

AUTOMATED LABELING OF OBJECTS IN RADAR IMAGES AND SIGNALS USING ELECTRO-OPTIC IMAGES

FIELD OF THE INVENTION

5 The present disclosure is in the field of data collection for training machine learning algorithms, in particular machine learning algorithms of non-Electro-Optic data such as RADAR images and signals.

BACKGROUND OF THE INVENTION

Machine learning is used to classify objects in images captured by electro-optic
10 (EO) sensors.

Non-electro-optic sensors, such as RADAR, capture image data associated with objects. The task of gathering large amount of RADAR data is difficult to accomplish.

GENERAL DESCRIPTION

15 In accordance with a first aspect of the presently disclosed subject matter, there is presented a method of labelling data, the method performed by a computerized system comprising a processing circuitry, the method being performed by the processing circuitry and comprising performing the following:

- 20 (a) receiving unlabeled non-Electro-Optical (non-EO) data, the non-EO data comprising a set of non-EO images or signals captured during a first time period, the data being indicative of a first region;
- (b) receiving unlabeled EO data,
 captured during a second time period corresponding to the first time period,
 the unlabeled EO data comprising at least a plurality of video image frames,
25 the images being indicative of a second region corresponding to the first region;
- (c) classifying at least one second object appearing on one or more of the plurality of video image frames,
 thereby generating at least one video-image label,
30 thereby generating labelled EO data from the unlabeled EO data; and

(d) labeling at least one first object appearing in the unlabeled non-EO data with a corresponding video-image label of the at least one video-image label, the labelling being performed based at least on:

- 5 (I) a correspondence of at least one first time of capture, of at least one non-EO image or signal of the set of non-EO images or signals, to at least one second time of capture, of at least one video image frame of the plurality of video image frames;
- (II) a correspondence of positions of the at least one first object and the at least one classified second object; and
- 10 (III) a respective geographic position, associated with the at least one video image frame, of the at least one classified second object, thereby generating labelled non-EO data from the unlabeled non-EO data, the labelled non-Electro-Optical (non-EO) data being configured to facilitate machine learning of non-Electro-Optical signatures of first objects.

15 In addition to the above features, the system according to this aspect of the presently disclosed subject matter can include one or more features (i) to (xl) listed below, in any desired combination or permutation which is technically possible:

- (i) the unlabeled EO data further comprising at least one still image, wherein the method further comprises, prior to said step (d), performing
20 a geo-registration of the at least one still image, wherein the performing the labelling of the at least one first object, based at least on the respective geographic position of the at least one classified second object, is based at least on the geo-registration.
- (ii) the geo-registration comprises matching features of the at least one still
25 image to orthophoto data.
- (iii) the geo-registration utilizes Digital Terrain Elevation Data (DTED) data.
- (iv) the non-Electro-Optical (non-EO) data comprises RADAR imaging data.
- (v) the RADAR imaging data comprises at least one of Synthetic Aperture
30 RADAR (SAR) imaging data and Ground Moving Target Indication (GMTI) signals.
- (vi) each GMTI signal comprises a range-Doppler map.

- (vii) the second period comprises the first period.
- (viii) The method of any one of claims 1 to 7, wherein the first period and the second period are the same.
- (ix) the second region comprises the first region.
- 5 (x) the first region and the second region are the same.
- (xi) the method further comprises, prior to said step (a), capturing the unlabeled non-EO data.
- (xii) the capturing of the unlabeled non-EO data utilizes a non-EO sensor mounted on a first airborne vehicle.
- 10 (xiii) the first airborne vehicle comprises one of a first quadcopter, a first octa-copter, a first drone, and a first other remote-piloted vehicle (RPV).
- (xiv) the non-EO sensor is configured for setting of a capture mode of a plurality of capture modes.
- (xv) the plurality of capture modes comprises SAR capture mode and GMTI capture mode.
- 15 (xvi) the method further comprises, prior to said step (b), capturing the unlabeled EO data.
- (xvii) the capturing of the unlabeled EO data utilizes an EO sensor mounted on a second airborne vehicle.
- 20 (xviii) the second airborne vehicle comprises one of a second quadcopter, a second octa-copter, a second drone, and a first other remote-piloted vehicle (RPV).
- (xix) the labelling of the at least one first object comprises a labelling of objects in each non-EO image or signal.
- 25 (xx) the unlabeled EO data are captured while the EO sensor is stationary.
- (xxi) the classifying the at least one second object comprises detecting the at least one second object.
- (xxii) the classifying the at least one second object is performed at least partially in an automated manner.
- 30 (xxiii) the classifying the at least one second object utilizes Machine Learning.
- (xxiv) the classifying the at least one second object is performed at least partially manually.

(xxv) the generating the at least one video-image label comprises assigning, to each classified second object of the at least one second object, a respective video-image label indicative of the respective classification.

5 (xxvi) the classifying the at least one second object comprises tracking the at least one second object across video image frames of the plurality of video image frames,

thereby facilitating an improved classification of the at least one second object in at least one video image frame of the plurality of video image frames, as compared to a second classification performed without tracking.

10 (xxvii) the generating the at least one video-image label comprises:
(A) generating a ground-truth data set assigning, to each classified second object of the at least one second object, a video-image label indicative of the classification,

15 wherein each data item of the ground-truth data set comprises:

- (i) the respective video-image label indicative of the classification of the each classified second object;
- (ii) an indication of a second time of capture, of a respective video-image frame of the plurality of video image frames, which comprises the each classified second object; and
- 20 (iii) an indication of the respective geographic position of the each classified second object corresponding to the second time of capture.

(xxviii) the ground-truth data set comprises a ground-truth table.

25 (xxix) the indication of the respective geographic position comprises geographic coordinates.

(xxx) the generating the at least one video-image label comprises performing, prior to the step (A), the following step:

30 (B) generating an interim ground-truth data set, which assigns, to each classified second object of the at least one second object, a video-image label indicative of the classification, wherein each data item of the interim ground-truth data set comprises:

- (i) the video-image label;
- (ii) the indication of a time of capture of the respective video-image frame; and
- (iii) an indication of a respective second relative position within the respective video-image frame, of the each classified second object corresponding to the time of capture,

wherein the performing of the labelling of the at least one first object comprises deriving the ground-truth data set from the interim ground-truth data set,

wherein the deriving the ground-truth data set is performed by deriving the indication of respective geographic position from the respective second relative position, based at least on the geo-registration of the at least one still image,

wherein the deriving the indication of respective geographic position comprises registering the respective video-image frame to the at least one still image.

(xxxi) the indication of the respective relative second position within the respective video-image frame comprises coordinates within the respective video-image frame.

(xxxii) the generating the ground-truth data set is performed for respective second video image frames, for which the second time of capture, and a first time of capture of a corresponding non-EO image or signal, meet a time-matching criterion.

(xxxiii) the time-matching criterion comprises at least one of the following:

- a. the second time of capture is equal to the first time of capture; and
- b. a difference between the first time of capture and a closest video frame is a maximum of half of a video frame time.

(xxxiv) the performing the labelling of the at least one non-EO image,

based at least on a correspondence of at least one first time of capture, of the at least one non-EO image or signal, to

the to at least one second time of capture, of the of at least one video image frame,

on the correspondence of positions of the at least one first object and the at least one classified second object, and

5 on the respective geographic position of the at least one classified second object,

comprises performing the following steps:

- (I) determine that a corresponding second time of capture of the respective video-image frame, and a first time of capture of the
10 at least one non-EO image or signal, meet a matching criterion;
- (II) select one or more data items of the ground-truth data set that comprise an indication of the corresponding second time of capture of the respective video-image frame;
the selected one or more data items comprising one or more
15 selected classified second objects;
- (III) based on one or more indications of respective geographic positions of the one or more selected classified second objects, determine one or more respective geographic first positions of localized respective first objects associated with the at least one
20 non-EO image or signal,
the one or more selected classified second objects constituting the at least one second object;
- (IV) label a respective localized first objects with corresponding respective video-image labels of the one or more selected
25 classified second objects,
the respective localized first object constituting the at least one first object.

(xxxv) said step (III) comprises:

30 based on the determined one or more respective geographic first positions of the localized respective first objects, determine for each localized respective first object a corresponding first relative position, within the at least one non-EO image or signal.

(xxxvi) the time-matching criterion comprises at least one of the following:

- a. the second time of capture is equal to the first time of capture;
and
- b. a difference between the first time of capture and a closest
video frame is a maximum of half of a video frame time.

(xxxvii) each data item of the ground-truth data set further comprises:

- (iv) an indication of movement, indicating whether
the each classified second object is moving during the time
of capture of the respective video-image frame,

wherein the indication whether the each classified second
object is moving is determined based on tracking the each
classified second object across video image frames of the
plurality of video image frames.

(xxxviii) the non-EO data comprises RADAR imaging data, wherein the

RADAR imaging data comprises Synthetic Aperture RADAR (SAR)
imaging data,

wherein the selecting of the one or more data items of the ground-
truth data set comprises selecting one or more records of the ground-
truth data set for which a corresponding indication of movement is
indicative of a stationary classified second object.

(xxxix) the non-EO data comprises RADAR imaging data, wherein the RADAR

imaging data comprises Ground Moving Target Indication (GMTI)
range-Doppler maps,

wherein the selecting of the one or more data items of the ground-
truth data set comprises selecting one or more data items of the ground-
truth data set for which the corresponding indication of movement is
indicative of a moving classified second object.

(xl) the method further comprising:

- (V) training models for Non-EO objects classification, based at least
on the labeled respective first objects.

In accordance with a second aspect of the presently disclosed subject
matter, there is presented a computerized system configured to label data, the

computerized system comprising a processing circuitry configured to perform the following method:

- 5 (a) receiving unlabeled non-Electro-Optical (non-EO) data, the non-EO data comprising a set of non-EO images or signals captured during a first time period, the data being indicative of a first region;
- (b) receiving unlabeled EO data,
captured during a second time period corresponding to the first time period,
the unlabeled EO data comprising at least a plurality of video image frames,
the images being indicative of a second region corresponding to the first
10 region;
- (c) classifying at least one second object appearing on one or more of the plurality of video image frames,
thereby generating at least one video-image label,
thereby generating labelled EO data from the unlabeled EO data; and
- 15 (d) labeling at least one first object appearing in the unlabeled non-EO data with a corresponding video-image label of the at least one video-image label,
the labelling being performed based at least on:
 - (I) a correspondence of at least one first time of capture, of at least
one non-EO image or signal of the set of non-EO images or
20 signals, to at least one second time of capture, of at least one
video image frame of the plurality of video image frames;
 - (II) a correspondence of positions of the at least one first object and
the at least one classified second object; and
 - (III) a respective geographic position, associated with the at least one
25 video image frame, of the at least one classified second object,
thereby generating labelled non-EO data from the unlabeled non-EO data,
the labelled non-Electro-Optical (non-EO) data being configured to facilitate
machine learning of non-Electro-Optical signatures of first objects.

30 In accordance with a third aspect of the presently disclosed subject matter, there is presented a non-transitory computer readable storage medium tangibly embodying a program of instructions which, when executed by a computerized system, cause the computerized system to perform a method of

labelling data, the method being performed by a processing circuitry of the computerized system, the method comprising performing the following:

- 5 (a) receiving unlabeled non-Electro-Optical (non-EO) data, the non-EO data comprising a set of non-EO images or signals captured during a first time period, the data being indicative of a first region;
- (b) receiving unlabeled EO data,
captured during a second time period corresponding to the first time period,
the unlabeled EO data comprising at least a plurality of video image frames,
the images being indicative of a second region corresponding to the first
10 region;
- (c) classifying at least one second object appearing on one or more of the plurality of video image frames,
thereby generating at least one video-image label,
thereby generating labelled EO data from the unlabeled EO data; and
- 15 (d) labeling at least one first object appearing in the unlabeled non-EO data with a corresponding video-image label of the at least one video-image label,
the labelling being performed based at least on:
 - (I) a correspondence of at least one first time of capture, of at least
one non-EO image or signal of the set of non-EO images or
20 signals, to at least one second time of capture, of at least one
video image frame of the plurality of video image frames;
 - (II) a correspondence of positions of the at least one first object and
the at least one classified second object; and
 - (III) a respective geographic position, associated with the at least one
25 video image frame, of the at least one classified second object,
thereby generating labelled non-EO data from the unlabeled non-EO data,
the labelled non-Electro-Optical (non-EO) data being configured to facilitate
machine learning of non-Electro-Optical signatures of first objects.

