(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
12 December 2024 (12.12.2024)

WIPO | PCT

(10) International Publication Number
WO 2024/254027 A2

(51) International Patent Classification:
Not classified

(21) International Application Number:
PCT/US2024/032342

(22) International Filing Date:
04 June 2024 (04.06.2024)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:
63/506,233 05 June 2023 (05.06.2023) US

(71) Applicant: **THE JOHNS HOPKINS UNIVERSITY** [US/US]; 3400 N. Charles Street, Baltimore, MD 21218 (US).

(72) Inventors: **BAI, Harrison**; c/o THE JOHNS HOPKINS UNIVERSITY, 3400 N. Charles Street, Baltimore, MD 21218 (US). **POMPER, Martin Gilbert**; c/o THE JOHNS HOPKINS UNIVERSITY, 3400 N. Charles Street, Baltimore, MD 21218 (US). **ROWE, Steven P.**; C/o The Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218 (US). **JIAO, Zhicheng**; 593 Eddy Street, Providence, RI 02903 (US). **LI, Yang**; 76 Goddard Street, Providence, RI 02908 (US). **CHEN, Junyu**; c/o THE JOHNS HOPKINS UNIVERSITY, 3400 N. Charles Street, Baltimore, MD 21218 (US). **LEAL, Jeff**; c/o THE JOHNS HOPKINS UNIVERSITY, 3400 N. Charles Street, Baltimore, MD 21218 (US).

(74) Agent: **SAPPENFIELD, Christopher C.**; MH2 Technology Law Group, 1951 Kidwell Drive, Suite 310, Tysons Corner, VA 22182 (US).

(81) Designated States *(unless otherwise indicated, for every kind of national protection available)*: AE, AG, AL, AM, AO, AT, AU, AZ, BA, BB, BG, BH, BN, BR, BW, BY, BZ, CA, CH, CL, CN, CO, CR, CU, CV, CZ, DE, DJ, DK, DM, DO, DZ, EC, EE, EG, ES, FI, GB, GD, GE, GH, GM, GT, HN, HR, HU, ID, IL, IN, IQ, IR, IS, IT, JM, JO, JP, KE, KG, KH, KN, KP, KR, KW, KZ, LA, LC, LK, LR, LS, LU, LY, MA, MD, MG, MK, MN, MU, MW, MX, MY, MZ, NA, NG, NI, NO, NZ, OM, PA, PE, PG, PH, PL, PT, QA, RO, RS, RU, RW, SA, SC, SD, SE, SG, SK, SL, ST, SV, SY, TH, TJ, TM, TN, TR, TT, TZ, UA, UG, US, UZ, VC, VN, WS, ZA, ZM, ZW.

(84) Designated States *(unless otherwise indicated, for every kind of regional protection available)*: ARIPO (BW, CV, GH, GM, KE, LR, LS, MW, MZ, NA, RW, SC, SD, SL, ST, SZ, TZ, UG, ZM, ZW), Eurasian (AM, AZ, BY, KG, KZ, RU, TJ, TM), European (AL, AT, BE, BG, CH, CY, CZ, DE, DK, EE, ES, FI, FR, GB, GR, HR, HU, IE, IS, IT, LT, LU, LV, MC, ME, MK, MT, NL, NO, PL, PT, RO, RS, SE, SI, SK, SM, TR), OAPI (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, KM, ML, MR, NE, SN, TD, TG).

Published:
— *without international search report and to be republished upon receipt of that report (Rule 48.2(g))*

(54) Title: MACHINE LEARNING METHODS AND RELATED ASPECTS FOR PROCESSING MEDICAL IMAGES

(57) Abstract: Examples may provide a method of segmenting test image data sets obtained from a test subject. The method includes using a first trained machine learning model to segment non-suspicious foci in the test image data set to produce segmented non-suspicious foci data. The test image data set comprises positron emission tomography (PET) images and/or computed tomography (CT) images. The method also includes applying a mask to the non-suspicious foci in the test image data set using the segmented non-suspicious foci data to produce masked and unmasked portions of the test image data set. In addition, the method also includes using a second trained machine learning model and the unmasked portions of the test image data set to segment suspicious foci in the test image data to produce segmented suspicious foci data. Related methods, systems, and computer readable media are also provided.

WO 2024/254027 A2

MACHINE LEARNING METHODS AND RELATED ASPECTS FOR PROCESSING MEDICAL IMAGES

Cross-Reference To Related Applications

[0001] This application claims priority to, and the benefit of, U.S. Provisional Patent Application Ser. No. 63/506,233, filed June 5, 2023, the disclosure of which is incorporated herein by reference.

Field

[0002] This disclosure relates generally to machine learning, e.g., in the context of medical applications, such as diagnostics.

Background

[0003] Prostatic cancer (PCa) is the most common male malignancy and the third most common cause of cancer mortality. The management of PCa has recently been revolutionized by the introduction of prostate-specific membrane antigen (PSMA) positron emission tomography/computed tomography (PET-CT) which is the most accurate staging tool now available. PSMA PET/CT is increasingly being used to assess treatment response in patients with biochemically recurrent or metastatic PCa. Accurate delineation of the lesions on PSMA PET/CT is a prerequisite for estimating whole body tumor burden and in guiding treatment approaches. However, currently, determining whole body tumor burden on PSMA PET-CT is cumbersome and requires the manual annotation by radiologists, which is not only time-consuming but also requires significant domain expertise. Furthermore, there is significant inter- and intra-reader variability that may affect the interpretation, and the subsequent treatment response assessment. Semi-automatic software for whole body tumor segmentation on PSMA-PET imaging has been developed but are still problematic when facing patients with multiple lesions. Such tools can take up to 20 minutes per patient when disease is widespread. Hence, computer aided lesion segmentation and classification are currently not feasible for immediate implementation in clinical practice. Thus, development of automatic segmentation and classification algorithms for lesion increased uptake on PSMA PET/CT scans is of great interest.

[0004] Artificial intelligence (AI) has shown promise in automating time-consuming image processing tasks such as segmentation. These convolutional neural network (CNN)-based approaches have proven highly effective and reproducible in segmenting anatomy and lesions on cross sectional images of all types and classifying uptake on PET/CT. Despite the existing success and the popularity of CNN-based algorithms, they have not yet been used widely in uptake segmentation and classification on [^{18}F] DCFPyL PET/CT. Each PET agent has its own biodistribution in normal tissue so that each AI must be uniquely trained for each agent. One specific challenge is that the model cannot distinguish normal organs from target lesions when both exhibit the same PET avidity such as the salivary or lacrimal glands.

[0005] Accordingly, there is an urgent need for novel and generalizable deep neural network (DNN) models which automatically segment and classify uptake patterns into foci suspicious and non-suspicious for disease states, such as cancer from whole body PET/CT images of patients with, for example, biochemically recurrent and/or metastatic prostate cancer.

Summary

[0006] The present disclosure provides, in certain aspects, an artificial intelligence (AI) system capable of segmenting and/or classifying image data sets that comprise positron emission tomography (PET) images and/or computed tomography (CT) images. These and other aspects will be apparent upon a complete review of the present disclosure, including the accompanying figures.

[0007] According to various embodiments, a method of segmenting a test image data set obtained from a test subject using a computer is presented. The method includes using, by the computer, a first trained machine learning model to segment one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject. The method also includes applying, by the computer, a mask to the non-suspicious foci in the test image data set using the segmented non-suspicious foci data to produce masked and unmasked portions

of the test image data set. In addition, the method also includes using, by the computer, a second trained machine learning model and the unmasked portions of the test image data set to segment one or more suspicious foci in the test image data to produce segmented suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject.

[0008] According to various embodiments, a method of segmenting a test image data set obtained from a test subject using a computer is presented. The method includes segmenting, by the computer, one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject; excluding, by the computer, the non-suspicious foci from the test image data set using the segmented non-suspicious foci data to produce suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject; and, segmenting, by the computer, the suspicious foci data, thereby segmenting the test image data set obtained from the subject using the computer.

[0009] According to various embodiments, a method of classifying a test image data set obtained from a test subject using a computer is presented. The method includes using, by the computer, a first trained machine learning model to produce a first set of predicted probabilities that selected foci in one or more positron emission tomography (PET) images from the test image data set are each individually classified as being either suspicious or non-suspicious; using, by the computer, a second trained machine learning model to produce a second set of predicted probabilities that the selected foci in one or more computed tomography (CT) images from the test image data set are each individually classified as being suspicious or non-suspicious; using, by the computer, a third trained machine learning model to produce a third set of predicted probabilities that the selected foci in one or more fusions of the one or more PET images and the one or more CT images from the test image data set are each individually classified as being suspicious or non-suspicious; and, generating, by the computer, a final pattern classification of the selected foci as each individually being suspicious or non-suspicious using averaged predicted probabilities of the first,

second, and third sets of predicted probabilities, thereby classifying the test image data set obtained from the test subject using the computer.

[0010] Various optional features of the above embodiments include the following. An electronic neural network comprises the first and/or second trained machine learning model. The electronic neural network comprises one or more 3D hybrid transformer-convolutional neural networks (CNNs). The first and/or second trained machine learning model has been trained on a set of training data that comprises a plurality of reference image data sets obtained from reference subjects that each comprise one or more PET images and/or one or more CT images, wherein a given reference image data set in the plurality of reference image data sets comprises at least one PET image and/or at least one CT image that is labeled with one or more ground truth suspicious foci and/or one or more ground truth non-suspicious foci. The one or more ground truth suspicious foci and/or the one or more ground truth non-suspicious foci are assigned by a consensus of at least two experts. The test image data set and the reference image data sets comprise prostate-specific membrane antigen (PSMA) PET and/or CT images. The method comprises classifying selected loci in the test image data set that exhibit increased imaging agent uptake relative to other regions in the test image data set as the non-suspicious foci or as the suspicious foci using a multi-modal decision fusion framework. The method further comprises estimating a whole-body disease burden value for the test subject by calculating suspicious foci volume. The method does not comprise setting a predefined threshold of a PET standardized uptake value (SUV) and/or applying any manual corrections. The test subject is a human subject. One or more organs comprise at least some of the non-suspicious foci. The method comprises administering a [^{18}F] DCFPyL imaging agent to the test subject prior to obtaining the test image data set from the test subject. The non-suspicious and suspicious foci in the test image data set exhibit an increased uptake of an imaging agent relative to other regions in the test image data set. The method comprises administering a therapy to the test subject to treat the disease state based at least in part upon the segmented test image data set. The method comprises segmenting multiple test image data sets obtained from the test subject at different time points to assess treatment response of the test subject to the therapy. The disease state comprises a cancer type. The cancer type comprises

prostate cancer. The selected foci exhibit increased imaging agent uptake relative to other regions in the test image data set. The method further comprises using, by the computer, a third trained machine learning model to segment one or more suspicious foci in the masked portion of the test image data set to produce additional segmented suspicious foci data. The method further comprises combining, by the computer, the additional segmented suspicious foci data with the segmented suspicious foci data produced using the second trained machine learning model.

[0011] According to various embodiments, a system for segmenting a test image data set obtained from a test subject using an electronic neural network is presented. The system includes a processor; and a memory communicatively coupled to the processor, the memory storing instructions which, when executed on the processor, perform operations including: using a first trained machine learning model to segment one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject; applying a mask to the non-suspicious foci in the test image data set using the segmented non-suspicious foci data to produce masked and unmasked portions of the test image data set; and using a second trained machine learning model and the unmasked portions of the test image data set to segment one or more suspicious foci in the test image data to produce segmented suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject. In some embodiments, the memory storing instructions which, when executed on the processor, further perform operations comprising: using a third trained machine learning model to segment one or more suspicious foci in the masked portion of the test image data set to produce additional segmented suspicious foci data; and, combining the additional segmented suspicious foci data with the segmented suspicious foci data produced using the second trained machine learning model.

[0012] According to various embodiments, a system for segmenting a test image data set obtained from a test subject using an electronic neural network is presented. The system includes a processor; and a memory communicatively coupled

to the processor, the memory storing instructions which, when executed on the processor, perform operations including: segmenting one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject; excluding the non-suspicious foci from the test image data set using the segmented non-suspicious foci data to produce suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject; and segmenting the suspicious foci data.

