Trace4

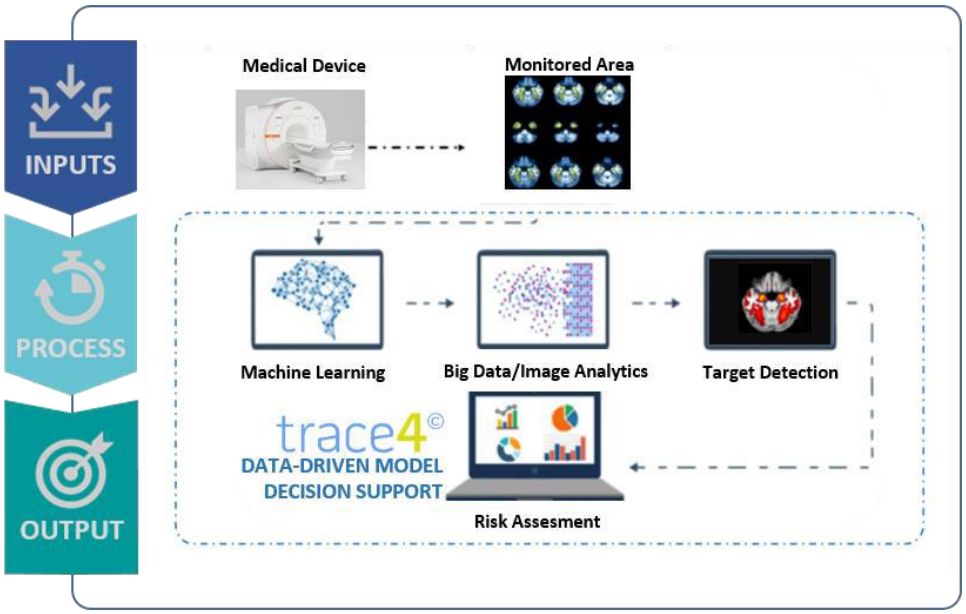
Deep Trace
TECHNOLOGIES

Trace4© is a DeepTrace-Technologies' proprietary Software-platform based on Artificial Intelligence (AI) and Data/Image Analytics for statistical analysis, automatic classification and development of AI predictive models in health at both single-subject and population level.

Trace4© includes proprietary and open-access software to enable:

1 - Analysis and archiving of Big Data and medical images
2 - Development of predictive models and customizable digital tools based on AI and health-medical data/images
3 - View data/images and health-clinical results explaining how AI work.

Highly specialized algorithms based on Machine Learning, Deep Learning, Transfer Learning, Big Data Analytics and Mining, Image Processing, Image Analytics and Mining (including Radiomics) are offered as support to intended users in a user-friendly and robust way, following a very simple workflow



The Trace4 © Platform offers an integrated, highly specialized decision support tool based on near real-time processing of health-medical data/images. Furthermore, thanks to the "privacy by default" security and privacy setting, patient/subject/user data are constantly protected during the development/implementation/use of the platform services.

The Trace4 © platform can be used on dedicated workstations (on Microsoft Windows operating systems) or in a server / cloud-client configuration.

Image Analytics/Radiomics and AI based tools at the level of single subject

INDICATION OF USE

Designed to be used by referring physicians for medical imaging studies (e.g. radiologists, nuclear doctors, ophthalmologists, dermatologists).
Physicians can use Trace4© and its derived predictive models to visualize, process, classify, archive, print medical images / data, in particular to extract quantitative characteristics from data and medical images, including 2D and 3D images acquired from optical cameras, ultrasound, X-ray radiographs, gamma cameras, PET, CT and MRI scanners also in multiple modality.

