(19)

Europäisches Patentamt

European Patent Office

Office européen des brevets

(11) **EP 3 889 894 A1**

(12) **EUROPEAN PATENT APPLICATION**

(43) Date of publication:
06.10.2021 Bulletin 2021/40

(51) Int Cl.:
G06T 7/11 (2017.01)

(21) Application number: 20167487.6

(22) Date of filing: 01.04.2020

(84) Designated Contracting States:
AL AT BE BG CH CY CZ DE DK EE ES FI FR GB GR HR HU IE IS IT LI LT LU LV MC MK MT NL NO PL PT RO RS SE SI SK SM TR
Designated Extension States:
BA ME
Designated Validation States:
KH MA MD TN

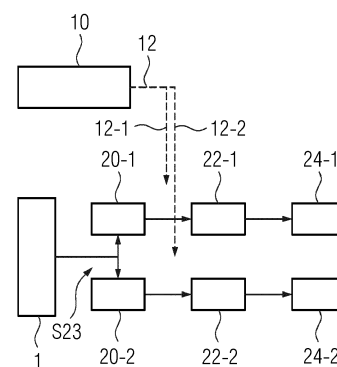(71) Applicant: **Siemens Healthcare GmbH**
91052 Erlangen (DE)

(72) Inventors:
• **Diallo, Mamadou**
Plainsboro, NJ 08536 (US)

• **Ezzi, Afshin**
Princeton, NJ 08540 (US)
• **Kamen, Ali**
Skillman, NJ 08558 (US)
• **Lou, Bin**
Princeton, NJ 08540 (US)
• **Meng, Tongbai**
Ellicott City, MD 21043 (US)
• **Shi, Bibo**
Monmouth Junction, NJ 08852 (US)
• **Winkel, David Jean**
4058 Basel (CH)
• **Yu, Xin**
Plainsboro, NJ 08536 (US)

(54) **IMPROVED ABNORMALITY DETECTION IN MEDICAL IMAGING DATA**

(57) The invention relates to improved methods for training artificial intelligence entities for abnormality detection as well as improved methods for abnormality detection. The training methods comprise:

providing (S10; S210) medical imaging data of the human organ as training data comprising training samples (1; 201), the medical imaging data comprising a plurality of imaging results (202-1; 202-2) from different types of imaging techniques for each training sample (1; 201) of the training data;

providing (S20; S21; S220) at least one pre-trained or randomly initialized artificial intelligence entity, AIE (20-1, 20-2); and

training (S40; S230) the AIE (20-1, 20-2) using the provided training samples (1; 201),

wherein in a first sub-step (S41; S231), for at least one training sample (1; 201), a first loss function (22-1; L2$_{DWI}$) for at least a sub-structure (20-1) of the at least one AIE is calculated independently of at least one first spatial region of the human organ, and

wherein in a second sub-step (S42; S232), for at least one training sample (1; 201), a second loss function (22-2; L2$_{T2W}$) for at least a sub-structure (20-2) of the at least one AIE is calculated independently of at least one second spatial region of the human organ;

wherein the AIE (20-1, 20-2) is trained (S40; S230) using at least the calculated first loss function (22-1; L$_{DWI}$) and the calculated second loss function (22-2; L$_{T2W}$).

FIG 2

Description

Technical Field

[0001]   The present invention relates to methods for training an artificial intelligence entity, AIE, for detecting abnormalities such as lesions in medical imaging data of a human organ, in particular a prostate. Moreover, the invention relates to methods for detecting abnormalities in such medical imaging data, in particular using an artificial intelligence entity, AIE. The invention also relates to devices for detecting abnormalities in such medical imaging data.

Background of the invention

[0002]   Imaging techniques such as magnetic resonance imaging, MRI, computed tomography, CT, and so on are more and more used for diagnostic tasks. Different imaging techniques are differently suitable for imaging different organs so that often for one type of organ one type of imaging technique is dominant, i.e. is predominantly used and/or is best used for diagnostics.
[0003]   Some organs, in particular the human prostate, have different regions which respond differently to different medical imaging techniques. For example, a human prostate is divided into, among others, a peripheral zone, PZ, and a transition zone, TZ. The transition zone, TZ, surrounds the prostatic urethra and enlarges in aging men as a result of benign prostatic hyperplasia. The peripheral zone, PZ, is situated on the posterior and lateral side of the prostate, surrounding the transition zone.
[0004]   Medical imaging techniques for detecting lesions within the prostate are usually different types (or: series) of MRI (Magnetic Resonance Imaging) scans. For the prostate, typical MRI scans comprise T2-weighted MRI scans (T2W), diffusion-weighted imaging (DWI) and dynamic contrast-enhanced (DCE) imaging. These types of MRI scans, in particular the first two, are also used in the PI-RADS assessment categories. Whenever PI-RADS is mentioned herein, it is referred to the "Prostate Cancer - PI-RADS v2" document by Rhiannon van Loenhout, Frank Zijta, Robin Smithuis and Ivo Schoots, publication date 2018-08-01, available e.g. at the URL: http://radiologyassitant.nl/abdomen/prostate-cancer-pi-rads-v2 . However, it shall be understood that also future version of PI-RADS may be similarly applied.
[0005]   A combination of the results of multiple types of MRI scans is sometimes designated as a "multiparametric MRI", wherein "multiparametric" refers to the different parameter settings with which the respective MRI scan is performed. PI-RADS, for example, designates a combination of T2-weighted, diffusion weighted and dynamic contrast-enhanced imaging as a "multiparametric MRI". Based on the results of multiparametric MRIs, PI-RADS assessment categories are based which are used for standardized acquisition, interpretation and reporting of prostate MRIs. The PI-RADS categories range from "PI-RADS 1" (very low - clinically significant cancer highly unlikely) to "PI-RADS 5" (clinically significant cancer highly likely), wherein a clinically significant prostate cancer is defined as a tumor with a Gleason score of 7 or more.
[0006]   T2W imaging is the primary determining sequence (i.e. scan) to assign the PI-RADS category in the transition zone, TZ. On the other hand, diffusion-weighted imaging (in particular DWI ADC) is the primary determining sequence to assign the PI-RADS assessment category for the peripheral zone, PZ.
[0007]   In the international publication WO 2019/238804 A1, various advantageous methods and variants for detecting and localizing abnormalities in medical images are described.
[0008]   However, hitherto known attempts of employing deep learning for organs that happen to comprise different spatial regions (e.g. transition zone, TZ and peripheral zone, PZ, in case of the prostate) responding differently to different types of medical imaging scans neglect information about these different spatial regions.

Summary of the invention

[0009]   The above-mentioned problems are solved by the subject-matter of the independent claims.
[0010]   According to a first aspect of the present invention, a computer-implemented method for training an artificial intelligence entity, AIE, for detecting abnormalities in medical imaging data of a human organ, in particular a prostate, is provided the method comprising the steps of:

providing medical imaging data of the human organ as training data comprising training samples, the medical imaging data comprising a plurality of imaging results from different types of imaging techniques for each training sample of the training data;
providing at least one pre-trained or randomly initialized artificial intelligence entity, AIE; and
training the AIE using the provided training samples; wherein in a first sub-step, for at least one training sample, a first loss function for at least a sub-structure of the AIE is calculated independently of at least one first spatial region of the human organ;

wherein in a second sub-step, for at least one training sample, a second loss function for at least a sub-structure of the AIE is calculated independently of at least one second spatial region of the human organ; and

wherein the AIE is trained using at least the calculated first loss function and the calculated second loss function.

[0011] Alternatively or additionally, the first loss function is calculated dependent on only the second spatial region and the second loss function is calculated dependent on only the first spatial region.

[0012] Preferably, the first loss function is calculated independently of the first spatial region and is further calculated only dependent on the second spatial region, and the second loss function is calculated independently of the second spatial region and is further calculated only dependent on the first spatial region.

[0013] Calculating a loss function independently of a spatial region may in particular be understood to mean that no change in any pixel/voxel corresponding to (or, in other words, within) said spatial region within the medical imaging data has any influence on said loss function. Correspondingly, if/when the training is then performed based on such a loss function alone, no change in a pixel/voxel corresponding to said spatial region makes a difference for the gradient descent.

[0014] The sub-structures trained in the first and the second sub-steps may be different or may be the same sub-structure. When a single artificial intelligence entity, AIE, is provided, both sub-structures may be identical to this AIE. When more than one AIE is provided, then each sub-structure may refer to a different one of the provided AIEs.

[0015] It is one of the core concepts of the present invention to use zonal information in abnormality detection to make an artificial intelligence entity, AIE, emphasize particular modalities (i.e. particular medical imaging techniques such as particular MRI series) when identifying lesions in particular zones. In the present context, a lesion shall in particular be understood to be a region in an organ or tissue which has suffered damage through injury or disease, such as a wound, ulcer, abscess, or tumor.

[0016] This is particularly advantageous for detecting and distinguishing malignant or benign lesions with a distinct feature appearance mainly in one of multiple imaging techniques, for example benign prostatic hyperplasia which mainly effects the transition zone, TZ of the prostate but not the peripheral zone, PZ. Incorporating zonal information can thus significantly reduce false positive detections and thus eliminate unnecessary biopsies.

[0017] The main ideas of the present invention are presented herein with the human prostate as an example. It shall be understood, however, that the same method may also be applied to other organs (e.g., brain, heart, liver, kidney, ...), either of humans or other organisms, as will become apparent. Therefore, whenever the human prostate is mentioned herein, it shall be understood that this mention be replaced by any other suitable organ according to the principles set out explicitly and/or implicitly herein.

[0018] As main examples for different regions within the human prostate which respond differently to different medical imaging scans, the peripheral zone, PZ, and the transition zone, TZ, are chosen. It shall be understood that the described methods and apparatuses can also be applied if the organ is virtually separated into more than two different regions. For example, for most applications it is sufficient if the central zone, CZ, of the prostate is grouped together with the transition zone, TZ, and this approach will be used herein. However, for other applications it may be more suitable to individually distinguish also the central zone, CZ, of the prostate.

[0019] In the event that more than two different spatial zones are defined/present, there may be provided one loss function per different spatial zone, and each loss function may be calculated as dependent on only one of the zones and thus independently of all of the other zones. For example, with spatial regions A, B, and C, then the first loss function may be calculated as dependent only on spatial region B and independently of regions A and C; the second loss function may be calculated as dependent only on spatial region C and independently of regions A and B; and the third loss function may be calculated as dependent only on spatial region A and independently of regions B and C.

[0020] The methods according to the invention adopt additional information from the specific anatomical zones, and in many particularly advantageous embodiments also further additional information from the relationship between the specific anatomical zones and medical imaging techniques, much in the same way as experienced physicians do.

