

Description

Method and System for Quality Assessment of Objects in an Industrial Environment

5

The present disclosure relates to a method, a system, and a computer program product used for assessing quality of one or more objects in an industrial environment. More particularly, the present disclosure relates to assessing quality of ob-
10 jects using a combination of image processing techniques and artificial intelligence methods in conjunction with analytical methods.

Rapid digitalization of industries is bringing a pivotal
15 change in the current industrial practices. For example, categorizing of an object in an industry is typically performed to certify that the object and/or associated product meets the defined grade and quality requirements.

20 Especially in some industrial environments such as container terminals, loading processes with the help of cranes are increasingly popular and automated, i.e., without manual intervention by operators. In such cases, quality sorting including picking and dropping of objects such as coils by cranes
25 requires a human intervention for assessing quality of objects and thereafter, picking, dropping, and categorized loading of the objects by cranes. Moreover, to ensure the safety of loading operations, especially for automated cranes, there is a great need for safety systems and protec-
30 tive devices that monitor the lanes in which the cranes are deployed or the environment during crane movement in order to avoid collisions with objects or persons in the proximity of the crane.

Typically, quality sorting is performed by trained human inspectors who assess the object by looking for a specific quality attribute. Such inspection process usually involves
5 further testing using laboratories, and is therefore time consuming. Moreover, such human driven object quality inspection requires other equipment for example, a scale for estimating the current weight of object, a laboratory for measuring traces and parameters such as humidity in the object,
10 etc.

Furthermore, the conventional quality inspection processes are not only subject to inconsistencies due to heavy reliance on human expertise but also expensive due to manual labor
15 costs, and cumbersome considering industrial scale and huge volumes of objects to be inspected. To address the aforementioned problems, machine learning techniques have been suggested and/or used for identifying defects in an object based on the object images, however the training time for such machine learning techniques is quite long. Typically, longer
20 training times are known to increase the CO₂ emissions from a machine employed in the training.

Furthermore, existing methods for identifying industrial object defects use an unsupervised learning approach wherein a
25 surface color and a surface texture of the object acts as a primary parameter in the object image to predict defects associated therewith. Image segmentation plays a central role in identifying defects from an image. Image segmentation includes classifying pixels of the image into random number of
30 clusters. However, selecting the number of cluster labels in an unsupervised image segmentation is a challenging task and

uses up processing power and memory of the system thus, rendering it resource intensive.

Accordingly, it is an object of the present disclosure to
5 provide a quality assessment system, a device and a method
for assessing quality of an object in an industrial environment,
that employ a supervised machine learning approach in
conjunction with a select set of image processing techniques
to identify quality of an object in an industrial environment
10 based on the object image(s) in a time and cost effective
manner while ensuring that the associated training time is
kept minimal.

The aforementioned object is achieved in that a method for
15 managing a crane system according to claim 1, a crane system
according to claim 8, and a computer program product for managing
the crane system according to claim 11 are provided.

Disclosed herein is a crane system capable of handling an object.
20 According to an embodiment, the crane system is an
overhead crane deployable in an industrial environment. As
used herein, the term "object" refers to an industrial object,
for example, metal coils that are heavy and therefore
require a crane system to be moved from one place to other.

25 The crane system comprises cameras positioned so as to capture
a plurality of images of the object. The cameras comprise,
for example, high definition light detection and ranging (LIDAR)
cameras capable of capturing high definition real
30 time images of the object. According to an embodiment, the
crane system comprises three cameras mounted at predefined
locations on the crane system. According to this embodiment,
a first camera is arranged at a first end of a gantry of the

crane system; a second camera is arranged at a second end of the gantry; and a third camera is arranged in proximity of a hoist on the gantry. Advantageously, each of these cameras is aligned for a predefined angle of capture with respect to the
5 object and/or the hoist that is capable of moving the object. Advantageously, the angle of capture is defined based on a size and an orientation of the crane system and the area in which the crane system is deployed.

10 It would be appreciated by a person skilled in the art that the two cameras positioned at either end of the gantry have similar functionality and are could easily replace one another. However, having both of these provides the required redundancy in cases where a field of view of one of the cam-
15 eras is obstructed due to unforeseen circumstances.

The crane system comprises a computing unit having an artificial neural network. The computing unit receives, from the cameras, the images of the object, advantageously in real
20 time. Advantageously, the artificial neural network is stored on the computing unit. The artificial neural network is a trained artificial neural network, for example, trained to analyze images for recognizing patterns in the images. For example, the artificial neural network is a Convolutional
25 Neural Network (CNN) that may include a Pyramid Scene Parsing network (PSPnet) with CNN, so as to capture both local and global information along with spatial information of an image, thereby enabling the crane system handling the object to determine not only defects in the object but also remaining
30 useful life of the object. It would be appreciated by a person skilled in the art that the remaining useful life factor is subject to the type of object. For example, the remaining useful life may be more relevant for perishable objects being

handled by the crane system, if any as compared with non-perishable objects.

The crane system includes a control unit. The control unit
5 may be in communication with a drive system of the crane system, such that the control unit either directly or via the drive system causes physical movement of the cameras and/or the crane system.

10 According to one embodiment, the control unit moves the crane system and thereby, causes the physical movement of the cameras. According to this embodiment, the cameras are triggered by the movement of the crane system to start capturing the images of the object.