[0013] According to various embodiments, a system for segmenting a test image data set obtained from a test subject using an electronic neural network is presented. The system includes a processor; and a memory communicatively coupled to the processor, the memory storing instructions which, when executed on the processor, perform operations including: using a first trained machine learning model to produce a first set of predicted probabilities that selected foci in one or more positron emission tomography (PET) images from the test image data set are each individually classified as being either suspicious or non-suspicious; using a second trained machine learning model to produce a second set of predicted probabilities that the selected foci in one or more computed tomography (CT) images from the test image data set are each individually classified as being suspicious or non-suspicious; using a third trained machine learning model to produce a third set of predicted probabilities that the selected foci in one or more fusions of the one or more PET images and the one or more CT images from the test image data set are each individually classified as being suspicious or non-suspicious; and generating a final pattern classification of the selected foci as each individually being suspicious or non-suspicious using averaged predicted probabilities of the first, second, and third sets of predicted probabilities.

[0014] According to various embodiments, a computer readable media is presented. The computer readable media comprises non-transitory computer executable instructions which, when executed by at least one electronic processor, perform at least: using a first trained machine learning model to segment one or more

non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject; applying a mask to the non-suspicious foci in the test image data set using the segmented non-suspicious foci data to produce masked and unmasked portions of the test image data set; and using a second trained machine learning model and the unmasked portions of the test image data set to segment one or more suspicious foci in the test image data to produce segmented suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject. In some embodiments, the non-transitory computer executable instructions which, when executed by the electronic processor, further perform at least: using a third trained machine learning model to segment one or more suspicious foci in the masked portion of the test image data set to produce additional segmented suspicious foci data; and, combining the additional segmented suspicious foci data with the segmented suspicious foci data produced using the second trained machine learning model.

[0015] According to various embodiments, a computer readable media is presented. The computer readable media comprises non-transitory computer executable instructions which, when executed by at least one electronic processor, perform at least: segmenting one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject; excluding the non-suspicious foci from the test image data set using the segmented non-suspicious foci data to produce suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject; and segmenting the suspicious foci data.

[0016] According to various embodiments, a computer readable media is presented. The computer readable media comprises non-transitory computer executable instructions which, when executed by at least one electronic processor,

perform at least: using a first trained machine learning model to produce a first set of predicted probabilities that selected foci in one or more positron emission tomography (PET) images from the test image data set are each individually classified as being either suspicious or non-suspicious; using a second trained machine learning model to produce a second set of predicted probabilities that the selected foci in one or more computed tomography (CT) images from the test image data set are each individually classified as being suspicious or non-suspicious; using a third trained machine learning model to produce a third set of predicted probabilities that the selected foci in one or more fusions of the one or more PET images and the one or more CT images from the test image data set are each individually classified as being suspicious or non-suspicious; and generating a final pattern classification of the selected foci as each individually being suspicious or non-suspicious using averaged predicted probabilities of the first, second, and third sets of predicted probabilities.

Drawings

[0017] The above and/or other aspects and advantages will become more apparent and more readily appreciated from the following detailed description of examples, taken in conjunction with the accompanying drawings, in which:

[0018] Fig. 1A is a flow chart that schematically shows exemplary method steps of segmenting a test image data set obtained from a test subject using a computer according to some aspects disclosed herein;

[0019] Fig. 1B is a flow chart that schematically shows exemplary method steps of segmenting a test image data set obtained from a test subject using a computer according to some aspects disclosed herein;

[0020] Fig. 1C is a flow chart that schematically shows exemplary method steps of classifying a test image data set obtained from a test subject using a computer according to some aspects disclosed herein;

[0021] Fig. 2 is a schematic diagram of an exemplary system suitable for use with certain aspects disclosed herein;

[0022] Fig. 3 schematically illustrates an anatomical prior guided deep learning-based framework to segment suspicious foci on PSMA PET according to some aspects disclosed herein;

[0023] Fig. 4 schematically illustrates a multi-modal decision fusion framework for uptake pattern classification as suspicious and non-suspicious according to some aspects disclosed herein;

[0024] Fig. 5 show distributions of segmentation metrics of non-suspicious, suspicious and global foci. (a)-(d) and (e)-(h) correspond to settings (a)-(d) and (e)-(h) in Table 2 for internal and external testing sets, respectively;

[0025] Fig. 6 are visualizations of segmentation results for non-suspicious and suspicious foci on the internal and external testing sets; and

[0026] Fig. 7 are ROC curves of models in classifying increased uptake on PSMA PET (a) internal testing set, (b) external testing set.

Definitions

[0027] In order for the present disclosure to be more readily understood, certain terms are first defined below. Additional definitions for the following terms and other terms may be set forth throughout the specification. If a definition of a term set forth below is inconsistent with a definition in an application or patent that is incorporated by reference, the definition set forth in this application should be used to understand the meaning of the term.

[0028] As used in this specification and the appended claims, the singular forms "a," "an," and "the" include plural references unless the context clearly dictates otherwise. Thus, for example, a reference to "a method" includes one or more methods, and/or steps of the type described herein and/or which will become apparent to those persons skilled in the art upon reading this disclosure and so forth.

[0029] It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only and is not intended to be limiting. Further, unless defined otherwise, all technical and scientific terms used herein have the same meaning as commonly understood by one of ordinary skill in the art to which this disclosure pertains. In describing and claiming the methods, systems, and computer readable media, the following terminology, and grammatical variants thereof, will be used in accordance with the definitions set forth below.

[0030] ***Cancer Type***: As used herein, "cancer type" refers to a type or subtype of cancer defined, e.g., by histopathology. Cancer type can be defined by any

conventional criterion, such as on the basis of occurrence in a given tissue (e.g., blood cancers, central nervous system (CNS), brain cancers, lung cancers (small cell and non-small cell), skin cancers, throat cancers, nose cancers, liver cancers, bone cancers, lymphomas, pancreatic cancers, thyroid cancers, bladder cancers, kidney cancers, mouth cancers, stomach cancers, breast cancers, prostate cancers, bowel cancers, rectal cancers, ovarian cancers, intestinal cancers, soft tissue cancers, neuroendocrine cancers, lung cancers, gastroesophageal cancers, urothelial cancers, solid state cancers, heterogeneous cancers, homogenous cancers), head and neck cancers, gynecological cancers, colorectal cancers, unknown primary origin and the like, and/or of the same cell lineage (e.g., carcinoma, sarcoma, lymphoma, cholangiocarcinoma, leukemia, mesothelioma, melanoma, or glioblastoma) and/or cancers exhibiting cancer markers, such as Her2, CA15-3, CA19-9, CA-125, CEA, AFP, PSA, HCG, hormone receptor and NMP-22. Cancers can also be classified by stage (e.g., stage 1, 2, 3, or 4) and whether of primary or secondary origin.

[0031] ***Classifier***: As used herein, "classifier" generally refers to algorithm computer code that receives, as input, test data and produces, as output, a classification of the input data as belonging to one or another class.

[0032] ***Data set***: As used herein, "data set" refers to a group or collection of information, values, or data points related to or associated with one or more objects, records, and/or variables. In some embodiments, a given data set is organized as, or included as part of, a matrix or tabular data structure. In some embodiments, a data set is encoded as a feature vector corresponding to a given object, record, and/or variable, such as a given test or reference subject. For example, a medical data set for a given subject can include one or more observed values of one or more variables associated with that subject.

[0033] ***Electronic neural network***: As used herein, "electronic neural network" refers to a machine learning algorithm or model that includes layers of at least partially interconnected artificial neurons (e.g., perceptrons or nodes) organized as input and output layers with one or more intervening hidden layers that together form a network that is or can be trained to classify data, such as test subject medical data sets (e.g., medical images or the like).

[0034] ***Expert***: As used herein, “expert” refers to an entity that is trained to at least to a selected threshold level regarding a given knowledge domain. In some embodiments, an expert is a “human expert” such as a healthcare provider (e.g., a pathologist, radiologist, oncologist, or the like). In some embodiments, an expert is a “machine expert” such as a machine learning model that has been trained as to one or more aspects of the given knowledge domain.

[0035] ***Labeled***: As used herein, “labeled” in the context of data sets or points refers to data that is classified as, or otherwise associated with, having or lacking a given characteristic or property.

[0036] ***Machine Learning Algorithm***: As used herein, "machine learning algorithm" generally refers to an algorithm, executed by computer, that automates analytical model building, e.g., for clustering, classification or pattern recognition. Machine learning algorithms may be supervised or unsupervised. Learning algorithms include, for example, artificial neural networks (e.g., back propagation networks), discriminant analyses (e.g., Bayesian classifier or Fisher’s analysis), multiple-instance learning (MIL), support vector machines, decision trees (e.g., recursive partitioning processes such as CART -classification and regression trees, or random forests), linear classifiers (e.g., multiple linear regression (MLR), partial least squares (PLS) regression, and principal components regression), hierarchical clustering, and cluster analysis. A dataset on which a machine learning algorithm learns can be referred to as "training data." A model produced using a machine learning algorithm is generally referred to herein as a “machine learning model.”

[0037] ***Subject***: As used herein, “subject” or “test subject” refers to an animal, such as a mammalian species (e.g., human) or avian (e.g., bird) species. More specifically, a subject can be a vertebrate, e.g., a mammal such as a mouse, a primate, a simian or a human. Animals include farm animals (e.g., production cattle, dairy cattle, poultry, horses, pigs, and the like), sport animals, and companion animals (e.g., pets or support animals). A subject can be a healthy individual, an individual that has or is suspected of having a disease or pathology or a predisposition to the disease or pathology, or an individual that is in need of therapy or suspected of needing therapy. The terms “individual” or “patient” are intended to be interchangeable with “subject.” A

"reference subject" refers to a subject known to have or lack specific properties (e.g., a known pathology, such as melanoma and/or the like).

[0038] ***Value***: As used herein, "value" generally refers to an entry in a dataset that can be anything that characterizes the feature to which the value refers. This includes, without limitation, numbers, words or phrases, symbols (e.g., + or -) or degrees.

Description of the Embodiments

[0039] Reference will now be made in detail to example implementations. These embodiments are described in sufficient detail to enable those skilled in the art to practice the invention and it is to be understood that other embodiments may be utilized and that changes may be made without departing from the scope of the invention. The following description is, therefore, merely exemplary.

[0040] I. Description of Example Embodiments

[0041] In some aspects, the present disclosure provides computer-implemented methods of segmenting a test image data set obtained from a test subject. To illustrate, Fig. 1A is a flow chart that schematically shows certain of these exemplary method steps. As shown, method 100 includes using a first trained machine learning model to segment one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data (step 102). The test image data set typically includes one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images. In addition, the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject. Method 100 also includes applying a mask to the non-suspicious foci in the test image data set using the segmented non-suspicious foci data to produce masked and unmasked portions of the test image data set (step 104). In addition, method 100 also includes using a second trained machine learning model and the unmasked portions of the test image data set to segment one or more suspicious foci in the test image data to produce segmented suspicious foci data (step 106). The suspicious foci are suspicious for the presence of the disease state in the test subject. Although not shown in Fig. 1A, in some embodiments, method 100 further includes using, by the computer, a third trained machine learning model to segment one or more suspicious foci in the masked

portion of the test image data set to produce additional segmented suspicious foci data (e.g., suspicious foci contained within the non-suspicious foci, such as normal organs). In some of these embodiments, method 100 further includes combining, by the computer, the additional segmented suspicious foci data with the segmented suspicious foci data produced using the second trained machine learning model.

[0042] To further illustrate, Fig. 1B is a flow chart that schematically shows certain of exemplary method steps of segmenting a test image data set obtained from a test subject using a computer. As shown, method 110 includes segmenting one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data (step 112). The test image data set typically comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images. In addition, the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject. Method 110 also includes excluding the non-suspicious foci from the test image data set using the segmented non-suspicious foci data to produce suspicious foci data (step 114). The suspicious foci are suspicious for the presence of the disease state in the test subject. In addition, method 110 also includes segmenting the suspicious foci data (step 116).

[0043] As an additional exemplary illustration, Fig. 1C is a flow chart that schematically shows certain of exemplary method steps of classifying a test image data set obtained from a test subject using a computer. As shown, method 120 includes using a first trained machine learning model to produce a first set of predicted probabilities that selected foci in one or more positron emission tomography (PET) images from the test image data set are each individually classified as being either suspicious or non-suspicious (step 122). Method 120 also includes using a second trained machine learning model to produce a second set of predicted probabilities that the selected foci in one or more computed tomography (CT) images from the test image data set are each individually classified as being suspicious or non-suspicious (step 124). Method 120 also includes using a third trained machine learning model to produce a third set of predicted probabilities that the selected foci in one or more fusions of the one or more PET images and the one or more CT images from the test image data set are each individually classified as being suspicious or non-suspicious (step 126). In addition, method 120 also includes generating a final pattern

classification of the selected foci as each individually being suspicious or non-suspicious using averaged predicted probabilities of the first, second, and third sets of predicted probabilities (step 128).