[0021] Whenever different medical imaging techniques are recited herein, this may refer to different modalities such as computed tomography, CT, magnetic resonance imaging, MRI, X-ray imaging, ultrasound imaging and so on, and/or to specific sub-types or "series" of these modalities, such as differently weighted MR imaging scans. In particular the latter will be mainly used as examples, and among these mainly the T2-weighted MR imaging ("T2W") and the diffusion-weighted MR imaging ("DWI"). The ideas and principles described are, however, equally applicable to other modalities and/or other sub-types, depending on the organ to be scanned, on the personal situation of the patient, on the equipment of a particular medical site and so on.

[0022] Results of imaging techniques in particular comprise spatially resolved images, for example bitmaps of pixels or voxels in which each pixel or voxel has a value that indicates a result of a particular imaging technique in a particular position of the human organ.

[0023] An artificial intelligence entity, AIE, may in particular be an artificial neural network, ANN, or a pipeline of artificial neural networks and/or other algorithms. Training in this context in particular may comprise updating parameters (for

example, weights and biases within an ANN) in order to optimize (usually minimize) a loss (or: cost) function. Whenever herein "at least a sub-structure" is recited, this may comprise the case that a true sub-structure is meant (i.e. a smaller part of the artificial intelligence entity, AIE, as a whole); the case that the whole artificial intelligence entity, AIE, is meant (in the same sense in that "at least part of" includes also the possibility of "all"); and also the case that more than one sub-structure is meant, i.e. two or more sub-substructures. Examples will be provided, in particular with respect to the drawings, to illustrate at least some of these cases.

[0024] The term "sub-step" is used herein for better explaining individual aspects of the method; the distinction into sub-steps may in some instances be a virtual distinction between simultaneous processes for easier explanation. Some sub-steps may be performed simultaneously, other sub-steps may be performed sequentially.

[0025] Further advantageous embodiments and variants will be found in the other independent claims, in the dependent claims, and in the description of the drawings in combination with the drawings themselves.

[0026] In some advantageous embodiments, refinements, or variants of embodiments, the first spatial region and the second spatial region are disjunct. Thus, there is no spatial region from which both of the first and second loss function are independent.

[0027] In some advantageous embodiments, refinements, or variants of embodiments, the first loss function is based only on pixels or voxels of the provided medical imaging data within the second spatial region and/or the second loss function is based only on pixels or voxels of the provided medical imaging data within the first spatial region. This increases the relationship between the respective loss function and a corresponding zone of the human organ. For example, for the case of the human prostate, it is preferred if the first loss function depends only on (i.e. is calculated only based on) pixels or voxels from the peripheral zone, PZ, and that the second loss function depends only on (i.e. is calculated only based on) pixels or voxels from the transition zone, TZ.

[0028] In some advantageous embodiments, refinements, or variants of embodiments, the first loss function is calculated independently of the imaging results from at least one first of the types of imaging techniques, and the second loss function is calculated independently of the imaging results from at least one second of the types of imaging techniques. Preferably, each loss function is calculated independently of all but one type of imaging technique, or, in other words, based on only one type of imaging technique of all of the types of imaging techniques for which results are provided in the training samples. This further strengthens the relationship between the respective loss function and a specific imaging technique. For example, for the case of the human prostate, it is preferred if the first loss function depends only (i.e. is only calculated based on) T2W scans, so that the first loss function is particularly associated with the transition zone, TZ, and that the second loss function depends only (i.e. is only calculated based on) DWI scans so that the second loss function is particularly associated with the peripheral zone, PZ.

[0029] Then, more preferably, the first loss function is calculated independently of at least one first spatial region of the human organ and also independently from at least one first of the types of imaging techniques, and the second loss function is calculated independently of at least one second spatial region of the human organ and also independently from at least one second of the types of imaging techniques. More preferably, each loss function is calculated based on only a single type of imaging technique and only based on a single spatial region (or: zone) of the human organ. This further strengthens the relationship between the respective loss function, a specific spatial region of the human organ, and a particular imaging technique which is particularly useful ("dominant") for the corresponding spatial region.

[0030] For example, for the case of the human prostate, it is preferred if the first loss function depends only on (i.e. is calculated only based on) pixels or voxels from the peripheral zone, PZ, and only on results of (one or more) DWI scans and/or that the second loss function depends only on (i.e. is calculated only based on) pixels or voxels from the transition zone, TZ and only on results of a T2W scan. In this way, the known strong relationships between said zones and said medical imaging techniques are exploited in order to provide the best analysis. In this example, the peripheral zone, PZ, would be defined as the "second spatial region" and the transition zone, TZ, would be defined as the "first spatial region".

[0031] In some advantageous embodiments, refinements, or variants of embodiments, in the first sub-step and in the second sub-step the first loss function and the second loss function are calculated for the same at least sub-structure of the AIE, in particular for the same artificial neural network, ANN. Most preferably, the artificial intelligence entity, AIE, consists of this artificial neural network, ANN, which is trained using (at least) both the first and the second loss function. First and second loss function, and optionally one or more additional loss functions, may be combined in a weighted sum to form a total loss function used in a back-propagation algorithm for training the artificial neural network, ANN. The weights of the weighted sum may be suitably chosen hyperparameters. The first and the second loss function can thus also be designated as first and second "loss terms", respectively.

[0032] In some advantageous embodiments, refinements, or variants of embodiments, for the first sub-step the imaging results from the at least one first type of imaging technique within each training sample of the at least one training sample are replaced by a pre-defined blank before being used for calculating the first loss function (or: first loss term), and for the second sub-step the imaging results from the at least one second type of imaging technique within each training sample of the same at least one training sample are replaced by the pre-defined blank before being used for calculating the second loss function (or: second loss term).

[0033] Using these blanks means that the information from a specific imaging technique is left out from a specific training sample, so that any loss function that is calculated based on an output which results from a training sample thus processed is independent of that specific imaging technique. Similarly, if the blank is applied to all but one results of different imaging techniques, then the loss function calculated based on an output which results from a thus processed training sample depends only on the single remaining imaging technique.

[0034] The pre-defined blank may depend on the format of the results of the imaging techniques and may, for example, consist of a vector, matrix or tensor with all "0", all "1", all "0.5" or the like.

[0035] In some advantageous embodiments, refinements, or variants of embodiments, in a third sub-step, a third loss function is calculated for each training sample of the same at least one training sample. In the third sub-step each training sample is left unaltered before being used for calculating the third loss function. In other words, each training sample is used "unaltered", i.e. as it is provided - in contrast to the first and second sub-step in this context, which may each make use of training samples that have been processed by replacing parts of the results with a pre-defined blank. The third loss function penalizes differences between the output of the at least sub-structure of the AIE and corresponding labels for each training sample, wherein for the training the labels preferably are ground truth labels provided by a physician annotating the results of the different imaging techniques as a whole for each training sample. The at least sub-structure of the AIE is then trained using a total loss function which is based on at least the first loss function, the second loss function and the third loss function. In particular, the total loss function may be a weighted sum of the first, second, and third loss functions, wherein the weight accorded to each loss function is a hyperparameter.

[0036] In some advantageous embodiments, refinements, or variants of embodiments, the training sample with the imaging results from the at least one first type of medical imaging technique replaced by the pre-defined blank is input into the at least sub-structure of the AIE in order to produce a first intermediate spatial feature map. The training sample with the medical imaging results from the at least one second type of medical imaging technique replaced by the pre-defined blank is input into the at least sub-structure of the AIE in order to produce a second intermediate spatial feature map. The unaltered training sample is input into the at least sub-structure of the AIE in order to produce a third intermediate spatial feature map. The intermediate spatial feature maps may be the direct product of the at least sub-structure of the AIE. Alternatively, the intermediate spatial feature maps may be the result of one or more further processing steps on these direct products, respectively. For example, a respective mask may be applied to the respective direct products.

[0037] In the present context, a spatial feature map may be understood to refer in particular to a feature map in which each feature indicates a property of a spatial location in the organ, and which comprises information about different spatial locations in the organ. For example, the spatial feature map may comprise a feature for each pixel or voxel of medical imaging data.

[0038] Preferably, the first loss function penalizes differences between the third intermediate spatial feature map and the first intermediate spatial feature map, and the second loss function penalizes differences between the third interme-diate spatial feature map and the second intermediate spatial feature map. In this way, the first and the second loss function push the at least sub-structure being trained towards a specific combination of spatial region of the human organ on the one hand, and medical imaging technique on the other hand.

[0039] In some advantageous embodiments, refinements, or variants of embodiments, the method comprises a step of pre-training at least a sub-structure of the artificial intelligence entity, AIE (or even the AIE as a whole). Pre-training may, for example, be performed using a set of labelled training data showing not only the organ of interest (e.g. the prostate) but showing a plurality of different organs with corresponding abnormality heat map labels so as to increase the number of training samples. Using pre-trained artificial intelligence entities, in particular pre-trained artificial neural networks, ANN, is an effective way to improve the starting performance of an artificial intelligence entity, AIE, to be trained.

[0040] In some advantageous embodiments, refinements, or variants of embodiments, the first sub-step comprises training a first copy (or: clone, or: instance) of the pre-trained sub-structure using the first loss function in order to obtain a first trained sub-structure. Moreover, the second sub-step comprises training a second copy (or: clone, or: instance) of the pre-trained sub-structure using the second loss function in order to obtain a second trained sub-structure. In other words, a common starting point, namely the pre-trained sub-structure, is provided, two copies of it are provided (or one original and one copy, to be precise), and the two copies are then separately trained using different loss functions.

[0041] Again, each loss function is independent of a certain spatial region (e.g. zone of the prostate), and preferably only depends on one specific spatial region. This may be achieved, for example, by using a spatial mask on both the ground truth label and the output of the sub-structure being trained, wherein the spatial mask comprises values of "1" for all pixels or voxels of the spatial region on which the loss function shall depend and further comprises values of "0" for all other pixels or voxels, specifically the pixels or voxels of the other spatial regions.

[0042] It shall be understood that, if more than two specific spatial regions with different properties and/or different responses to different medical imaging techniques are present with the human organ of interest, correspondingly more copies of the pre-trained sub-structure may be provided. Each of these copies is then trained separately with a corre-sponding loss function, wherein each loss function only depends on one of the spatial regions and is independent on the other spatial regions. For each spatial region a corresponding spatial mask may be provided.

[0043] In these embodiments, then (fine-tuned) trained sub-structures are provided by the method, each sub-structure fine-tuned to one specific spatial region of the human organ of interest. For the production mode (or: inference stage), then input samples can be provided to each of these sub-structures separately, and a total abnormality heat map may be generated by taking from each trained sub-structure the information about the spatial region for which it is fine-tuned, i.e. to which it corresponds (or: with which it is associated).

[0044] The invention also provides, according to a second aspect, a method for detecting abnormalities in medical imaging data of a human organ, in particular a prostate, the method comprising:

providing the first trained sub-structure and the second trained sub-structure, in particular according to any embodiment of the first aspect of the present invention;
inputting a data sample (or: input sample) comprising a plurality of imaging results for the human organ from different types of imaging techniques into the first trained sub-structure in order to generate a first intermediate spatial feature map;
inputting the same data sample into the second trained sub-structure in order to generate a second intermediate spatial feature map; and
generating an abnormality heat map (or: total abnormality heat map) based on the generated first intermediate spatial feature map and the generated second intermediate spatial feature map.