30 In accordance with a fourth aspect of the presently disclosed subject matter, there is presented a method of labelling data, the method performed by a computerized system comprising a processing circuitry, the method being performed by the processing circuitry and comprising performing the following:

- (a) receiving unlabeled non-Electro-Optical (non-EO) data, the non-EO data comprising a set of non-EO images or signals captured during a first time period, the data being indicative of a first region;
- (b) receiving unlabeled EO data,
5 captured during a second time period corresponding to the first time period, the unlabeled EO data comprising at least a plurality of video image frames, the images being indicative of a second region corresponding to the first region;
- (c) classifying at least one second object appearing on one or more of the
10 plurality of video image frames, thereby generating at least one video-image label, thereby generating labelled EO data from the unlabeled EO data, wherein the classifying the at least one second object comprises tracking the at least one second object across video image frames of the
15 plurality of video image frames, thereby facilitating an improved classification of the at least one second object in at least one video image frame of the plurality of video image frames, as compared to a second classification performed without tracking; and
- (d) labeling at least one first object appearing in the unlabeled non-EO data with
20 a corresponding video-image label of the at least one video-image label, the labelling being performed based at least on:
- (I) a correspondence of at least one first time of capture, of at least one non-EO image or signal of the set of non-EO images or
25 signals, to at least one second time of capture, of at least one video image frame of the plurality of video image frames;
 - (II) a correspondence of positions of the at least one first object and the at least one classified second object; and
 - (III) a respective geographic position, associated with the at least one
30 video image frame, of the at least one classified second object,
- thereby generating labelled non-EO data from the unlabeled non-EO data,

the labelled non-Electro-Optical (non-EO) data being configured to facilitate machine learning of non-Electro-Optical signatures of first objects.

5 In accordance with a fifth aspect of the presently disclosed subject matter, there is presented a method of labelling data, the method performed by a computerized system comprising a processing circuitry, the method being performed by the processing circuitry and comprising performing the following:

- 10 (a) receiving unlabeled non-Electro-Optical (non-EO) data, the non-EO data comprising a set of non-EO images or signals captured during a first time period, the data being indicative of a first region;
- (b) receiving unlabeled EO data,
captured during a second time period corresponding to the first time period, the unlabeled EO data comprising at least a plurality of video image frames,
15 the images being indicative of a second region corresponding to the first region;
- (c) classifying at least one second object appearing on one or more of the plurality of video image frames,
thereby generating at least one video-image label,
20 thereby generating labelled EO data from the unlabeled EO data,

wherein the generating the at least one video-image label comprises generating an indication of movement, indicating whether the each classified second object is moving during the time of capture of the respective video-image frame,

25 wherein the indication whether the each classified second object is moving is determined based on tracking the each classified second object across video image frames of the plurality of video image frames;
and

- (d) labeling at least one first object appearing in the unlabeled non-EO data with
30 a corresponding video-image label of the at least one video-image label,
the labelling being performed based at least on:

- (I) a correspondence of at least one first time of capture, of at least one non-EO image or signal of the set of non-EO images or signals, to at least one second time of capture, of at least one video image frame of the plurality of video image frames;
- 5 (II) a correspondence of positions of the at least one first object and the at least one classified second object; and
- (III) a respective geographic position, associated with the at least one video image frame, of the at least one classified second object,
- thereby generating labelled non-EO data from the unlabeled non-EO data,
- 10 the labelled non-Electro-Optical (non-EO) data being configured to facilitate machine learning of non-Electro-Optical signatures of first objects.

- In accordance with a sixth aspect of the presently disclosed subject matter, there is presented a method of labelling data, the method performed by a computerized system comprising a processing circuitry,
- 15 the method being performed by the processing circuitry and comprising performing the following:
- (a) receiving unlabeled non-Electro-Optical (non-EO) data, the non-EO data comprising a set of non-EO images or signals captured during a first time
- 20 period, the data being indicative of a first region;
- (b) receiving unlabeled EO data,
- captured during a second time period corresponding to the first time period, the unlabeled EO data comprising at least a plurality of video image frames, the images being indicative of a second region corresponding to the first
- 25 region,
- wherein the unlabeled EO data further comprising at least one still image;
- (c) classifying at least one second object appearing on one or more of the plurality of video image frames,
- 30 thereby generating at least one video-image label,
- thereby generating labelled EO data from the unlabeled EO data; and
- (d) performing a geo-registration of the at least one still image,

(e) labeling at least one first object appearing in the unlabeled non-EO data with a corresponding video-image label of the at least one video-image label, the labelling being performed based at least on:

- 5 (I) a correspondence of at least one first time of capture, of at least one non-EO image or signal of the set of non-EO images or signals, to at least one second time of capture, of at least one video image frame of the plurality of video image frames;
- (II) a correspondence of positions of the at least one first object and the at least one classified second object; and
- 10 (III) a respective geographic position, associated with the at least one video image frame, of the at least one classified second object, thereby generating labelled non-EO data from the unlabeled non-EO data, the labelled non-Electro-Optical (non-EO) data being configured to facilitate machine learning of non-Electro-Optical signatures of first
15 objects,

wherein the performing the labelling of the at least one first object, based at least on the respective geographic position of the at least one classified second object, is based at least on the geo-registration.

The second to sixth aspects can optionally further comprise one or more of
20 features (i) to (xl) listed above, *mutatis mutandis*, in any technically possible combination or permutation.

BRIEF DESCRIPTION OF THE DRAWINGS

25 In order to understand the invention and to see how it may be carried out in practice, some specific embodiments will now be described, by way of non-limiting example only, with reference to the accompanying drawings, in which:

Fig. 1 is a schematic illustration of a scenario of capturing electro-optic (EO) and non-EO data, according to some embodiments of the presently disclosed subject
30 matter;

Fig. 2 is a schematic illustration of a scenario of capturing electro-optic (EO) and non-EO data, according to some embodiments of the presently disclosed subject matter;

5 **Fig. 3** is a schematic illustration of a non-limiting example of a generalized schematic diagram of a vehicle comprising computerized EO imaging system, in accordance with some embodiments of the presently disclosed subject matter;

Fig. 4 is a schematic illustration of a non-limiting example of a generalized schematic diagram of a vehicle comprising computerized non-EO imaging system, in accordance with some embodiments of the presently disclosed subject matter;

10 **Fig. 5** is a schematic illustration of a non-limiting example of a generalized schematic diagram of a system comprising computerized non-electro-optic (EO) labeling system, in accordance with some embodiments of the presently disclosed subject matter.;

15 **Fig. 6** is a schematic illustration of a non-limiting example of a generalized flow diagram for labeling of non-EO images, in accordance with some embodiments of the presently disclosed subject matter;

Fig. 7 is a schematic illustration of a non-limiting example of a geo-registration, in accordance with some embodiments of the presently disclosed subject matter;

20 **Fig. 8** is a schematic illustration of a non-limiting example of EO object detection, tracking, and classification, in accordance with some embodiments of the presently disclosed subject matter;

Fig. 9A is a schematic illustration of a non-limiting example of an interim ground truth data set, in accordance with some embodiments of the presently disclosed subject matter;

25 **Fig. 9B** is a schematic illustration of a non-limiting example of a ground truth data set, in accordance with some embodiments of the presently disclosed subject matter;

30 **Fig. 10** is a schematic illustration of a non-limiting example of object localization, in accordance with some embodiments of the presently disclosed subject matter;

Fig. 11 is a schematic illustration of a non-limiting example of GMTI, in accordance with some embodiments of the presently disclosed subject matter;

Fig. 12 is a schematic illustration of a non-limiting example of position matching, in accordance with some embodiments of the presently disclosed subject matter;

Fig. 13 is a schematic illustration of a non-limiting example of labeling, in accordance with some embodiments of the presently disclosed subject matter; and

Fig. 14 is a schematic illustration of a non-limiting example of a generalized Graphical User Interface, in accordance with some embodiments of the presently disclosed subject matter.

10 DETAILED DESCRIPTION OF THE INVENTION

In the following description the invention will be illustrated with reference to specific embodiments of a system and method in accordance with the invention. It may be appreciated, that the invention is not limited to such use and the illustrated embodiment is illustrative to the full scope of the invention as described and claimed herein.

It is appreciated that certain features of the invention, which are, for clarity, described in the context of separate embodiments, may also be provided in combination in a single embodiment. Conversely, various features of the invention, which are, for brevity, described in the context of a single embodiment, may also be provided separately or in any suitable sub-combination. While the invention has been shown and described with respect to particular embodiments, it is not thus limited.

As used herein, the phrase "for example," "such as" and variants thereof describing exemplary implementations of the present invention are exemplary in nature and not limiting.

As used herein, the terms "one or more" and "at least one" aim to include one as well as any number greater than one e.g. two, three, four, etc.

Fig. 1 is a schematic illustration of a scenario **100** of capturing electro-optic (EO) and non-EO data, according to some embodiments of the presently disclosed subject matter. In this example, the scenario **100** illustrates capture of images of objects **150, 160, 170, 180** by EO sensors comprised on platform **110**, and by non-EO sensors comprised on platform **120**.

In the field of image processing and object detection, recognition and classification, there is a need to detect, recognize and classify, correctly and accurately,

objects appearing in images and other data captured by non-electro-optic sensors such as Radio Detection and Ranging (RADAR) systems. One way of detecting, recognizing and classifying objects on images or other type of signals is to use machine learning methods using for example neural networks. Machine learning is used to classify
5 objects in images and other signals.

For example, objects on EO images can be labelled with their classification, the classification models can be trained using the labelled image data, and the models then can be used to classify objects on unlabeled images. In the domain of electro-optic (EO) sensors there are a large number of labeled data sources, and the labeling procedure is
10 considered a task of which human beings are capable.

However, when developing classifiers and detectors in non-EO domains, e.g. for Synthetic Aperture RADAR (SAR) images and Ground Moving Target Indication (GMTI) signals, for use operationally when imaging or tracking e.g. enemy forces and vehicles, such a machine learning model has at least certain technical challenges.

First, in such domains, a human cannot label certain objects in an easy manner. While in EO images, the human can easily look at the image and determine, based on visual inspection, "this is a bus, this is a car, this is a tank", non-EO signatures such as SAR and GMTI signatures cannot be easily determined. People's senses, e.g. the human eye, do not "see" the frequencies associated with such RADAR sensors. Also, the bus
15 visually looks like a bus in the EO image, intuitively, no matter from what angle it is viewed, and whether or not it is partly hidden by another object. However, this is not true for e.g. RADAR signatures.

Second, in the domain of non-electro-optic sensors such as RADAR there is a serious lack of labeled data sources. In order do effective machine learning, yielding an
25 accurate model, it is in many cases necessary to train the model on thousands, tens or hundreds of thousands, and even millions, of labelled instances of each class. While millions of optical/visual images are already available, publicly etc., for "a bus", or "a car ", such large volumes of data are not available for e.g. SAR images and GMTI signals.

Third, the human being is incapable of performing the labeling procedure without executing a fully-controlled experiment such that the fully ground-truth is known and controlled. Typically, labeling of non-EO data is done using methods such as the following: for e.g. SAR images, a sterile area or region is allocated, in which no

changes are expected to be made during the image capture, and in which foreign objects are not expected to enter or pass, and thus will not interfere with the imaging process. The targets or other objects to be classified are placed there, e.g. various types of tanks, cars and other vehicles, or non-vehicle objects. The objects are placed in a specific arrangement, so that they do not block or occlude each other when capturing the non-EO data. The RADAR or other sensor(s) is flown past the region and/or around it, on an airplane or similar platform, to obtain e.g. the RADAR signatures. The scenario is a very controlled one. This single static scenario, with one particular arrangement/configuration of vehicles/objects, yields perhaps hundreds of images. However, all images are from a single scenario, over a single background, which is typically not sufficient for training a robust model. The objects are manually labelled, based on their known locations, based on the very controlled conditions in which the objects were placed and the flights were performed.

For GMTI signals, in some cases the scenario is an even more complex one. A stretch of "sterile" road is allocated, on which it is known that other vehicles and objects will not appear or will rarely appear. A specific vehicle, e.g. a car of a specific model, is driven back and forth along this road. A Global Positioning System (GPS) or other location device and a dashboard camera are mounted on the vehicle during its travel. An airplane or other similar platform flies in the area, and captures GMTI signals. The position of the driven vehicle at every point of the capture scenario is known. The label or tag for the GMTI capture is that of the known vehicle. The location on the GMTI capture is matched to that of the GPS. Thus the GMTI signatures can be labeled. Again, a very large effort is expended, to capture a relatively small amount of training data for GMTI signatures.

In both cases, a comparatively large amount of personnel resources is expended, as well as resources of time and money. The payoff for this effort is comparatively low, and the cost per labelled data item can be very high. Thus, such a fully controlled experiment is a high cost high effort event, and eventually the amount of data derived from such an event is very small relative to the cost and effort invested, deriving only a relatively small number of events.

It becomes cost prohibitive to perform such labelling for a wide variety of scenarios, covering a wide variety of placements of the objects and a wide variety of angles of the object relative to the imaging sensor. Similarly, it is difficult, using such

methods, to obtain labels for a relatively large variety of objects and targets (natural objects, structures and various types of vehicles).

The presently disclosed subject matter enables the acquiring of large amount of data without the conduction of such fully controlled expensive experiments, but, instead, conducting many cheap un-controlled imaging occasions of scenarios of opportunity.

Having considered some disadvantages, problems and challenges with the current methods, attention is now reverted to **Fig. 1**. There is disclosed herein a method of labelling data, a computerized system comprising a processing circuitry which performs the method, and a non-transitory computer readable storage medium tangibly embodying a program of instructions which, when executed by a computerized system, cause the computerized system to perform the method.