[0044] Typically, an electronic neural network comprises the first, second, and/or third trained machine learning model. In some embodiments, electronic neural network comprises one or more 3D hybrid Transformer-convolutional neural networks (CNNs). Other electronic neural network architectures, such as the U-net, are also optionally used. In some embodiments, the first and/or second trained machine learning model has been trained on a set of training data that comprises a plurality of reference image data sets obtained from reference subjects that each comprise PET images and/or CT images. A given reference image data set in the plurality of reference image data sets comprises at least one PET image and/or at least one CT image that is labeled with ground truth suspicious foci and/or ground truth non-suspicious foci. In some embodiments, the ground truth suspicious foci and/or the ground truth non-suspicious foci are assigned by a consensus of at least two subject matter experts. In some embodiments, the test image data set and the reference image data sets comprise prostate-specific membrane antigen (PSMA) PET and/or CT images.

[0045] In some embodiments, the methods of the present disclosure include classifying selected loci in the test image data set that exhibit increased imaging agent uptake relative to other regions in the test image data set as the non-suspicious foci or as the suspicious foci using a multi-modal decision fusion framework. The methods of the present disclosure also further include estimating a whole-body disease burden value for the test subject by calculating suspicious foci volume. Typically, the methods of the present disclosure do not include setting a predefined threshold of a PET standardized uptake value (SUV) and/or applying any manual corrections. In some embodiments, the test subject is a human subject. In some embodiments, organs (e.g., the spleen, liver, kidneys, prostate, bowel, etc.) comprise at least some of the non-suspicious foci.

[0046] In some embodiments, the methods of the present disclosure include administering a [^{18}F] DCFPyL imaging agent to the test subject prior to obtaining the test image data set from the test subject. In some embodiments, the non-suspicious

and suspicious foci in the test image data set exhibit an increased uptake of an imaging agent relative to other regions in the test image data set. In some embodiments, the methods of the present disclosure include administering a therapy to the test subject to treat the disease state based at least in part upon the segmented test image data set. In some embodiments, the methods of the present disclosure include segmenting multiple test image data sets obtained from the test subject at different time points to assess treatment response of the test subject to the therapy. In some embodiments, the disease state comprises a cancer type. In some embodiments, the cancer type comprises prostate cancer.

[0047] Fig. 2 is a schematic diagram of a hardware computer system 200 suitable for implementing various embodiments. For example, Fig. 2 illustrates various hardware, software, and other resources that can be used in implementations of any of methods disclosed herein, including, e.g., method 100 and/or one or more instances of an electronic neural network. System 200 includes training corpus source 202 and computer 201. Training corpus source 202 and computer 201 may be communicatively coupled by way of one or more networks 204, e.g., the internet.

[0048] Training corpus source 202 may include an electronic clinical records system, such as an LIS, a database, a compendium of clinical data, or any other source of images suitable for use as a training corpus as disclosed herein. Due to hardware volatile memory storage limitations, each constituent image of an image may be broken down into a number of tiles, which may be, e.g., 128 pixels by 128 pixels. Such tiles are examples of “components” as that term is used herein. According to some embodiments, each component is implemented as a vector, such as a feature vector, that represents a respective tile. Thus, the term “component” refers to both a tile and a feature vector representing a tile.

[0049] Computer 201 may be implemented as any of a desktop computer, a laptop computer, can be incorporated in one or more servers, clusters, or other computers or hardware resources, or can be implemented using cloud-based resources. Computer 201 includes volatile memory 214 and persistent memory 212, the latter of which can store computer-readable instructions, that, when executed by electronic processor 210, configure computer 201 to perform any of the methods disclosed herein, including method 100, 110, or 120, and/or form or store any

electronic neural network, and/or perform any classification technique as described herein. Computer 201 further includes network interface 208, which communicatively couples computer 201 to training corpus source 202 via network 204. Other configurations of system 200, associated network connections, and other hardware, software, and service resources are possible.

[0050] Certain embodiments can be performed using a computer program or set of programs. The computer programs can exist in a variety of forms both active and inactive. For example, the computer programs can exist as software program(s) comprised of program instructions in source code, object code, executable code or other formats; firmware program(s), or hardware description language (HDL) files. Any of the above can be embodied on a transitory or non-transitory computer readable medium, which include storage devices and signals, in compressed or uncompressed form. Exemplary computer readable storage devices include conventional computer system RAM (random access memory), ROM (read-only memory), EPROM (erasable, programmable ROM), EEPROM (electrically erasable, programmable ROM), and magnetic or optical disks or tapes.

[0051] II. Example: An automated deep learning-based framework for uptake segmentation and classification on PSMA PET/CT imaging of patients with prostate cancer

[0052] 1. Materials and Methods

[0053] 1.1 Datasets and annotation

[0054] PSMA PET/CT scans comprising our training and testing cohorts have been previously described elsewhere. Briefly, 137 men with biochemically recurrent prostate cancer underwent [^{18}F] DCFPyL PET/CT scans. 8.0±0.9 mCi of [^{18}F] DCFPyL was injected intravenously and images were acquired 120 minutes post-injection on a GE Discovery MI DR TOF scanner with a voxel size of (2.73 mm, 2.73 mm, 3.27 mm). This group was randomly divided into a training and test cohort (96 and 41 patients, respectively). To assess the generalizability of the model, an external testing set was generated from 56 consecutive patients who underwent [^{18}F] DCFPyL PET/CT scans at another institution since October, 2021. Images for the external testing set were acquired 60 minutes post-injection with an GE-Discovery RX or Siemens Biograph

mCT TOF scanner and had an inter-slice pixel spacing varying from 2.73 to 4.69 mm and a slice thickness in the range of 1.5 to 4 mm. Patient characteristics for each cohort are summarized in Table 1.

[0055] Suspicious and non-suspicious foci of uptake were manually segmented on each 3D PET Volume in MIRADA Medical Imaging Software by a radiologist (HXB) with 4 years of practice experience. Information on both PET and CT was used in determining whether any PET avid area (>2.5 SUV-bw) was suspicious versus non-suspicious. Segmentations were checked by a second radiologist (LZ) with 3 years of practice experience. If errors were found, the segmentations were manually corrected by the first radiologist. Any disagreement was resolved in consensus by consulting a third radiologist (SR) with 7 years of practice experience. The manually delineated masks were utilized as ground truths for AI model training and performance evaluation.

Table 1. Patient characteristics of the training, internal and external testing sets.

	Training set	Internal testing set	External testing set	P-value[a]
Total Patients(n)	96	41	56	
Total foci (n)	4815	1950	1342	
Non-suspicious n, (%)	3366, 70%	1469, 75%	954, 71%	
Suspicious n, (%)	1449, 30%	481, 25%	388, 29%	
Age	66±7	65±7	69±8	0.046
Race				
White n, (%)	71, 74%	31, 76%	43, 77%	0.54[b]
African American n, (%)	23, 24%	9, 22%	9, 16%	
Asian n, (%)	2, 2%	1, 2%	3, 5%	
Other n, (%)	0, 0%	0, 0%	1, 2%	
Height (cm)	176±7	177±6	174±10	0.11
Weight (kg)	91±16	93±13	89±15	0.44
PSA at imaging	110±580	50±190	6±10	0.33
Gleason Score (n)	8.1±1.1	7.7±1.1	7.9±1.0	0.08

(a. Reported p-value from one-way ANOVA unless otherwise specified. b. P-value from χ-squared test)

[0056] **1.2 Segmentation method**

[0057] An automated segmentation method consisting of two cascaded 3D hybrid Transformer-CNN networks (i.e., UNETR) was developed. The first network focuses on segmenting non-suspicious foci (e.g., normal organs) leveraging the

consistency of organ shapes and relative positions within the human body across patients, while the second one aims to segment suspicious foci (e.g., lesions) after non-suspicious foci are masked. Segmentation of suspicious foci on PSMA PET was achieved in three steps: 1. Segment non-suspicious foci from the PET image; 2. Exclude non-suspicious foci from the PET image (anatomical prior guidance); 3. Segment suspicious foci. An illustration of our framework is shown in Fig. 3. Quality check on step 1 was performed by a radiologist (HXB).

[0058] The 3D hybrid Transformer-CNN network was formulated as an encoder-decoder shape bridged by skip connection, which is similar to UNet. However, in the hybrid network, the encoder is structured with the Transformer and regular convolutional layer. The Transformer consists of layer normalization (LN), multilayer perceptron (MLP) and multi-head attention. The multi-head attention was comprised of parallel self-attention that achieves both local detail and global dependencies. The skip connection can propagate fine-grained details learned in the encoder to decoder. The Transformer-based encoder-decoder architecture increases the model's ability to learn long-range dependences so that feature representations have global context with multi-scales. This promotes the accurate segmentations of PET uptake in various sizes, shapes and locations. Inspired by a recent investigation into segmentation loss functions, which demonstrated the superiority of Dice-related compound loss functions, we opted to use a combination of soft Dice loss and Cross-Entropy loss as the overall loss function to optimize the segmentation model.

[0059] **1.3 Classification method**

[0060] A new multi-modal decision fusion framework was developed based on deep learning to distinguish foci with increased PET uptake as suspicious or non-suspicious. The classification framework (Fig. 4) includes three sub-networks constructed with a 3D DenseNet as backbone CNNs, a classifier predicting probability and prediction with binary classes, and a multi-modal Decision fusion structure. In this framework, two of the sub-networks are single-modal CNNs conducting classification solely using PET or CT images, and the other one is a multi-modal CNN classifying suspicious and non-suspicious foci by early fusing CT and PET images with concatenation operation. During model training, the Cross-Entropy loss was used to optimize the three sub-networks separately. When testing, a multi-modal Decision

fusion structure was built to predict the final uptake pattern classification results as suspicious and non-suspicious by implying an argmax operation on the averaged predicted probabilities from the trained single- and multi-modal CNNs.

[0061] **1.4 Experimental setup**

[0062] In this study, we trained the proposed segmentation framework using 96 cases. To evaluate the performance and generalizability of the segmentation framework, we performed testing on the remaining 41 cases from the internal dataset and 56 cases from the external dataset. PET images from both datasets underwent identical preprocessing steps, with intensity converted to SUVbw and subsequent normalization to the range of [0,1]. Furthermore, the slice thickness of the volumes was resampled to 2 mm. During training, we cropped image volumes to a size of 96×96×96 voxels, containing regions of increased uptake and neighboring background.

[0063] The segmentation model was trained and tested in Pytorch and MONAI [26] using a NVIDIA GeForce RTX 3090 GPU. The models for steps 1 and 3 were both trained with batch sizes of 1, using AdamW optimizer [27] with a learning rate of 0.0001 for 500 and 800 epochs, respectively, and the final models were chosen as the optimal model. In testing, we adopted a sliding window inference strategy with a window overlap of 0.8 to get the segmented volumes.

[0064] For uptake classification, we trained the framework using 96 cases including 1449 suspicious and 3366 non-suspicious foci. To evaluate the performance the classification framework, we performed testing on the remaining 41 cases (481 suspicious and 1469 non-suspicious) from the internal dataset and 56 cases (388 suspicious and 954 non-suspicious) from the external dataset. PET and CT images from both datasets were preprocessed in the same way: 1. Getting the foci regions with 3D bounding boxes from CT and SUVbw PET images; 2. Resampling the bounding boxes to [1 mm, 1 mm, 1 mm]; 3. Resizing the bounding boxes to a size of 48×48×48 voxels. Random rotation was conducted in model training to avoid over fitting.

[0065] The classification model was trained and tested in Pytorch and MONAI using a NVIDIA GeForce RTX 3090 GPU, using AdamW optimizer with a learning rate of 0.00001, batch size of 10, epoch of 1000. The final models were chosen with a

validation set divided from the training set (15 cases with 183 suspicious and 472 non-suspicious) achieving the highest classification accuracy.

[0066] 1.5 Performance metrics

[0067] The performance of the segmentation method was evaluated by Dice, IoU, Precision and Recall. Whole-body lesion volume based on model segmentations was calculated and compared with ground truth by radiologists. Model performance in classifying increased PET uptake into suspicious versus non-suspicious foci was evaluated with ACC, AUC, F1-Score, Precision and Recall.

[0068] 2. Results

[0069] 2.1 Demographics Data

[0070] As shown in Table 1, patients in the external testing set were older than those in the internal dataset ($p<0.05$). There was no other significant difference in race, weight, height, PSA at imaging, and Gleason score among the datasets.

