, all models benefit from large batch sizes, with the YOLOv5n being the most efficient. However, there is a point where increasing the batch size does not result in further decreases in inference time, with an optimal configuration of batch size around 16.

A hyperparamter tuning approach using genetic algorithms was applied to YOLOv5n. The results provided an optimal configuration of training hyperparameters with a 0.0102 learning rate, 0.98 momentum and an alternative configuration for the augmentation. Table II shows the performance of the model trained with these alternative training hyperparameters. Results indicate that the performance of the model improved by only 1.38% in the metric $AP_{50:95}$. This suggests that while the automated tuning approach was able to find a better configuration of hyperparameters, the performance improvement was limited. However, the slight increase in performance makes the performance of YOLOv5n when considering $AP_{50:95}$ comparable to that of the default configuration of the more complex model, YOLOv8n, and even surpasses it when considering F_1.

TABLE II
PERFORMANCE COMPARISON BETWEEN YOLOV5N TRAINED WITH DIFFERENT HYPERPARAMETERS.

	P	R	F_1	AP_{50}	$AP_{50:95}$
Default	0.993	0.991	0.992	0.995	0.869
Tuned	0.995	0.990	0.993	0.995	0.881
Improvement	+0.20%	-0.10%	0.10%	0.00%	+1.38%

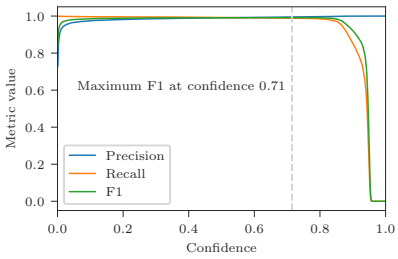


Fig. 7. Performance of the resulting model for different confidence configurations.

The final step in the deep learning pipeline is to evaluate the performance of the resulting model on the test subset, which represents a completely independent and unseen subset of data. To evaluate the model, various metrics such as precision (P), recall (R), and F_1 score are calculated for different confidence thresholds. Figure 7 shows the results of these evaluations. The figure shows a large plateau where all three metrics are maximized simultaneously, indicating excellent performance. The confidence threshold that maximizes the F_1 score is 0.71, which represents the ideal setting for reducing false positives and missed detections.

The performance achieved by the model indicates that the model is able to accurately identify the positive instances while also keeping the number of false positives low. This combination of high precision and high recall represents the optimal balance as it indicates that the model can make accurate predictions while also identifying as many relevant instances as possible.

D. Tracking and identification

Tracking and identification are evaluated using a laboratory prototype where different pipeline configurations are tested, introducing synthetic false positive and missed detections. This prototype is used to tune and evaluate the proposed approach. The system is then evaluated with a large pipeline system in an industrial facility.

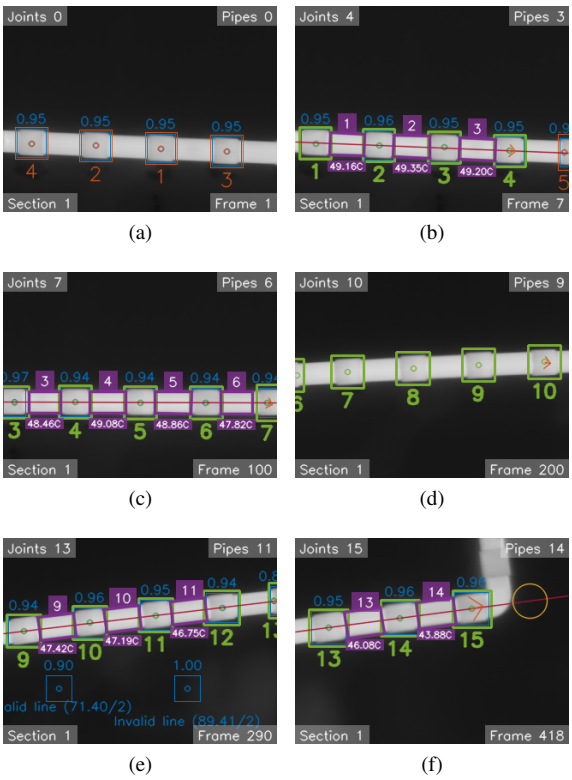


Fig. 8. Evolution of the tracking and identification approach. (a) Four pipeline joints are detected and marked as tentative joints in the first frame. (b) After some consecutive matches, the pipeline joints are confirmed and the pipe identification number is assigned considering the movement direction. (c)-(f) The process continues as the drones move along the pipeline path dealing with missed detections and false positive detections.