The method comprises performing the following:

- (e) receiving unlabeled non-Electro-Optical (non-EO) data, the non-EO data comprising a set of non-EO signals/images captured during a first time period. The data is indicative of a first region **125**. In the presently disclosed subject matter, the term “captures” is used to refer to both images and signals;
- (f) receiving unlabeled EO data, captured during a second time period corresponding to the first time period. The unlabeled EO data comprises a plurality of video image frames e.g. captured by sensor **380**. The images are indicative of a second region **115** corresponding to the first region **125**; In some examples, also still EO frames are received, e.g. captured by sensor **370**;
- (g) classifying at least one second object **170** appearing on one or more of the plurality of video image frames. This thereby generates at least one video-image label. This thereby generates labelled EO data from the unlabeled EO data; and
- (h) labeling at least one first object **170** appearing in the unlabeled non-EO data with a corresponding video-image label of the at least one video-image label.

In some examples, the labelling is performed based at least on the following:

- (IV) a correspondence of at least one first time of capture, of at least one non-EO image of the set of non-EO captures/images, to at least one second time of capture, of at least one video image frame of the plurality of video image frames;
- 5 (V) a correspondence of positions of the at least one first object and the at least one classified second object; and
- (VI) a respective geographic position, associated with the at least one video image frame, of the at least one classified second object.

This method thereby generates labelled non-EO data from the unlabeled non-EO data. In some examples, the labelled non-Electro-Optical (EO) data is configured to facilitate machine learning of non-Electro-Optical images\signatures of first objects.

In some examples of the method, a platform **120**, e.g. an airborne vehicle **120**, is flown **128** around, over and/or past a region or area **125**, referred to herein also as first region **125**. In some examples, first airborne vehicle **120** comprises one of a first quadcopter, a first octa-copter, a first drone, and another remote-piloted vehicle (RPV) **120**. More details on vehicle **120** are disclosed further herein with reference to **Fig. 4**.

The platform **120** is referred to herein also as first platform or first airborne vehicle **120**. The platform comprises a non-EO sensor, which covers **123** and thus captures e.g. RADAR images or signals of region **125**. As indicated above, these images or signals/tracks are referred to herein, more generally, as captures. The capturing of the unlabeled non-EO data thus utilizes a non-EO sensor mounted on e.g. this first airborne vehicle. The imaging is performed over a period of time, referred to herein also as a first period. Each captured non-EO capture has a corresponding time tag, time stamp or other time indication, referred to herein also as a first time of capture.

The region is for example a parking lot, or an airport or airfield. Other examples include an army base, a port, forests and open areas. In the example, several objects are located in region **125**: a car or taxi **180**, which is in motion, and several stationary objects: building or other structure **150**, parked airplane **160** and parked bus **170**.

As was indicated before, in some examples the non-Electro-Optical (EO) sensor **480** comprises RADAR imaging sensor **480**. Thus, in some examples, the non-Electro-Optical (EO) data comprises RADAR imaging data. In some examples, the RADAR sensor comprises at least one of a Synthetic Aperture RADAR (SAR) sensor **480** and Ground Moving Target Indication (GMTI) sensor **480**. In some examples, the non-EO

sensor is configured for the setting of a capture mode to be used in the capture – e.g. a SAR capture mode and a GMTI capture mode.

In some examples, there are multiple GMTI tracks or captures, and each such GMTI capture comprises, or is indicative of, a range-Doppler map. Note that each such GMTI capture has its associated first time indication of non-EO capture.

The example of **Fig. 1** is that of capture of SAR imaging data.

In some examples, first platform **120** is relatively heavy, e.g. an eight-rotor octa-copter or other flying vehicle, in order to be able to carry the comparatively heavy non-EO sensor **480**, e.g. a SAR sensor.

In some examples, it is advantageous to capture the non-EO images and signals of region **125** from multiple angles, that is from various azimuths and elevations, as well as possibly seeing each object **170** with a different type of clutter in its vicinity. This can provide additional training data for the machine learning, out of the same session of image capture. In addition, RADAR signatures, for example, of a particular object **170**, appear different in images captured from different angles etc. Thus the platform **120** can in some cases fly or otherwise travel many kilometers, during the SAR capture. SAR images are typically captured while platform **120** is moving.

The capture from additional angles is illustrated in the figure by another view of the non-EO platform, shown with reference **120B**, flying or otherwise moving on an additional path **128B**.

In some examples of the method, another platform **110**, e.g. a second airborne vehicle **110**, captures images of another region or area **115**, referred to herein also as second region **115**. In some examples, second airborne vehicle **110** comprises one of a second quadcopter, a second opta-copter, a second drone **120** and a second other remotely piloted vehicle (RPV). ~~In a second thought, I really don't know.~~

As indicated before, in some examples the second platform **110** comprises one or more video EO imaging sensors **380**, for example a video camera. In some examples the second platform **110** comprises one or more still EO imaging sensors **370**, for example a still camera. In many typical examples the same sensor captures both video and still EO images.

These sensors "see" **113**, and thus capture, EO images of region **115**. Each image has a corresponding time tag, time stamp or other time indication, referred to herein also as a second time of capture (to distinguish it from the first capture time of

the non-EO sensor **480**). The second region **115** corresponds to the first region **125**, in that there is at least partial overlap between the two regions. In some cases, the second region comprises the first region. In some cases, the first region and the second region are the same.

5 In some examples, platform **110** stays stationary, and captures at least one still image, e.g. using still EO sensor **370**, while stationary. This can provide the example technical advantage, of enabling accurate geo-registration of this still image(s).

 In some examples, the platform **110** stays stationary also when capturing the video images using video sensor **380**. The term "stationary" in the presently disclosed
10 subject matter refers to the fact that the relevant vehicle/platform **110** hovers substantially in one place, and is not programmed to move from position to position so as to capture images from multiple positions. The platform in some cases moves only in that it is disturbed only by winds. In some examples, the movement of the relevant "stationary" platform **110** is limited to approximately in the range of centimeters, or
15 several meters.

 The capturing of the unlabeled EO data thus utilizes at least one EO sensor mounted on e.g. this second airborne vehicle. The imaging is performed over a period of time, referred to herein also as a second time period, which corresponds to the first time period. For example, sensors **380**, **370** can start capturing images a certain amount
20 of time before sensor **480**, but the first and second time periods overlap at least partially. In some cases, the second period comprises the first period. In some cases, the first period and the second period are the same. The overlap time period can enable time matching of video and non-EO frames, as disclosed further herein.

 In one implementation, the EO platform **110** remains stationary, and a still EO
25 image **740** is captured, and then video images **770** are captured throughout the second period during which the non-EO sensor **480** is capturing data/images/signals/tracks.

 More details on second vehicle **110** are disclosed further herein with reference to **Fig. 3**. In some examples, second vehicle **110** is a lighter vehicle than is vehicle **120**, e.g. a four-rotor quadcopter, since EO payloads/sensors are in some cases lighter than
30 e.g. SAR payloads.

 The terms first platform and first vehicle **120** and second platform and second vehicle **110** are used herein, to distinguish between the two platforms.

In some examples, vehicle **110** stands still (within defined tolerances) while performing the imaging **113**.

Methods for using the unlabeled EO data and non-EO data, to label the non-EO data, are disclosed further herein.

5 **Fig. 2** is a schematic illustration of a scenario **200** of capturing electro-optic (EO) and non-EO data, according to some embodiments of the presently disclosed subject matter. This figure differs from **Fig. 1**, in that the imaged regions are part of a highway or other road or street **250**, on which vehicles are moving. Shown are car **260**, bus **170** and taxi **180**, all moving **265**, **275**, **285**. The EO platform **110** images **213**
10 second region **215**, and non-EO platform **120**, **220** captures **223** data of first region **225**. In the non-limiting example of the figure, regions **215** and **225** are the same (while in the non-limiting example of **Fig. 1**, the corresponding regions **115** and **125** are not the same, although they do have partial overlap, and they do correspond).

15 This figure is relevant, in some examples, for GMTI tracks/signals capture, using a platform **220** configured for at least GMTI capture.

In some cases, non-EO platform **120**, **220** is in motion **245**, e.g. for SAR capture.

20 This may be required, for example, for SAR image capturing. In other cases, platform **220** is not in motion, that is it is substantially stationary, moving e.g. only a few centimeters or meters while hovering, e.g. due to the influence of winds. Such a stationary/hovering vehicle **220** can be used to capture GMTI signals.

25 In still other example, not illustrated herein, the imaged regions **115**, **125**, **215**, **225** include both stationary objects such as in a parking lot, as well as part of a road. In such a case, one scenario of imaging, performed by the three platforms **110**, **120**, and **220** can capture data for both SAR (captured by **120**) and GMTI (captured by **220**). In still another scenario, two platforms **110** and **120** are used. In one stage of the process, non-EO platform **120** is set to SAR mode, and moves during capture, while in another stage the same platform **120** is set to GMTI mode, and hovers during capture.

30 The method disclosed herein can label the non-EO data in a comparatively efficient, inexpensive, automated and quick manner, for a large number of variables and classifications, and for large numbers of non-EO signals and images. In addition, the solution facilitates use of the artificial intelligence / deep learning capabilities of the EO domain, which are comparatively more developed, and with large volumes of data already available for use, to label non-EO domains such as GMTI and SAR.

Attention is now drawn to **Fig. 3**, schematically illustrating a non-limiting example of a generalized schematic diagram of a vehicle **110** comprising computerized EO imaging system **310**, in accordance with some embodiments of the presently disclosed subject matter. Second vehicle or platform **110** in some examples is
5 configured to capture EO images.

Vehicle **110** comprises one or more electro-optical (EO) sensor(s) – e.g. EO video sensor **380** and EO stills sensor **370** -- configured e.g. to perform the imaging disclosed above with reference to **Figs. 1** and **2**. In some examples, sensors **370** and **380** are combined, in that one single EO sensor (e.g. **370**) captures both still and video
10 images.

In some examples, platform **110** also comprises computerized EO imaging system **310**. In some examples, EO sensors **380**, **370** are operatively coupled to imaging system **310**.

In some non-limiting examples, computerized EO imaging system **310** includes
15 a computer. It way, by way of non-limiting example, comprise a processing circuitry **315**. This processing circuitry may comprise a processor **330** and a memory **350**. This processing circuitry **315** may be, in some non-limiting examples, general-purpose computers specially configured for the desired purpose by a computer program stored in a non-transitive computer-readable storage medium. They may be configured to
20 execute several functional modules in accordance with computer-readable instructions. One non-limiting example is a personal computer (PC).

In other non-limiting examples, this processing circuitry **315** may be a computer(s) specially constructed for the desired purposes.

Processor **330** may comprise, in some examples, at least one or more functional
25 modules. In some examples it may perform at least functions, such as those disclosed herein with reference to **Figs. 1, 2** and **6-13**.

In some examples, processor **330** comprises communications system module **333**. In some examples, this module is configured to receive EO data from one or more EO sensors **370**, **380**, and to transmit them to non-EO labeling system **510**.

30 In some examples, memory **350** of processing circuitry **315** is configured to store data at least associated with the image capture and transfer.

Attention is now drawn to **Fig. 4**, schematically illustrating a non-limiting example of a generalized schematic diagram of a vehicle **120**, **220** comprising

computerized non-EO imaging system **410**, in accordance with some embodiments of the presently disclosed subject matter. First vehicle or platform **120** in some examples is configured to capture non-EO images.

Vehicle **120, 220** comprises one or more non-electro-optical (EO) sensor(s) **480**,
5 configured e.g. to perform the imaging disclosed above with reference to **Figs. 1** and **2**.

In some examples, platform **120** also comprises computerized non-EO imaging system **410**. In some examples, non-EO sensors **480** are operatively coupled to imaging system **410**.

In some non-limiting examples, computerized non-EO imaging system **410**
10 includes a computer. It way, by way of non-limiting example, comprise a processing circuitry **415**. This processing circuitry may comprise a processor **430** and a memory **450**. This processing circuitry **415** may be, in some non-limiting examples, general-purpose computers specially configured for the desired purpose by a computer program stored in a non-transitive computer-readable storage medium. They may be configured
15 to execute several functional modules in accordance with computer-readable instructions. In other non-limiting examples, this processing circuitry **415** may be a computer(s) specially constructed for the desired purposes.

Processor **430** may comprise, in some examples, at least one or more functional modules. In some examples it may perform at least functions, such as those disclosed
20 herein with reference to **Figs. 1, 2** and **6-13**.

In some examples, processor **430** comprises communications system module **433**. In some examples, this module is configured to receive non-EO data from one or more non-EO sensors **480**, and to transmit them to non-EO labeling system **510**.

In some examples, processor **430** comprises RADAR mode selector module
25 **433**. In some examples, this module is configured to select the mode in which the non-EO sensor **480** will perform imaging, if the sensor **480** is capable of imaging in multiple modes, e.g. RADAR and GMTI. This example module is relevant in some cases in which the non-EO imaging **123** is RADAR imaging **123**.

In some examples, memory **450** of processing circuitry **415** is configured to
30 store data at least associated with the non-EO image capture and transfer.