[0071] 2.2 Segmentation results for internal and external testing sets

[0072] The segmentation results for the internal test set are shown in Table 2 (a)-(d). Our model achieved mean Dice, IoU, Precision and Recall of 0.957, 0.921, 0.952, 0.966 on the segmentation of non-suspicious foci and 0.962, 0.929, 0.962, and 0.965 in segmenting increased uptake on PET (includes both suspicious and non-suspicious foci). For the segmentation of suspicious foci, our model achieved mean Dice, IoU, Precision and Recall of 0.700, 0.566, 0.809 and 0.660 when non-suspicious foci were masked first, compared to 0.461, 0.315, 0.697 and 0.423 when segmenting the suspicious foci directly ($p<0.001$). The mean total volumes of the segmented suspicious foci with masking vs. ground truth were as 90.60 mm^3 and 108.76 mm^3 ($p=0.783$). Fig. 5 (a)-(d) summarized the patient-level of Dice, IoU, Precision and Recall of the segmentations on the internal test set. Fig. 4 shows the representative visualizations of the non-suspicious and suspicious foci segmentations, compared with the ground truth annotations.

[0073] The segmentation results for the external test set are shown in Table 2 (e)-(h). By excluding non-suspicious foci segmentations from the PET images, the mean Dice, IoU, Precision and Recall increased from 0.206 to 0.680, 0.126 to 0.548, 0.476 to 0.749, and 0.182 to 0.740, respectively ($p<0.001$), when compared to model segmentation of suspicious foci directly. The mean lesion volumes of the segmented

suspicious lesions and the ground truth were not significantly different (186.56 mm^3 vs. 198.10 mm^3, p=0.942). Fig. 5 (e)-(h) and Fig. 4 show patient-level Dice, IoU, Precision and Recall of the segmentations and visualization of example segmentation in the external testing set.

Table 2. Segmentation results for internal and external testing set with evaluation metrics. (mean)

Dataset	Settings			Metrics			
		Object	Prior guidance	Dice	IoU	Precision	Recall
Internal testing set	(a)	non-suspicious	N/A	0.957	0.921	0.952	0.966
	(b)	suspicious	N/A	0.461	0.315	0.697	0.423
	(c)	suspicious	non-suspicious	0.700	0.566	0.809	0.660
	(d)	foci	N/A	0.962	0.929	0.962	0.965
External testing set	(e)	non-suspicious	N/A	0.945	0.900	0.947	0.947
	(f)	suspicious	N/A	0.206	0.126	0.476	0.182
	(g)	suspicious	non-suspicious	0.680	0.548	0.749	0.740
	(h)	foci	N/A	0.961	0.930	0.961	0.965

(Non-suspicious: Non-suspicious foci, such as normal organs on PET image as segmentation object; Suspicious: suspicious lesion uptake on PET image as segmentation object; foci: Both suspicious and suspicious foci of uptake as segmentation object.)

[0074] 2.3 Classification of suspicious versus non-suspicious uptake pattern

[0075] There were 481 suspicious and 1469 non-suspicious foci of increased uptake on the 41 PSMA PET/CT scans from the internal testing set. The classification results for the internal test set are shown in Table 3 and Fig. 7a. The multi-modal Decision fusion framework classified foci in PET image as suspicious versus non-suspicious with the highest values of ACC, AUC, F1-Score, Precision and Recall of 0.764, 0.863, 0.844, 0.841 and 0.847 when compared to single-modal framework using only PET or CT or multi-modal framework using both PET and CT.

[0076] To evaluate the generalizability of the AI-based classification framework, we tested the model on the external dataset comprising 56 cases (includes 388 suspicious and 954 non-suspicious). The performance metrics of our model in

distinguishing suspicious from non-suspicious foci are shown in Table 3 and Fig. 7b. Compared with the single-modal CNNs using CT or PET image and multi-modal CNN early fusing CT and PET images, the multi-modal decision fusion framework achieved the highest performance metrics of 0.796, 0.851, 0.865, 0.814 and 0.923 in ACC, AUC, F1-Score, Precision and Recall.

Table 3. Classification results for internal and external testing set with evaluation metrics.

Dataset	Settings		Metrics				
	CNN	Image	ACC	AUC	F1-Score	Precision	Recall
Internal testing set	Single-modal	CT	0.724	0.814	0.811	0.840	0.784
	Single-modal	PET	0.745	0.834	0.832	0.828	0.835
	Multi-modal	CT, PET	0.732	0.818	0.821	0.827	0.815
	Multi-modal Decision fusion	CT, PET	0.764	0.863	0.844	0.841	0.847
External testing set	Single-modal	CT	0.774	0.831	0.847	0.819	0.876
	Single-modal	PET	0.717	0.800	0.806	0.785	0.829
	Multi-modal	CT, PET	0.774	0.805	0.851	0.803	0.904
	Multi-modal Decision fusion	CT, PET	0.796	0.851	0.865	0.814	0.923

[0077] 3. Discussion

[0078] Whole body tumor burden on PSMA PET/CT can predict patient outcomes in men with prostate cancer. However, currently estimates of tumor burden are difficult to obtain and therefore not widely used. Traditionally, such analysis requires both expertise and is time consuming and thus, cannot easily be incorporated into routine clinical workflow. Although implementation of automatic segmentation and classification approaches based on DNN has been introduced for several tumors (e.g., liver, brain, lung cancer et al.) on different imaging modalities (such as CT, MRI and FDG-PET/CT) with promising results, there is limited research on the AI-based segmentation and classification of uptake patterns on the whole-body [^{18}F] DCFPyL PET/CT. To address this and facilitate further research, we developed a novel AI

framework based on 3D DNNs for automating segmentation and classification of uptake patterns on PSMA PET.

[0079] Our initial attempts to segment suspicious foci directly from the PET images yielded unacceptably poor segmentation compared to the ground truth with mean Dice scores of 0.461 and 0.206 in the internal and external testing sets, respectively (Table 2). This was largely due to physiologic tracer uptake within unaffected organs. To avoid spurious segmentation of non-suspicious areas, such as normal organs, we developed an anatomical prior-guidance strategy to first train the algorithm to recognize non-suspicious foci of uptake (e.g., normal organs) and then excluded these areas in the PET image to allow the subsequent network to focus solely on suspicious PSMA-avid lesions. In the first segmentation step, our method attained high mean Dice scores of 0.95 for the internal dataset and >0.94 for the external testing dataset (Table 2). This demonstrates that our method was able to accurately segment non-suspicious foci on PSMA PET, with high concordance to manual annotations provided by domain experts (Fig. 6) and shows strong generalizability. Following guiding and masking of the segmented non-suspicious foci in the second step, the segmentation of suspicious foci in the third step of the framework attained mean Dice scores of 0.70 on the internal test set and 0.68 on the external test set (Table 2). There was a statistically insignificant difference between the segmentation volumes produced by our method and the delineations drawn by experts, with p-values of 0.78 and 0.94 in the internal and external test sets respectively (Figs. 5 and 6). The comparative results presented in Table 2, Fig. 5, and Fig. 6, which pertains to the internal and external testing sets, indicate the robust generalizability of the proposed segmentation framework across PSMA PET/CT data obtained using varied imaging parameters.

[0080] Currently, there is no existing tool that can segment the suspicious lesions on PSMA PET/CT fully automatically. Two examples of previously published tools are qPSMA and aPROMISE. qPSMA is a semi-automated segmentation framework that uses bone and normal organ uptake masks based on SUV thresholds and machine-learning based segmentation in combination with thresholding based on background liver uptake to detect PSMA-avid lesions. However, manual registration of PET and CT scans and correction of the normal organ map may be necessary if

there is poor co-registration of the PET and CT scans (e. g. patient movement). The use of a liver threshold reduces the sensitivity for lesions with lower avidity as well as in patients with extensive hepatic metastases. Furthermore, the segmentation procedures for normal organs and lesions in qPSMA are developed in isolation, leading to a convoluted workflow that poses challenges for a smooth workflow and restricts its applicability in actual clinical settings. qPSMA requires user time and may require time consuming manual corrections. Compared with qPSMA, our model is more automated and represents an end-to-end solution for both normal organ and lesion segmentation and is largely immune from co-registration difficulties.

[0081] A second segmentation framework for PSMA PET/CT is aPROMISE. In the aPROMISE framework, the co-registered CT is first used to delineate normal organs and skeletal structures, much like qPSMA. In the accompanying PET/CT images, the radiologist is required to manually identify each lesion. These selected hot-spots are automatically segmented and a whole-body analysis is generated. While adequate for a patient with a limited number of lesions, lesion selection can become tedious in a patient with extensive metastatic disease and, as it relies solely on the reading physician's judgement, can introduce significant inter-reader variability. Our detection framework circumvents this limitation by directly segmenting all of the lesions without physician input.

[0082] On the task of uptake classification, our multi-modal decision fusion framework outperformed the three sub-networks, including the single-modal CNNs constructed with 3D DenseNet solely using CT or PET images and the multi-modal CNNs structed by 3D DenseNet integrated with an early fusion of CT and PET images. We hypothesized that the anatomic information available from the CT of the combined PET/CT examination may be useful for improving the accuracy of this CNN method, as organs with non-fixed anatomical location may also display PET avidity intermittently or in only a portion of the organ (i.e., intestines and ureters). Thus, even experienced radiologists need to understand the contextual information (often from CT) about potential lesion locations in relation to the entire body, when classifying uptake on PET as suspicious or non-suspicious. By using information from both CT and PET, a more accurate decision can be made regarding the determination of

physiologic [^{18}F] DCFPyL uptake (e.g., bowel activity, ureter), which further improves the diagnostic value of [^{18}F] DCFPyL PET/CT.

[0083] Some further aspects are also defined in the following clauses:

[0084] Clause 1: A method of segmenting a test image data set obtained from a test subject using a computer. The method includes using, by the computer, a first trained machine learning model to segment one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject; applying, by the computer, a mask to the non-suspicious foci in the test image data set using the segmented non-suspicious foci data to produce masked and unmasked portions of the test image data set; and, using, by the computer, a second trained machine learning model and the unmasked portions of the test image data set to segment one or more suspicious foci in the test image data to produce segmented suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject, thereby segmenting the test image data set obtained from the subject using the computer.

[0085] Clause 2: The method of Clause 1, wherein an electronic neural network comprises the first and/or second trained machine learning model.

[0086] Clause 3: The method of Clause 1 or Clause 2, wherein the electronic neural network comprises one or more 3D hybrid transformer-convolutional neural networks (CNNs).

[0087] Clause 4: The method of any of Clauses 1-3, wherein the first and/or second trained machine learning model has been trained on a set of training data that comprises a plurality of reference image data sets obtained from reference subjects that each comprise one or more PET images and/or one or more CT images, wherein a given reference image data set in the plurality of reference image data sets comprises at least one PET image and/or at least one CT image that is labeled with one or more ground truth suspicious foci and/or one or more ground truth non-suspicious foci.

[0088] Clause 5: The method of any of Clauses 1-4, wherein the one or more ground truth suspicious foci and/or the one or more ground truth non-suspicious foci are assigned by a consensus of at least two experts.

[0089] Clause 6: The method of any of Clauses 1-5, wherein the test image data set and the reference image data sets comprise prostate-specific membrane antigen (PSMA) PET and/or CT images.

[0090] Clause 7: The method of any of Clauses 1-6, comprising classifying selected loci in the test image data set that exhibit increased imaging agent uptake relative to other regions in the test image data set as the non-suspicious foci or as the suspicious foci using a multi-modal decision fusion framework.

[0091] Clause 8: The method of any of Clauses 1-7, further comprising estimating a whole-body disease burden value for the test subject by calculating suspicious foci volume.

[0092] Clause 9: The method of any of Clauses 1-8, wherein the method does not comprise setting a predefined threshold of a PET standardized uptake value (SUV) and/or applying any manual corrections.

[0093] Clause 10: The method of any of Clauses 1-9, wherein the test subject is a human subject.

[0094] Clause 11: The method of any of Clauses 1-10, wherein one or more organs comprise at least some of the non-suspicious foci.

[0095] Clause 12: The method of any of Clauses 1-11, comprising administering a [^{18}F] DCFPyL imaging agent to the test subject prior to obtaining the test image data set from the test subject.

[0096] Clause 13: The method of any of Clauses 1-12, wherein the non-suspicious and suspicious foci in the test image data set exhibit an increased uptake of an imaging agent relative to other regions in the test image data set.

[0097] Clause 14: The method of any of Clauses 1-13, comprising administering a therapy to the test subject to treat the disease state based at least in part upon the segmented test image data set.

[0098] Clause 15: The method of any of Clauses 1-14, comprising segmenting multiple test image data sets obtained from the test subject at different time points to assess treatment response of the test subject to the therapy.

[0099] Clause 16: The method of any of Clauses 1-15, wherein the disease state comprises a cancer type.

[00100] Clause 17: The method of any of Clauses 1-16, wherein the cancer type comprises prostate cancer.