Figure 8 shows a tracking and identification example using the laboratory prototype. In the first image, the system detects the pipeline joints with a high level of confidence (number above the bounding box). These newly detected objects are marked as tentative joints (red boxes), and a tentative identification is assigned. At this moment, the system does not know the direction of advance. Thus, these numbers are assigned in random consecutive order. After some frames, the tentative joints are confirmed (green boxes) as they are consistently detected over multiple consecutive frames. The identification of the pipes takes into account the direction of advance (to the right), assigning numbers from left to right. Semantic information is now extracted from the joints for further processing including the relative distance between joints and direction advance. This information is used to estimate the position of the next pipeline joint, protecting the inspection system from false positive detections in other regions of the image. Once the pipes are identified, they are segmented, providing the boundaries of the objects in the images. This results in an area where various measurements can be performed. In this case, the average temperature is shown below the pipe. The process continues as the drone moves along the pipeline path. The inspection process is robust to small rotations and changes in trajectory, and can handle objects of different scales as the drone changes its flying altitude. The difference in size between the joints in the first and last image is noticeable. The extracted semantic information is constantly adjusted based on the confirmed pipeline joints.

Figure 8d depicts a scenario where object detection fails to detect objects, leading to missing detections. Despite this, the tracking method continues to update the position of confirmed pipeline joints estimating future positions given the current velocity. When the issue is resolved, tracking resumes without any incorrect identification. Figure 8e displays a situation where the object detector produces false position detections. These false positives are easily detected based on the semantic information about the pipeline. In this case, these detections violate multiple constraints, such as distance to the pipeline line. Constraints are evaluated using a short-circuit strategy, where the failure of one constraint results in false position categorization.

As the drone progresses along the pipeline, the same pipe becomes visible in successive images, depending on the flight speed. With each image, the system utilizes the tracking approach to identify the pipe number and measure its temperature. As the drone moves forward, the tracking procedure consistently identifies the same pipe across different images. Consequently, temperature measurements can be obtained for the same pipe multiple times. In this work, our proposal is to combine all these measurements to generate a robust final value.

To characterize the temperature within the pipes, various temperature intensity features are calculated. These include the mean, median, standard deviation, mode, maximum, minimum, quartiles, moments, skewness, kurtosis, and peak height. Furthermore, temperature measurements are compared at different positions within the pipe, such as between the center and the edges. Since these values are calculated for numerous images during each capture, multiple measurements are available for each pipe.

To enhance accuracy and mitigate the impact of outliers, the multiple measurements are combined. Inconsistent measurements are identified and eliminated, ensuring a more reliable and representative estimation of the true temperature value. By leveraging this combination of measurements, the proposed approach provides improved accuracy and a more robust assessment of the measured temperature within the pipes.

The proposed pipeline inspection method is being used for more than a year on a large industrial pipeline system. During this operation, a test dataset has been created with carefully annotated data. This dataset includes 71 flights, with 1.18 million frames recorded, equating 11 hours of video. These inspections evaluated 1164 pipeline sections, totaling 385 km. Throughout these tests, the drone's altitude fluctuated between 15 and 20 meters. The results showed that the detection and tracking system correctly identified all pipes in every section, which can be confirmed as the number of pipes per section is known beforehand. The inspections were conducted in various weather conditions and with different types of pipeline joints, while the sun's position also posed challenges due to its reflections on the pipelines. Despite these challenges, the proposed inspection system performed exceptionally well, providing

accurate temperature measurements for each individual pipe.

Figure 9 shows the temperature map for a section of the inspected pipeline. The acquired data is divided into different segments according to the waypoints, each containing a different number of pipes according to their length. In this case, individual pipes are 5 meters in length. The figure represents the average temperate in each pipe. Temperature comparison between the pipes reveals that pipe 7.31 (segment 7, number 31) has an unusual temperature compared to the others. Further investigation into this pipe confirmed the presence of a defect, specifically insulation problems.