15

According to another embodiment, the control unit, independently causes movement of the cameras and thereby, triggering them to capture the images of the object.

20 Advantageously, the cameras capture the images of the object, along a longitudinal axis A-A' of gantry tracks. The control unit may move one of the cameras arranged in proximity of the hoist on the gantry along a lateral axis perpendicular to the longitudinal axis. According to this embodiment, the computing unit is operably coupled to the control unit, for example, via a wired or a wireless communication network including, the internet, an intranet, a wired network, a wireless network, and/or any other suitable communication network capable of establishing a strong and secure communication. According to another embodiment, the control unit may include
25 the computing unit as a part or as a whole. According to yet another embodiment, the computing unit may include the control unit as a part or as a whole.

According to an embodiment, the computing unit receives images from the cameras. According to this embodiment, the computing unit generates an image data stream from the images by performing pre-processing of the images including but not limited to reducing noise in the images, enhancing contrast of the images, for example, by applying a median filter followed by histogram equalization followed by another median filter, and/or stitching the images together to form an image data stream. According to this embodiment, the computing unit determines from the images, a foreground associated with object, for example, by performing background subtraction. According to this embodiment, the computing unit annotates the foreground(s) from the images. According to another embodiment, the computing unit receives an image data stream generated based on the images captured by the cameras.

The computing unit analyzes the image data stream generated based on the images using an artificial neural network. The computing unit determines one or more object properties associated with the object based on the analysis of the image data stream. The object properties comprise at least a quality of the object. The quality of the object being defined based on presence of defect(s) in the object. The object properties may also comprise a position, an orientation, a dimension of the object, a type of the object, a surface of the object, an edge of the object, etc.

The computing unit detects, from the image data stream, presence of one or more abnormalities in the object. The abnormalities include, for example, a defect in the object, a human in close proximity of the object, etc. The computing unit segments the images based on the artificial neural network

and identifies markers in the object, wherein a distance between the markers corresponds to an extent of the abnormality in the object.

5 The artificial neural network comprises labelled images based on which the computing unit segments the images into multiple areas, that is, pixel clusters based on color, contours, etc. Advantageously, the computing unit determines a distance between the markers. The computing unit based on the distance
10 between the markers derives primary object properties comprising, for example, a length, a height, an area, etc., of the object. Based on these primary object properties, the computing unit derives using the trained artificial neural network, secondary object properties comprising, for example,
15 an approximate weight, a remaining useful life, a size, and a defect area of the object.

According to an embodiment, the computing unit detects from the image data stream, presence of a human in proximity of
20 the object. The computing unit segments the images based on the artificial neural network and identifies markers in the images corresponding to humans, wherein a distance between the markers corresponds to a proximity of the human with respect to the object.

25

The computing unit using the trained artificial neural network thus determines based on the distance between the markers and the segmented images, presence of the abnormalities.

30 The control unit automatically operates a hoist of the crane system for handling the object based on the object properties. For example, the control unit causes the crane system to pick and drop an object in separate areas based on the

quality of the object thereby, categorizing the objects. In another example, the control unit causes the crane system to not pick the object when a presence of a human is detected in proximity of the object.

5

Also disclosed herein, is a computing unit having an artificial neural network for managing a crane system.

According to an embodiment of the present disclosure, the
10 computing unit is deployable in a cloud computing environment and is capable of communicating with the crane system. As used herein, "cloud computing environment" refers to a processing environment comprising configurable computing physical and logical resources, for example, networks, servers,
15 storage, applications, services, etc., and data distributed over the cloud platform. The cloud computing environment provides on-demand network access to a shared pool of the configurable computing physical and logical resources.

20 According to another embodiment of the present disclosure, the computing unit is deployable in an industrial environment, where the crane system is physically located, as an edge device capable of communicating with one or more crane systems. According to this embodiment, the computing unit
25 comprises a processor, a memory unit, a network interface, and/or an input/output unit to function as an edge device. For example, in order to function as an edge device, the aforementioned hardware components of the computing unit are deployed in the industrial environment in operable communication with the crane system(s), and the data, that is, images
30 collected from the cameras mounted on the crane system, are communicated via the network interface to a cloud-based server wherein the artificial neural network is stored for

processing the images thus received for managing the crane system. Moreover, according to this embodiment, there may exist more than one computing units as edge devices deployed in the industrial environment and the edge devices may communicate with one another for managing one or more crane systems simultaneously. Advantageously, according to this embodiment, the computing units may share the processing loads therebetween while managing the crane systems.

10 According to yet another embodiment, the computing unit is deployable in a distributed architecture where parts of the computing unit are deployable in the industrial environment in proximity of the crane system(s) as an edge device and parts of the computing unit are deployable in the cloud computing environment.

The computing unit disclosed herein may comprise one or more software modules such as a data acquisition module for receiving the image data stream from the cameras, a data processing module for processing the image data stream, and a data analytics and management module for analysing the image data stream with the artificial neural network and for causing the crane system to handle the object. However, it will be appreciated by a person skilled in the art, that the functionalities offered by each of these modules may be combined into a single module.

The computing unit analyzes an image data stream generated based on the images with help of the artificial neural network and determines object properties associated with the object based on the analysis of the image data stream, wherein the object properties comprise at least a quality of the object.

