## Claims

1. A system comprising:

a) an input/output (I/O) unit (202) configured to receive an input image (110) that comprises at least one object of interest (112); and

b) a processor (204), configured to perform sequentially the steps of (i) to (v):

(i) obtaining using a first deep neural network, DNN1 (304), from the input image at least a first Region of Interest (RoI) (304A) that includes the at least one object of interest, and

(ii) classifying using a second deep neural network, DNN2 (306), executed on the first RoI (304A), the at least one object of interest in the first RoI (304A) by returning a class (306A) for the at least one object of interest, and

(iii) detecting using a third deep neural network, DNN3 (308), executed on the at least one object of interest classified with the class (306A), the at least one object of interest within the first RoI (304A) by returning a bounding box (308A) around the at least one object of interest classified with the class (306A), and

(iv) segmenting, using a binary mask (310A) generated by a fourth deep neural network, DNN4 (310), executed on the at least one object of interest within the bounding box (308A), the at least one object of interest classified with the class (306A) and detected within the first RoI (304A), and

(v) extracting features (312) from the said at least one object of interest classified with the class (306A), detected within the first RoI (304A) and segmented using the binary mask (310A) , wherein the features comprise shape, size and/or distribution of the objects.

2. The system according to claim 1, wherein DNN1 (304) comprises a Region Proposal Network and/or

wherein DNN3 (308) comprises a Fast R-CNN or a Faster R-CNN.

3. The system according to any of the claims 1 or 2, wherein through the binary mask (310A), each pixel in the bounding box (308A) is classified as belonging to the object of interest or the background in the bounding box (308A).

4. A system comprising:

a) an input/output (I/O) unit (202) configured to receive an input image (602) that comprises at least one object of interest (602A); and

b) a processor (204) configured to perform the steps of:

(i) using a deep neural network, DNN7 (604), which comprises a Mask R-CNN, on the input image and thereby performing the sub-steps of

1. generating a region proposal (604A) within the image, where the at least one object of interest (602A) can be located,
2. classifying the at least one object of interest in the region proposal (604A) by returning a class (604B) for the at least one object of interest, and
3. detecting the at least one classified object of interest the region proposal (604A) by returning a bounding box (604C) around the at least one classified object of interest,
4. segmenting of the at least one object of interest in the region proposal (604A),

and wherein the sub-step 2. for object classification and the sub-set 3. for object detection are performed in parallel to the sub-step 4. for object segmentation,

(ii) extracting features (606) from the said at least one classified, detected and segmented object of interest, wherein the features comprise shape, size and/or distribution of the objects,

5. The system according to claim 4, wherein DNN7 comprises a Region Proposal Network.

6. The system according to claim 4 or 5, wherein the object segmentation returns a binary mask (604D) through which each pixel in the bounding box (604C) is classified as belonging to the at least one object of interest or the background.

7. The system according to any one of claims 1 to 6, wherein the input image comprises a plurality of isolated or overlapping objects of interest; and/or

wherein the input image is a digital gastrointestinal histological image.

8. The system according to any of the claims 1 to 7, wherein the one or more objects of interest are lesions associated with a gastrointestinal disorder.

9. The system according to any of the claims 1 to 8, wherein the comparison of the extracted features of the same object in images related to different points in time allow for assessing the temporal evolution of biomarkers.

10. The system according to any of the claims 1-9, wherein the step of extracting the object features is performed by a data-driven decision apparatus (104) communicatively coupled with the data processing apparatus (102), wherein the data processing apparatus (102) comprises the processor (204).

11. A computer-implemented method comprising the steps of:

(i) receiving an input image from a database comprising at least one object of interest, and

(ii) obtaining, using a first deep neural network, DNN1 (804), from the input image at least a first Region of Interest (RoI) that includes the at least one object of interest, and

(iii) classifying, using a second deep neural network, DNN2 (806), executed on the first RoI the at least one object of interest in the first RoI by returning a class for the at least one object of interest, and

(iv) detecting, using a third deep neural network, DNN3 (808), executed on the at least one object of interest classified with the class, the at least one object of interest within the first RoI by returning a bounding box around the at least one object of interest classified with the class, and

(v) segmenting, using a binary mask generated by a fourth deep neural network, DNN4 (810), executed on the at least one object of interest within the bounding box, the at least one classified and detected object of interest, and

(vi) extracting features (812) from the said at least one classified, detected and segmented object of interest using the binary mask, wherein the features comprise shape, size and/or distribution of the objects.

12. The method according to claim 11, wherein the input image comprises a plurality of isolated and/or overlapping objects of interest and/or

wherein the input image is a digital gastrointestinal histological image.

13. The method according to any of the claims 11 or 12, wherein the one or more objects of interest are lesions associated with a gastrointestinal disorder.

14. The method according to any of the claims 11 to 13, wherein the comparison of the extracted features of the same object in images related to different points in time allow for assessing the temporal evolution of biomarkers.

15. The method according to any of the claims 11-15, wherein the step of classification and detection of the at least one object of interest is performed using a fifth deep neural network, DNN5 (906), or

wherein the steps of classification and detection of the at least one object of interest are performed in parallel with the step of segmentation of the at least one object of interest, whereby the steps of classification and detection are performed using two DNNs, and whereby the step of segmentation is performed using a sixth deep neural network, DNN6 (1008) on the Region of Interest (RoI) or

wherein the steps of classification, detection and segmentation of the at least one object of interest are performed using a seventh deep neural network, DNN7 (1104) on the input image, and optionally wherein the deep neural network DNN7 (1104) comprises a Mask R-CNN.

16. The method according to any of the claims 11 to 15, wherein a step of post-processing (1206) of the at least one detected and segmented object of interest is performed.