[00101] Clause 18: The method of any of Clauses 1-17, further comprising: using, by the computer, a third trained machine learning model to segment one or more suspicious foci in the masked portion of the test image data set to produce additional segmented suspicious foci data.

[00102] Clause 19: The method of any of Clauses 1-18, further comprising: combining, by the computer, the additional segmented suspicious foci data with the segmented suspicious foci data produced using the second trained machine learning model.

[00103] Clause 20: A method of segmenting a test image data set obtained from a test subject using a computer. The method comprises segmenting, by the computer, one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject; excluding, by the computer, the non-suspicious foci from the test image data set using the segmented non-suspicious foci data to produce suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject; and, segmenting, by the computer, the suspicious foci data, thereby segmenting the test image data set obtained from the subject using the computer.

[00104] Clause 21: The method of Clause 20, wherein the test image data set comprises prostate-specific membrane antigen (PSMA) PET and/or CT images.

[00105] Clause 22: The method of Clause 20 or Clause 21, comprising classifying selected loci in the test image data set that exhibit increased imaging agent uptake relative to other regions in the test image data set as the non-suspicious foci or as the suspicious foci using a multi-modal decision fusion framework.

[00106] Clause 23: The method of any of Clauses 20-22, further comprising estimating a whole-body disease burden value for the test subject by calculating suspicious foci volume.

[00107] Clause 24: The method of any of Clauses 20-23, wherein the method does not comprise setting a predefined threshold of a PET standardized uptake value (SUV) and/or applying any manual corrections.

[00108] Clause 25: The method of any of Clauses 20-24, wherein the test subject is a human subject.

[00109] Clause 26: The method of any of Clauses 20-25, wherein one or more organs comprise at least some of the non-suspicious foci.

[00110] Clause 27: The method of any of Clauses 20-26, comprising administering a [^{18}F] DCFPyL imaging agent to the test subject prior to obtaining the test image data set from the test subject.

[00111] Clause 28: The method of any of Clauses 20-27, wherein the non-suspicious and suspicious foci in the test image data set exhibit an increased uptake of an imaging agent relative to other regions in the test image data set.

[00112] Clause 29: The method of any of Clauses 20-28, comprising administering a therapy to the test subject to treat the disease state based at least in part upon the segmented test image data set.

[00113] Clause 30: The method of any of Clauses 20-29, comprising segmenting multiple test image data sets obtained from the test subject at different time points to assess treatment response of the test subject to the therapy.

[00114] Clause 31: The method of any of Clauses 20-30, wherein the disease state comprises a cancer type.

[00115] Clause 32: The method of any of Clauses 20-31, wherein the cancer type comprises prostate cancer.

[00116] Clause 33: A method of classifying a test image data set obtained from a test subject using a computer. The method comprises using, by the computer, a first trained machine learning model to produce a first set of predicted probabilities that selected foci in one or more positron emission tomography (PET) images from the test image data set are each individually classified as being either suspicious or non-suspicious; using, by the computer, a second trained machine learning model to

produce a second set of predicted probabilities that the selected foci in one or more computed tomography (CT) images from the test image data set are each individually classified as being suspicious or non-suspicious; using, by the computer, a third trained machine learning model to produce a third set of predicted probabilities that the selected foci in one or more fusions of the one or more PET images and the one or more CT images from the test image data set are each individually classified as being suspicious or non-suspicious; and, generating, by the computer, a final pattern classification of the selected foci as each individually being suspicious or non-suspicious using averaged predicted probabilities of the first, second, and third sets of predicted probabilities, thereby classifying the test image data set obtained from the test subject using the computer.

[00117] Clause 34: The method of Clause 33, wherein the selected foci exhibit increased imaging agent uptake relative to other regions in the test image data set.

[00118] Clause 35: A system for segmenting a test image data set obtained from a test subject using an electronic neural network. The system comprises a processor; and a memory communicatively coupled to the processor, the memory storing instructions which, when executed on the processor, perform operations comprising: using a first trained machine learning model to segment one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject; applying a mask to the non-suspicious foci in the test image data set using the segmented non-suspicious foci data to produce masked and unmasked portions of the test image data set; and using a second trained machine learning model and the unmasked portions of the test image data set to segment one or more suspicious foci in the test image data to produce segmented suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject.

[00119] Clause 36: The system of Clause 35, wherein the memory storing instructions which, when executed on the processor, further perform operations comprising: using a third trained machine learning model to segment one or more suspicious foci in the masked portion of the test image data set to produce additional

segmented suspicious foci data; and, combining the additional segmented suspicious foci data with the segmented suspicious foci data produced using the second trained machine learning model.

[00120] Clause 37: A system for segmenting a test image data set obtained from a test subject. The system comprises: a processor; and a memory communicatively coupled to the processor, the memory storing instructions which, when executed on the processor, perform operations comprising: segmenting one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject; excluding the non-suspicious foci from the test image data set using the segmented non-suspicious foci data to produce suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject; and segmenting the suspicious foci data.

[00121] Clause 38: A system for classifying a test image data set obtained from a test subject using an electronic neural network. The system comprises: a processor; and a memory communicatively coupled to the processor, the memory storing instructions which, when executed on the processor, perform operations comprising: using a first trained machine learning model to produce a first set of predicted probabilities that selected foci in one or more positron emission tomography (PET) images from the test image data set are each individually classified as being either suspicious or non-suspicious; using a second trained machine learning model to produce a second set of predicted probabilities that the selected foci in one or more computed tomography (CT) images from the test image data set are each individually classified as being suspicious or non-suspicious; using a third trained machine learning model to produce a third set of predicted probabilities that the selected foci in one or more fusions of the one or more PET images and the one or more CT images from the test image data set are each individually classified as being suspicious or non-suspicious; and generating a final pattern classification of the selected foci as each individually being suspicious or non-suspicious using averaged predicted probabilities of the first, second, and third sets of predicted probabilities.

[00122] Clause 39: A computer readable media comprising non-transitory computer executable instructions which, when executed by at least one electronic processor, perform at least: using a first trained machine learning model to segment one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject; applying a mask to the non-suspicious foci in the test image data set using the segmented non-suspicious foci data to produce masked and unmasked portions of the test image data set; and using a second trained machine learning model and the unmasked portions of the test image data set to segment one or more suspicious foci in the test image data to produce segmented suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject.

[00123] Clause 40: The computer readable media of Clause 39, wherein the non-transitory computer executable instructions which, when executed by the electronic processor, further perform at least: using a third trained machine learning model to segment one or more suspicious foci in the masked portion of the test image data set to produce additional segmented suspicious foci data; and, combining the additional segmented suspicious foci data with the segmented suspicious foci data produced using the second trained machine learning model.

[00124] Clause 41: A computer readable media comprising non-transitory computer executable instructions which, when executed by at least one electronic processor, perform at least: segmenting one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject; excluding the non-suspicious foci from the test image data set using the segmented non-suspicious foci data to produce suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject; and segmenting the suspicious foci data.

[00125] Clause 42: A computer readable media comprising non-transitory computer executable instructions which, when executed by at least one electronic processor, perform at least: using a first trained machine learning model to produce a first set of predicted probabilities that selected foci in one or more positron emission tomography (PET) images from the test image data set are each individually classified as being either suspicious or non-suspicious; using a second trained machine learning model to produce a second set of predicted probabilities that the selected foci in one or more computed tomography (CT) images from the test image data set are each individually classified as being suspicious or non-suspicious; using a third trained machine learning model to produce a third set of predicted probabilities that the selected foci in one or more fusions of the one or more PET images and the one or more CT images from the test image data set are each individually classified as being suspicious or non-suspicious; and generating a final pattern classification of the selected foci as each individually being suspicious or non-suspicious using averaged predicted probabilities of the first, second, and third sets of predicted probabilities.

[00126] While the invention has been described with reference to the exemplary embodiments thereof, those skilled in the art will be able to make various modifications to the described embodiments without departing from the true spirit and scope. The terms and descriptions used herein are set forth by way of illustration only and are not meant as limitations. In particular, although the method has been described by examples, the steps of the method can be performed in a different order than illustrated or simultaneously. Those skilled in the art will recognize that these and other variations are possible within the spirit and scope as defined in the following claims and their equivalents. All patents, patent applications, other publications or documents, and the like cited herein are incorporated by reference in their entirety for all purposes to the same extent as if each individual item were specifically and individually indicated to be so incorporated by reference.

What is claimed is:

1. A method of segmenting a test image data set obtained from a test subject using a computer, the method comprising:

using, by the computer, a first trained machine learning model to segment one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject;

applying, by the computer, a mask to the non-suspicious foci in the test image data set using the segmented non-suspicious foci data to produce masked and unmasked portions of the test image data set; and,

using, by the computer, a second trained machine learning model and the unmasked portions of the test image data set to segment one or more suspicious foci in the test image data to produce segmented suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject, thereby segmenting the test image data set obtained from the subject using the computer.

2. The method of claim 1, wherein an electronic neural network comprises the first and/or second trained machine learning model.

3. The method of claim 2, wherein the electronic neural network comprises one or more 3D hybrid transformer-convolutional neural networks (CNNs).

4. The method of claim 1, wherein the first and/or second trained machine learning model has been trained on a set of training data that comprises a plurality of reference image data sets obtained from reference subjects that each comprise one or more PET images and/or one or more CT images, wherein a given reference image data set in the plurality of reference image data sets comprises at least one PET image and/or at least one CT image that is labeled with one or more ground

truth suspicious foci and/or one or more ground truth non-suspicious foci.

5. The method of claim 4, wherein the one or more ground truth suspicious foci and/or the one or more ground truth non-suspicious foci are assigned by a consensus of at least two experts.

6. The method of claim 4, wherein the test image data set and the reference image data sets comprise prostate-specific membrane antigen (PSMA) PET and/or CT images.

7. The method of claim 1, comprising classifying selected loci in the test image data set that exhibit increased imaging agent uptake relative to other regions in the test image data set as the non-suspicious foci or as the suspicious foci using a multi-modal decision fusion framework.

8. The method of claim 1, further comprising estimating a whole-body disease burden value for the test subject by calculating suspicious foci volume.

9. The method of claim 1, wherein the method does not comprise setting a predefined threshold of a PET standardized uptake value (SUV) and/or applying any manual corrections.

10. The method of claim 1, wherein the test subject is a human subject.

11. The method of claim 1, wherein one or more organs comprise at least some of the non-suspicious foci.

12. The method of claim 1, comprising administering a [^{18}F] DCFPyL imaging agent to the test subject prior to obtaining the test image data set from the test subject.

13. The method of claim 1, wherein the non-suspicious and suspicious foci in the

test image data set exhibit an increased uptake of an imaging agent relative to other regions in the test image data set.

14. The method of claim 1, comprising administering a therapy to the test subject to treat the disease state based at least in part upon the segmented test image data set.

15. The method of claim 14, comprising segmenting multiple test image data sets obtained from the test subject at different time points to assess treatment response of the test subject to the therapy.

16. The method of claim 1, wherein the disease state comprises a cancer type.

17. The method of claim 16, wherein the cancer type comprises prostate cancer.

18. The method of claim 1, further comprising:

using, by the computer, a third trained machine learning model to segment one or more suspicious foci in the masked portion of the test image data set to produce additional segmented suspicious foci data.

19. The method of claim 18, further comprising:

combining, by the computer, the additional segmented suspicious foci data with the segmented suspicious foci data produced using the second trained machine learning model.

20. A method of segmenting a test image data set obtained from a test subject using a computer, the method comprising:

segmenting, by the computer, one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test

subject;

excluding, by the computer, the non-suspicious foci from the test image data set using the segmented non-suspicious foci data to produce suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject; and,

segmenting, by the computer, the suspicious foci data, thereby segmenting the test image data set obtained from the subject using the computer.

21. The method of claim 20, wherein the test image data set comprises prostate-specific membrane antigen (PSMA) PET and/or CT images.

22. The method of claim 20, comprising classifying selected loci in the test image data set that exhibit increased imaging agent uptake relative to other regions in the test image data set as the non-suspicious foci or as the suspicious foci using a multi-modal decision fusion framework.

23. The method of claim 20, further comprising estimating a whole-body disease burden value for the test subject by calculating suspicious foci volume.

24. The method of claim 20, wherein the method does not comprise setting a predefined threshold of a PET standardized uptake value (SUV) and/or applying any manual corrections.

25. The method of claim 20, wherein the test subject is a human subject.

26. The method of claim 20, wherein one or more organs comprise at least some of the non-suspicious foci.

27. The method of claim 20, comprising administering a [^{18}F] DCFPyL imaging agent to the test subject prior to obtaining the test image data set from the test subject.

28. The method of claim 20, wherein the non-suspicious and suspicious foci in the test image data set exhibit an increased uptake of an imaging agent relative to other regions in the test image data set.