[0045] The abnormality heat map is generated by taking pixels or voxels for the second spatial region from the first intermediate spatial feature map and by taking pixels or voxels for the first spatial region from the second intermediate spatial feature map such that the abnormality heat map comprises pixels or voxels for the first and the second spatial region. Correspondingly, for N>2 two spatial regions to be distinguished, N (fine-tuned) trained sub-structures may be provided, each associated with a specific spatial region, and the (total) abnormality heat map is then generated by taking from each intermediate spatial feature map the pixels or voxels belonging to that associated specific spatial region only. In this way, it is ensured that every pixel or voxel of the input data sample is analyzed by a fine-tuned sub-structure trained specifically for the spatial region to which the pixel or voxel belong.

[0046] In some advantageous embodiments, refinements, or variants of embodiments, a segmentation (e.g. comprising, or consisting of, one or more segmentation masks represented by a corresponding segmentation mask channel) is provided for identifying at least the first spatial region and the second spatial region within the input data sample, and wherein the pixels or voxels are taken from the first intermediate spatial feature map and the second intermediate spatial feature map, respectively, based on the segmentation. The segmentation may be automatically provided by inputting the data sample into a trained artificial intelligence segmentation entity (e.g. a segmentation artificial neural network). Alternatively, the data sample may already comprise a segmentation, e.g. one or more segmentation mask (or: segmentation mask channels).

[0047] According to a third aspect, the invention provides a system for detecting abnormalities in medical imaging data of a human organ, in particular a prostate, the system comprising: an input interface for receiving a data sample comprising a plurality of imaging results for the human organ from different types of imaging techniques;
a computing device configured to implement a trained artificial intelligence entity, AIE, (in particular a trained artificial neural network, ANN), comprising a first trained sub-structure, a second trained sub-structure, and a fusion module;
wherein the first trained sub-structure is configured and trained to receive the data sample from the input interface and to provide, as its output based thereon, a first intermediate spatial feature map;
wherein the second trained sub-structure is configured and trained to receive the data sample from the input interface and to provide, as its output based thereon, a second intermediate spatial feature map;
wherein the fusion module is configured to generate an abnormality heat map by taking pixels or voxels for the second spatial region from the first intermediate spatial feature map and by taking pixels or voxels for the first spatial region from the second intermediate spatial feature map such that the abnormality heat map comprises pixels or voxels for the first and the second spatial region; and
an output interface for outputting at least the generated abnormality heat map.

[0048] The first and second trained sub-structure are preferably trained according to any embodiment of the method of the second aspect of the present invention.

[0049] The input interface and/or the output interface may be realized in hardware and/or software. The input interface and/or the output interface can each comprise one or more different communication channels using one or more different communication protocols (e.g. HTTP). Each of the input interface and/or the output interface can be configured to connect to a cable-bound data connection and/or to a wireless data connection such as Bluetooth, Zigbee, WiFi and so on. The input interface and/or the output interface may also be configured to connect to Ethernet networks.

[0050] The computing device may be realized as any device, or any means, for computing. For example, the computing device may comprise at least one data processing (or: calculating) unit such as at least one central processing unit, CPU, and/or at least one graphics processing unit, GPU, and/or at least one field-programmable gate array, FPGA,

and/or at least one application-specific integrated circuit, ASIC, and/or any combination of the foregoing. The computing device may comprise a working memory operatively coupled to the at least one processing unit and/or a non-transitory memory operatively connected to the at least one processing unit and/or the working memory.

[0051] The computing device may be partially or completely realized as a remote device, for example as a cloud computing platform.

[0052] In some advantageous embodiments, variants, or refinements of embodiments, the computing device is further configured to implement a trained artificial intelligence segmentation entity configured and trained for receiving the data sample as its input and to output a segmentation as its output based thereon, and wherein the fusion module is configured to take the pixels or voxels based on the segmentation for the generating of the abnormality heat map.

[0053] According to a fourth aspect, the invention provides a system for detecting abnormalities in medical imaging data of a human organ, in particular a prostate, the system comprising: an input interface for receiving a data sample comprising a plurality of imaging results for the human organ from different types of imaging techniques;

a computing device configured to implement an artificial intelligence entity, AIE, trained according to a method according to any embodiment of the first aspect of the present invention,

wherein the trained AIE, preferably a trained artificial neural network, is configured and trained to receive the data sample from the input interface and to provide, as its output based thereon, an abnormality heat map; and

an output interface for outputting at least the generated abnormality heat map.

According to a fifth aspect, the invention also provides a use of an artificial intelligence entity, AIE, trained according to a method according to any embodiment of the first aspect of the present invention, for detecting abnormalities in medical imaging data of a human organ, in particular a prostate. The trained AIE is preferably a trained artificial neural network ANN. Preferably, the use of the trained AIE, particularly trained ANN, is in a method according to any embodiment of the second aspect of the invention.

[0054] According to a sixth aspect, the invention also provides a use of an artificial intelligence entity, AIE, trained according to a method according to any embodiment of the first aspect of the present invention, in a system for detecting abnormalities in medical imaging data of a human organ, in particular a prostate. The trained AIE is preferably a trained artificial neural network ANN. The system is preferably a system according to any embodiment of the third or fourth aspect.

[0055] According to a seventh aspect of the present invention, a computer program product is provided comprising executable program code configured to, when executed by a computing device, to perform the method according to any embodiment of the first aspect of the present invention and/or according to any embodiment of the second aspect of the present invention.

[0056] According to an eighth aspect of the present invention, a non-transitory, computer-readable data storage medium is provided, which comprises program code configured to, when executed by a computing device, executes the method according to an embodiment of the first aspect of the present invention and/or according to any embodiment of the second aspect of the present invention. Such a data storage medium may be, for example, a USB stick, a CD ROM, a DVD ROM, a Blue-Ray Disc, a hard drive, a solid state drive and/or the like.

[0057] According to a ninth aspect of the present invention, a data stream is provided, which comprises program code (or which is configured to generate program code), configured to, when executed by a computing device, perform the method according to any embodiment of the first aspect of the present invention and/or according to any embodiment of the second aspect of the present invention. The data stream may, for example, be provided by a server and be downloaded by a client or a user terminal or by a personal computer or a laptop.

[0058] Additional advantageous variants, refinements, embodiments and aspects of the invention will become more obvious in connection with the following description with reference to the drawings.

[0059] It will be understood that specific features, options or variants as are described with respect to specific aspects of the present invention may be equally, similarly or analogously also be implement in embodiments according to other aspects of the present invention.

Brief Description of the Drawings

[0060] The invention will be explained in yet greater detail with reference to exemplary embodiments depicted in the drawings as appended.

[0061] The accompanying drawings are included to provide a further understanding of the present invention and are incorporated in and constitute a part of the specification. The drawings illustrate the embodiments of the present invention and together with the description serve to illustrate the principles of the invention. Other embodiments of the present invention and many of the intended advantages of the present invention will be readily appreciated as they become better understood by reference to the following detailed description. Like reference numerals designate corresponding similar parts.

[0062] The numbering of method steps is intended to facilitate understanding and should not be construed, unless explicitly stated otherwise, or implicitly clear, to mean that the designated steps have to be performed according to the

numbering of their reference signs. In particular, several or even all of the method steps may be performed simultaneously, in an overlapping way or sequentially.

[0063] In the drawings:

Fig. 1   shows a schematic flow diagram illustrating a method for training an artificial intelligence entity according to an embodiment;

Fig. 2   schematically illustrates the steps of the method of Fig. 1;

Fig. 3   shows a schematic flow diagram illustrating a method for detecting abnormalities within medical imaging data;

Fig. 4   schematically illustrates the steps of the method of Fig. 3;

Fig. 5   shows a schematic flow diagram illustrating a method for training an artificial intelligence entity according to another embodiment;

Fig. 6   schematically illustrates the steps of the method of Fig. 5;

Fig. 7   shows a schematic flow diagram illustrating another method for detecting abnormalities within medical imaging data;

Fig. 8   shows a schematic block diagram illustrating a system according to another embodiment;

Fig. 9   shows a schematic block diagram illustrating a system according to yet another embodiment;

Fig. 10   shows a schematic block diagram illustrating a computer program according to another embodiment; and

Fig.11   shows a schematic block diagram illustrating a data storage medium according to yet another embodiment.

[0064] Although specific embodiments have been illustrated and described herein, it will be appreciated by those of ordinary skill in the art that the variety of alternate and/or equivalent implementations may be substituted for the specific embodiments shown and described without departing from the scope of the present invention. Generally, this application is intended to cover any adaptations or variations of the specific embodiments discussed herein.

Detailed Description of the Invention

[0065] Fig. 1 is a schematic flow diagram illustrating a method for training an artificial intelligence entity, AIE, for detecting abnormalities in medical imaging data of a human organ, in particular a human prostate. The general concepts will be described with respect to the human prostate and with respect to its peripheral zone, PZ, and its transition zone, TZ, but it will be apparent that these concepts are not in any way limited to this particular application and can be readily applied to numerous other organs and/or zones thereof. The method is in particular applicable to organs with distinct and disjunct zones (or: regions) which respond differently to (or: are scanned differently by) different medical imaging techniques.

[0066] Fig. 2 schematically illustrates the steps of the method of Fig. 1.

[0067] The method according to Fig. 1 provides an artificial intelligence entity, AIE, in particular artificial neural network, ANN, which essentially and effectively mimics the decision-making procedure of a trained radiologist. The main idea is to add additional inputs from the T2W scans for the information within the transition zone, TZ, and to add additional inputs from the DWI scans for the information within the peripheral zone, PZ.

[0068] In a step S10, medical imaging data of the human organ, in particular the prostate, are provided as training data. The training data is organized in training samples, wherein each training sample comprises a plurality of imaging results from different types of imaging techniques. These imaging results all show the same organ, i.e. they all comprise pixels or voxels that relate to spatial positions of the organ. The training samples are also provided with labels as ground truths for the training, i.e. with additional information, preferably for each pixels or voxels, about which result of the artificial intelligence entity, AIE, is to be achieved.

[0069] In the examples described herein, the result is usually an abnormality heat map, although other possible results will be readily appreciated by the skilled person. The heat map can be a feature map of entries, wherein each entry corresponds to one pixel of the sample used as input. The entries may for example be binary, e.g. 0 - no abnormality likely or 1 - abnormality likely, or percentile, e.g. 0 - abnormality highly unlikely, 0.5 - equal chances of abnormality or

no abnormality, 1 - abnormality highly likely. In the latter case, the heat map may also be designated as a probability heat map, since its entries indicate a probability for each pixels or voxels to belong to a particular abnormality, e.g. a cancerous growth or lesion. Other types of heat maps may also be envisaged.