Attention is now drawn to **Fig. 5**, schematically illustrating a non-limiting example of a generalized schematic diagram of a system **505** comprising computerized non-electro-optic (EO) labeling system **510**, in accordance with some embodiments of

the presently disclosed subject matter. The system **505** provides an example of an overall solution, comprising the labeling system, as well as the image capturing platforms, relevant data stores and data sources, and machine learning systems.

5 Solution system **505** comprises, platforms **110**, **120**, **220**, examples which are disclosed above with reference to **Figs. 3** and **4**.

In some examples, solution **505** comprises orthophoto data system **580**, and/or Digital Terrain Elevation Data (DTED) system **580**.

10 In some examples, solution **505** comprises EO machine learning system **590**. In some implementations this system **590** utilizes known per se machine learning architectures, methods and techniques.

In some examples, solution **505** comprises a non-EO signature database (DB) **595**. This DB **595** can be used to train non-EO models using an (e.g. external) non-EO machine learning system. Labelled signatures can be store in this DB.

15 Example functions of systems **580**, **590** and **595** are disclosed further herein with reference to flow chart **Figs. 6**.

In some examples, solution **505** comprises a computerized non-EO labeling system **510**. In some non-limiting examples, computerized non-EO labeling system **510** includes a computer. It way, by way of non-limiting example, comprise a processing circuitry **515**. This processing circuitry may comprise a processor **530** and a memory **520**. This processing circuitry **515** may be, in some non-limiting examples, general-purpose computers specially configured for the desired purpose by a computer program stored in a non-transitive computer-readable storage medium. They may be configured to execute several functional modules in accordance with computer-readable instructions. In other non-limiting examples, this processing circuitry **515** may be a computer(s) specially constructed for the desired purposes.

20

25

Processor **530** may comprise, in some examples, at least one or more functional modules. In some examples it may perform at least functions, such as those disclosed herein with reference to **Figs. 6-14**.

30 In some examples, processor **530** comprises EO object detector module **532**. In some examples, this module is configured to detect objects in the EO images. Examples of how this is performed are disclosed with reference to **Fig. 6**, and/or using known per se methods.

In some examples, processor **530** comprises EO object labeler module **542**. In some examples, this module is configured to label detected objects in the EO images. Examples of how this is performed are disclosed with reference to **Fig. 6**, and/or using known per se methods.

5 In some examples, processor **530** comprises non-EO object labeler module **544**. In some examples, this module is configured to label localized objects in the non-EO images. Examples of how this is performed are disclosed with reference to **Fig. 6**.

In some examples, processor **530** comprises Graphical User Interface (GUI) module **539**. In some examples, this module is configured to allow a human user to
10 control the labeling process, e.g. entering which data will be used. An example GUI provided by this module is disclosed further herein with reference to **Fig. 14**.

In some examples, processor **530** comprises cross-registration module **550**. In some examples, this module is configured to register video images to the still EO image frame(s).

15 In some examples, processor **530** comprises video-object geographic-position calculation module **556**. In some examples, this module is configured to determine the geographic position of objects identified in video image frames. In some implementations, this is performed for selected video image frames, as will be disclosed herein.

20 In some examples, processor **530** comprises localization module **555**. In some examples, this module is configured to localize of objects detected in EO images (e.g. video images) on non-EO captures (images or signals/maps).

In some examples, processor **530** comprises geo-registration module **557**. In some examples, this module is configured to geo-register still images, e.g. per block
25 **640** of **Fig. 6**.

In some examples, processor **530** comprises tracking module **547**, referred to herein also as tracker module or tracker **547**. In some examples, this module is configured to perform tracking of video images **810**, **812**, **814**, **816**, e.g. as disclosed further herein with reference to block **646** and **Fig. 8**.

30 In some examples, processor **530** comprises time matching module **547**. In some examples, this module is configured to match first times of capture of non-EO images **1050**, **1250**, and second times of capture of video EO images **770**.

In some examples, processor **530** comprises orchestrator module **537**. In some examples, this module is configured to orchestrate and control the processed performed by the other modules of process **530**.

5 In some examples, processor **530** comprises input/output module **560**. In some examples, this module is configured to interface with the systems external to system **510**, e.g. via communications interface **570**.

10 In some examples, memory **520** of processing circuitry **515** is configured to store data at least associated with processes such as position and time matching, matching criteria parameters, lists of which video and non-EO images correspond to each other etc.

In some examples, computerized non-EO labeling system **510** includes communication interface **570**. This interface can be used to interface to vehicles/platforms **110**, **120**, **220**, and to e.g. systems **580**, **590**, and **595**.

15 In some examples, computerized non-EO labeling system **510** includes data store **575**. This can be used, for example, to store data which is more persistent than that stored in memory **520**. Non-limiting examples of such data include ground truth data set **990** and interim ground truth data set **900**. In some examples, database **595** can reside in data store **575**.

20 The example of **Fig. 5** is non-limiting. In other examples, other divisions of data storage between storage **575** and memory **520** may exist.

Fig. 5 illustrates only a general schematic of the system architecture, describing, by way of non-limiting example, certain aspects of the presently disclosed subject matter in an informative manner, merely for clarity of explanation. It will be understood that the teachings of the presently disclosed subject matter are not bound by what is described with reference to **Fig. 5**.

25 Only certain components are shown, as need to exemplify the presently disclosed subject matter. Other components and sub-components, not shown, may exist. Systems such those described with reference to **Fig. 5** may be capable of performing all, some, or part of the methods disclosed herein.

30 Each system component and module in **Fig. 5** can be made up of any combination of software, hardware and/or firmware, as relevant, executed on suitable device or devices, which performed the functions as defined and explained herein. The hardware can be digital and/or analog. Equivalent and/or modified functionality, as

described with respect to each component and module, can be consolidated or divided in any other manner. Thus, in some embodiments of the presently disclosed subject matter, the system may include fewer, more, modified and/or different components, modules and functions than those shown in **Fig. 5**. To provide one non-limiting example of this, in some examples input/output module **560** is replaced by separate input and output modules. Similarly, in some examples video-object geographic-position calculation module **556** and localization module **555** are combined into one module.

One or more of these components and modules can be centralized in one location, or dispersed and distributed over more than one locations, as is relevant.

Each component of **Fig. 5** may represent a plurality of the particular component, possibly in a distributed architecture, which are adapted to independently and/or cooperatively operate to process various data and electrical inputs, and for enabling operations related to optical detection and localization. In some cases, multiple instances of component may be utilized for reasons of performance, redundancy and/or availability. Similarly, in some cases, multiple instances of a component may be utilized for reasons of functionality or application. For example, different portions of the particular functionality may be placed in different instance of the component.

Communication between the various components of the systems of **Fig. 5**, in case where they are not located entirely in one location or in one physical components, can be realized by any signaling system or communication components, modules, protocols, software languages and drive signals, and can be wired and/or wireless, as appropriate. The same applies to interfaces such as **560**, **570**.

Attention is drawn to **Fig. 6**, schematically illustrating a non-limiting example of a generalized flow diagram for labeling of non-EO images, in accordance with some embodiments of the presently disclosed subject matter. Flow **600** starts with block **605**.

According to some examples, systems are provided to capture EO data (block **605**). For example, second airborne vehicle **110** is provided.

According to some examples, unlabeled EO data is captured (block **610**). In some examples, airborne vehicle **110** and EO sensors **380**, **370** are utilized. Examples of such a capture are disclosed with reference to **Figs. 1** and **2**.

According to some examples, the captured unlabeled EO data is received (block **620**). In some examples, this block utilizes communications system **333** on second vehicle **110**, transferring the data to non-EO labeling system **510**. The transfer can also

utilize interface components **570, 560**. System **510** thus receives the data. This block serves as an input **A** to blocks **642, 646** and **640** on this figure, and as an input **G** to block **655** on Fig. **6B**, as well as being an input to block **628**.

According to some examples, systems are provided to capture non-EO data
5 (block **607**). For example, first airborne vehicle **120, 220** is provided.

According to some examples, unlabeled non-EO data is captured (block **617**). In some examples, airborne vehicle **120, 220** and non-EO sensor **480** are utilized. Examples of such a capture are disclosed with reference to **Figs. 1** and **2**.

According to some examples, the captured unlabeled non-EO data is received
10 (block **623**). In some examples, this block utilizes communications system **433** on vehicle **120, 220**, transferring the data to non-EO labeling system **510**. The transfer can also utilize interface components **570, 560**. System **510** thus receives the data.

Note that the sequences **605, 620** and **607, 623** are not necessarily synchronized with each other. This is illustrated with them appearing on two parallel paths. However,
15 as disclosed for example with reference to **Fig. 1**, the non-EO and EO captures are in first and second periods, which should correspond to each other, e.g. overlap with each other.

In some examples, the data to be labeled is chosen (block **625**). In some examples, this block utilizes GUI module **539** of processor **520** of processing circuitry
20 **515**. A human user (not shown in **Fig. 5**) can use a computer with the GUI (not shown in **Fig. 5**) to make this choice. The computer screen can be part of labeling system **510**, or on an external computer operatively coupled to system **510** with the communication interface **570**.

To provide more illustration of this step, attention is drawn to **Fig. 14**,
25 schematically illustrating a non-limiting example of a generalized Graphical User Interface, in accordance with some embodiments of the presently disclosed subject matter. Example Graphical User Interface (GUI) **1400** can appear on a computer screen. In a simplified illustrative situation, the user uses a selection mechanism, e.g. radio buttons **1430, 1440**, to choose whether the captured non-EO data is of type SAR or
30 GMTI, e.g. in which mode the vehicle **120** operated during the particular capture.

The path entry area **1410** can, for example, be where the user, using a mouse and keyboard (not shown in **Fig. 5**) to type in the path, folder and/or filename in which the non-EO data (e.g. RADAR, as shown) resides for the relevant capture session. In

other cases, a drop down menu lets the user make the choice. Other known per se methods can be used.

The second path entry area **1420** can, for example, be where the user, using a mouse and keyboard to type in the path, folder and/or filename in which the EO data resides for the relevant capture session. In other cases, a drop down menu lets the user make the choice. Other known per se methods can be used.

The start button, soft-key etc. **1470** can be pressed or otherwise selected, to start the process of system **510** in processing the EO and non-EO data so as to generate the non-EO data labels.

Reverting now to **Fig. 6**, in some examples, times of capture of video image frames, and non-EO images/signals, are matched (block **628**). Based on this match, video frames to detect/classify/register are identified. In some examples, this block utilizes time matching module **547**. The match is based on the two types of frames meeting one or more time-matching criteria. In some examples, the criterion or criteria comprise at least one of the following:

- c. the second time of capture, of the video image **1010**, is equal to the first time of capture, of the corresponding non-EO image **1050**; and
- d. the second time of capture is closer to the first time of capture than other second times of capture of the plurality of video image frames – or as close to the first time of capture are as other second times of capture. That is, if the first time of capture, of non-EO data $t_{\text{non-EO-1}}$, falls in the time interval T_{vid} between two second times of capture of video frames t_{v1} and t_{v2} , the difference between $t_{\text{non-EO-1}}$ and its closest video frame t_{v1} should be a maximum of half of a frame time, that is a maximum of $T_{\text{vid}}/2$.

Note that $T_{\text{vid}} = 1 / \text{fps}$, where fps = frames per second.

If such a match is made for each captured non-EO image or signal **1050**, the method of **Fig. 6** can in some examples facilitate labelling all captured non-EO captures **1050** - thus providing as large a labeled non-EO training set as possible.

In some examples, this time matching is an input **E** to block **655** on **Fig. 6B**.

Note that, although block **628** can be performed after block **625**, once EO and non-EO data are available. However, it can be performed at various later stages, as long as it is performed prior to blocks **655** and/or **660**.

One example calculation of the relevant video frame is as follows:

5 $T_n = T_o + n \times (\text{frames per second rate}) \qquad (1)$

Where T_n is the time of capture of e.g. SAR image **1050**, and T_o is the start time of the video capture by vehicle **110**. Manipulating this equation gives:

$$n = (T_n - T_o) / (\text{fps rate})$$

10 where n is the index of the video frame that matches the SAR image captured at first time T_n .

According to some examples, a geo-registration of at least one still image is performed (block **640**). This is performed, in some implementations, using geo-registration module **557**, on e.g. one or more still images captured by still EO sensor **370**. In some examples, known per se geo-registration methods and algorithms are
15 utilized.

For example, orthophoto infrastructure can be used **580**, as an input to the method. In some implementations, each pixel in the infrastructure has coordinates assigned. Features in the still photo (e.g. structures, trees, roads/streets, peaks etc.) can be matched to that in the orthophoto infrastructure. Note that in some examples, the
20 more transient features in the still images, e.g. the vehicles, are not matched to the orthophoto infrastructure.