29. The method of claim 20, comprising administering a therapy to the test subject to treat the disease state based at least in part upon the segmented test image data set.

30. The method of claim 29, comprising segmenting multiple test image data sets obtained from the test subject at different time points to assess treatment response of the test subject to the therapy.

31. The method of claim 20, wherein the disease state comprises a cancer type.

32. The method of claim 31, wherein the cancer type comprises prostate cancer.

33. A method of classifying a test image data set obtained from a test subject using a computer, the method comprising:

using, by the computer, a first trained machine learning model to produce a first set of predicted probabilities that selected foci in one or more positron emission tomography (PET) images from the test image data set are each individually classified as being either suspicious or non-suspicious;

using, by the computer, a second trained machine learning model to produce a second set of predicted probabilities that the selected foci in one or more computed tomography (CT) images from the test image data set are each individually classified as being suspicious or non-suspicious;

using, by the computer, a third trained machine learning model to produce a third set of predicted probabilities that the selected foci in one or more fusions of the one or more PET images and the one or more CT images from the test image data set are each individually classified as being suspicious or non-suspicious; and,

generating, by the computer, a final pattern classification of the selected foci as each individually being suspicious or non-suspicious using averaged predicted

probabilities of the first, second, and third sets of predicted probabilities, thereby classifying the test image data set obtained from the test subject using the computer.

34. The method of claim 33, wherein the selected foci exhibit increased imaging agent uptake relative to other regions in the test image data set.

35. A system for segmenting a test image data set obtained from a test subject using an electronic neural network, the system comprising:
a processor; and
a memory communicatively coupled to the processor, the memory storing instructions which, when executed on the processor, perform operations comprising:

using a first trained machine learning model to segment one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject;

applying a mask to the non-suspicious foci in the test image data set using the segmented non-suspicious foci data to produce masked and unmasked portions of the test image data set; and

using a second trained machine learning model and the unmasked portions of the test image data set to segment one or more suspicious foci in the test image data to produce segmented suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject.

36. The system of claim 35, wherein the memory storing instructions which, when executed on the processor, further perform operations comprising:

using a third trained machine learning model to segment one or more suspicious foci in the masked portion of the test image data set to produce additional segmented suspicious foci data; and,

combining the additional segmented suspicious foci data with the segmented suspicious foci data produced using the second trained machine learning model.

37. A system for segmenting a test image data set obtained from a test subject, the system comprising:
a processor; and
a memory communicatively coupled to the processor, the memory storing instructions which, when executed on the processor, perform operations comprising:

segmenting one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject;

excluding the non-suspicious foci from the test image data set using the segmented non-suspicious foci data to produce suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject; and

segmenting the suspicious foci data.

38. A system for classifying a test image data set obtained from a test subject using an electronic neural network, the system comprising:
a processor; and
a memory communicatively coupled to the processor, the memory storing instructions which, when executed on the processor, perform operations comprising:

using a first trained machine learning model to produce a first set of predicted probabilities that selected foci in one or more positron emission tomography (PET) images from the test image data set are each individually classified as being either suspicious or non-suspicious;

using a second trained machine learning model to produce a second set of predicted probabilities that the selected foci in one or more computed tomography (CT) images from the test image data set are each individually classified as being suspicious or non-suspicious;

using a third trained machine learning model to produce a third set of predicted probabilities that the selected foci in one or more fusions of the one or

more PET images and the one or more CT images from the test image data set are each individually classified as being suspicious or non-suspicious; and

generating a final pattern classification of the selected foci as each individually being suspicious or non-suspicious using averaged predicted probabilities of the first, second, and third sets of predicted probabilities.

39. A computer readable media comprising non-transitory computer executable instructions which, when executed by at least one electronic processor, perform at least:

using a first trained machine learning model to segment one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject;

applying a mask to the non-suspicious foci in the test image data set using the segmented non-suspicious foci data to produce masked and unmasked portions of the test image data set; and

using a second trained machine learning model and the unmasked portions of the test image data set to segment one or more suspicious foci in the test image data to produce segmented suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject.

40. The computer readable media of claim 39, wherein the non-transitory computer executable instructions which, when executed by the electronic processor, further perform at least:

using a third trained machine learning model to segment one or more suspicious foci in the masked portion of the test image data set to produce additional segmented suspicious foci data; and,

combining the additional segmented suspicious foci data with the segmented suspicious foci data produced using the second trained machine learning model.

41. A computer readable media comprising non-transitory computer executable instructions which, when executed by at least one electronic processor, perform at least:

segmenting one or more non-suspicious foci in the test image data set to produce segmented non-suspicious foci data, wherein the test image data set comprises one or more positron emission tomography (PET) images and/or one or more computed tomography (CT) images and wherein the non-suspicious foci are non-suspicious for a presence of a disease state in the test subject;

excluding the non-suspicious foci from the test image data set using the segmented non-suspicious foci data to produce suspicious foci data, wherein the suspicious foci are suspicious for the presence of the disease state in the test subject; and

segmenting the suspicious foci data.

42. A computer readable media comprising non-transitory computer executable instructions which, when executed by at least one electronic processor, perform at least:

using a first trained machine learning model to produce a first set of predicted probabilities that selected foci in one or more positron emission tomography (PET) images from the test image data set are each individually classified as being either suspicious or non-suspicious;

using a second trained machine learning model to produce a second set of predicted probabilities that the selected foci in one or more computed tomography (CT) images from the test image data set are each individually classified as being suspicious or non-suspicious;

using a third trained machine learning model to produce a third set of predicted probabilities that the selected foci in one or more fusions of the one or more PET images and the one or more CT images from the test image data set are each individually classified as being suspicious or non-suspicious; and

generating a final pattern classification of the selected foci as each individually being suspicious or non-suspicious using averaged predicted probabilities of the first, second, and third sets of predicted probabilities.

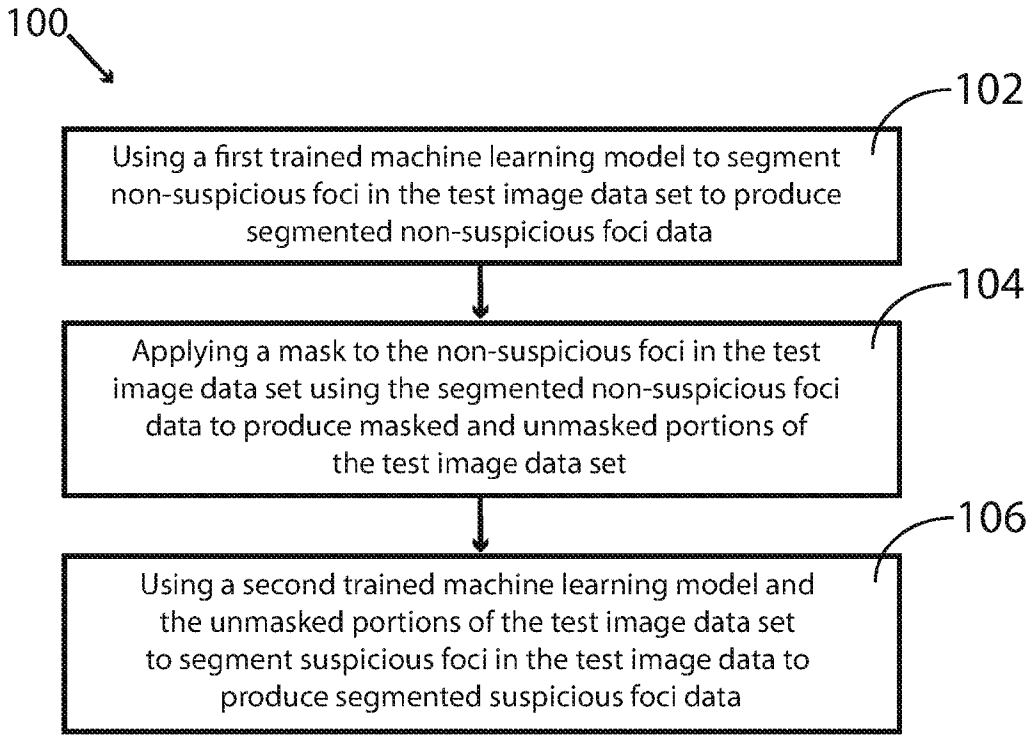
100
102
Using a first trained machine learning model to segment non-suspicious foci in the test image data set to produce segmented non-suspicious foci data
104
Applying a mask to the non-suspicious foci in the test image data set using the segmented non-suspicious foci data to produce masked and unmasked portions of the test image data set
106
Using a second trained machine learning model and the unmasked portions of the test image data set to segment suspicious foci in the test image data to produce segmented suspicious foci data

FIG. 1A

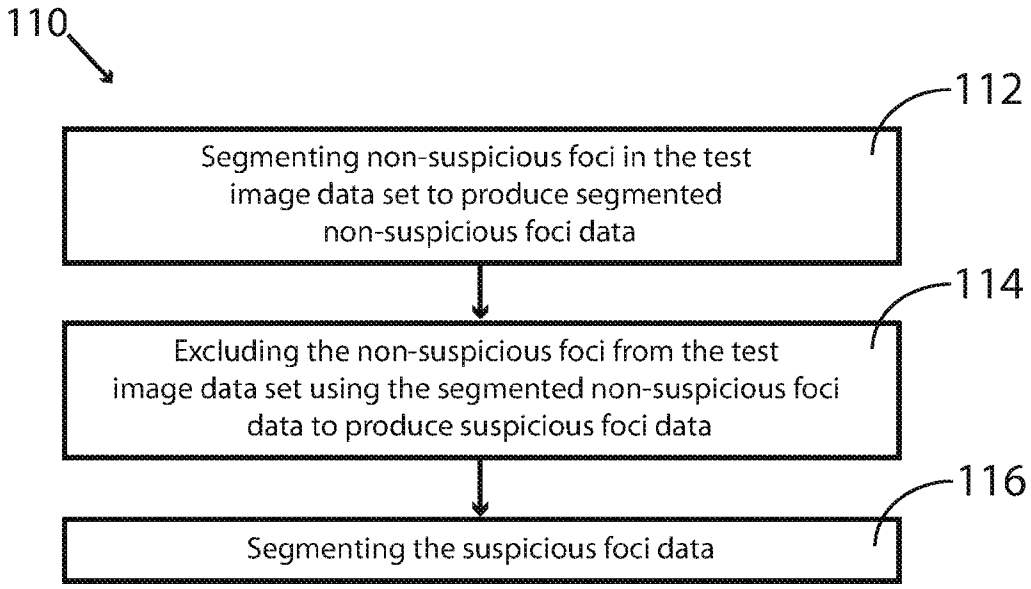
110
112
Segmenting non-suspicious foci in the test image data set to produce segmented non-suspicious foci data
114
Excluding the non-suspicious foci from the test image data set using the segmented non-suspicious foci data to produce suspicious foci data
116
Segmenting the suspicious foci data