Also disclosed herein is a method for managing a crane system capable of handling, for example, physically displacing such as lifting, picking, dropping, loading, unloading, etc., an
5 object. It would be appreciated by a person skilled in the art, that aforementioned managing of the crane system may also be extended to positioning of the crane system in proximity of the object in order to handle the object with maximal accuracy and with minimal effort. The object being an industrial object in an industrial environment handled by a
10 crane system, for example, a hoist of the crane system. The method comprises generating an image data stream based on a plurality of images of the object captured by cameras, arranged at predefined positions and/or angles, of the crane
15 system. The method receives the plurality of images from the cameras and pre-processes the images to form an image data stream. It would be appreciated by a person skilled in the art, that the image data stream may even have a single high resolution image. The method preprocesses each of the images
20 captured by the cameras by reducing noise in the images and/or enhancing contrast of the images. The method determines from the images a foreground associated with the object and annotates the foreground from the images, for example, by applying markers on the foreground of the image.

25

The method analyzes the image data stream by employing a computing unit of the crane system using an artificial neural network. The method analyzes the image data stream to detect presence of defect(s) in the object by segmenting the images,
30 that is the annotated images, based on the artificial neural network, advantageously an artificial neural network trained for identifying markers from an image data stream corresponding to the various object properties of various objects. The

method identifies markers in the images corresponding to defects in the object such that a distance between the markers corresponds to an extent of the defects in the object.

5 According to an embodiment, the method analyzes the image data stream by employing a computing of the crane system using the artificial neural network to detect presence of a human in proximity of the object from the image data stream. The method segments the images, that is the annotated images,
10 based on the artificial neural network and identifies markers in the images corresponding to humans, wherein a distance between the markers corresponds to a proximity of the human with respect to the object.

15 The method determines by employing the computing unit, object properties associated with the object based on the analysis of the image data stream, wherein the object properties comprise at least a quality of the object. The quality of the object is based on the presence of defect(s) in the object.

20 According to an embodiment, the method positions the crane system based on the quality of the object. For example, the crane system may be positioned so as to handle the object(s) of a certain predefined quality at a given time instant. Advantageously, this expedites sorting of the objects.
25

The method automatically operates the crane system for handling the object based on the object properties. The method operates the hoist of the crane system in order to handle the
30 object based on predefined handling parameters defined based on the object properties during training the artificial neural network. The predefined handling parameters include predefined actions to be performed by the hoist including, for

example, pick, drop, load, unload, etc., movement of the gantry of the crane system, etc., based on the object properties. Advantageously, controlling of at least one working parameter of the hoist and/or the crane system depends on positions of markers in the images of the image data stream. For
5 example, if a distance between the markers in an annotated image having a defect is greater than a predefined defect threshold for the object, then the crane system is automatically made to pick the object and drop the object at a predefined
10 location, thereby sorting the object out. Such a predefined threshold is available in the trained artificial neural network. In another example, if a distance between the markers in an annotated image corresponds to a minimum distance indicating a human being in proximity of the object, then the
15 crane system is automatically stopped from moving the object. Such a minimum distance is available in the trained artificial neural network.

Also disclosed herein is a method for training the artificial
20 neural network. Advantageously, said training of the untrained artificial neural network is required to be performed only once so as to enable the trained artificial neural network to identify markers from an image data stream. The method generates a training image data stream by obtaining
25 the images of the object, in an industrial environment, and of surroundings of the object, the object being illuminated at predefined angles by illumination source(s) of the crane system, captured by cameras of the crane system. The method extracts, that is scales and balances, using the artificial
30 neural network, from the images of the training image data stream, the object properties associated with the object and surroundings data associated with the surroundings of the ob-

ject, based on reference images of the object and the surroundings stored in a memory unit accessible to the artificial neural network. As used herein surroundings data refers to data associated with surroundings of the object and comprises, for example, data of humans when present in proximity of the object. Also used herein, reference images of the object include multiple images with and without the object, with and without defects in the object, with and without humans in proximity of the object, etc. The method generating a training database comprising the object properties and the surroundings data for training the artificial neural network. Advantageously, the artificial neural network is trained using a supervised deep learning method for identifying remaining useful life of the object and an unsupervised deep learning method for identifying defect in the object.

Also disclosed herein is a computer program product comprising a non-transitory computer readable storage medium that stores computer program codes comprising instructions executable by at least one processor of the aforementioned computing unit for managing the crane system capable of handling an object. The computer program codes comprise instructions for performing the aforementioned method for managing the crane system.

The above summary is merely intended to give a short overview over some features of some embodiments and implementations and is not to be construed as limiting. Other embodiments may comprise other features than the ones explained above.

30

BRIEF DESCRIPTION OF THE DRAWINGS

The above and other elements, features, steps and characteristics of the present disclosure will be more apparent from the following detailed description of embodiments with reference to the following figures:

5

FIG 1A illustrates a process flow chart of a method for managing a crane system, according to an embodiment of the present disclosure;

10

FIG 1B illustrates a process flow chart of a method for training an artificial neural network capable of identifying markers from an image data stream, according to an embodiment of the present disclosure;

15

FIG 2 illustrates a crane system capable of handling an object in an industrial environment, according to an embodiment of the present disclosure;

20

FIGS 3A-3C illustrate images of an object analysed by the computing unit shown in FIG 2, according to an embodiment of the present disclosure; and

25

FIG 4 illustrates a representation of the object being captured by one of the cameras of the crane system shown in FIG 2.