This infrastructure typically has accurate position/coordinate data for each such feature. Accurate information for elevation can also be provided, e.g. from Digital Terrain Elevation Data (DTED) **580**, or more generally Digital Elevation Models (DEM) such as a Digital Terrain Model (DTM) **580**.
25

To provide more illustration of this step, attention is drawn to **Fig. 7**, schematically illustrating a non-limiting example of a geo-registration, in accordance with some embodiments of the presently disclosed subject matter. Illustration **700** depicts, in its left portion, a conceptual schematic view of the earth **710**. Orthophoto infrastructure **580** shows, for a particular building, bridge, or other feature **720**, that the absolute geographic position of coordinates x , y , z . The figure illustrates the geo-registering of a corresponding feature **750** on still image **740**. Feature **750** is determined to be the same feature as feature **720** in the orthophoto infrastructure. Feature **750** has
30

relative positions i-still, j-still within the image **740**. Based on the geo-registration, the feature is assigned the absolute geographic position **752**, having the same coordinates x, y, z as point **720**.

5 In real examples, multiple features **750** are matched to features **720** on the orthophoto infrastructure.

This accurate position data can be used to assign accurate position information to each pixel on the still image, in some cases with an accuracy of 2-5 pixels.

The same method can apply to all pixels on the still frame **740**.

10 Reverting now to **Fig. 6**, in some examples, pixel coordinates are transformed to geographic coordinates (block **641**). This block is performed, in some implementations, using EO object detection and classification module **532**. Thus, each pixel i-still, j-still on still image **740** of **Fig. 7** is converted to geographic (absolute) coordinates x, y, z, making use of the geo-registration of features in still image **740** which was performed in block **640**.

15 In some examples, based on the geo-registration of the features **750**, and on relative positions of objects **170**, **180** within the still image **740**, absolute coordinates can be determined for these objects as well. For example, the absolute coordinates x-170, y-170, z-170 of object **170** are determined. In some examples, a data structure, e.g. a list or table of detected objects and their absolute geographic positions/coordinates, is
20 stored, e.g. in memory **520** or in data store **570**.

As will be disclosed further herein, in cases where block **640** is performed, the labelling of the first object(s) in the non-EO images or signals, based at least on the respective geographic position of the classified second object(s), can be performed based at least on the geo-registration.

25 Also as disclosed further herein, steps **640** is **641** are optional. Other implementations do not use it. However, utilizing a still image with geo-registration can in some cases give more accurate positioning of objects captured in the video images, using block **655** (register video frames to stills), disclosed further herein.

30 In some examples, one or more second objects **170**, **180** are detected in the video frames **770** (block **642**). This block is performed, in some implementations, using EO object detection and classification module **532**. It is performed in some cases using known per se methods. In some examples, this makes use of EO Machine Learning

system **590** of system **500**, or by an external ML system. These ML systems can interface to labeling system **510** using, for example, interfaces **570**, **560**.

Note that in the example of the figure, blocks **642**, **646**, **648**, performed for video images, are shown occurring parallel to blocks **640**, **641** performed for the still image, where each of the parallel flows depend **A** on block **620**. In other example implementations, these blocks could instead be performed serially with each other.

To provide more illustration of this step, attention is drawn to **Fig. 8**, schematically illustrating a non-limiting example of EO object detection, tracking, and classification, in accordance with some embodiments of the presently disclosed subject matter. Illustration **800** depicts, in its left portion, a schematic view of image **810**, captured at time T-23. The objects **830** and **820**, a car and a tree, are detected.

In some examples, one or more of the second objects **170**, appearing in the video frames **770**, are classified (block **648**). Classifications can be performed frame by frame. The classifications are used to facilitate labelling of the corresponding second object. This is performed, in some implementations, using EO object labeler module **542**. It is performed in some cases using known per se methods.

Note that this step can provide at least some example advantages, in that classification of video images is a well-developed field, and thus can provide accurate classification. For example, a considerable amount of open-source solutions exist for classification of objects in videos. In some examples, machine learning, e.g. deep learning networks, are utilized to classify various types of vehicles, trees, structures of different types (bridges, buildings, tower etc.) etc. In some implementations, this is performed utilizing EO Machine Learning (ML) system **590**, of system **500**, or by an external ML system. These ML systems can interface to labeling system **510** using, for example, interfaces **570**, **560**. Machine learning classification is an example of performing classification in an automated manner. Thus, the classification of second objects **170** can be performed in some cases at least partly in an automated manner.

In other examples, the classification is performed by comparing the video frame **770**, **810** to example pictures of similar objects.

In still other examples, the classification of the second objects is performed at least partially manually. For example, a human can look at the video image, and input that object **830** is an automobile, and object **820** is a maple tree.

In some examples, the same machine learning networks usable for classification in block **648** are also used for object detection in block **642**.

Thus, for example, reverting again to **Fig. 8**, the objects **830**, **832**, **834**, **836** are all labeled as "car" or "car A", while the objects **820** of video image **810**, and its
5 corresponding objects also in the other video image frames **812**, **814**, **816**, are all labelled "tree", or "oak tree".

Classification using such networks in some examples provides accuracy of about 90%. Using also the tracking methods of optional block **646**, described next, the classification process is assisted by the tracking process, and the accuracy of the
10 classification can, in some cases, be increased above such levels.

In some examples, the classifying of the second objects comprises performing classification in each video image frame of the plurality of video image frames. In other examples, as indicated further herein, the block of **628** can facilitate classification of second objects, only in that sub-set of video frames which meet one or more time-
15 matching criteria relative to corresponding captured non-EO images.

In some examples, tracking at least one second object **830** is tracked across video image frames **810**, **814** (block **646**). In some examples, this step is performed utilizing EO object labeler module **542**, part of which functions as a tracker. In some other examples, there is a separate tracking module **545**, which is utilized for this task,
20 and functions as the tracker.

This block **646** utilizes as inputs the classifications **647** performed in block **648**, as well as the unlabeled EO data **645** provided **A** by block **620**.

Reverting again to **Fig. 8**, a non-limiting example of tracking is illustrated. The frames **810**, **812**, **814**, **816** are added to the tracker module **545**. The tracker compares
25 the frames, and looks for similar objects repeating in each. In the example of the figure, the tracker notes that objects **830**, **832**, **834**, **836** are visually similar, and that their changed position from frame to frame indicate that the object is moving during the video capture. Each such object was classified in block **648**.

This can provide at least two example advantages. If the classification of each
30 frame (in block **648**) gives different results, for the same object across multiple frames, the tracking can increase the accuracy of the classification. Consider a case in which objects **830**, **834**, **836** are all classified as "car", while object **832** is mistakenly classified as "bus". Since the tracker **545** knows that object **832** is the same as the other three, it

can (in block **648**) correct the classification of **832**, and determine that it is in fact a car, and not a bus. In another case, with a larger number of frames, the object is classified 80% or 90% as "car", and 20% or 10% as "bus". These numeric examples are non-limiting.

5 A second example advantage is to help classify, accurately, partly occluded objects. Consider video frame **814**, in which the car **834** appears behind the tree **820**, and is partly blocked/hidden by the tree. It may be hard for e.g. an automated system to classify object **834**. However, the tracker **545** has determined that **834** is the same object as **830**, **832** and **836**, and is in movement. If these other objects have been determined
10 to e.g. be "car", the labeler module **542** (for example) can then assign object **834** the classification of "car".

 In some examples, such a step can facilitate an improved classification of the at second object **834** in at least one video image frame **814**, as compared to a second classification, one performed without tracking. Again, note that tracking is an optional
15 process to the method, and it can be skipped, at the cost of somewhat decreased classification accuracy.

 A third example advantage is to assist in the determination and population of movement indicators **960**. More on this is disclosed further herein, with reference to block **960**.

20 The flow of **Fig. 6A** proceeds **B** to **Fig. 6B**, to block **650**.

 In some examples, one or more of the second objects **170**, appearing in the video frames **770**, are labelled, based on the classification (block **650**). One or more video-image labels are generated. For example, to each classified second object **830**, a respective video-image label, indicative of the respective classification, is assigned.

25 Labelling can be performed frame by frame. This is performed, in some implementations, using EO object labeler module **542**. This block can generate, in some implementations, an interim ground truth data set **900**. In some examples, this ground truth data set is stored in memory **520**, and/or in data store **570**.

 To provide more illustration of this step, attention is drawn to **Fig. 9A**,
30 schematically illustrating a non-limiting example of an interim-ground truth data set, in accordance with some embodiments of the presently disclosed subject matter. In some examples, this interim ground-truth data set assigns, to each classified second object, a

video-image label indicative of the classification. Each data item of the interim ground-truth data set **900** comprises:

- (v) the indication **920** of a time of capture of the respective video-image frame, e.g. a time tag or time stamp in Coordinated Universal Time (UTC); and
- (vi) an indication **930**, **935** of a respective second relative position within the respective video-image frame, of each classified second object corresponding to the time of capture – for example columns i-vid and j-vid, corresponding to relative coordinates i-video and j-video in video image **770** of **Fig. 7**.

In some non-limiting examples, the data set is a table, as in **Fig. 9A**, or a list or other data structure. In the example table **900**, the field "record" **910** is a record number.

In the example of the figure, there is a single table, showing various times of capture, and the objects detected and classified at each time. In another example implementation, each time of video capture has its own associated table, and in each table are those objects detected at that time. These are non-limiting examples, for exposition purposes only.

Note that in the example table, the same video frame **770** can have more than once associated record. For example, the video frame capture at time T-3 contains both a taxi and a bus, and thus the frame of T-3 has two associated records.

In some implementations, the indication of the respective relative second position within the respective video-image frame comprises relative coordinates within the respective video-image frame. For example, "(0.315, 0.723)" shown in columns **930**, **935**, for the first record **915**, denotes that this object lies in normalized pixel-coordinates ($i=0.315*nor$ $j=0.723*noc$), where *nor* is the number of rows in the frame and *noc* is the number of columns in the frame. This methodology of indicating coordinates enables changing the scale of the image (which is very common in the deep-learning field), without the necessity of modifying the coordinates of the objects detected within the video image frames.

Reverting to **Fig. 6B**, in some examples, a movement indicator is added (block **653**). This can be added to the interim ground truth data set. This can be performed, for example, by EO object labeler module **542**, or by some other component. The

indication, whether the classified second object is moving, can be determined based on tracking the classified second object across video image frames, e.g. as shown in **Fig. 8**. In some examples, in which a tracker **545** is used per block **646**, the system can know whether an object is stationary or moving. If an object does not move in a substantial manner across the frame, from frame to frame (e.g. more than a defined number of pixels), it is considered stationary. Thus, in **Fig. 8**, tree **820** is stationary, while vehicle **820, 832, 834, 836** is moving. The same applies to moving vehicle **180** and stationary vehicles **160, 170** in **Fig. 1**. This information can be recorded/stored in the interim table **900**, as "movement indicator" **960**. In the example, the indications are "moving/stationary". In other examples, the indications could be of the type "yes/no" for the status "moving".

As will be seen further herein, this indication can be used for selecting the relevant objects related to e.g. the RADAR mode of operation: in SAR mode only the stationary objects will be selected, and in GMTI mode only the moving objects will be selected.

A purpose of a ground truth table or other ground truth data set is to provide information about the "what, where, when" (a "W³ database") – which/what classes of objects, appear in what position, at each capture time in the video capture process. In the example table **900**, each row or other record represents a particular object at a particular capture time – for example the taxi captured at the i-video relative position 0.315 at time T-3, and having moved to the i-video relative position 0.317 by T-4. For the stationary bus, the table shows that at both T-3 and T-4, its positions **930, 935** within video frames are the same: (0.678, 0.214).

Typically, the generating the interim ground-truth data set is performed for substantially all, or all, of the second video image frames **770**. This will enable the correct determination of the movement indicator **960**, especially if tracking **646** is performed. This indicator **960** is in many cases important for the process of labeling of non-EO data. As disclosed further herein with reference to block **660**, portions of the non-EO data are matched, based on position, to objects seen in the video. Considering, for example, the case of SAR data, if a video image is captured of a moving **180** car **185**, and the car is not labeled as moving, the system might, having seen this car in a video frame, attempt to label SAR data in a corresponding SAR image. Since moving objects are not seen in SAR images, the label "car" will be applied to a portion of the

SAR image that does not in fact contain data corresponding to the presence of a car. Such a process will cause incorrect labeling, and incorrect data will thus be fed into the non-EO machine learning, resulting in erroneous ML.

Therefore, typically, a tracker (e.g. a tracking module **545**) is not needed, and
5 the use of the movement indicator **960** is not needed, ONLY if the imaging is performed in an area where it is known *a priori* that all objects are stationary (for the case of SAR imaging). In GMTI mode, by contrast, the tracker is still required in order to estimate object velocity and therefore its location in the GMTI signal.

Note that in many implementations, blocks **950** and **953** are combined.

10 As indicated above, the matching of times of capture of video image frames and non-EO data, disclosed with reference to block **628**, can serve as an input **E** to both block **655** and block **660**. Both of these blocks are now disclosed. In some examples, they can be performed in parallel.

In some examples, video image frames are registered to the still EO image
15 frame(s) (block **655**). This can be performed, for example, by cross-registration module **550**. In some cases, the indication of respective geographic position **940**, **942**, **944** is derived from the respective second relative position **930**, **935**, based at least on the geo-registration of the at least one still image performed in block **640**. The deriving the indication of respective geographic position **940**, **942**, **944** in some implementations
20 comprises registering the respective video-image frame **770** to a corresponding still EO image **740**. As a result, in some cases, each pixel of the video image frame **770** is thus matched to a corresponding pixel on the still image frame **740**. E.g. a table of points is created.