FIG. 1B

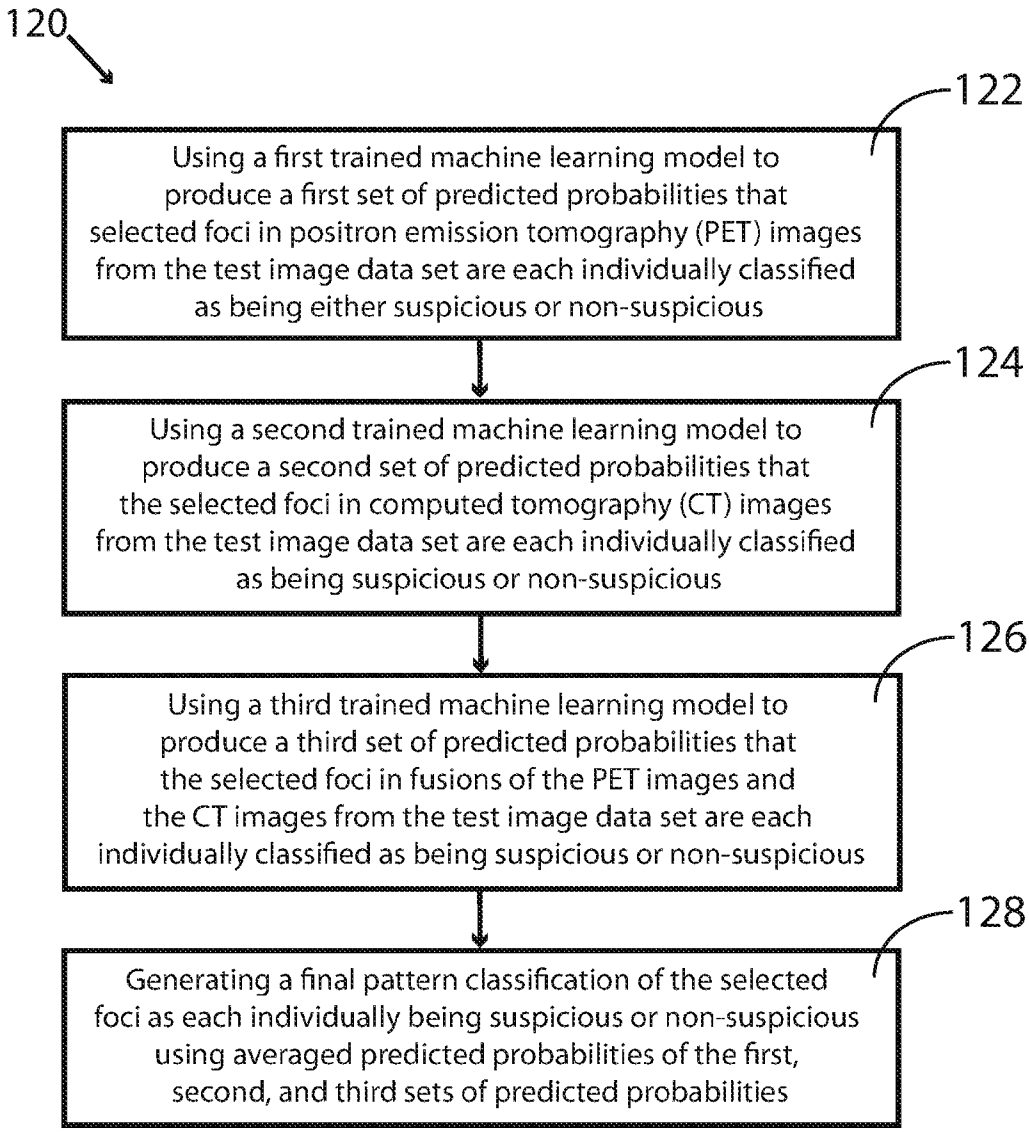
120
122
Using a first trained machine learning model to produce a first set of predicted probabilities that selected foci in positron emission tomography (PET) images from the test image data set are each individually classified as being either suspicious or non-suspicious
124
Using a second trained machine learning model to produce a second set of predicted probabilities that the selected foci in computed tomography (CT) images from the test image data set are each individually classified as being suspicious or non-suspicious
126
Using a third trained machine learning model to produce a third set of predicted probabilities that the selected foci in fusions of the PET images and the CT images from the test image data set are each individually classified as being suspicious or non-suspicious
128
Generating a final pattern classification of the selected foci as each individually being suspicious or non-suspicious using averaged predicted probabilities of the first, second, and third sets of predicted probabilities

FIG. 1C

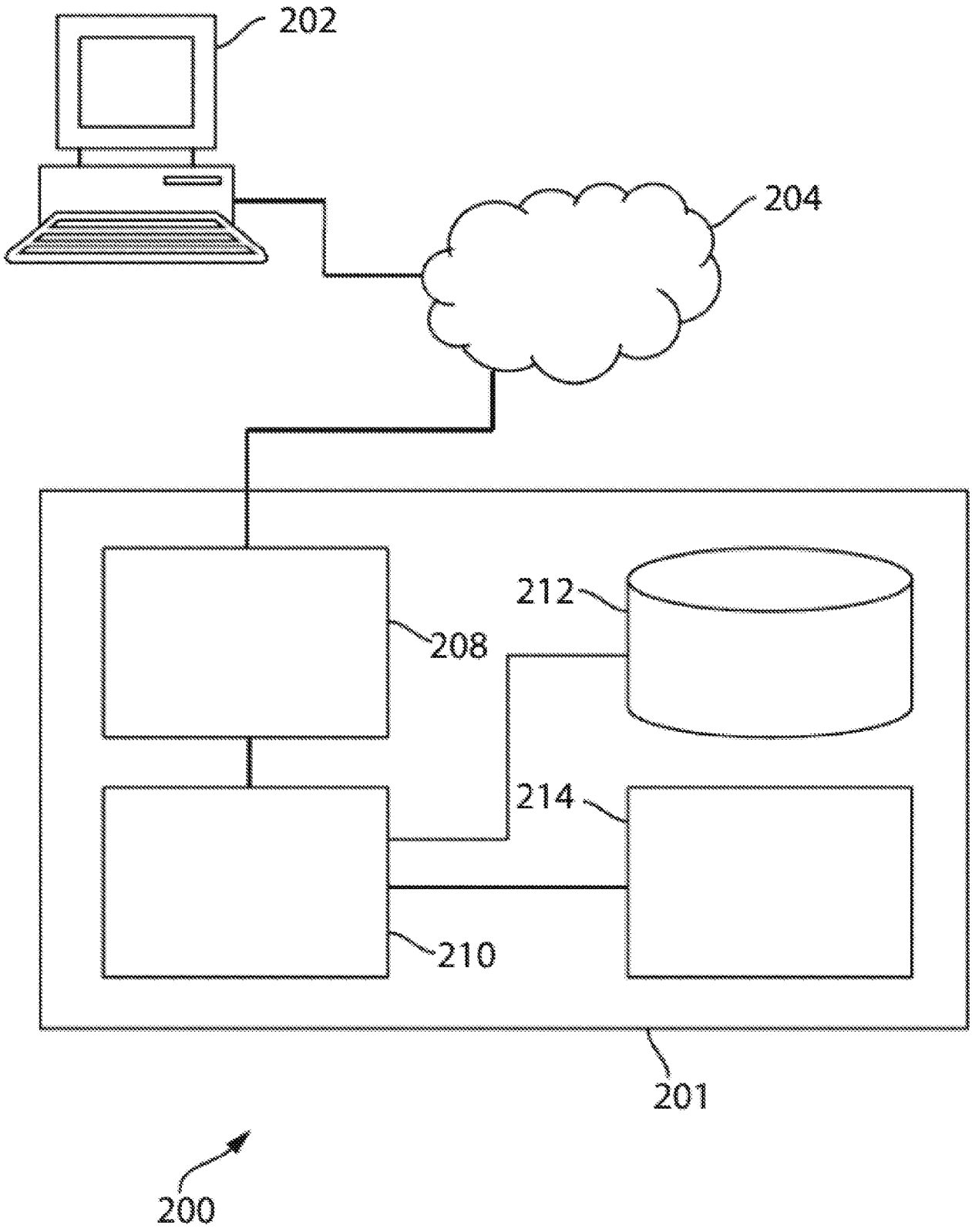

FIG. 2

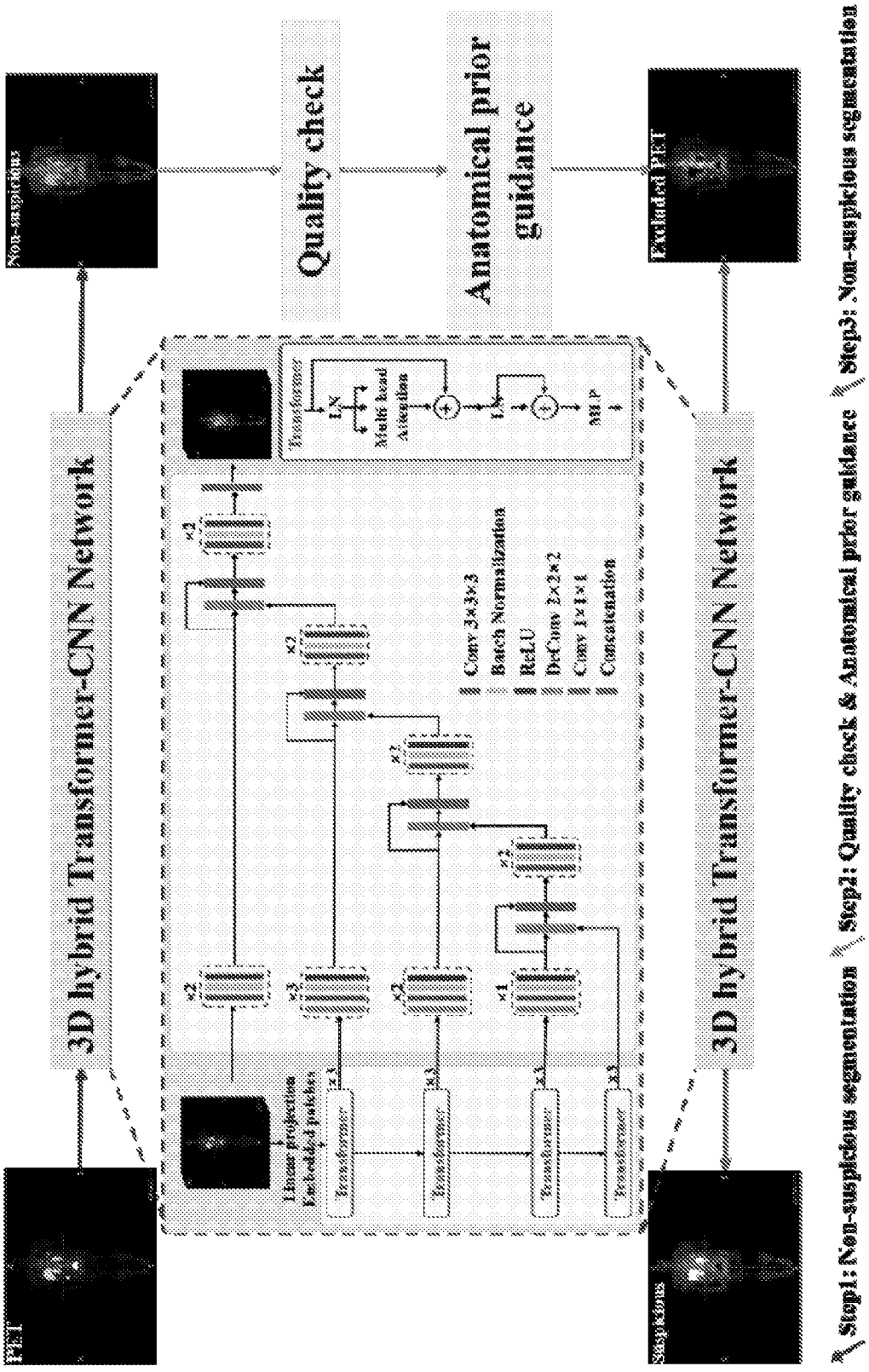
Non-suspicious
Quality check
Anatomical prior guidance
Excluded PET
3D hybrid Transformer-CNN Network
3D hybrid Transformer-CNN Network
PET
Suspicious
Conv 3×3×3
Batch Normalization
ReLU
DeConv 2×2×2
Conv 1×1×1
Concatenation
Step1: Non-suspicious segmentation
Step2: Quality check & Anatomical prior guidance
Step3: Non-suspicious segmentation

FIG. 3

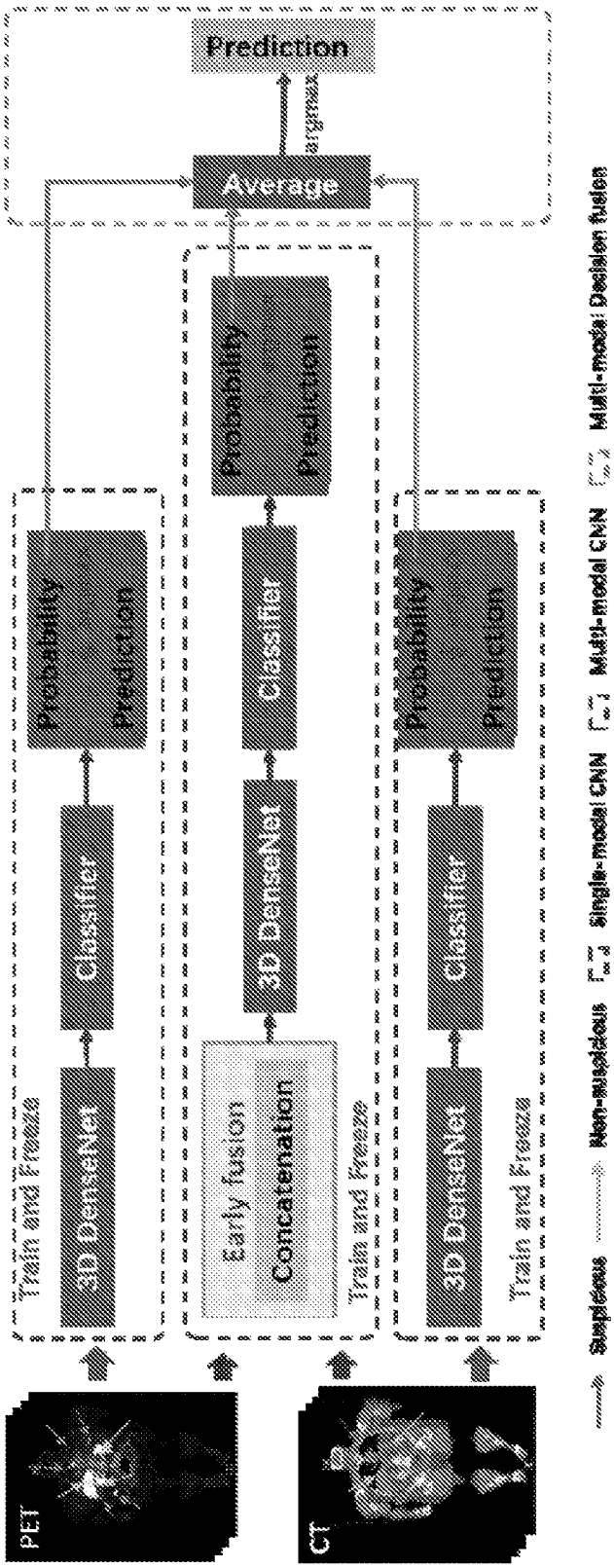
Prediction
argmax
Average
Probability
Prediction
Classifier
3D DenseNet
Early fusion
Concatenation
Train and Freeze
PET
CT
Suspicious
Non-suspicious
Single-modal CNN
Multi-modal CNN
Multi-modal Decision fusion