DETAILED DESCRIPTION OF EMBODIMENTS

30

In the following, embodiments of the disclosure will be described in detail with reference to the accompanying drawings. It is to be understood that the following description of embodiments is not to be taken in a limiting sense.

5

The drawings are to be regarded as being schematic representations and elements illustrated in the drawings, which are not necessarily shown to scale. Rather, the various elements are represented such that their function and general purpose
10 become apparent to a person skilled in the art. Any connection or coupling between functional blocks, devices, components, or other physical or functional units shown in the drawings or described herein may also be implemented by an indirect connection or coupling. A coupling between compo-
15 nents may also be established over a wireless connection. Functional blocks may be implemented in hardware, firmware, software, or a combination thereof.

FIG 1A illustrates a process flow chart 100A of a method for
20 managing a crane system, according to an embodiment of the present disclosure. The method employs a trained artificial neural network and/or a computing unit of the crane system for managing the crane system capable of handling an object in an industrial environment.

25

The method, at step 101, generates an image data stream based on plurality of images of the object captured by cameras of the crane system. The image data stream being a processed image data stream having processed images that can be inter-
30 preted accurately by a trained artificial neural network which the computing unit employs.

At step 101A, the method receives the images captured by the cameras either directly from the cameras or from the crane system that may have a memory unit or a database into which the images are temporarily stored.

5

At step 101B, the method preprocesses each of the images captured by the cameras. The preprocessing comprises cropping the images, reducing noise in the images and/or enhancing contrast and/or brightness of the images. The image preprocessing is required due to the intensity variations, low contrast and a high rate of noise in images that may be captured, for example, using low resolution web cameras. At first, a median filter is applied to the images for noise reduction. After noise reduction, an adaptive histogram equalization is performed on the images for contrast enhancement, considering exponential distribution of histogram. Even though, adaptive histogram equalization provides improvement of the image contrast with no destructive effects on the areas with higher contrast, it may increase the noise on the image. As a result, after adaptive histogram equalization, a median filter is reapplied on the images to reduce the noise, if added any.

At step 101C, the method determines from the images a foreground associated with the object. The foreground may be determined by a variety of image processing techniques including a simple background subtraction, a global grey-level or gradient thresholding or segmentation, statistical classification and/or color classification. The foreground thus determined enables the method, and in turn the trained artificial neural network to identify objects in the images. For example, a steel coil on a factory floor of a steel plant, a human in proximity of the steel coil, etc.

At step 101D, the method annotates the foreground(s) from the image. The method applies image mask(s) for annotating the foreground(s) from the image.

5

At step 102, the method analyzes the image data stream by employing the computing unit using the trained artificial neural network. The method analyzes the image data stream to detect, one or more abnormalities from the image data stream.

10 The abnormalities include, for example, presence of defect(s) in the object, presence of a human in proximity of the object, etc.

At step 102A, the method segments the images based on the
15 trained artificial neural network. Image segmentation plays a crucial role in identifying abnormalities from the image. The artificial neural network is trained, for example, with deep learning based unsupervised and supervised segmentation techniques using U-Net architecture, described in the detailed
20 description of FIG 1B. The image segmentation yields cluster labels as an output wherein a label is assigned to each pixel in the image such that pixels with the same label are connected with respect to some visual or semantic property.

25 At step 102B, the method identifies a distance between markers in the images. The markers are available from the annotated images. The markers represent visual indications applied to the foreground(s) in the image that define the area of the foreground therewithin. The distance between the markers
30 indicates a length of the foreground, a width of the foreground, an area of the foreground, etc.

At step 102C, the method determines based on the distance presence of the abnormalities associated with the object. The distance corresponds to an extent of the defect, a level of proximity of a human with respect to the object, etc. The artificial neural network is trained such that it is capable of identifying, based on the distance between the markers, the abnormalities associated with the object.

At step 103, the method determines object properties associated with the object based on the analysis of the image data stream. The object properties comprise at least a quality of the object. The quality of the object is associated with the abnormalities such as presence of defect(s) in the object. The quality of the object may also be associated with a remaining useful life of the object which is derivable based on the extent of the defect in the object, size of the object, weight of the object, average life of the object, etc. The object properties may also comprise, a position, an orientation, physical dimensions, etc., of the object.

At step 104, the method automatically operates the crane system for handling the object based on the object properties. The method operates the hoist of the crane system for handling the object based on predefined handling parameters used in training the artificial neural network. The predefined handling parameters include a set of actions such as pick, drop, stall, etc., mapped to the object properties. For example, if there exists a quality issue with the object then the hoist is made to pick up the object and drop it at a specified location so as to sort the defective object. The method automatically operates the crane system by controlling at least one working parameter of the hoist and/or the crane system depending on positions of markers in the images of the

image data stream, that is, the distance between the markers and therefore, the object properties derived therefrom which are used to train the artificial neural network.