[0070] Thus, in the case of an abnormality heat map, the training samples are provided with labels for each pixels or voxels which indicates the desired entry of the heat map. During training, the parameters of the artificial intelligence entity, AIE, or at least sub-structures or parts of it, will be adjusted so as to minimize the losses resulting from differences between the actually output abnormality heat map and the abnormality heat map according to the provided labels.

[0071] The label abnormality heat maps are provided, for example, by trained physicians which review the medical imaging results comprised in each training sample and mark which pixels, according to their expertise, belong to abnormalities and which do not.

[0072] As has been described in detail, various results from various imaging techniques may be used to form the training samples (and will, accordingly, also be used as input during the production, or inference, phase of the artificial intelligence entity, AIE).

[0073] Since the present examples focus mostly on the human prostate, training samples comprising results of a T2W imaging scan and further comprising results of a DWI scan will now be discussed for the sake of simplicity, although it shall be understood that training samples may include any other types of imaging scan results as well, in particular for other organs. It shall also be understood that DWI here may signify different sub-types of DWI scans, for example a DWI/ADC sequence or a DWI High B sequence.

[0074] In a step S20, a pre-trained or randomly initialized artificial intelligence entity, AIE, is provided. In particular, the artificial intelligence entity, AIE, may comprise, or consist of, an artificial neural network, ANN, in particular a deep artificial neural network, DNN. It will be apparent to the skilled person that different types of artificial neural networks, ANN, in particular DNNs, may be employed as or within the artificial intelligence entity, AIE, in particular a U-net, a U-net with Res-Net blocks or any of the other types of artificial neural network, ANN, as described in WO 2019/238804 A1.

[0075] The providing S20 of the artificial intelligence entity, AIE, may comprise a sub-step S21 of providing a randomly initialized artificial intelligence entity, AIE (for example an ANN with randomly initialized weights and biases) and pre-training S22 the artificial intelligence entity, AIE, for example using a set of labelled training data showing not only the organ of interest (herein: the prostate) but showing a plurality of different organs with corresponding abnormality heat map labels so as to increase the number of training samples. Providing a pre-trained artificial intelligence entity, AIE, is preferred since this always improves the performance of the artificial intelligence entity, AIE.

[0076] In the following, as an example, the case will be described that at least a pre-trained artificial neural network, ANN, is provided.

[0077] In a sub-step S23, the provided pre-trained artificial neural network, ANN, is then cloned (i.e. copied once) so that two identical pre-trained artificial neural networks, ANN, are obtained: a first pre-trained artificial neural network, ANN, and a second pre-trained artificial neural network, ANN.

[0078] In a step S30, for each of the training samples a segmentation is obtained, i.e. information about which pixels or voxels of the medical imaging scan results belong to which of a plurality of pre-specified regions of the human organ. In the case of the prostate, in particular the transition zone, TZ, and the peripheral zone, PZ, are useful to distinguish from one another, so that the segmentation for each training sample may comprise one or more mask channels indicating which pixels or voxels belong to the peripheral zone, PZ, and which pixels or voxels belong to the transition zone, TZ. Preferably, a first mask channel indicates for each pixel whether it belong to the peripheral zone, PZ, by a corresponding "1" in the channel, or whether it does not belong to the peripheral zone, PZ, by a corresponding "0" in the channel, and a second mask channel indicates for each pixel whether it belong to the transition zone, TZ, by a corresponding "1" in the channel, or whether it does not belong to the transition, TZ, by a corresponding "0" in the channel.

[0079] For the training samples, the segmentations may be provided simply as part of the training data, just as the labels are provided. For example, the physician who has classified the pixels or voxels of the medical imaging scan results into pixels or voxels containing abnormalities or no abnormalities may also have provided a segmentation of the same medical imaging scan results into (at least) transition zone, TZ, and peripheral zone, PZ.

[0080] As an alternative, in a sub-step S35 of step S30, the medical imaging scan results of the training samples may be input into a segmentation network (i.e. an artificial neural network, ANN, trained for producing the desired segmentation). Advantageously, since usually in the later production (or: inference) phase no segmentation is a priori available in any case, the same segmentation network may also be used then.

[0081] In one or more (generally a large number of) first sub-steps S41 of training S40, the first pre-trained artificial neural network, ANN, is then trained to obtain a PZ-specific sub-structure, and in one or more (generally a large number of) second sub-steps S42 of the training S40, the second pre-trained artificial neural network, ANN, is trained to obtain a PZ-specific sub-structure.

[0082] For the first sub-step S41, a first loss function is calculated for each training sample, wherein the first loss function only depends on pixels or voxels of the peripheral zone, PZ, according to the segmentation, and is independent from pixels or voxels of the transition zone, TZ. In more abstract words, the calculation S41 of the first loss function is

independent of a first spatial region (the transition zone, TZ), and is only dependent on a second spatial region (the peripheral zone, PZ).

[0083] This may e.g. be achieved in that only differences between the output of the PZ-specific sub-structure and the ground truth label in the pixels or voxels of the peripheral zone, PZ, are penalized during training. In other words, the first loss function is only changed when any of the pixels or voxels within the peripheral zone, PZ, is changed but remains unchanged for any change in any of the pixels or voxels within the transition zone, TZ.

[0084] One simple way to realize this is, for example, to use a mask that gives a 1 for all pixels or voxels of the peripheral zone, PZ. This mask may then be applied automatically to both the output of the PZ-specific sub-structure and to the ground truth for the respective training sample. The first loss function may be designed to penalize pixel- (or voxel-) wise difference between these two heat maps, e.g. using an L2 metric. Now, since all pixels or voxels not of the peripheral zone, PZ, will have a value of zero both in the output of the PZ-specific sub-structure and the ground truth, the difference between them will be zero and there will be no penalization by the first loss function. This also means that when the PZ-specific sub-structure is trained using the first loss function, it will learn specifically to find accurate results for the peripheral zone, PZ, while its results for the transition zone, TZ, will be unsupervised.

[0085] The training of the PZ-specific sub-structure is done in the usual way based on the first loss function, i.e. in an iterative process parameters (e.g. weights and biases) of the sub-structure are updated based on a back-propagation algorithm in order to minimize the first loss function. Thus, one iteration of the first sub-step S41 may comprise inputting one or more training samples into the PZ-specific sub-structure, generating an output of the sub-structure based thereon, calculating the first loss function for one or more training samples, and then use a back-propagation algorithm to change parameters of the PZ-specific sub-structure such as to best minimize the first loss function (e.g. gradient descent).

[0086] For the second sub-step S42, a second loss function is calculated for each training sample, wherein the second loss function only depends on pixels or voxels of the transition zone, TZ, according to the segmentation, and is independent from pixels or voxels of the peripheral zone, PZ. In other words, the calculation S42 of the second loss function is independent of the second spatial region (the transition zone, TZ), and is only dependent on the first spatial region (the peripheral zone, PZ).

[0087] This may e.g. be achieved in that only differences between the output of the TZ-specific sub-structure and the ground truth label in the pixels or voxels of the transition zone, TZ, are penalized during training. In other words, the second loss function is only changed when any of the pixels or voxels within the transition zone, TZ, is changed but remains unchanged for any change in any of the pixels or voxels within the peripheral zone, PZ. This can be realized in the same way as has been described in the foregoing for the first sub-step S41.

[0088] As a result of the repeated performing of the first and second sub-steps S41, S42, then a trained PZ-specific sub-structure and a training TZ-specific sub-structure are provided. Each of these sub-structures is fine-tuned for detecting abnormalities within their respective zone and produce less reliable (or even random) results for the pixels or voxels outside of the respective zone.

[0089] These sub-structures may then be used individually, or they may be combined in an artificial intelligence entity, AIE, or in a pipeline, to produce even better or more complete results, as will be explained in more detail in the following.

[0090] Fig. 2 shows a schematic block diagram illustrating one possible realization of the method of Fig. 1. A training sample 1 (e.g. multi-parametric MRI images) which have been provided in step S10 are input into a segmentation network 10 in order to generate the segmentation 12 for the training sample 1. The segmentation 12 may comprise a first segmentation mask 12-1 with a "1" for each pixels or voxels that belongs to the peripheral zone, PZ, and with a "0" for each pixel or voxel that belongs to the transition zone, TZ, and a second segmentation mask 12-2 with a "1" for each pixel or voxel that belongs to the transition zone, TZ, and with a "0" for each pixel or voxel that belongs to the peripheral zone, PZ.

[0091] The training sample 1 is also input into each of two identical copies (clones) 20-1, 20-2 of the same pre-trained artificial neural network, ANN. One copy 20-1 is the starting point for the PZ-specific sub-structure, and one copy 20-2 is the starting point for the TZ-specific sub-structure.

[0092] Based on the input training sample 1, each of the sub-structures 20-1, 20-2 produces an output. Using the segmentation 12, for the PZ-specific sub-structure a first loss function 22-1 is calculated which only takes into account pixels or voxels from the peripheral zone, PZ, and which is independent of the pixels or voxels from the transition zone, TZ.

[0093] Similarly, using the segmentation 12, for the TZ-specific sub-structure a second loss function 22-2 is calculated which only takes into account pixels or voxels from the transition zone, TZ, and which is independent of the pixels or voxels from the peripheral zone, PZ. As has been described in the foregoing, the training samples 1 are provided with labels (more specifically: label abnormality heat maps), and the first and/or second loss function may be realized as any known loss function for penalizing differences between the values of the label abnormality heat maps for each pixels or voxels and the corresponding output heat map of the sub-structures 20-1, 20-2. For example, binary cross entropy (BCE) loss functions may be used.

[0094] By repeated performing of the respective sub-steps S41, S42 then final results for the PZ-specific sub-structure 24-1 and for the TZ-specific sub-structure 24-2 are provided.

[0095] Fig. 3 shows a schematic flow diagram for illustrating a method for detecting abnormalities in medical imaging data of a human organ, in particular of a prostate. Fig. 4 shows a schematic block diagram illustrating one possible realization of the method of Fig. 3. The methods of Fig. 3 and/or Fig. 4 may be seen as realizations of an inference stage (or: production mode) of the artificial intelligence entity, AIE, which has been trained with the method as described with respect to Fig. 1 and/or Fig. 2.

[0096] In a step S110, a first and a second trained sub-structure are provided, namely a trained PZ-specific sub-structure 24-1 and a trained TZ-specific sub-structure 24-2, in particular in the way as has been described in the foregoing with respect to Fig. 1 and/or Fig. 2.

[0097] In a step S120, a data sample 2 comprising a plurality of imaging results for the human organ (in particular prostate) from different types of imaging techniques, e.g. a real-world multi-parametric MRI image, is input into the trained PZ-specific sub-structure 24-1 in order to generate a first intermediate spatial feature map 26-1 (here specifically: abnormality heat map).