Real-world testing reveals that the drone's velocity changes as it moves from one waypoint to the next. Figure 10 shows the estimated velocity of the drone during the inspection of a pipeline section. In this experiment, the drone was programmed to fly at a speed of 15 meters per second and 18 meters above the ground. However, a decrease in velocity is observed as the drone switches direction at waypoints. The same pattern is evident throughout the duration of the flight. The inspection process is designed to address this type of variability, adapting both the tracking method and the semantic information extracted from the video footage.

Performance testing with the full inspection system, which combines detection, tracking, and measurements, demonstrates a processing speed of 240 fps. Given that the image acquisition rate of the drone is 30 fps, this means processing is 8 times faster than image acquisition. As a result, a 30-minute video recording can be processed in just 3.75 minutes. Given the processing speed and the real-time operating capability of the designed inspection procedure, the proposed system has the potential to process data directly on drone hardware. This capability would eliminate the need for post-processing, providing quicker and more efficient inspections with real-time feedback if necessary.

IV. Conclusions

Overall, pipeline inspection is a critical aspect of maintaining and operating a safe and reliable pipeline system. This work proposes a novel solution that presents a comprehensive operational approach that covers the entire inspection process. The proposed work combines the benefits of automated flight and high-speed infrared data acquisition with advanced data processing techniques, addressing the challenge of scale in pipeline inspection.

The proposed inspection procedure has undergone extensive testing, both in the laboratory and in real-world conditions. These tests have confirmed its robustness and exceptional performance in various challenging conditions. The results of these tests demonstrate the proposed inspection procedure's reliability and effectiveness, making it a valuable tool for pipeline inspections in industrial facilities.

The proposed method addresses important challenges commonly neglected in the literature and offers a novel approach to overcoming them. The focus on the effective management of large datasets, high-speed data acquisition and GPS unreadability, makes this approach a significant contribution to the field of pipeline inspection. The use of deep learning algorithms and an active learning procedure enhances the ability to detect hidden defects and monitor temperature patterns over time, making it a reliable solution for large-scale pipeline inspection.

This work presents important contributions that collectively advance the field of pipeline inspection and offer innovative solutions to the challenges faced in this domain.

References

[1] G. A. Antaki, *Piping and pipeline engineering: design, construction, maintenance, integrity, and repair*. CRC Press, 2003.

[2] E. Z. Hou, "Performance optimization of harmonized flexible printed coils of axial magnetized magnetostrictive patch transducers for pipeline inspection," *Measurement*, vol. 199, p. 111478, 2022.

[3] H. A. Kishawy and H. A. Gabbar, "Review of pipeline integrity management practices," *International Journal of Pressure Vessels and Piping*, vol. 87, no. 7, pp. 373–380, 2010.

[4] B. F. Spencer Jr, V. Hoskere, and Y. Narazaki, "Advances in computer vision-based civil infrastructure inspection and monitoring," *Engineering*, vol. 5, no. 2, pp. 199–222, 2019.

[5] M. Syifa, S.-J. Park, and C.-W. Lee, "Detection of the pine wilt disease tree candidates for drone remote sensing using artificial intelligence techniques," *Engineering*, vol. 6, no. 8, pp. 919–926, 2020.

[6] *Leveraging drone based imaging technology for pipeline and RoU monitoring survey*. OnePetro, 2019.

[7] J. Seo, L. Duque, and J. Wacker, "Drone-enabled bridge inspection methodology and application," *Automation in Construction*, vol. 94, pp. 112–126, 2018.

[8] R. Ashour, T. Taha, F. Mohamed, E. Hableel, Y. A. Kheil, M. Elsalamouny, M. Kadadha, K. Rangan, J. Dias, L. Seneviratne *et al.*, "Site inspection drone: A solution for inspecting and regulating construction sites," in *2016 IEEE 59th International midwest symposium on circuits and systems (MWSCAS)*. IEEE, 2016, pp. 1–4.

[9] P. Nooralishahi, C. Ibarra-Castanedo, S. Deane, F. López, S. Pant, M. Genest, N. P. Avdelidis, and X. P. Maldague, "Drone-based non-destructive inspection of industrial sites: A review and case studies," *Drones*, vol. 5, no. 4, p. 106, 2021.

[10] S. Deane, N. P. Avdelidis, C. Ibarra-Castanedo, A. A. Williamson, S. Withers, A. Zolotas, X. P. Maldague, M. Ahmadi, S. Pant, M. Genest *et al.*, "Development of a thermal excitation source used in an active thermographic uav platform," *Quantitative InfraRed Thermography Journal*, pp. 1–32, 2022.