5 At step 105, the method stores into a database or a memory unit of the crane system the images received from the cameras, the image data stream generated using processed images, the segmented images, the annotated images, the distances between the markers, and/or the object properties. These stored
10 values may in turn be used by the trained artificial neural network to continuously enhance the identification of markers from the images thereby leading to an improved quality assessment of the object and an effective and efficient management of the crane system based on the object properties.

15

FIG 1B illustrates a process flow chart 100B of a method for training an artificial neural network, that is an untrained artificial neural network, capable of identifying markers from an image data stream as disclosed in the detailed description of FIG 1A, according to an embodiment of the present disclosure.
20

At step 106, the method generates a training image data stream by obtaining the images of the object and of surroundings of the object, for example as much as allowed by a field
25 of view of each of the cameras of the crane system, illuminated at predefined angles by illumination source(s) of the crane system. The training image data stream includes multiple images with and without the object in the field of view,
30 with and without the same object in the field of view, with and without multiple objects in the field of view, with and without human(s) in proximity of the object in the surroundings, etc.

At step 107, the method extracts using the untrained artificial neural network, from the images of the training image data stream, the object properties associated with the object and surroundings data associated with the surroundings of the object (201), based on reference images of the object and the surroundings. The method scales and balances the images and extracts the object properties such as defects in the object, remaining useful life of the object, and/or weight, size, physical dimensions, orientation, etc. of the object. The surroundings data includes other objects that are usually found in and around the object such as conveyor belts, factory floor markings, humans, etc. The reference images are pre-labelled and are fed to the artificial neural network for learning purposes. For example, these reference images may be labelled with the object, the defects in the object, a human in proximity of the object, etc. It would be appreciated by a person skilled in the art that while training the artificial neural network, the labeling of reference images is performed as a one time activity. This allows the trained artificial neural network to automatically identify markers from an image data stream without the need to re-label the images captured.

At step 108, the method generates a training database comprising the object properties and the surroundings data for training the artificial neural network. The training database may also include predefined handling parameters for the crane system that enable the method disclosed in FIG 1A to automatically operate the crane system based on the object properties.

The untrained artificial neural network is a Convolutional Neural Network (CNN) that may include a Pyramid Scene Parsing network (PSPnet) with CNN, so as to capture both local and global information along with spatial information of an image, thereby enabling the crane system handling the object to determine not only defects in the object but also remaining useful life of the object.

The artificial neural network is trained using machine learning methods such as a supervised deep learning method for estimating the remaining useful life and/or an unsupervised deep learning method for identifying abnormalities such as presence of a defect in the object. For example, a swish-ReLU activation function is used together with PSP-net and CNN for estimating remaining useful life to increase the smoothness of the learning curve of the artificial neural network. This helps to optimize and generalize the artificial neural network and therefore, the training becomes faster which leads to less CO₂ emissions. The method disclosed herein employs a central processing unit (CPU) and not a graphics processing unit (GPU) thereby, being easy to be deployed in any industrial environment and also being economical.

In an example used for training the artificial neural network, the unsupervised deep learning method computes a d-dimensional feature map $\{a_m\}$ from three different image planes $\{u_m\}$ including, for example, the RGB image planes, through the N convolutional module, the swish ReLU activation function, and a batch normalization function, where a batch corresponds to M pixels of a single input image from the training image data stream. The batch normalization is used for generating a feature map $\{y_m\}$ before assigning cluster labels via argmax classification into the deep learning architecture

used. Next, a large margin soft-max loss between the network responses $\{y_m\}$ and the refined cluster labels $\{c_m\}$ is calculated. Next, the error signals are backpropagated to update the parameters of convolutional filters $\{w_{tm}, b_m\}_{Mm=1}$ as well as the parameters of the classifier $\{w_{tc}, b_c\}$. Here, a stochastic gradient descent with momentum is used for parameter up-dation. Advantageously, setting the learning rate of the artificial neural network to 0.1 (with a momentum of 0.9) yields optimal learning results.

10

The identification of age or remaining useful life is correlated to the object properties. The determination of these object properties is carried out by image processing and deep learning techniques. Thus, rapid, intelligent, and non-destructive techniques are required in training of the artificial neural network. The method formulates calculation of the remaining useful life of an object from an image as a classification problem, for example, each new image from the training image data stream is classified into a class from classes 1-N such that each class corresponds to a time duration indicating the remaining useful life. Basically, the convolutional neural network (CNN) is used for distinguishable feature representation of an image with age information and is trained on those features, including visual and non-visual features, with support vector machine. Furthermore, the CNN's output layer also known as a probability layer consists of 'n' number values for 'n' age classes such as "1-10 units of time", "11-20 units of time", and so on.

30

The identification of defects from an image heavily relies on an external appearance of the object as available in the image, especially when applied to quality inspection and defect sorting applications such as sorting of steel coils based on

their quality. Image segmentation, that is, a process of assigning a label to each pixel in the image such that pixels with the same label are connected with respect to some visual or semantic property, plays a central role in identifying defects from image. The method for training the artificial neural network to identify markers in an image and derive whether or not there exists a defect in the object, employs unsupervised segmentation techniques using U-Net architecture.

10

The problem formulation that is solved for image segmentation is represented using below equation

$$\{a_m \in \mathbb{R}^q\}_{m=1}^M \quad (1)$$

15 for a set of q dimensional feature vectors of image pixels, where M denotes the number of pixels in an input image.