[0098] In a step S130, the data sample 2 is input also into the trained TZ-specific sub-structure 24-2 to generate a second intermediate spatial feature map 26-2 (here specifically: abnormality heat map).

[0099] In a step S140, a total abnormality heat map 28 is generated based on the first intermediate spatial feature map 26-1 and the second intermediate spatial feature map 26-2, specifically such that the abnormality heat map 28 comprises the pixels or voxels for the peripheral zone, PZ, from the first intermediate spatial feature map 26-1 and the pixels or voxels for the transition zone, TZ, from the second intermediate spatial feature map 26-2. The term "total" here refers to the fact that the total abnormality heat map 28 comprises the best possible results not only for one zone (or: region) of the human organ but for all of its zones.

[0100] The total abnormality heat map 28 can, for example, be generated as illustrated with Fig. 4:
The data sample 2 may be input into a segmentation network 10, preferably the same segmentation network 10 that has already been used for providing the PZ-specific sub-structure 24-1 and the TZ-specific sub-structure 24-2. The first segmentation mask 12-1 may be convolved (pixel-wise or voxel-wise multiplied) with the output of the trained PZ-specific sub-structure 24-1 in order to generate a processed first intermediate spatial feature map 26-1, and the second segmentation mask 12-2 be convolved (pixel-wise or voxel-wise multiplied) with the output of the trained TZ-specific sub-structure 24-2 in order to generate a processed second intermediate spatial feature map 26-2. In this way, the thus processed intermediate spatial features maps 26-1, 26-2 each comprise non-zero values only for the pixels or voxels of their respective corresponding zone (or: region) and zero values otherwise.

[0101] Then, these processed intermediate spatial feature maps 26-1, 26-2 can simply be added (or: fused) to generate the total abnormality heat map 28. Since in the present example the total abnormality heat map indicates the possibility (binary entries) or likelihood (percentile entries) malignant lesions, it may also be designated as a lesion malignancy map.

[0102] Fig. 5 is a schematic flow diagram illustrating a further method for training an artificial intelligence entity, AIE, for detecting abnormalities in medical imaging data of a human organ, in particular a human prostate. Again, the general concepts will be described with respect to the human prostate and with respect to its peripheral zone, PZ, and its transition zone, TZ, but it will be apparent that these concepts are not in any way limited to this particular application and can be readily applied to numerous other organs and/or zones thereof. The method is in particular applicable to organs with distinct and disjunct zones (or: regions) which respond differently to (or: are scanned differently by) different medical imaging techniques.

[0103] Fig. 6 schematically illustrates the steps of the method of Fig. 5.

[0104] The method according to Fig. 5 provides an end-to-end artificial intelligence entity, AIE, in particular end-to-end artificial neural network, ANN, which essentially and effectively mimics the decision-making procedure of a trained radiologist. The main idea is to add additional inputs from the T2W scans for the information within the transition zone, TZ, and to add additional inputs from the DWI scans for the information within the peripheral zone, PZ.

[0105] In a step S210, medical imaging data of the human organ, in particular the prostate, are provided as training data 201, as illustrated in Fig. 6. The training data 201 is organized in training samples, wherein each training sample comprises a plurality of imaging results 202-1, 202-2 from different types of imaging techniques. In Fig. 6, again for the example of the human prostate, the training data 201 comprises first imaging results 202-1 from DWI scans, and second imaging results 202-2 from T2W scans.

[0106] The training data 201 also comprise, in this example, at least one mask channel 202-3 which provides segmentation information, i.e. information about which pixels or voxels belong to the transition zone, TZ, or to the peripheral zone, PZ, respectively. In some variants, as has been described in the foregoing, the at least one mask channel may be automatically generated based on the other channels with the imaging results 202-1, 202-2 in a foregoing method step, in particular using a segmentation network 10 as has been described in the foregoing.

[0107] Preferably, a first mask channel indicates for each pixel whether it belong to the peripheral zone, PZ, by a corresponding "1" in the channel, or whether it does not belong to the peripheral zone, PZ, by a corresponding "0" in the channel, and a second mask channel indicates for each pixel whether it belong to the transition zone, TZ, by a corresponding "1" in the channel, or whether it does not belong to the transition, TZ, by a corresponding "0" in the channel.

[0108] The training data 201 are also provided with labels in form of a label abnormality heat map for each of the training samples as has been described in the foregoing with respect to Fig. 1 and Fig. 2.

[0109] In a step S220, a pre-trained or randomly initialized artificial intelligence entity, AIE, is provided. In particular, the artificial intelligence entity, AIE, may comprise, or consist of, an artificial neural network, ANN, in particular a deep artificial neural network, DNN. It will be apparent to the skilled person that different types of artificial neural networks, ANN, in particular DNNs, may be employed as or within the artificial intelligence entity, AIE, in particular a U-net, a U-net with Res-Net blocks or any of the other types of artificial neural network, ANN, e.g. as described in WO 2019/238804 A1.

[0110] The providing S220 of the artificial intelligence entity, AIE, may comprise a sub-step S221 of providing a randomly initialized artificial intelligence entity, AIE (for example an ANN with randomly initialized weights and biases) and pre-training S222 the artificial intelligence entity, AIE, for example using a set of labelled training data showing not only the organ of interest (herein: the prostate) but showing a plurality of different organs with corresponding abnormality heat map labels so as to increase the number of training samples. Providing a pre-trained artificial intelligence entity, AIE, is preferred since this always improves the performance of the artificial intelligence entity, AIE.

[0111] In the following, as an example, the case will be described that at least a pre-trained artificial neural network, ANN 204, is provided.

[0112] Then, during training S230, this provided pre-trained artificial neural network, ANN 204, is further trained with the provided training data 201 using a total loss function L_{total} which comprises, or consists of, three separate loss function terms that are added, preferably using a weighted sum:

$$L_{total}=\lambda_{BCE}{}^{*}L_{BCE} \ + \ \lambda_{DWI}{}^{*}L_{DWI} \ + \ \lambda_{T2W}{}^{*}L_{T2W} \qquad\qquad (1)$$

[0113] These terms will be explained in detail in the following.

[0114] Referring to Fig. 6, three separate outputs are produced with the same ANN 204 in the same training state (i.e. with the same current set of parameters, e.g. weights and biases), based on three different inputs:

a first output 206-1 (or intermediate spatial feature map) is generated by the ANN 204 as a result of inputting a training sample of the training data 201, wherein the first imaging results, from the DWI scan (DWI channel), have been replaced by a blank 202-4;

a second output 206-2 (or intermediate spatial feature map) is generated by the ANN 204 as a result of inputting the same training sample of the training data 201, wherein the second imaging results, from the T2W scan (T2W channel), have been replaced by a blank 202-4; and

a third output 206-3 (or intermediate spatial feature map) is generated by the ANN 204 as a result of inputting the same training sample of the training data 201, without any previous alterations.

[0115] A "blank" 202-4 may herein be a channel of the same spatial dimension as the imaging result channels but with all of its pixels or voxels set to the same value, e.g. to "0", to "1", to "0.5" or so, depending on the specific realization and the values used in the imaging results.

[0116] As a result, the first output 206-1 will in principle be more suitable for detecting abnormalities in the peripheral zone, PZ, as it contains, apart from the - for this region dominant - first imaging results 202-1 from the DWI scan less "noise", i.e. less (or none at all) non-dominant imaging results. Similarly, the second output 206-2 will in principle be more suitable for detecting abnormalities in the transition zone, TZ, as it contains, apart from the - for this region dominant - second imaging results 202-2 from the T2W scan less "noise", i.e. less (or none at all) non-dominant imaging results. Finally, the third output 206-3 will be balanced, taking all of the imaging results 202-1, 202-2 into account.

[0117] These outputs 206-1, 206-2, 206-3 are then used to calculate S234 the total loss function L_{total} in order to improve, via back-propagation 208, the parameters of the ANN 204 in an iterative process.

[0118] Returning to the formula (1) above for the total loss function L_{total}, the term L_{BCE}, multiplied with a weighting factor λ_{BCE} as one hyperparameter, is a binary cross entropy term which simply penalizes, in the usual way, the difference (or: loss) between the ground truth according to the labels of the training data 201 and the third output 206-3, preferably pixel-wise (or voxel-wise, when the medical imaging results 202-1, 202-2 are provided as voxels instead of pixels, or when in a pre-processing steps pixels are transformed into voxels or the like). L_{BCE} is calculated in a sub-step S233.

[0119] Moving on to the next term, the symbol L_{DWI} symbolizes a first loss function which penalizes differences (again, preferably pixel- or voxel-wise) between the first output 206-1 and the third output 206-3 but only as far as the peripheral zone, PZ, is concerned. In other words, the first loss function is independent of any pixels or voxels within the transition zone, TZ (or, for that matter, in any other region of the prostate).

[0120] Thus, L_{DWI} may be calculated S231 as follows, with the first output 206-1 designated as Out_{DWI} and with the third output 206-3 designated as Out_{total}:

$$L_{DWI} = ||y_{PZ} * (Out_{total} - Out_{DWI})||^2$$

with the L2-norm $||.||^2$, with * here a elementwise (i.e. pixel-wise) product, and with y_{PZ} a mask (first mask channel) having 1's for pixels or voxels that belong to the peripheral zone, PZ, and with 0's for pixels or voxels that do not belong to the peripheral zone, PZ, or which in particular belong to the transition zone, TZ. As in the case of λ_{BCE}, λ_{DWI} is a hyperparameter giving a relative weight to the first loss function within the total loss function L_{total}.

[0121] Again moving on to the next term, the symbol L_{T2W} symbolizes a second loss function which penalizes differences (again, preferably pixel- or voxel-wise) between the second output 206-2 and the third output 206-3 but only as far as the transition zone, TZ, is concerned. In other words, the second loss function is independent of any pixels or voxels within the peripheral zone, PZ (or, for that matter, in any other region of the prostate).

[0122] Thus, L_{T2W} may be calculated S232 as follows, with the second output 206-2 designated as Out_{T2W} and with the other symbols as defined above:

$$L_{T2W} = ||y_{TZ} * (Out_{total} - Out_{T2W})||^2,$$

with y_{TZ} a mask (second mask channel) having 1's for pixels or voxels that belong to the transition zone, TZ, and with 0's for pixels or voxels that do not belong to the transition zone, TZ, or which in particular belong to the peripheral zone, PZ. As in the case of λ_{BCE}, λ_{T2W} is a hyperparameter giving a relative weight to the second loss function within the total loss function L_{total}.

[0123] Advantageously, the first and second mask channels (y_{PZ} and y_{TZ}) are not only used for the calculation of the respective loss functions but also as further input into the ANN 204 besides the medical imaging data themselves.

[0124] In some variants, λ_{T2W} is equal to λ_{DWI}, and it is preferred that $\lambda_{BCE} \geq 2 \lambda_{DWI}$, more preferably $\lambda_{BCE} \geq 3 \lambda_{DWI}$, still more preferably $\lambda_{BCE} \geq 4 \lambda_{DWI}$. However, the hyperparameters λ_{T2W}, λ_{BCE} and λ_{DWI} may also be the result of a grid search to find optimal hyperparameters on a training dataset.