Note that the matching/mapping, performed in block **655** of this flow, is a pixel
25 to pixel mapping, between frames **740** and **770**. However, the geographic information associated with each pixel on video frame **770** is not performed in block **655**, in this example implementation. Rather, in the example of flow chart **600**, geographic positions associated with the pixels of video frame **770** are determined in block **635**, disclosed further herein.

30 This registration of block **655** is not geo-registration. In many cases, each video frame will be very similar to the still frame. Consider, for example, the case in which quadcopter **110** is hovering stationary during video capture, and (as illustrated in **Fig. 1**) in which few of the vehicles **180** are moving. Similarly, the video frame **770** and still

frame **740** are captured perhaps several minutes, or an hour or two apart, while the orthophoto and DTED databases **580** typically have data based on images captured years ago.

An example of the registration process of block **655** is illustrated in **Fig. 7**.

5 For example, the objects **750** and **780** are matched, using a similarity criterion, e.g. as having the same visual appearance, or very similar (within a defined amount), visual appearance. The absolute geographic position **x, y, z**, previously assigned **752** to object **730, 750**, can thus be assigned **782** to object **780** as well.

10 Note that associated with the video frames themselves there is only an approximation of the geographic location of each second object, based e.g. on the GPS location of the platform **110** at that moment, and on angles at that moment of the video camera/sensor **380**. Thus the registration of video frames to still frames can improve the accuracy of the geographic positioning of the associated second classified objects.

15 As was indicated, in many cases, for a particular EO sensor **380**, the still images are of a higher resolution than the video images captured by that same sensor. Thus, it can be advantageous that the resource-intensive process of geo-registration (in block **640**) should be performed specifically on a still image. Block **655** thus enables the video frame **770** to inherit the position accuracy of the geo-registered still frame(s) **740**.

20 The input **E** to block **655** illustrates that in some example implementations, block **655** is performed responsive to performance of block **628**. In such a case, the video-to-still-EO registration is performed only for those video frames **770** for which the second video capture time and a first non-EO image capture time match one or more time matching criteria. This can provide the example technical advantage of performing the video-still registration process, which in some cases is resource-intensive, for a relatively small portion of the video frames, that is for only those needed for the labeling
25 of the non-EO images or signals **1050**. For example, in some cases SAR images are capture every 10 seconds, while video capture is done at 50 frames per second (50fps). At 50 fps, 500 frames are captured between two SAR captures. In this example, we are required to perform registration for only 1 out of every 500 video frames, thus making
30 the registration process hundreds of time faster.

Similarly, if a tracker **545** is not being used to enhance performance of classification, and determination of movement indicators **960**, the same resource and time savings can apply also to performing block **628** before e.g. blocks **642, 648, 650**

etc. That is, if only a small portion of the video frames need to be detected, classified, labelled, and/or stored to the data set, there is an increase in efficiency. This can also decrease the storage needs of the interim ground truth data sets **900**. Again, this will typically be done only if it is known a priori that all vehicles are stationary (for the SAR case).

According to some examples, one or more data items **915**, **937** of the interim ground-truth data set **900** are selected (block **660**). This can be performed, for example, by time matching module **547**. The selected one or more data items comprise one or more selected classified second objects **170**, **160**.

As indicated above, in some examples this block is preceded **E** by block **628**, in which it was determined that a corresponding second time T-3 of capture of the respective video-image frame **812**, and a first time of capture of a non-EO capture/image **1050**, meet a time matching criterion. In such a case, in block **660** the system **510** selects one or more data items of the interim ground-truth data set that comprise an indication of the corresponding second time of capture of the respective video-image frame **812**.

In some examples, the time-matching criterion comprises at least one of the following:

- a. the second time T-3000 of capture, of a video frame, is equal to the first time of capture, e.g. T-100, of a non-EO image/capture; and
- b. the second time T-3000 of capture is closer to the first time of capture than other second times of capture T-3001 of the plurality of video image frames, and the difference in time is equal to or less than $(\leq) \frac{1}{2} * T_{\text{vid}}$. (Recall that $T_{\text{vid}} = 1 / \text{fps}$, where fps = frames per second.)

Thus, in one example of case "b" above, a video frame rate of 50 fps, that is an interval $T_{\text{vid}} = 0.02$ sec, is assumed. In the example, there is no non-EO timestamp, which exactly matches EO capture timestamp T-3000. However, there is a non-EO image with timestamp T-100, which differs from T-3000 by 0.005 seconds, which is less than $\frac{1}{2} * 0.02 = 0.01$ sec. By contrast, the nearest other second capture times T-3001 differs from T-100 by at least 0.015 seconds.

The selected frame **812** is used in blocks **655** and **635**. By selecting only certain frames for which blocks **655** and **635** are performed, the performance of the system can be greatly improved, as disclosed further herein.

5 In some examples, not shown in the figure, blocks **660** and **655** can be performed sequentially rather than in parallel, e.g. performing **655** after **660**.

In some examples, a ground truth data set is derived (block **635**). The ground-truth data set assigns, to each classified second object **170**, a video-image label indicative of the classification. In the non-limiting example of the figure, the ground-truth data set **990** is derived from the interim ground-truth data set **900**. This can be
10 performed, for example, by EO object labeler module **542**, and by video-object geographic-position calculation module **556**.

As indicated by the input **E** to blocks **655** and **660**, in many implementations, in which tracking is performed (per block **646**), generating the ground-truth data set **990** is performed only for respective second video image frames, for which the second
15 time of capture **T-3000**, and the first time of capture of a corresponding non-EO signal/image **1050** of the set of non-EO signals/images, meet a time-matching criterion. Since only the video frames that match the non-EO captures are registered to the still image(s), and are selected from interim ground truth data set **900**, the ground truth data set **990** can be much reduced in size, compared to a case in which every single video
20 frame is recorded in the data set **990** (and in some cases, in data set **900**). This can also reduce resource usage, and can in some cases speed up the process.

Recall that if tracking is not required, e.g. when a SAR capture is made and it is known that all vehicles are stationary, it is not strictly required that block **635** be performed only on selected records of interim ground truth data set **900**. However, even
25 if no tracking is performed, the above-mentioned performance improvements are achievable if block **635** is performed only on selected records.

Note that another input into block **635** is **C**, the geo-registration of still image(s) of block **640** and the transform from pixels to geographic coordinates in block **641**. This can provide the geographic coordinates **x, y, z** to add when deriving ground truth data
30 set **990**.

Thus, inputs to block **635** include:

Block **660**, a filtered set of objects on video frames, including object labels;

Block **655**, the mapping of pixels on the video and still images; and

C (block **641**), providing geographic information associated with the still images.

To provide more illustration of this step, attention is drawn to **Fig. 9B**, schematically illustrating a non-limiting example of a ground truth data set, in accordance with some embodiments of the presently disclosed subject matter. Each data
5 item or record of the interim ground-truth data set **990** comprises:

- (i) the video-image label **970**, indicative of the classification of the particular object;
- (ii) the indication **920** of a time of capture of the respective
10 video-image frame; and
- (iii) an indication **940, 942, 944** of the respective geographic position, of each classified second object **170**, corresponding to the second time of capture. In some implementations, the indication of the respective geographic position comprises
15 geographic coordinates, e.g. an absolute position, e.g. longitude, latitude and altitude/elevation, e.g. x, y, z as shown in the figure.

In some examples, such as in the figure, the ground-truth data set comprises a ground-truth table **990**. In other examples, instead of a table, the ground-truth data set
20 (and the interim ground-truth data set **900**) can be another data structure.

In some cases, deriving the ground-truth data set **990** is performed by deriving the indication x, y, z of respective geographic position from the respective second relative position i-video, j-video, based at least on the geo-registration of the still image(s), e.g. per the method of block **655**. See e.g. **Fig. 7**.

Reverting to **Fig. 6B**, according to some examples, based on indication(s) of a
25 respective geographic position(s) of the selected classified second object(s) **1020, 1028**, respective geographic first position(s) of localized respective first object(s), associated with the at least one non-EO signal or image **1050**, are determined (block **665**). This can be performed, for example, by localization module **555**.

To provide more illustration of this step, attention is drawn to **Fig. 10**, schematically illustrating a non-limiting example of object localization, in accordance
30 with some embodiments of the presently disclosed subject matter. Illustration **1000**

depicts an example of a video frame **1010** and a non-EO capture **1050**. In the example of the figure, the non-EO image is a SAR image **1050**.

In some examples, based on one or more indications $x_8, y_8, z_8, x_9, y_9, z_9$ of respective geographic positions of selected classified second object(s), the system **510** determines one or more respective geographic first positions $x_8, y_8, z_8, x_9, y_9, z_9$, of localized respective first objects in the corresponding non-EO capture/image **1050**.

In the example of the figure, the system knows (e.g. per its ground truth table **990**) that the relative position/coordinates of three objects in video frame **1010** are as follows: car A **1020** is at coordinates x_8, y_8, z_8 ; taxi **1024** is at x_7, y_7, z_7 ; and bus **1028** is positioned at x_9, y_9, z_9 . The system has matched the times of capture of the video image frame **1010** and the SAR image **1050**. Thus, the system determines that the expected first position of car A **1020**, in SAR image **1050**, will also be at absolute geographic coordinates x_8, y_8, z_8 . The same is done for the object **1028**, a stationary bus in the video frame **1010**, determined to be at geographic coordinates x_9, y_9, z_9 .

Non-EO captures, such as SAR images, are associated with various metadata convert from geographic coordinates to coordinates in the frame of reference of the non-EO capture. In some example, based on the determined one or more respective geographic first positions $x_8, y_8, z_8, x_9, y_9, z_9$ of the localized respective first objects **1020, 1028**, the system is configured to determine, for each localized respective first object a corresponding first relative position $i\text{-SAR-8}, j\text{-SAR-8}, i\text{-SAR-9}, j\text{-SAR-9}$, within the at least one non-EO signal/image.

Thus the x, y, z of the object in SAR **1050** image can be known, as well as the relative position in the frame of reference of the captured non-EO data. The object can be localized on the SAR **1050** image via the absolute geo-positioning.

In one non-limiting example, a point at the bottom center of the object is chosen on the video image, and a bounding box **1062, 1068** is built based on that point. The SAR signatures associated with the set of pixels within that bounding box are taken to be indicative of the relevant object. Other methods for object detection, e.g. using segmentation or other known *per se* methods, can be used.

. Recall that, in some implementations, in the SAR case, it can be technically advantageous to match the video and SAR images, only for stationary objects. Similarly, in e.g. GMTI, it can be advantageous to match the video frames and the

GMTI map, only for moving objects. The ground truth data set **990** can be used to accomplish this.

In this way, the system performs fusion of the non-EO and EO data, matching times of capture, and applying EO-associated labels to portions of the non-EO data
5 corresponding to the same geographic position.

Thus, where the data set **990** includes an indication **960** of movement, indicating whether or not each classified second object **1024** is moving during the time of capture of the respective video-image frame **1010**, this indicator can be used. When selecting data items of the ground-truth data set in block **660**, only records/data items with the
10 movement indicator **960** value that is relevant to the non-EO imaging mode, will be selected for localization on non-EO image or signal **1050**. Thus, in some implementations, in which the non-EO data comprises Synthetic Aperture RADAR (SAR) imaging data **1050**, the selecting of the data item(s) **1020**, **1028** of the ground-truth data set **990** comprises selecting records of the ground-truth data set for which a
15 corresponding indication **960** of movement is indicative of a stationary classified second object. Similarly, in some implementations, in which the non-EO data comprises Ground Moving Target Indication (GMTI) captures, the selecting of data item(s) **1024** of the ground-truth data set **990** comprises selecting data items of the ground-truth data set for which the corresponding indication of movement is indicative
20 of a moving classified second object.

If such a method is applied, then, in the example of **Fig. 10**, the system **510** will not attempt to localize taxi **1024** on SAR image **1050**, since the SAR image is not expected to show information for such a moving object. Thus, in some examples, using e.g. SAR imaging, only stationary objects **170** are localized. Moving **185** objects **180**
25 may appear smeared in the SAR capture, and therefore they are not being localized in the SAR image. Similarly, stationary object **170** are in some examples not being localized in e.g. GMTI maps.

Before continuing exposition, the methods of **665**, which were disclosed with respect to the example of SAR radar images, are now exemplified with respect to the
30 example of e.g. Ground Moving Target Indication (GMTI) tracks or signals. To provide more illustration of this, attention is first drawn to **Fig. 11**, schematically illustrating a non-limiting example of GMTI, in accordance with some embodiments of the presently disclosed subject matter. Illustration **1100** depicts an example graph of a GMTI capture.

The graph shows a plot **1130** of signal amplitude as a function of range, for one detected object. Such a graph will exist for each detected object.

Attention is now drawn to **Fig. 12**, schematically illustrating a non-limiting example of position matching, in accordance with some embodiments of the presently disclosed subject matter. This figure, for GMTI capturing, is similar to **Fig. 10**, which disclosed the example of SAR imaging, mutatis mutandis.