The cluster labels used for segmentation are represented using the equation given below

20

$$\{c_m \in \mathbb{Z}\}_{m=1}^M \quad (2)$$

The labels are assign to all of the pixels by representing a mapping function given below

25

$$c_m = f(a_m), \text{ where } f: \mathbb{R}_q \rightarrow \mathbb{Z} \quad (3)$$

Here, f returns the number of the cluster centroid which is nearer to a_m (using k -means clustering) among k centroids.

30 Therefore, in an unsupervised technique, c_m can be derived/predicted using fixed value of f and a_m whereas in a supervised technique, f and a_m are trainable and c_m are fixed. However, f and a_m can be optimized using different optimization like

stochastic gradient descent etc. Therefore, spatially continuous pixels of similar features are desired to be assigned the same label.

5 The method assigns the same label to pixels of similar features. Next, linear classifier is applied that classifies the features of each pixel into d classes.

Let's assume, for an RGB image

$$10 \quad I = \{u_m \in \mathbb{R}^3\}_{m=1}^M \quad (4)$$

after image normalization. We compute a d -dimensional feature map $\{a_m\}$ from three different image planes $\{u_m\}$ through N convolutional module, swish ReLU activation function, and a batch
15 normalization function, where a batch corresponds to M pixels of a single input image. Here, q filters of region size 3×3 for all the N components are set. Next, a mapping function is obtained by applying a linear classifier

$$20 \quad \{y_m = wt_c a_m + b_c\}_{m=1}^M \quad (5)$$

where $wt_c \in \mathbb{R}^{d \times q}$ and $b_c \in \mathbb{R}^d$.

The response map to $\{y'_m\}$ is normalized such that $\{y'_m\}_{m=1}^M$ has
25 mean (= 0) and variance (=1). Finally, the cluster label c_m is obtained for each pixel by selecting the dimension that has the maximum value in y'_m . This type of classification is referred as argmax classification. To make it more meaningful, an additional constraint that supports cluster labels that are
30 the same as those of adjacent pixels is added.

First, L fine super pixels $\{S_l\}_{l=1}^L$ (with a large L) from the input image I are extracted, where S_l denotes a set of the

indices of pixels that belong to the l -th super pixel. Then, all of the pixels in each super pixel are assigned to have the same cluster label. Here, simple minimum spanning tree based iterative clustering is used with L ($= 32$) for the super pixel
 5 extraction.

Selecting the number of cluster labels (d) in an unsupervised image segmentation is a challenging task. As describe above, the strategy is to classify pixels into random number
 10 d' ($1 \leq d' \leq d$) of clusters. The large and small values of d' indicates over and under segmentation. To prevent this kind of under segmentation failure a batch normalization is incorporated (where a batch corresponds to M pixels of a single input image) for generating feature map $\{y_m\}$ before assigning cluster
 15 labels via argmax classification into the used deep learning architecture.

The method auto-trains the artificial neural network for unsupervised image segmentation in the following manner. Once a
 20 target image, that is an image of the object, is input, there are two alternatives namely prediction of cluster labels with fixed network parameters which corresponds to the forward process of a network followed by the super pixel refinement described above and/or training of network parameters with
 25 the fixed cluster labels which corresponds to the backward process of a network based on gradient descent. As with the case of supervised learning, the method calculates the large margin soft-max loss between the network responses $\{y'_m\}$ and the refined cluster labels $\{c'_m\}$. Next, the method backpropa-
 30 gates the error signals to update the parameters of convolutional filters $\{wt_m, b_m\}_{m=1}^M$ as well as the parameters of the classifier $\{wt_c, b_c\}$ using stochastic gradient descent with momentum for parameters up-dation.

The artificial neural network trained in aforementioned manner not only allows computation of abnormalities, that is, defects in the object but also remaining useful life and has the ability to add more features in future as required. PSPnet with CNN, helps to capture both local and global information of an image which enables to find the remaining useful life along with defects present if any in the object from its images.

10

FIG 2 illustrates a crane system 200 capable of handling an object 201 in an industrial environment, according to an embodiment of the present disclosure. The crane system 200 comprises a hoist 208 for moving the object 201, for example, a steel coil in a steel plant. The crane system comprises cameras 202, 203, and 204 positioned at predefined positions and at predefined angles of capture so as to capture high-definition, real time images of the object 201. The camera 202 is arranged at a first end 211A of a gantry 211 of the crane system 200. The camera 203 is arranged at a second end 211B of the gantry 211 and the camera 204 is arranged in proximity of the hoist 208 on the gantry 211.

The crane system 200 comprises a control unit 209. The control unit 209 moves the cameras 202, 203, and 204, for capturing the images of the object 201, along a longitudinal axis A-A' of gantry tracks 210A, 210B.

The crane system 200 comprises a computing unit 205 in operable communication with the control unit 209 via a wired or a wireless communication network 206. The computing unit 205 may also communicate with the cameras 202, 203, and 204 via

30

the wired or a wireless communication network 206. The computing unit stores therein the artificial neural network.

5 The crane system 200 comprises a training database 207. The artificial neural network accesses the training database to train itself for identifying markers in the images of the object 201 and determined object properties therewith. The training database may also store therein the images captured by the cameras 202, 203, and 204.