[0125] Again, the described is a specific example for the case of the human prostate and for the case of exactly two zones of the prostate and for exactly two medical imaging scans, T2W and DWI. The principle is easily applied to any other type of problem: the basic idea is that in addition to the third output 206-3 which is based on all the information additional outputs 206-1, 206-2 are produced in which selectively some types of imaging techniques are removed from the underlying input. Specifically, each output is associated with a specific region of the human organ (here: first output 206-1 with the peripheral zone, PZ, and second output 206-2 with the transition zone, TZ).

[0126] When then each such specific region is associated with a particular imaging technique, then at least some of the other imaging techniques, and preferably all of the other imaging techniques, which are not specifically associated with that region, are removed from the underlying input, i.e. are replaced by blanks. Then, for each such specific region a corresponding loss function term of the total loss function is calculated which takes into account only the pixels or voxels within the corresponding specific region and which penalizes differences between the third output 206-3 and the output associated with the specific region.

[0127] Thus, the L_{BCE} term will drive the updates of the ANN 204 towards identity with the ground truth, wherein the other, additional loss functions within the total loss function will drive the ANN 204 towards becoming better at determining abnormalities within the respective corresponding regions based on the dominant imaging technique for the regions.

[0128] The updating of the parameters of the ANN 204 is repeated for as long as desired in order to then provide the ANN 204 as an artificial intelligence entity, AIE, for determining lesions.

[0129] Fig. 7 shows a schematic flow diagram illustrating a method for detecting abnormalities in medical imaging data of a human organ, in particular a prostate.

[0130] In a step S310, an artificial intelligence entity, AIE, is provided according to the method as it has been described with respect to Fig. 5 and Fig. 6. In a step S320, an input sample is input into the artificial intelligence entity, AIE, such as trained ANN 204 wherein the input sample comprises first imaging results 202-1 from a DWI scan, and second imaging results 202-2 from a T2W scans, as well as mask information 202-3, e.g. in form of two mask channels, one for the transition zone, TZ, and one for the peripheral zone, PZ, as described in the foregoing. As a result, an abnormality heat map 228 is generated as the output of the artificial intelligence entity, AIE, which may in particular be a probability heat map of the entire prostate. In this variant, no late fusion is needed to combine different feature maps.

[0131] Fig. 8 shows a schematic block diagram illustrating a system 100 according to an embodiment of the third aspect of the present invention, i.e. a system 100 for detecting abnormalities such as lesions in medical imaging data of a human organ, in particular a prostate. The system 100 is in particular suited for performing the method as it has

been described with respect to Fig. 3 and Fig. 4. Thus, the system 100 can be realized with any of the details, variants or modifications as have been described with respect to the method according to Fig. 3 and Fig. 4 and vice versa.

[0132] The system 100 comprises an input interface 110 for receiving a data sample 2 comprising a plurality of imaging results for the human organ, in particular prostate, from different types of imaging techniques. The data sample 2 may be provided to the input interface 110, and received thereby, directly from a medical imaging device such as an MR scanner and/or it may be provided from a medical imaging results repository, such as a picture archiving and communication system, PACS.

[0133] The system further comprises a computing device 150 configured to implement a trained artificial intelligence entity, AIE, (in particular a trained artificial neural network, ANN), comprising a first trained sub-structure 24-1, a second trained sub-structure 24-2, and a fusion module 151. The first and the second trained sub-structures 24-1, 24-2 may in particular have been provided by the training method as has been described with respect to Fig. 1 and Fig. 2.

[0134] The first trained sub-structure 24-1 is configured and trained to receive the data sample 2 from the input interface 110 and to provide, as its output based thereon, a first intermediate spatial feature map 26-1. The second trained sub-structure 24-2 is configured and trained to receive the same data sample 2 from the input interface 110 and to provide, as its output based thereon, a second intermediate spatial feature map 26-2.

[0135] The fusion module 151 is configured to generate S140 an abnormality heat map 28 by taking pixels or voxels for the second spatial region from the first intermediate spatial feature map 26-1 and by taking pixels or voxels for the first spatial region from the second intermediate spatial feature map 26-2 such that the abnormality heat map 28 comprises pixels or voxels for the first and the second spatial region.

[0136] The system 100 also comprises an output interface 190 for outputting at least the generated abnormality heat map 28.

[0137] Optionally, the computing device 150 is further configured to implement a trained artificial intelligence segmentation entity 10 configured and trained for receiving the data sample 2 as its input and to output a segmentation 12 as its output based thereon. The fusion module 151 may then be configured to take the pixels or voxels based on the segmentation 12 for the generating S140 of the abnormality heat map 28.

[0138] Fig. 9 shows a schematic block diagram illustrating a system 100 according to an embodiment of the fourth aspect of the present invention, i.e. a system 100 for detecting abnormalities such as lesions in medical imaging data of a human organ, in particular a prostate. The system 100 is in particular suited for performing the method as it has been described with respect to Fig. 7. Thus, the system 100 can be realized with any of the details, variants or modifications as have been described with respect to the method according to Fig. 7 and vice versa.

[0139] Accordingly, the system 100 according to Fig. 9 comprises an input interface 110, an output interface 190 and a computing device 150 as it has been described with respect to Fig. 8, with the difference that the computing device 150 is configured to implement the trained ANN 204 trained using the method as described with respect to Fig. 5 and 6.

[0140] The trained ANN 204 is configured and trained to receive the data sample 2 from the input interface 110 and to provide, as its output based thereon, an abnormality heat map 28.

[0141] Fig. 10 shows a schematic block diagram illustrating a computer program product 800 according to an embodiment of the seventh aspect of the present invention, i.e. a computer program product 800 comprising executable program code 850 configured to, when executed by a computing device 150 perform the method according to any embodiment of the first aspect of the present invention or according to any embodiment of the second aspect of the present invention, in particular any of the methods as described with respect to Fig. 1 to 7.

[0142] Fig. 11 shows a schematic block diagram illustrating a non-transitory, computer-readable data storage medium 900 according to an embodiment of the eighth aspect of the present invention, i.e. a data storage medium 900 which comprises program code 950 configured to, when executed by a computing device 150, execute the method according to any embodiment of the first aspect of the present invention or according to any embodiment of the second aspect of the present invention, in particular any of the methods as described with respect to Fig. 1 to 7. Such a data storage medium may be, for example, a USB stick, a CD ROM, a DVD ROM, a Blue-Ray Disc, a hard drive, a solid state drive and/or the like.

[0143] In the foregoing detailed description, various features are grouped together in the examples with the purpose of streamlining the disclosure. It is to be understood that the above description is intended to be illustrative and not restrictive. It is intended to cover all alternatives, modifications and equivalence. Many other examples will be apparent to one skilled in the art upon reviewing the above specification, taking into account the various variations, modifications and options as described or suggested in the foregoing.

Claims

1. A computer-implemented method for training an artificial intelligence entity, AIE, for detecting abnormalities in medical imaging data of a human organ, in particular a prostate, the method comprising the steps of :

providing (S10; S210) medical imaging data of the human organ as training data comprising training samples (1; 201), the medical imaging data comprising a plurality of imaging results (202-1; 202-2) from different types of imaging techniques for each training sample (1; 201) of the training data;

providing (S20; S21; S220) at least one pre-trained or randomly initialized artificial intelligence entity, AIE (20-1, 20-2); and

training (S40; S230) the AIE (20-1, 20-2) using the provided training samples (1; 201),

wherein in a first sub-step (S41; S231), for at least one training sample (1; 201), a first loss function (22-1; $L2_{DWI}$) for at least a sub-structure (20-1) of the at least one AIE is calculated independently of at least one first spatial region of the human organ, and

wherein in a second sub-step (S42; S232), for at least one training sample (1; 201), a second loss function (22-2; $L2_{T2W}$) for at least a sub-structure (20-2) of the at least one AIE is calculated independently of at least one second spatial region of the human organ;

wherein the AIE (20-1, 20-2) is trained (S40; S230) using at least the calculated first loss function (22-1; L_{DWI}) and the calculated second loss function (22-2; L_{T2W}).

2. The method of claim 1,
wherein the first spatial region and the second spatial region are disjunct.

3. The method of claim 2,
wherein the first loss function (22-1; $L2_{DWI}$) is based only on pixels or voxels of the provided medical imaging data within the second spatial region; and
wherein the second loss function (22-2; $L2_{T2W}$) is based only on pixels or voxels of the provided medical imaging data within the first spatial region.

4. The method of any of claims 1 to 3,
wherein in the first sub-step (S231) and in the second sub-step (S232) subsequently the corresponding loss function for the same at least sub-structure (204) of the AIE is calculated.

5. The method of any of claims 1 to 4,
wherein for the first sub-step (S231) the imaging results from the at least one first type (202-2) of imaging technique within each training sample (201) of the at least one training sample are replaced by a pre-defined blank (202-4) before being used for calculating (S231) the first loss function ($L2_{DWI}$), and wherein for the second sub-step (S232) the imaging results from the at least one second type (202-1) of imaging technique within each training sample (201) of the same at least one training sample are replaced by the pre-defined blank (202-4) before being used for calculating (S232) the second loss function.

6. The method of claim 5,
wherein in a third sub-step (S233), a third loss function (L_{BCE}) is calculated for each training sample (201) of the same at least one training sample;
wherein in the third sub-step (S233) each training sample (201) is left unaltered before being used for calculating the third loss function;
wherein the third loss function (L_{BCE}) when calculated penalizes differences between the output (206-3) of the at least sub-structure (204) of the AIE and corresponding labels for each training sample (201); and
wherein the at least sub-structure of the AIE is trained using a total loss function (L_{total}) which is based on at least the first loss function ($L2_{DWI}$), the second loss function ($L2_{T2W}$) and the third loss function (L_{BCE}).

7. The method of claim 6,
wherein the training sample (201) with the imaging results from the at least one first type (202-2) of imaging technique replaced by the pre-defined blank (202-3) is input into the at least sub-structure (204) of the AIE in order to produce a first intermediate spatial feature map (206-1);
the training sample with the imaging results from the at least one second type of imaging technique replaced by the pre-defined blank (202-4) is input into the at least sub-structure (204) of the AIE in order to produce a second intermediate spatial feature map (206-2); and
the unaltered training sample (201) is input into the at least sub-structure (204) of the AIE in order to produce a third intermediate spatial feature map (206-3);
wherein the first loss function ($L2_{DWI}$) when calculated penalizes differences between the third intermediate spatial feature map (206-3) and the first intermediate spatial feature map (206-1), and the second loss function ($L2_{T2W}$) penalizes differences between the third intermediate spatial feature map (206-3) and the second intermediate spatial

feature map (206-2).