As an option, recall that the data set **990** includes an indication **960** of movement, indicating whether or not each classified second object **1024** is moving during the time of capture of the respective video-image frame **1010**. This indicator can be used. If a GMTI method is applied, then, for example, the system **510** will not localize objects in first positions in GMTI map **1250** that correspond to the stationary car A and bus **1020**, **1028**, since the GMTI track is not expected to show information for such a stationary object.

Example GMTI track **1250** is a range-Doppler map, with one axis showing Doppler and the other showing range. This image **1250** corresponds to the first time of capture T-3. Other range-Doppler maps (not shown) correspond to different times of GMTI capture.

The points **1263**, **1264** indicate the high resolution Range (HRR) profiles. These profiles are spread into a single Doppler bin, in a case where the moving object can be considered as a rigid body, or in a few Doppler bins, in a case where there are moving parts in the moving object or the object is making a turn. These profiles are spread into several range bins as a function of the object effective length relative to the line-of-sight. For example, if the object is 5 meters long and the angle between the object main axis and the line-of-sight is 30 degrees, then the effective length equals $5 * \cos(30) = 4.33$ meters.

The x, y, z coordinates of an object, together with the non-EO sensor e.g. RADAR in GMTI mode, can be mapped into Range-Doppler values. Thus, the object can be localized as the first position within the Range-Doppler map. Using the methods of **664**, **666**, the system **510** can localize the object on the GMTI map, and determine one or more respective geographic first positions of localized respective first objects associated with the GMTI map. and can compare these expected first positions to the Doppler bins **1263**, **1264**. For example, a particular R and D value on the map is determined to correspond to geographic position x6, y6, z6 of vehicle (car B) **1223**. A

bounding box is constructed around that point, or otherwise in the vicinity of the point, and the GMTI signature of this box is taken to be that of object **1223**.

Reverting to **Fig. 6**, one or more respective localized first objects **1062**, **1068** are labeled with a respective video-image label **970** of the respective classified second object(s) **1020**, **1028** (block **680**). This can be performed, for example, by non-EO object labeling module **544**. In each SAR, GMTI or other non-EO image/signal **1050**, **1250**, the label for each first object **1062**, **1264** is the label of the corresponding object **1020**, **1223** in video image **1010**. Thus, in the non-limiting example of **Fig. 10**, the bounding boxes are labelled.

10 In some non-limiting example, the labeled non-EO signature data per localized first object, is stored in non-EO signature database (DB) **595**.

To provide more illustration of this, attention is drawn to **Fig. 13**, schematically illustrating a non-limiting example of labeling, in accordance with some embodiments of the presently disclosed subject matter. Illustration **1300** depicts a labeled image **1350**.
15 The non-limiting illustrative example is that of a SAR image **1350**. It and video image **1310** were both captured at time T-6. The images look very much like the images **1010**, **1050** of **Fig. 10**. However, in illustration **1300**, SAR image **1350** is now labeled, as indicated in bold. Label **1372** indicates that SAR-image object **1062** is in fact car A. Label **1378** indicates that SAR-image object **1068** is in fact a bus.

20 In the example of the figure, the image itself is labeled or tagged. In other examples, a data structure, such as a table or list, is used, in which a label or tag for each non-EO object, in each non-EO image (which in turn is associated with a time T-6).

In the case of GMTI signals capture, labels can be applied/assigned, in a similar manner as in example **1300**, to Range-Doppler maps such as that of **Fig. 12**.
25

Note that the labeling process, culminating in block **680**, can be an automated one, performed by system **510**.

The classified first objects, including the non-EO signature information and the classification labels, are archived or stored in a non-EO signature database **595** (block
30 **685**). In some examples, this is performed using Non-EO Object Labeler Module **544**, and/or input/output module **560**, of system **505**.

In the non-limiting example of the figure, blocks **660**, **635**, **665**, **680**, **685** are shown as being done in parallel for all non-EO captured images/signals **1050**, **1250**. In

other non-limiting example implementations, the blocks **660**, **635**, **665**, **680**, **685** are performed in sequence, separately for each non-EO capture. That is, they are e.g. performed for SAR capture at time T1, then they are performed for the SAR capture at time T2, and so on. The process loops back, and performs the sequence again and again,
5 until it has been performed for all of the N non-EO captures, from times T1 to T-N. At this point, all of the non-EO images or signals have been labeled.

Reverting to **Fig. 6**, Non-EO Object classification model(s) are trained, based at least on the labelled non-EO data, that is on the labeled respective first objects **1062**, **1068** (block **690**). In some examples, this is part of the presently disclosed subject
10 matter, and is performed utilizing non-EO machine learning system **593** of system **505**, using, for example, known *per se* Deep Learning or other Machine Learning techniques. These ML systems can interface to labeling system **510** using, for example, interfaces **570**, **560**.

In other examples, the output of the system **505** is the labeled non-EO images,
15 and the training itself is performed by an external ML system (not shown), which utilizes e.g. the non-EO signature database **595**. Typically, there is no local Machine Learning system.

The presently disclosed subject matter has been implemented, and the method works. As indicated above, the presently disclosed methods and systems provide at least
20 certain example advantages. For each set up of a data gathering scenario (i.e for each arrangement of objects e.g. per Fig. 1), a large number of SAR and/or GMTI images or signals can be captured in one session. The video image data (and still image data, in cases where it is captured) is used to label the large number of examples of SAR/GMTI signatures captured. A sufficient number of labelled data items can be more easily
25 generated, to train high-quality machine learning classification models for the various types of non-EO signatures (signatures of SAR images of a bus, signatures of GMTI RD maps of a particular model of taxi etc.). Once such high-quality models are feasible, they can be used to identify objects captured in future non-EO images. The various amount of data available, for classifying EO can be leveraged to easily gather sufficient
30 amount of training dataset to train these non-EO classification models.

Note that the above descriptions of process **600** are non-limiting examples only. In some embodiments, one or more steps of the flowcharts exemplified herein may be performed automatically. The flow and function illustrated in the flowchart figures may

be implemented in systems **110, 120, 220, 310, 410, 505, 510** and processing circuitries **315, 415, 515**, and may make use of components described with reference to **Figs. 3, 4** and **5**. It is also noted that whilst the flowcharts is described with reference to system elements that realize steps, such as for example systems **110, 120, 220, 310, 410, 505,**
5 **510** and processing circuitries **315, 415, 515**, this is by no means binding, and the operations can be carried out by elements other than those described herein.

It is noted that the teachings of the presently disclosed subject matter are not bound by the flowcharts illustrated in the various figures. The operations can occur out the illustrated order. One or more stages illustrated in the figures can be executed in a
10 different order and/or one or more groups of stages can be executed simultaneously. As one non-limiting example, steps **650** and **653**, shown in succession, can be executed together, as one block. As another non-limiting examples in some cases, blocks **640** and **641**, and/or blocks **642, 648, 646**, for instance can be performed before **628**.

Similarly, some of the operations and steps can be integrated into a
15 consolidation, or can be broken down into several operations, and/or other operations can be added.

In some embodiments of the presently disclosed subject matter, fewer, more and/or different stages than those shown in the figures can be executed. As one non-limiting example, certain implementations may not include blocks **640** and **641**.

20 In the claims that follow, alphanumeric characters and Roman numerals used to designate claim elements such as components and steps, are provide for convenience only, and do not imply any particular order of performing the steps.

It should be noted that the word "compromising" as used throughout the appended claims should be interpreted to mean "including but not limited to".

25 While there has been shown disclosed examples in accordance with the presently disclosed subject matter, it will be appreciated that many changes may be made herein without departing from the spirit of the presently disclosed subject matter.

It is to be understood that the presently disclosed subject matter is not limited in its application to the details set forth in the description contained herein or illustrated
30 in the drawings. The presently disclosed subject matter is capable of other embodiments and of being practiced and carried out in various ways. Hence, it is to be understood that the phraseology and terminology employed herein are for the purpose of description and should not be regarded as limiting. As such, those skilled in the art will

appreciate that the conception upon which the disclosure is based may readily utilized as a basis for designing other structures, methods and systems for carrying out the several purposes of the presently disclosed subject matter.

It will also be understood that the system according to the presently disclosed
5 subject matter may be, at least partly, a suitable programmed computer. Likewise, the presently disclosed subject matter contemplates a computer program product being readable by a machine or computer, for executing the method of the presently disclosed subject matter, or any part thereof. The presently disclosed subject matter further contemplates a non-transitory machine-readable or computer-readable memory
10 tangibly embodying a program of instructions executable by the machine or computer, for executing the method of the presently disclosed subject matter, or any part thereof. The presently disclosed subject matter further contemplates a non-transitory machine-readable storage medium having a computer-readable program code embodied therein, configured to execute the method of the presently disclosed subject matter, or any part
15 thereof.

Those skilled in the art will readily appreciate that various modifications and changes can be applied to the embodiments of the invention as hereinbefore described, without departing from its scope, defined in and by the appended claims.

CLAIMS:

1. A method of labelling data, the method performed by a computerized system comprising a processing circuitry, the method being performed by the processing circuitry and comprising performing the following:
 - 5 (a) receiving unlabeled non-Electro-Optical (non-EO) data, the non-EO data comprising a set of non-EO images or signals captured during a first time period, the data being indicative of a first region;
 - (b) receiving unlabeled EO data,
captured during a second time period corresponding to the first time period,
10 the unlabeled EO data comprising at least a plurality of video image frames,
the images being indicative of a second region corresponding to the first region;
 - (c) classifying at least one second object appearing on one or more of the plurality of video image frames,
thereby generating at least one video-image label,
15 thereby generating labelled EO data from the unlabeled EO data; and
 - (d) labeling at least one first object appearing in the unlabeled non-EO data with a corresponding video-image label of the at least one video-image label,
the labelling being performed based at least on:
 - 20 (I) a correspondence of at least one first time of capture, of at least one non-EO image or signal of the set of non-EO images or signals, to at least one second time of capture, of at least one video image frame of the plurality of video image frames;
 - (II) a correspondence of positions of the at least one first object and the at least one classified second object; and
 - 25 (III) a respective geographic position, associated with the at least one video image frame, of the at least one classified second object,thereby generating labelled non-EO data from the unlabeled non-EO data,
the labelled non-Electro-Optical (non-EO) data being configured to facilitate machine learning of non-Electro-Optical signatures of first objects.
- 30 2. The method of claim 1,
wherein the unlabeled EO data further comprising at least one still image,
wherein the method further comprises, prior to said step (d), performing a geo-registration of the at least one still image,

wherein the performing the labelling of the at least one first object, based at least on the respective geographic position of the at least one classified second object, is based at least on the geo-registration.

3. The method of claim 2, wherein the geo-registration comprises matching
5 features of the at least one still image to orthophoto data.
4. The method of any one of claims 2 to 3, wherein the geo-registration utilizes Digital Terrain Elevation Data (DTED) data.
5. The method any one of claims 1 to 4, wherein the non-Electro-Optical (non-EO) data comprises RADAR imaging data.
- 10 6. The method of claim 5, wherein the RADAR imaging data comprises at least one of Synthetic Aperture RADAR (SAR) imaging data and Ground Moving Target Indication (GMTI) signals.
7. The method of claim 6, wherein each GMTI signal comprises a range-Doppler map.
- 15 8. The method of any one of claims 1 to 7, wherein the second period comprises the first period.
9. The method of any one of claims 1 to 7, wherein the first period and the second period are the same.
10. The method of any one of claims 1 to 9, wherein the second region comprises
20 the first region.
11. The method of any one of claims 1 to 9, wherein the first region and the second region are the same.
12. The method of any one of claims 1 to 11, wherein the method further comprises, prior to said step (a), capturing the unlabeled non-EO data.
- 25 13. The method of claim 12, wherein the capturing of the unlabeled non-EO data utilizes a non-EO sensor mounted on a first airborne vehicle.
14. The method of claim 13, wherein the first airborne vehicle comprises one of a first quadcopter, a first octa-copter, a first drone, and a first other remote-piloted vehicle (RPV).
- 30 15. The method of any one of claims 13 to 14, wherein the non-EO sensor is configured for setting of a capture mode of a plurality of capture modes.
16. The method of claim 15, wherein the plurality of capture modes comprises SAR capture mode and GMTI capture mode.