10

FIGS 3A-3C illustrate images 300A, 300B, 300C of an object 201 analysed by the computing unit 205 shown in FIG 2, according to an embodiment of the present disclosure.

15 FIG 3A shows an image 300A of the object 201 captured by one or more of the cameras 202, 203, 204. The image 300A may also represent an image collated based on individual images captured by each of the cameras 202, 203, 204. As shown in FIG 3A, the object 201 which is a steel coil has defects, that is, cracks 201A.

20

FIG 3B shows an image 300B that is pre-processed, segmented, and annotated by the computing 205 based on the trained artificial neural network. The image 300B shows annotations
25 around the object 201, that is the steel coil, as a whole and around each of the areas in the image 300B that differ significantly from other areas in the image 300B, for example, the cracks 201A and the center 201C of the coil which is an empty space at the center of the steel coil roll. The computing unit 205 based on the trained artificial neural network
30 identifies markers 201B on each of the aforementioned areas. The computing unit 205 based on the trained artificial neural network then identifies a distance 'D' between each of these

markers and thereafter presence of a defect in the object 201.

FIG 3C shows an image 300C that is an image as seen by the trained artificial neural network wherein the foreground that is the steel coil, the background that is the surrounding of the steel coil, the defects that is the cracks 201A, and/or a non-defect area such as an empty space at the center 201C of the steel coil are clearly differentiated.

Based on the above derivations, that is a presence of a defect in the object 201, and predefined handling parameters used in training of the artificial neural network, the computing unit 205 via the control unit 209 causes the hoist 208 of the crane system shown in FIG 2 to handle the object 201.

FIG 4 illustrates a representation of the object 201 being captured by one of the cameras 204 of the crane system 200 shown in FIG 2. The camera 204 is positioned on the gantry in proximity of the hoist 208 as shown in FIG 2. The camera 204 is positioned at a height 'Z' from the ground level on which the object 201 is positioned. The camera has an angle of capture θ . The computing unit 205 based on the images captured by the camera 204 and the cameras 202 and 203 determines a size of the object, that is, a height 'h', a width 'w', and a breadth 'b' as shown in FIG 4. The computing unit 205 determines the width 'w' based on images captured by the cameras 202 or 203. The breadth 'b' is determined based on the images captured by the camera 204. A field of view 'a' and the height 'h' are determined as explained in the equations provided below:

$$a = 2 * Z * \tan(\theta/2)$$

$$\tan \alpha = \frac{b}{2Z} = \frac{\left(\frac{b-w}{2}\right)}{h}$$

$$\text{therefore, } h = \left[\frac{\frac{(b-w)}{2}}{\frac{b}{2Z}} \right]$$

5

$$\text{and therefore, } h = \frac{(b-w)Z}{b}$$

Where α is angle of capture of the camera 204.

10 Thus, the cameras 203, 203, and 204 captured images of the object 201 so as to enable the computing unit 205 to determined physical dimensions of the object, an orientation of the object and thereby, defects associated with the object with help of the trained artificial neural network.

15

Where databases are described such as the training database 207, it will be understood by one of ordinary skill in the art that (i) alternative database structures to those described may be readily employed, and (ii) other memory structures besides databases may be readily employed. Any illustrations or descriptions of any sample databases disclosed herein are illustrative arrangements for stored representations of information. Any number of other arrangements may be employed besides those suggested by tables illustrated in the drawings or elsewhere. Similarly, any illustrated entries of the databases represent exemplary information only; one of ordinary skill in the art will understand that the number and content of the entries can be different from those disclosed herein. Further, despite any depiction of the databases as tables, other formats including relational databases, object-

20

25

30

based models, and/or distributed databases may be used to store and manipulate the data types disclosed herein. Likewise, object methods or behaviors of a database can be used to implement various processes such as those disclosed

5 herein. In addition, the databases may, in a known manner, be stored locally or remotely from a device that accesses data in such a database. In embodiments where there are multiple databases in the system, the databases may be integrated to communicate with each other for enabling simultaneous updates

10 of data linked across the databases, when there are any updates to the data in one of the databases.

The present disclosure can be configured to work in a network environment comprising one or more computers that are in communication with one or more devices via a network. The computers may communicate with the devices directly or indirectly, via a wired medium or a wireless medium such as the Internet, a local area network (LAN), a wide area network (WAN) or the Ethernet, a token ring, or via any appropriate

20 communications mediums or combination of communications mediums. Each of the devices comprises processors, some examples of which are disclosed above, that are adapted to communicate with the computers. In an embodiment, each of the computers is equipped with a network communication device, for example,

25 a network interface card, a modem, or other network connection device suitable for connecting to a network. Each of the computers and the devices executes an operating system, some examples of which are disclosed above. While the operating system may differ depending on the type of computer, the operating system will continue to provide the appropriate communications protocols to establish communication links with

30 the network. Any number and type of machines may be in communication with the computers.