[11] E. Ciampa, L. De Vito, and M. R. Pecce, "Practical issues on the use of drones for construction inspections," in *Journal of Physics: Conference Series*, vol. 1249. IOP Publishing, 2019, p. 012016.

[12] A. Shukla, H. Xiaoqian, and H. Karki, "Autonomous tracking and navigation controller for an unmanned aerial vehicle based on visual data for inspection of oil and gas pipelines," in *2016 16th international conference on control, automation and systems (ICCAS)*. IEEE, 2016, pp. 194–200.

[13] A. Alharam, E. Almansoori, W. Elmadeny, and H. Alnoiami, "Real time ai-based pipeline inspection using drone for oil and gas industries in bahrain," in *2020 International Conference on Innovation and Intelligence for Informatics, Computing and Technologies (3ICT)*. IEEE, 2020, pp. 1–5.

[14] J. Bian, X. Hui, X. Zhao, and M. Tan, "A novel monocular-based navigation approach for uav autonomous transmission-line inspection," in *2018 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*. IEEE, 2018, pp. 1–7.

[15] W. Minkina, "Theoretical basics of radiant heat transfer–practical examples of calculation for the infrared (ir) used in infrared thermography measurements," *Quantitative InfraRed Thermography Journal*, vol. 18, no. 4, pp. 269–282, 2021.

[16] J. Zhang, Y. Yao, and B. Deng, "Fast and robust iterative closest point," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2021.

[17] R. Usamentiaga, D. Lema, O. Pedrayes, and D. Garcia, "Automated surface defect detection in metals: a comparative review of object detection and semantic segmentation using deep learning," *IEEE Transactions on Industry Applications*, vol. 58, no. 3, pp. 4203–4213, 2022.

[18] O. D. Pedrayes, D. G. Lema, R. Usamentiaga, P. Venegas, and D. F. García, "Semantic segmentation for non-destructive testing with step-heating thermography for composite laminates," *Measurement*, vol. 200, p. 111653, 2022.

Fig. 9. Temperature map for a section of the inspected pipeline. All values a given in Celsius.

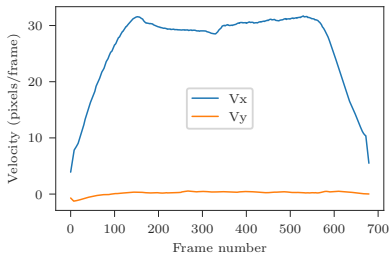


Fig. 10. Estimated velocity of the drone during the inspection of a pipeline section.

[19] N. Wojke, A. Bewley, and D. Paulus, "Simple online and realtime tracking with a deep association metric," in *2017 IEEE international conference on image processing (ICIP)*. IEEE, 2017, pp. 3645–3649.

[20] M. De Lange, R. Aljundi, M. Masana, S. Parisot, X. Jia, A. Leonardis, G. Slabaugh, and T. Tuytelaars, "A continual learning survey: Defying forgetting in classification tasks," *IEEE transactions on pattern analysis and machine intelligence*, vol. 44, no. 7, pp. 3366–3385, 2021.

[21] R. Usamentiaga, M. A. Fernandez, A. F. Villan, and J. L. Carus, "Temperature monitoring for electrical substations using infrared thermography: architecture for industrial internet of things," *IEEE transactions on industrial informatics*, vol. 14, no. 12, pp. 5667–5677, 2018.

[22] R. Usamentiaga, P. Venegas, J. Guerediaga, L. Vega, J. Molleda, and F. G. Bulnes, "Infrared thermography for temperature measurement and non-destructive testing," *Sensors*, vol. 14, no. 7, pp. 12 305–12 348, 2014.

R. Usamentiaga (M'14–SM'19) is a Full Professor in the Department of Computer Science and Engineering at the University of Oviedo. He received his M.S. and Ph.D. degrees in Computer Science from the University of Oviedo in 1999 and 2005 with the Extraordinary Doctorate Award. His research interests include machine vision, real-time imaging systems and infrared thermography, where he has published over a hundred papers in JCR journals with five Prize Paper Awards. He has been a visiting professor at the Universities of Laval and Pennsylvania. He is also a senior member of the IEEE and the IAS society.