8. The method of any of claims 1 to 3,
comprising a step of pre-training (S22) at least a sub-structure of the AIE and a step of cloning (S23) the at least sub-structure of the AIE so as to obtain a first pre-trained sub-structure (20-1) and a second pre-trained sub-structure (20-2) ;
wherein the first sub-step (S41) comprises training the first pre-trained sub-structure (20-1) using the first loss function (22-1) in order to obtain a first trained sub-structure (24-1); and
wherein the second sub-step (S42) comprises training the second pre-trained sub-structure (20-2) using the second loss function (22-2) in order to obtain a second trained sub-structure (24-1).

9. A computer-implemented method for detecting abnormalities in medical imaging data of a human organ, in particular a prostate, the method comprising:

providing (S110) a first trained sub-structure (24-1) and a second trained sub-structure (24-2), in particular according to the method of claim 8;
inputting (S120) a data sample (2) comprising a plurality of imaging results for the human organ from different types of imaging techniques into the first trained sub-structure (24-1) in order to generate a first intermediate spatial feature map (26-1);
inputting (S130) the same data sample (2) into the second trained sub-structure (24-2) in order to generate a second intermediate spatial feature map (26-2); and
generating (S140) an abnormality heat map (28) based on the generated first intermediate spatial feature map (26-1) and the generated second intermediate feature map (26-2);
wherein the abnormality heat map (28) is generated by taking pixels or voxels for the second spatial region from the first intermediate spatial feature map (26-1) and by taking pixels or voxels for the first spatial region from the second intermediate spatial feature map (26-2) such that the abnormality heat map (28) comprises pixels or voxels for the first and the second spatial region.

10. The method of claim 9,
wherein a segmentation (12) is provided for identifying at least the first spatial region and the second spatial region within the input data sample (2), and wherein the pixels or voxels are taken from the first intermediate spatial feature map (26-1) and the second intermediate spatial feature map (26-2), respectively, based on the segmentation (12).

11. The method of claim 10,
wherein the segmentation (12) is automatically provided by inputting the data sample (2) into a trained artificial intelligence segmentation entity (10), or wherein the data sample (2) comprises at least one segmentation mask.

12. A system (100) for detecting abnormalities in medical imaging data of a human organ, in particular a prostate, the system comprising:

an input interface (110) for receiving a data sample (2) comprising a plurality of imaging results for the human organ from different types of imaging techniques;
a computing device (150) configured to implement a trained artificial intelligence entity, AIE, in particular a trained artificial neural network, ANN, comprising a first trained sub-structure (24-1), a second trained sub-structure (24-2), and a fusion module (151);
wherein the first trained sub-structure (24-1) is configured and trained to receive the data sample (2) from the input interface (110) and to provide, as its output based thereon, a first intermediate spatial feature map (26-1);
wherein the second trained sub-structure (24-2) is configured and trained to receive the same data sample (2) from the input interface (110) and to provide, as its output based thereon, a second intermediate spatial feature map (26-2);
wherein the fusion module (151) is configured to generate (S140) an abnormality heat map (28) by taking pixels or voxels for the second spatial region from the first intermediate spatial feature map (26-1) and by taking pixels or voxels for the first spatial region from the second intermediate spatial feature map (26-2) such that the abnormality heat map (28) comprises pixels or voxels for the first and the second spatial region; and
an output interface (190) for outputting at least the generated abnormality heat map (28).

13. The system (100) of claim 12,
wherein the computing device (150) is further configured to implement a trained artificial intelligence segmentation

entity (10) configured and trained for receiving the data sample (2) as its input and to output a segmentation (12) as its output based thereon, and wherein the fusion module (151) is configured to take the pixels or voxels based on the segmentation (12) for the generating (S140) of the abnormality heat map (28).

14. A computer program product (800) comprising executable program code (850) configured to, when executed, perform the method according to any of claims 1 to 11.
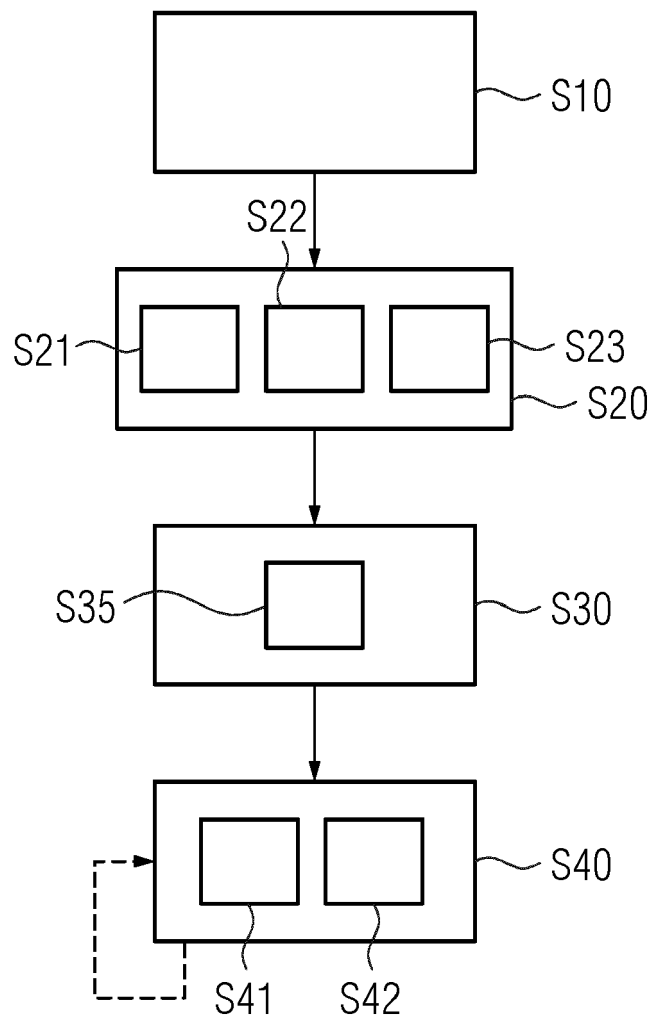
15. A computer-readable, non-transitory data storage medium (900) comprising executable program code (950) configured to, when executed, perform the method according to any of claims 1 to 11.