The present disclosure is not limited to a particular computer system platform, processor, operating system, or network. One or more aspects of the present disclosure may be distributed among one or more computer systems, for example, servers configured to provide one or more services to one or more client computers, or to perform a complete task in a distributed system. For example, one or more aspects of the present disclosure may be performed on a client-server system that comprises components distributed among one or more server systems that perform multiple functions according to various embodiments. These components comprise, for example, executable, intermediate, or interpreted code, which communicate over a network using a communication protocol. The present disclosure is not limited to be executable on any particular system or group of systems, and is not limited to any particular distributed architecture, network, or communication protocol.

The foregoing examples have been provided merely for the purpose of explanation and are in no way to be construed as limiting of the present disclosure disclosed herein. While the disclosure has been described with reference to various embodiments, it is understood that the words, which have been used herein, are words of description and illustration, rather than words of limitation. Further, although the disclosure has been described herein with reference to particular means, materials, and embodiments, the disclosure is not intended to be limited to the particulars disclosed herein; rather, the disclosure extends to all functionally equivalent structures, methods and uses, such as are within the scope of the appended claims. Those skilled in the art, having the benefit of the teachings of this specification, may affect

numerous modifications thereto and changes may be made without departing from the scope of the disclosure in its aspects.

Patent Claims

1. A method (100A) for managing a crane system (200) capable of handling an object (201), the method comprising:
 - 5 - generating (101) an image data stream based on a plurality of images (300A) of the object (201) captured by cameras (202, 203, 204) of the crane system (200);
 - analyzing (102) the image data stream by employing a computing unit (205) of the crane system (200) using an
10 artificial neural network, wherein the artificial neural network is trained for identifying markers (201B) from the image data stream;
 - determining (103), by the computing unit (205), object properties associated with the object (201) based on
15 the analysis of the image data stream, wherein the object properties comprise at least a quality of the object (201); and
 - automatically operating (104) the crane system (200) for handling the object (201) based on the object prop-
20 erties.
2. The method according to claim 1, wherein generating the image data stream comprises performing at least one of:
 - 25 - preprocessing each of the images (300A) captured by the cameras (202, 203, 204), wherein preprocessing comprises one or more of reducing noise in the images (300A) and enhancing contrast of the images (300A);
 - determining from the images (300A) a foreground associated with the object (201);
 - 30 - annotating the foreground from the images (300A).
3. The method according to claim 1 or 2, wherein analyzing the image data stream using the artificial neural network

comprises detecting, from the image data stream, presence of one or more abnormalities associated with the object (201), and wherein the abnormalities comprise one or more of a defect in the object and a human in proximity of the object.

4. The method according to any one of the preceding claims, wherein analyzing the image data stream using the artificial neural network comprises:
- 10 - segmenting the images (300A) based on the artificial neural network;
 - identifying a distance between markers (201B) in the images; and
 - determining, based on the distance and the segmented
15 images, presence of the abnormalities.
5. The method according to any of the preceding claims, wherein automatically operating the crane system (200) comprises operating a hoist (208) of the crane system (200) for handling the object (201) based on predefined
20 handling parameters defined based on the object properties during training of the artificial neural network.
6. The method according to claim 5, further comprising controlling at least one working parameter of one of the
25 hoist (208) and the crane system (200) depends on positions of markers (201B) in the images (300A) of the image data stream.
- 30 7. A method (100B) for training the artificial neural network, in particular according to any one of the claims 1-6, for identifying markers (201B) from an image data stream, comprising:

- generating (106) a training image data stream by obtaining the images (300A) of the object (201) and of surroundings of the object (201), illuminated at predefined angles by one or more illumination sources of the crane system (200), captured by cameras (202, 203, 204) of the crane system (200);
- extracting (107) using the artificial neural network, from the images (300A) of the training image data stream, the object properties associated with the object (201) and surroundings data associated with the surroundings of the object (201), based on reference images of the object (201) and the surroundings; and
- generating (108) a training database (207) comprising the object properties and the surroundings data for training the artificial neural network.

8. A crane system (200) capable of handling an object (201), comprising:

- cameras (202, 203, 204) positioned so as to capture a plurality of images (300A) of the object (201); characterized by:
- a computing unit (205) having an artificial neural network configured to:
 - o analyze an image data stream generated based on the images (300A) using an artificial neural network; and
 - o determine object properties associated with the object (201) based on the analysis of the image data stream, wherein the object properties comprise at least a quality of the object (201); and

- a control unit (209) configured to automatically operate a hoist (208) of the crane system (200) for handling the object (201) based on the object properties.

5 9. The crane system (200) according to claim 8, wherein the control unit (209) is configured to move the cameras (202, 203, 204), for capturing of the images (300A) of the object (201), along an axis (A-A') of gantry tracks (210A, 210B).

10

10. The crane system (200) according to any one of the claims 8 and 9, comprises a first camera (202) arranged at a first end (211A) of a gantry (211) of the crane system (200), a second camera (203) arranged at a second end
15 (211B) of the gantry (211), and a third camera (204) arranged in proximity of the hoist (208) on the gantry (211).

11. A computing unit (205) having an artificial neural network for managing a crane system (200) according to any of
20 the claims 8-10.

12. A computer program product comprising a non-transitory computer readable storage medium that stores computer program codes comprising instructions executable by at least
25 one processor of a computing unit (205) for managing a crane system (200) according to any of the claims 8-10, capable of handling an object (201), wherein the computer program codes comprise instructions for performing the
30 method according to any of the claims 1-6.