17. The method of any one of claims 1 to 16, wherein the method further comprises, prior to said step (b), capturing the unlabeled EO data.
18. The method of claim 17, wherein the capturing of the unlabeled EO data utilizes an EO sensor mounted on a second airborne vehicle.
- 5 19. The method of claim 18, wherein the second airborne vehicle comprises one of a second quadcopter, a second octa-copter, a second drone, and a first other remote-piloted vehicle (RPV).
20. The method of any one of claims 1 to 19, wherein the labelling of the at least one first object comprises a labelling of objects in each non-EO image or signal.
- 10 21. The method of any one of claims 1 to 20, wherein the unlabeled EO data are captured while the EO sensor is stationary.
22. The method of any one of claims 1 to 21, wherein the classifying the at least one second object comprises detecting the at least one second object.
23. The method of any one of claims 1 to 22, wherein the classifying the at least one second object is performed at least partially in an automated manner.
- 15 24. The method of claim 23, wherein the classifying the at least one second object utilizes Machine Learning.
25. The method of any one of claims 1 to 24, wherein the classifying the at least one second object is performed at least partially manually.
- 20 26. The method of any one of claims 1 to 25, wherein the generating the at least one video-image label comprises assigning, to each classified second object of the at least one second object, a respective video-image label indicative of the respective classification.
27. The method of any one of claims 1 to 26, wherein the classifying the at least one second object comprises tracking the at least one second object across video image frames of the plurality of video image frames, thereby facilitating an improved classification of the at least one second object in at least one video image frame of the plurality of video image frames, as compared to a second classification performed without tracking.
- 25 28. The method of any one of claims 1 to 27, wherein the generating the at least one video-image label comprises:
- 30

(A) generating a ground-truth data set assigning, to each classified second object of the at least one second object, a video-image label indicative of the classification,

wherein each data item of the ground-truth data set comprises:

- 5 (i) the respective video-image label indicative of the classification of the each classified second object;
- (ii) an indication of a second time of capture, of a respective video-image frame of the plurality of video image frames, which comprises the each classified second object; and
- 10 (iii) an indication of the respective geographic position of the each classified second object corresponding to the second time of capture.

29. The method of claim 28, wherein the ground-truth data set comprises a ground-truth table.

30. The method of any one of claims 28 to 29, wherein the indication of the
15 respective geographic position comprises geographic coordinates.

31. The method of any one of claims 28 to 30, wherein the generating the at least one video-image label comprises performing, prior to the step (A), the following step:

(B) generating an interim ground-truth data set, which assigns, to each classified
20 second object of the at least one second object, a video-image label indicative of the classification, wherein each data item of the interim ground-truth data set comprises:

- (i) the video-image label;
- (ii) the indication of a time of capture of the respective video-image
25 frame; and
- (iii) an indication of a respective second relative position within the respective video-image frame, of the each classified second object corresponding to the time of capture,

wherein the performing of the labelling of the at least one first object comprises deriving
30 the ground-truth data set from the interim ground-truth data set,

wherein the deriving the ground-truth data set is performed by deriving the indication of respective geographic position from the respective second relative position, based at least on the geo-registration of the at least one still image,

wherein the deriving the indication of respective geographic position comprises registering the respective video-image frame to the at least one still image.

32. The method of claim 31, wherein the indication of the respective relative second position within the respective video-image frame comprises coordinates within the respective video-image frame.

33. The method of any one of claims 28 to 32, wherein the generating the ground-truth data set is performed for respective second video image frames, for which the second time of capture, and a first time of capture of a corresponding non-EO image or signal, meet a time-matching criterion.

34. The method of claim 33, wherein the time-matching criterion comprises at least one of the following:

- a. the second time of capture is equal to the first time of capture; and
- b. a difference between the first time of capture and a closest video frame is a maximum of half of a video frame time.

35. The method of any one of claims 1 to 34, wherein the performing the labelling of the at least one non-EO image, based at least on a correspondence of at least one first time of capture, of the at least one non-EO image or signal, to the to at least one second time of capture, of the of at least one video image frame, on the correspondence of positions of the at least one first object and the at least one classified second object, and on the respective geographic position of the at least one classified second object, comprises performing the following steps:

(I) determine that a corresponding second time of capture of the respective video-image frame, and a first time of capture of the at least one non-EO image or signal, meet a matching criterion;

(II) select one or more data items of the ground-truth data set that comprise an indication of the corresponding second time of capture of the respective video-image frame;

the selected one or more data items comprising one or more selected classified second objects;

(III) based on one or more indications of respective geographic positions of the one or more selected classified second objects,

determine one or more respective geographic first positions of localized respective first objects associated with the at least one non-EO image or signal,
the one or more selected classified second objects constituting the at least one second object;

- 5 (IV) label a respective localized first objects with corresponding respective video-image labels of the one or more selected classified second objects, the respective localized first object constituting the at least one first object.

36. The method of claim 35, wherein said step (III) comprises:
based on the determined one or more respective geographic first positions of the
10 localized respective first objects, determine for each localized respective first object a corresponding first relative position, within the at least one non-EO image or signal.

37. The method of any one of claims 35 to 36, wherein the time-matching criterion comprises at least one of the following:

- 15 a. the second time of capture is equal to the first time of capture; and
 b. a difference between the first time of capture and a closest video frame is a maximum of half of a video frame time.

38. The method of any one of claims 28 to 37, wherein the each data item of the ground-truth data set further comprises:

- 20 (iv) an indication of movement, indicating whether the each classified second object is moving during the time of capture of the respective video-image frame,

wherein the indication whether the each classified second object is moving is determined based on tracking the each classified second object across video image frames of the plurality of video image frames.

25 39. The method of claim 38,
wherein the non-EO data comprises RADAR imaging data, wherein the RADAR imaging data comprises Synthetic Aperture RADAR (SAR) imaging data,
wherein the selecting of the one or more data items of the ground-truth data set comprises selecting one or more records of the ground-truth data set for which a
30 corresponding indication of movement is indicative of a stationary classified second object.

40. The method of any one of claims 38 to 39,

wherein the non-EO data comprises RADAR imaging data, wherein the RADAR imaging data comprises Ground Moving Target Indication (GMTI) range-Doppler maps,

wherein the selecting of the one or more data items of the ground-truth data set comprises selecting one or more data items of the ground-truth data set for which the
5 corresponding indication of movement is indicative of a moving classified second object.

41. The method of any one of claims 1 to 40, the method further comprising:

(V) training models for Non-EO objects classification, based at least on the
10 labeled respective first objects.

42. A computerized system configured to label data, the computerized system comprising a processing circuitry configured to perform the following method:

(a) receiving unlabeled non-Electro-Optical (non-EO) data, the non-EO data comprising a set of non-EO images or signals captured during a first time
15 period, the data being indicative of a first region;

(b) receiving unlabeled EO data,
captured during a second time period corresponding to the first time period,
the unlabeled EO data comprising at least a plurality of video image frames,
the images being indicative of a second region corresponding to the first region;

(c) classifying at least one second object appearing on one or more of the
20 plurality of video image frames,

thereby generating at least one video-image label,
thereby generating labelled EO data from the unlabeled EO data; and

(d) labeling at least one first object appearing in the unlabeled non-EO data with
25 a corresponding video-image label of the at least one video-image label,
the labelling being performed based at least on:

(I) a correspondence of at least one first time of capture, of at least
one non-EO image or signal of the set of non-EO images or
signals, to at least one second time of capture, of at least one
30 video image frame of the plurality of video image frames;

(II) a correspondence of positions of the at least one first object and
the at least one classified second object; and

(III) a respective geographic position, associated with the at least one video image frame, of the at least one classified second object, thereby generating labelled non-EO data from the unlabeled non-EO data, the labelled non-Electro-Optical (non-EO) data being configured to facilitate machine learning of non-Electro-Optical signatures of first objects.

43. A non-transitory computer readable storage medium tangibly embodying a program of instructions which, when executed by a computerized system, cause the computerized system to perform a method of labelling data, the method being performed by a processing circuitry of the computerized system, the method comprising performing the following:

(a) receiving unlabeled non-Electro-Optical (non-EO) data, the non-EO data comprising a set of non-EO images or signals captured during a first time period, the data being indicative of a first region;

(b) receiving unlabeled EO data, captured during a second time period corresponding to the first time period, the unlabeled EO data comprising at least a plurality of video image frames, the images being indicative of a second region corresponding to the first region;

(c) classifying at least one second object appearing on one or more of the plurality of video image frames,

thereby generating at least one video-image label, thereby generating labelled EO data from the unlabeled EO data; and

(d) labeling at least one first object appearing in the unlabeled non-EO data with a corresponding video-image label of the at least one video-image label, the labelling being performed based at least on:

(I) a correspondence of at least one first time of capture, of at least one non-EO image or signal of the set of non-EO images or signals, to at least one second time of capture, of at least one video image frame of the plurality of video image frames;

(II) a correspondence of positions of the at least one first object and the at least one classified second object; and

(III) a respective geographic position, associated with the at least one video image frame, of the at least one classified second object, thereby generating labelled non-EO data from the unlabeled non-EO data,

the labelled non-Electro-Optical (non-EO) data being configured to facilitate machine learning of non-Electro-Optical signatures of first objects.

44. A method of labelling data, the method performed by a computerized system comprising a processing circuitry, the method being performed by the processing circuitry and comprising performing the following:

- (a) receiving unlabeled non-Electro-Optical (non-EO) data, the non-EO data comprising a set of non-EO images or signals captured during a first time period, the data being indicative of a first region;
- (b) receiving unlabeled EO data,
 - captured during a second time period corresponding to the first time period, the unlabeled EO data comprising at least a plurality of video image frames, the images being indicative of a second region corresponding to the first region;
 - (c) classifying at least one second object appearing on one or more of the plurality of video image frames, thereby generating at least one video-image label, thereby generating labelled EO data from the unlabeled EO data, wherein the classifying the at least one second object comprises tracking the at least one second object across video image frames of the plurality of video image frames, thereby facilitating an improved classification of the at least one second object in at least one video image frame of the plurality of video image frames, as compared to a second classification performed without tracking; and
 - (d) labeling at least one first object appearing in the unlabeled non-EO data with a corresponding video-image label of the at least one video-image label, the labelling being performed based at least on:
 - (I) a correspondence of at least one first time of capture, of at least one non-EO image or signal of the set of non-EO images or signals, to at least one second time of capture, of at least one video image frame of the plurality of video image frames;
 - (II) a correspondence of positions of the at least one first object and the at least one classified second object; and
 - (III) a respective geographic position, associated with the at least one video image frame, of the at least one classified second object,thereby generating labelled non-EO data from the unlabeled non-EO data,

the labelled non-Electro-Optical (non-EO) data being configured to facilitate machine learning of non-Electro-Optical signatures of first objects.

45. A method of labelling data, the method performed by a computerized system comprising a processing circuitry, the method being performed by the processing circuitry and comprising performing the following:

- 5 (a) receiving unlabeled non-Electro-Optical (non-EO) data, the non-EO data comprising a set of non-EO images or signals captured during a first time period, the data being indicative of a first region;
- (b) receiving unlabeled EO data,
 - 10 captured during a second time period corresponding to the first time period, the unlabeled EO data comprising at least a plurality of video image frames, the images being indicative of a second region corresponding to the first region;
 - (c) classifying at least one second object appearing on one or more of the plurality of video image frames,
 - 15 thereby generating at least one video-image label, thereby generating labelled EO data from the unlabeled EO data, wherein the generating the at least one video-image label comprises generating an indication of movement, indicating whether the each classified second object is moving during the time of capture of the respective video-image frame,
 - 20 wherein the indication whether the each classified second object is moving is determined based on tracking the each classified second object across video image frames of the plurality of video image frames; and
 - (d) labeling at least one first object appearing in the unlabeled non-EO data with a corresponding video-image label of the at least one video-image label,
 - 25 the labelling being performed based at least on:
 - (I) a correspondence of at least one first time of capture, of at least one non-EO image or signal of the set of non-EO images or signals, to at least one second time of capture, of at least one video image frame of the plurality of video image frames;
 - 30 (II) a correspondence of positions of the at least one first object and the at least one classified second object; and
 - (III) a respective geographic position, associated with the at least one video image frame, of the at least one classified second object,

thereby generating labelled non-EO data from the unlabeled non-EO data,
the labelled non-Electro-Optical (non-EO) data being configured to facilitate machine
learning of non-Electro-Optical signatures of first objects.

5 46. A method of labelling data, the method performed by a computerized system
 comprising a processing circuitry, the method being performed by the
 processing circuitry and comprising performing the following:

 (a) receiving unlabeled non-Electro-Optical (non-EO) data, the non-EO data
 comprising a set of non-EO images or signals captured during a first time
 period, the data being indicative of a first region;

10 (b) receiving unlabeled EO data,
 captured during a second time period corresponding to the first time period,
 the unlabeled EO data comprising at least a plurality of video image frames,
 the images being indicative of a second region corresponding to the first region,
 wherein the unlabeled EO data further comprising at least one still image;

15 (c) classifying at least one second object appearing on one or more of the
 plurality of video image frames,

 thereby generating at least one video-image label,

 thereby generating labelled EO data from the unlabeled EO data; and

 (d) performing a geo-registration of the at least one still image,

20 (e) labeling at least one first object appearing in the unlabeled non-EO data with
 a corresponding video-image label of the at least one video-image label,

 the labelling being performed based at least on:

 (I) a correspondence of at least one first time of capture, of at least
 one non-EO image or signal of the set of non-EO images or
25 signals, to at least one second time of capture, of at least one
 video image frame of the plurality of video image frames;

 (II) a correspondence of positions of the at least one first object and
 the at least one classified second object; and

 (III) a respective geographic position, associated with the at least one
30 video image frame, of the at least one classified second object,

 thereby generating labelled non-EO data from the unlabeled non-EO data,

 the labelled non-Electro-Optical (non-EO) data being configured to facilitate machine
 learning of non-Electro-Optical signatures of first objects,

wherein the performing the labelling of the at least one first object, based at least on the respective geographic position of the at least one classified second object, is based at least on the geo-registration.