FIG 1

FIG 2
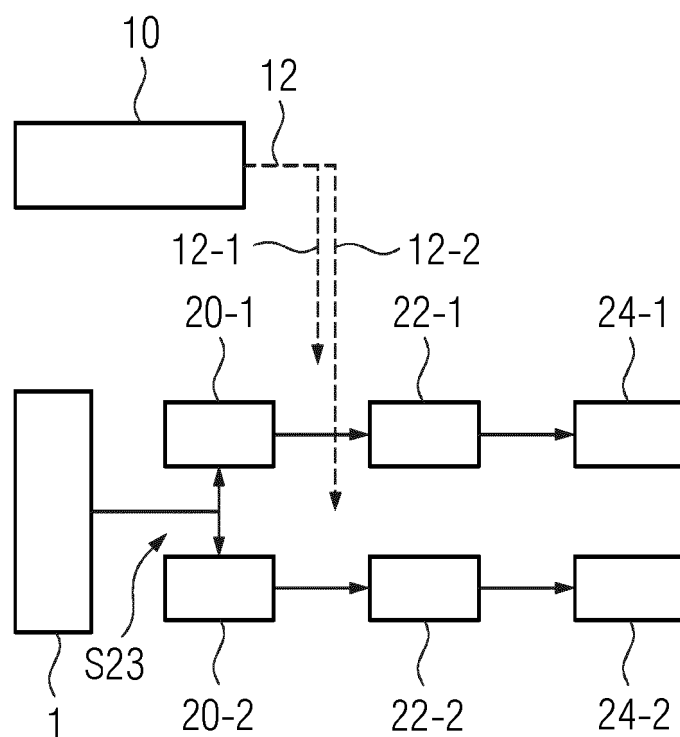
FIG 3
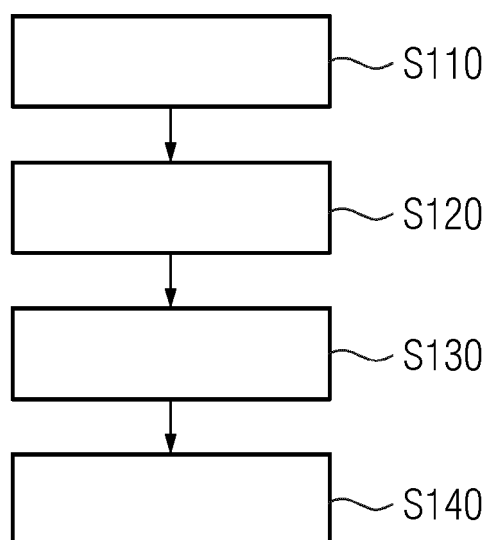
FIG 4
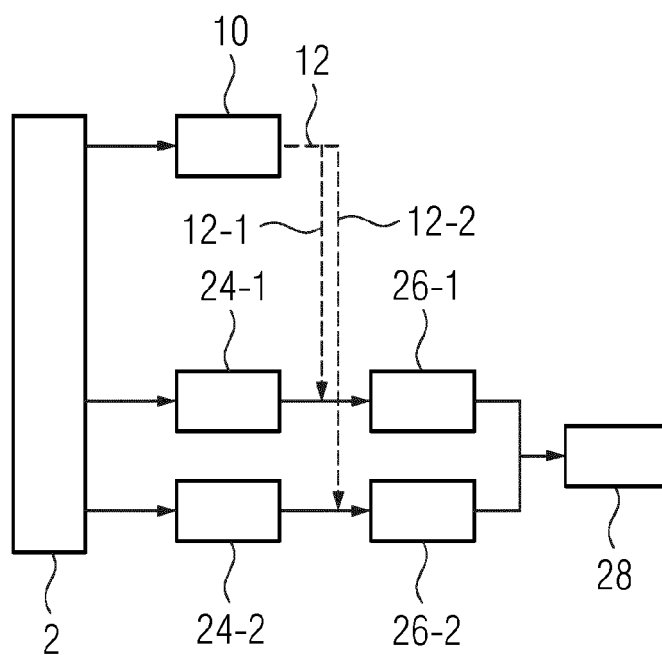
FIG 5
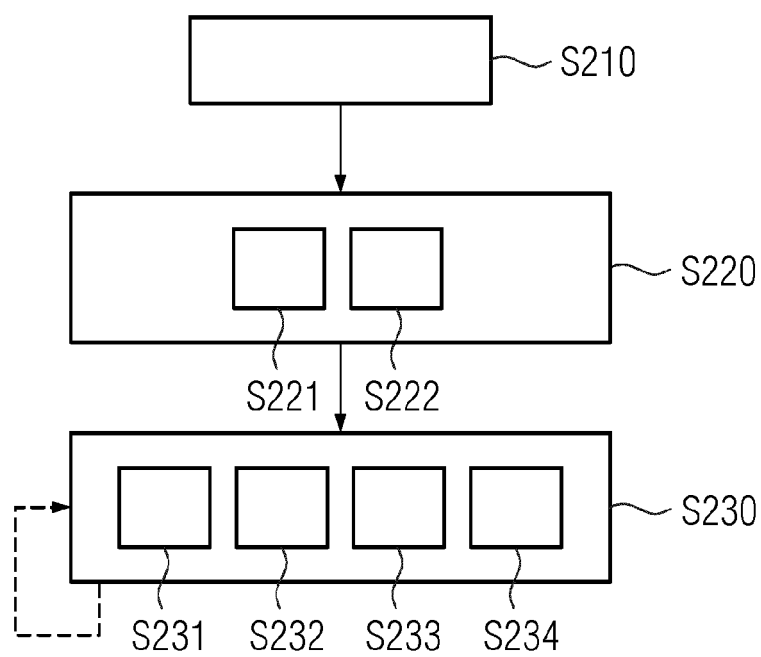
FIG 6
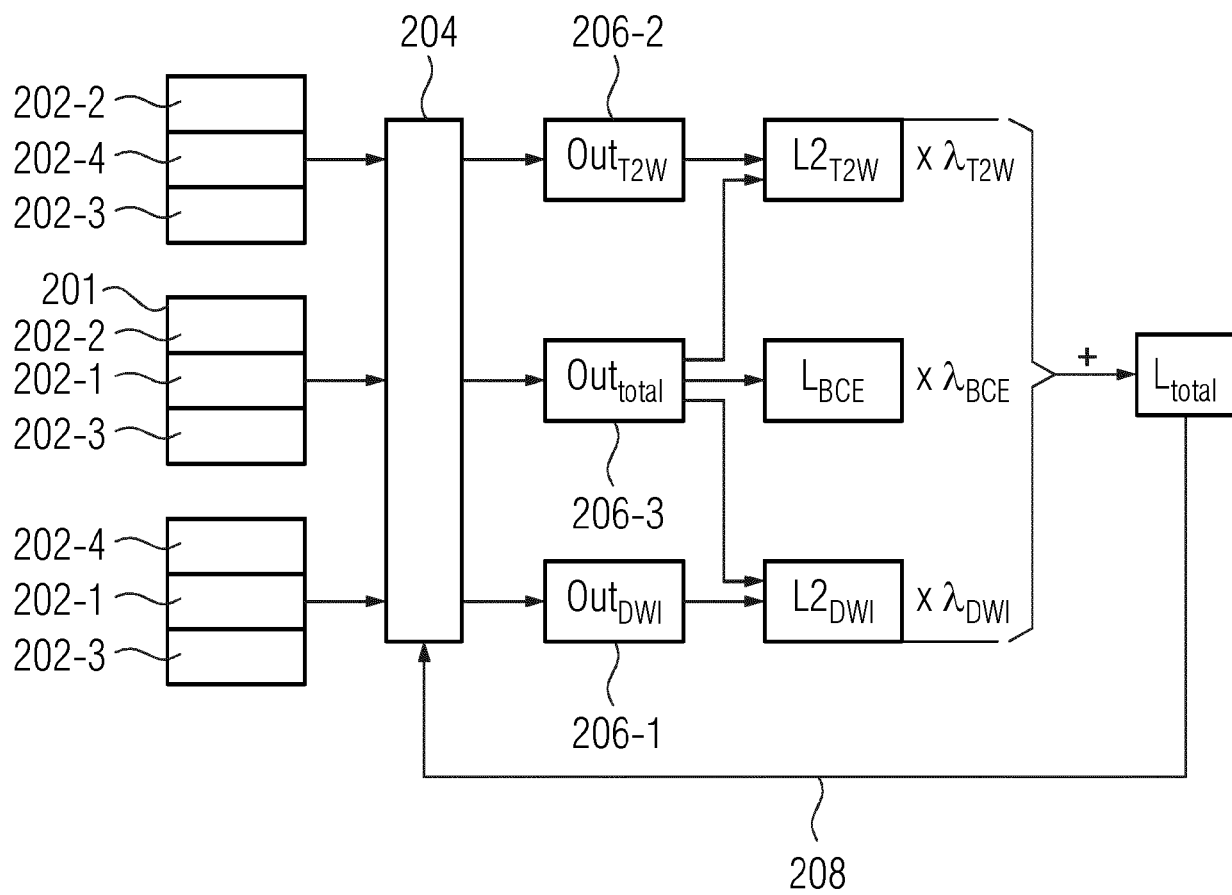
FIG 7
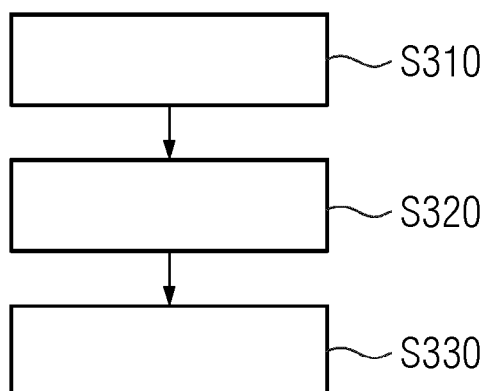
FIG 8

FIG 9

FIG 10

FIG 11

Europäisches Patentamt
European Patent Office
Office européen des brevets
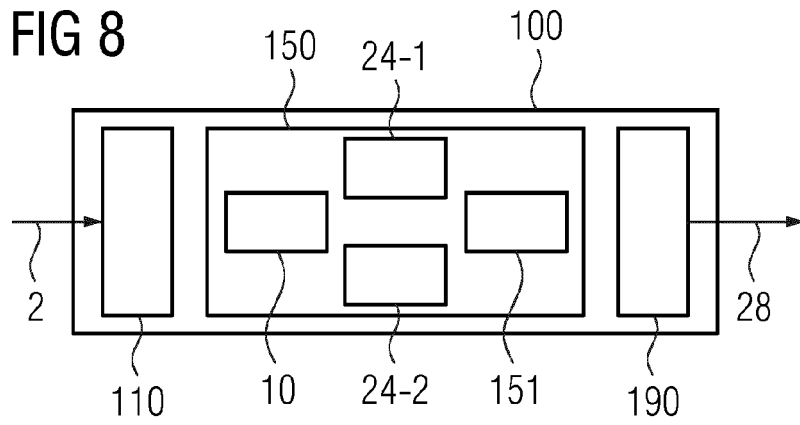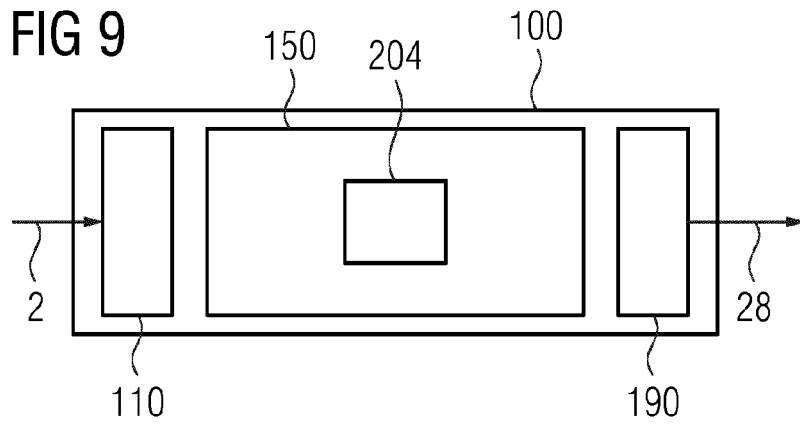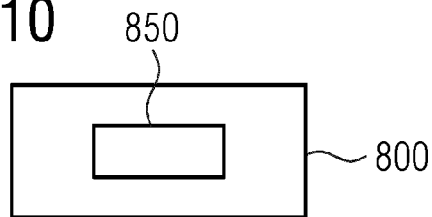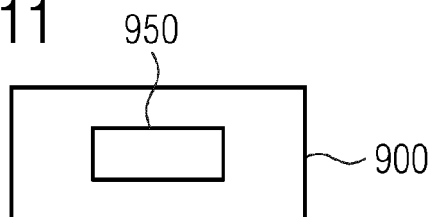
EUROPEAN SEARCH REPORT

Application Number

EP 20 16 7487

DOCUMENTS CONSIDERED TO BE RELEVANT

Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (IPC)
X	YANG XIN ET AL: "Co-trained convolutional neural networks for automated detection of prostate cancer in multi-parametric MRI", MEDICAL IMAGE ANALYSIS, vol. 42, 24 August 2017 (2017-08-24), pages 212-227, XP085240557, ISSN: 1361-8415, DOI: 10.1016/J.MEDIA.2017.08.006 * abstract * * left column; page 216, line 22; figures 1, 3, 5, 6 * * section 2 *	1-15	INV. G06T7/11
A	SAED ASGARI TAGHANAKI ET AL: "Deep Semantic Segmentation of Natural and Medical Images: A Review", ARXIV.ORG, CORNELL UNIVERSITY LIBRARY, 201 OLIN LIBRARY CORNELL UNIVERSITY ITHACA, NY 14853, 16 October 2019 (2019-10-16), XP081516776, * the whole document *	1-15	
A	US 2020/012761 A1 (EL-BAZ AYMAN S [US] ET AL) 9 January 2020 (2020-01-09) * the whole document *	1-15	TECHNICAL FIELDS SEARCHED (IPC) G06T
A,D	WO 2019/238804 A1 (SIEMENS HEALTHCARE GMBH [DE]) 19 December 2019 (2019-12-19) * the whole document *	1-15	

The present search report has been drawn up for all claims

Place of search	Date of completion of the search	Examiner
Munich	10 September 2020	Blaszczyk, Marek

ANNEX TO THE EUROPEAN SEARCH REPORT
ON EUROPEAN PATENT APPLICATION NO.

EP 20 16 7487

This annex lists the patent family members relating to the patent documents cited in the above-mentioned European search report.
The members are as contained in the European Patent Office EDP file on
The European Patent Office is in no way liable for these particulars which are merely given for the purpose of information.

10-09-2020

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
US 2020012761 A1	09-01-2020	NONE	
WO 2019238804 A1	19-12-2019	NONE	

EPO FORM P0459

For more details about this annex : see Official Journal of the European Patent Office, No. 12/82

REFERENCES CITED IN THE DESCRIPTION

This list of references cited by the applicant is for the reader's convenience only. It does not form part of the European patent document. Even though great care has been taken in compiling the references, errors or omissions cannot be excluded and the EPO disclaims all liability in this regard.

Patent documents cited in the description

- WO 2019238804 A1 **[0007] [0074] [0109]**

Non-patent literature cited in the description

- **RHIANNON VAN LOENHOUT ; FRANK ZIJTA ; ROBIN SMITHUIS ; IVO SCHOOTS.** *Prostate Cancer - PI-RADS v2,* 01 August 2018, http://radiologyassitant.nl/abdomen/prostate-cancer-pi-rads-v2 **[0004]**