FIG. 4

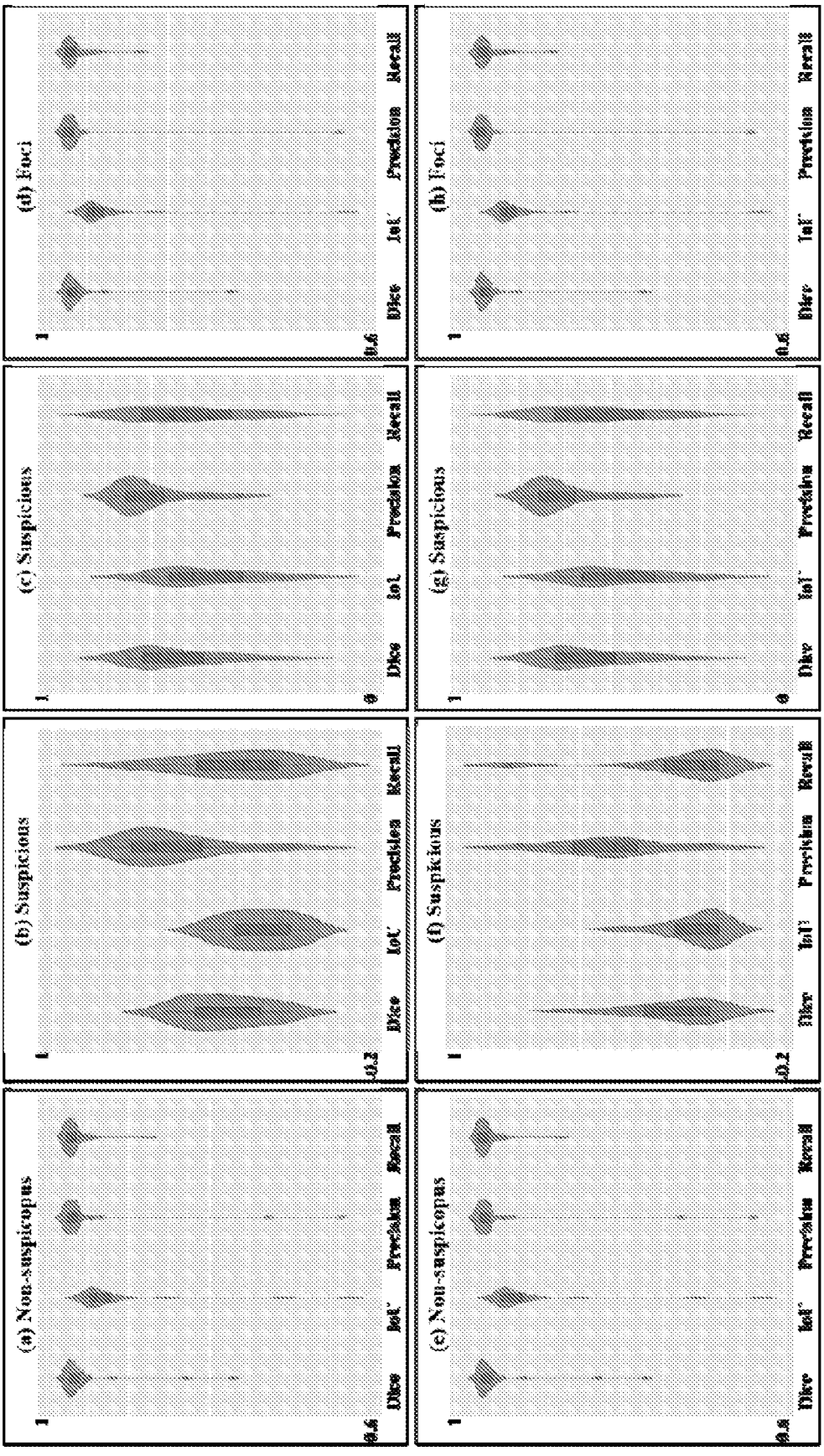
(a) Non-suspicopus
(b) Suspicious
(c) Suspicious
(d) foci
(e) Non-suspicopus
(f) Suspicious
(g) Suspicious
(h) foci
Dice
IoU
Precision
Recall

FIG. 5

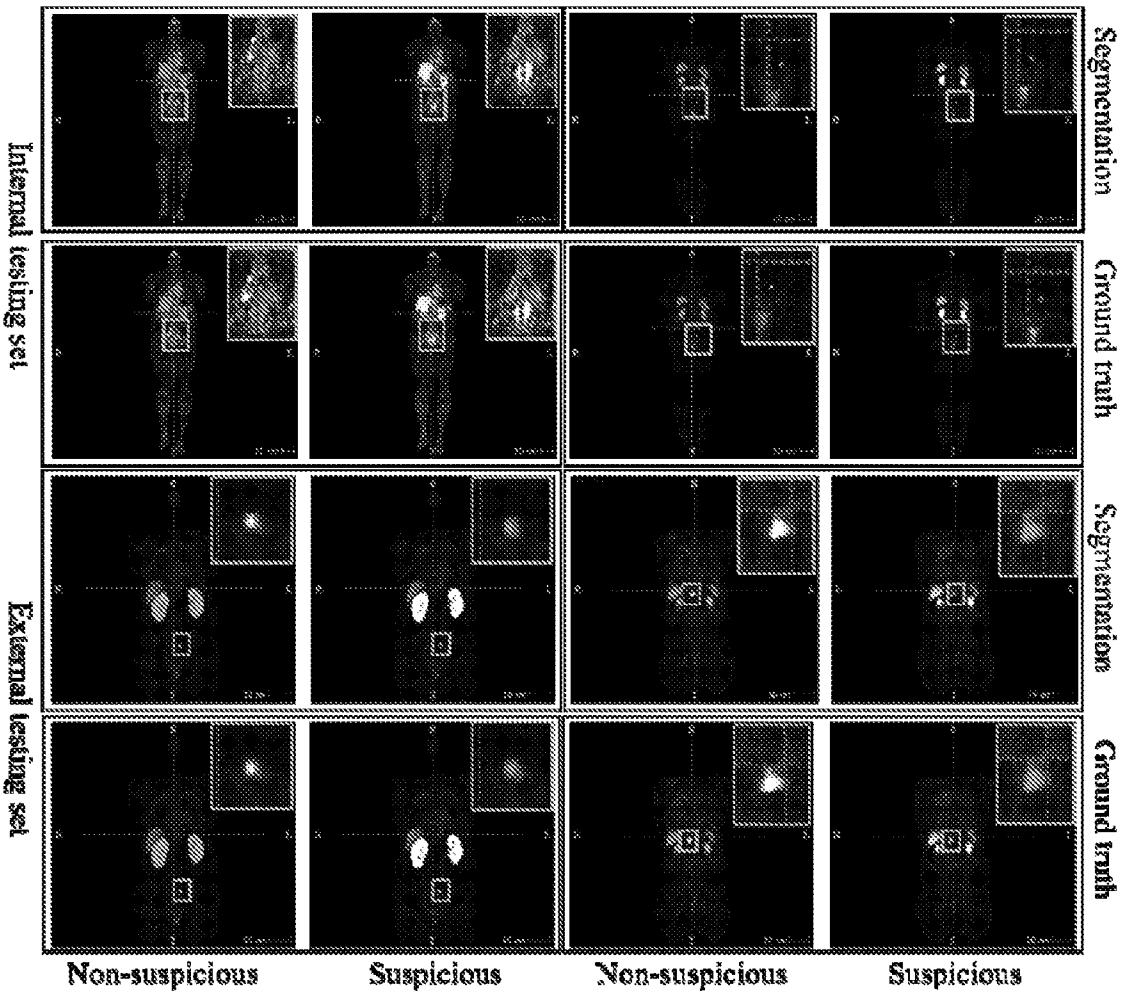
Internal testing set
External testing set
Segmentation
Ground truth
Segmentation
Ground truth
Non-suspicious
Suspicious
Non-suspicious
Suspicious

FIG. 6

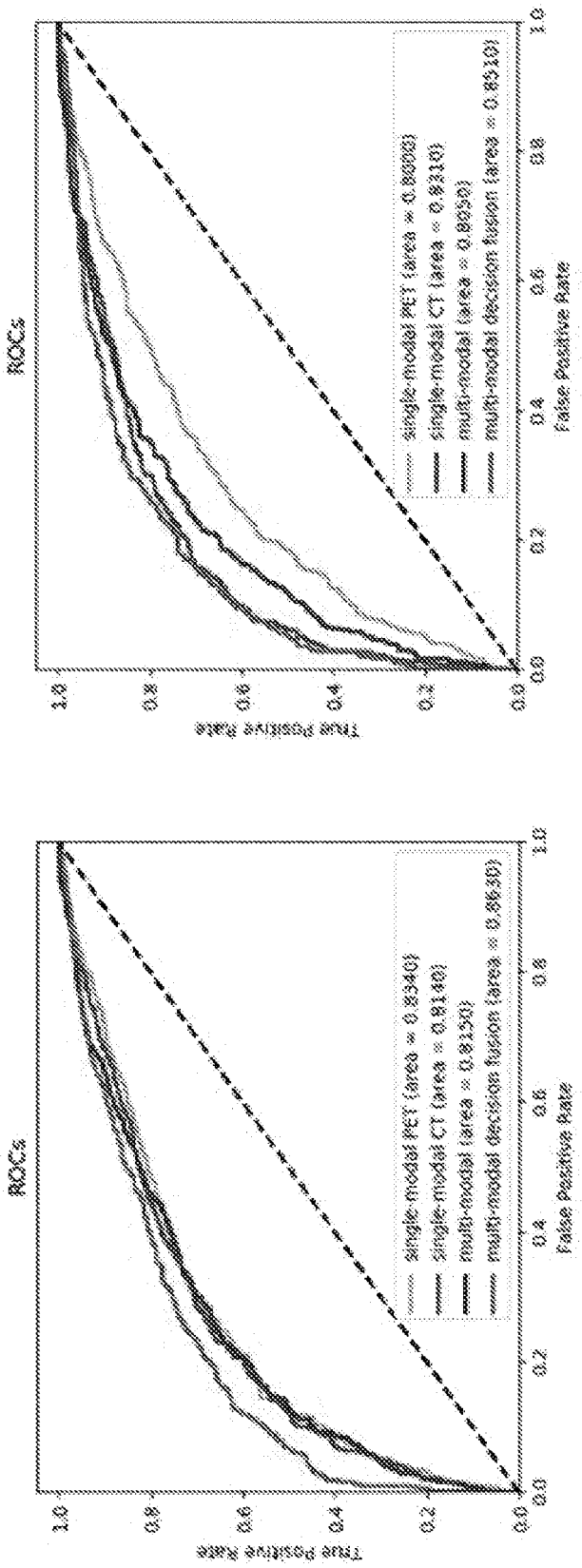
ROCs
True Positive Rate
False Positive Rate
single-modal PET (area = 0.89340)
single-modal CT (area = 0.81140)
multi-modal (area = 0.91530)
multi-modal decision fusion (area = 0.96310)
(a)
ROCs
True Positive Rate
False Positive Rate
single-modal PET (area = 0.88020)
single-modal CT (area = 0.82110)
multi-modal (area = 0.90560)
multi-modal decision fusion (area = 0.95310)
(b)

FIG. 7