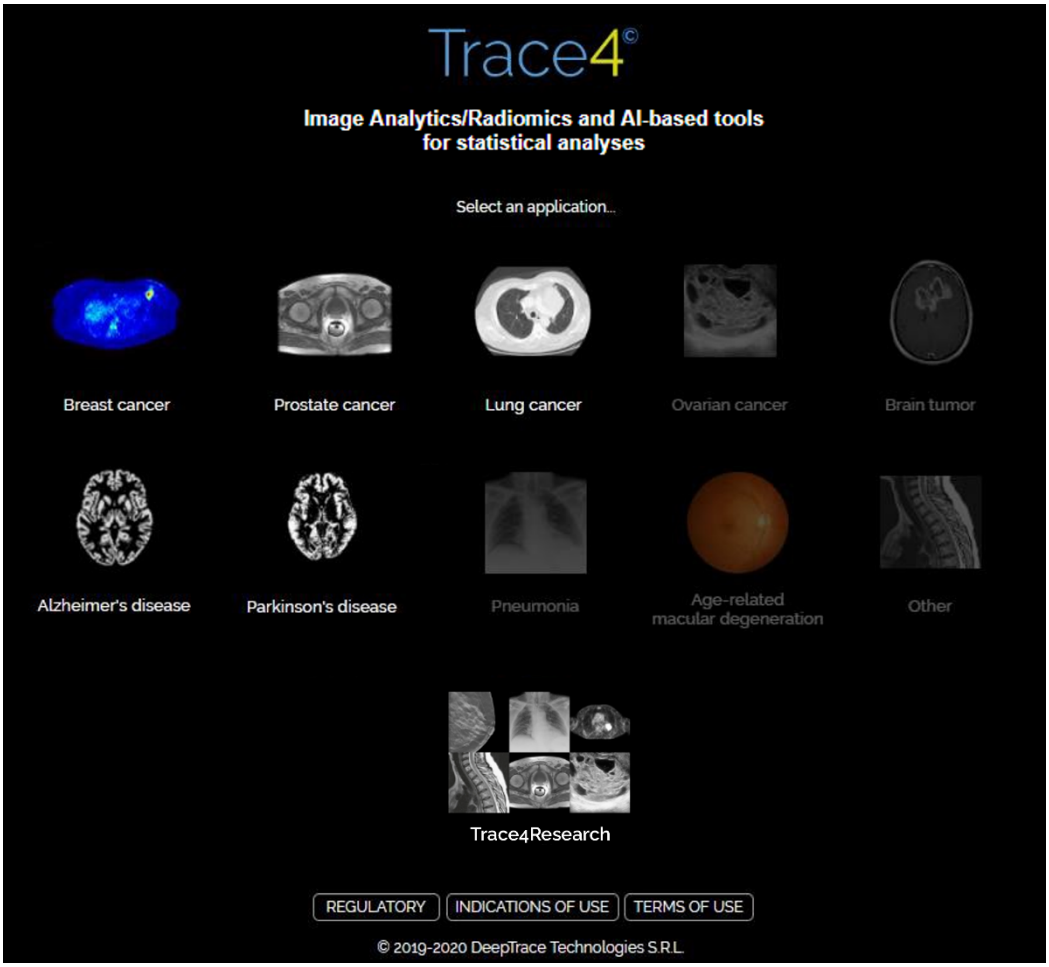


Image Analytics/Radiomics and AI based tools at the level of single subject

TECHNICAL DESCRIPTION

Images acquired by imaging systems (such as optical cameras in the visible and near visible spectra, ultrasonographs, radiographs, gamma cameras, PET, CT and MRI scanners), or statistical data can be analyzed with the Trace4© feature extraction methods. For medical images, the extracted features, possibly from Regions or Volumes of Interest (ROI / VOI), can be used to represent, localize and / or quantify the heterogeneity and morphology of the anatomical / functional structures scanned even in the presence of contrast media or radiopharmaceuticals, for research studies or statistical predictions.

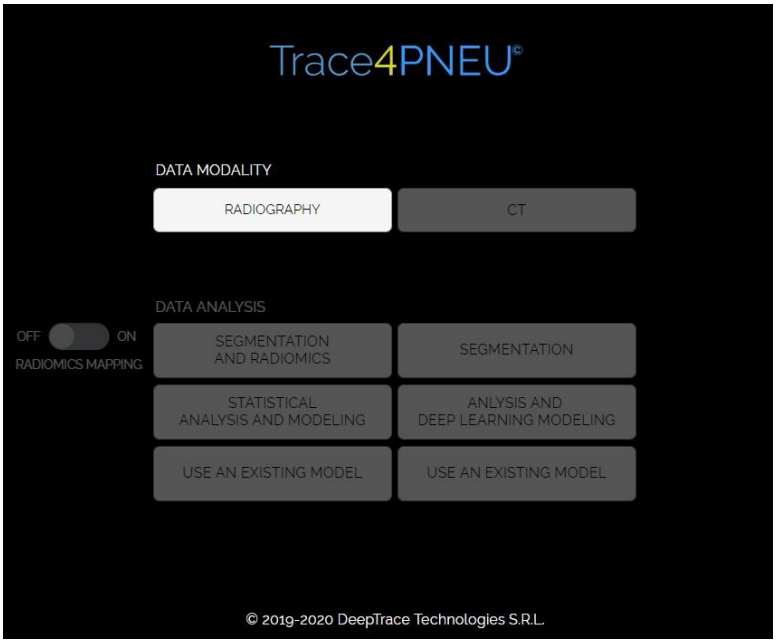


Image Analytics/Radiomics and AI based tools at the level of single subject

SEGMENTATION AND ANALYSIS OF THE IMAGE

The "Segmentation" option of Trace4© allows the extraction and storage and display of a segmented portion of a 2D or 3D image based on ROI / VOI.

The "Segmentation and Radiomics / Image Analytics" option of Trace4© allows the quantification, storage and display of quantitative characteristics from a segmented portion of a 2D or 3D image based on ROI / VOI.

The characteristics of the analysis of the images extracted from Trace4© are quantitative measures of tissue heterogeneity and of the morphology of the segmented region / volume, expressed in terms of morphological descriptors, statistical descriptors derived from the intensity histogram, from the co-occurrence matrix gray levels (GLCM), and other statistical matrices such as: Gray-Level Run Length Matrix (GLRLM), Gray-Level Size Zone Matrix (GLSZM), Neighborhood Gray Tone Difference Matrix (NGTDM), Gray-Level Distance Zone Matrix (GLDZM), Neighboring Gray Level Dependence Matrix (NGLDM).

The definition of the characteristics of Trace4©, their calculation and nomenclature comply with the guidelines of the International Biomarker Standardization Initiative (IBSI) [https://arxiv.org/abs/1612.07003].

In the case of 2D images, the morphological descriptors, originally defined by IBSI for 3D images, are replaced with equivalent 2D characteristics (for example, the volume and surface in 3D are replaced with the area and perimeter in 2D, respectively).

Isotropic voxel resampling and intensity discretization are performed prior to computation of the quantitative image descriptors.

Image Analytics/Radiomics and AI based tools at the level of single subject

STATISTICAL ANALYSIS AND MULTIVARIATE MODELS

The "Statistical Analysis and Modeling" option of Trace4 © allows the multivariate statistical analysis of either characteristics of medical images (extracted from ROI / VOI) at the level of image groups or single images, or statistical data based on population, and the development of predictive models based on these values.
For the modeling based on characteristics of medical images, Trace4 © makes a selection of stable image characteristics between groups of images with respect to different segmentation modes and repeatable in test-retest studies, This selection is performed by statistically comparing the obtained characteristics and selecting those characteristics with intra-class correlation coefficient (ICC)> 0.80, as a result of several image processing that avoids the need for multiple operators: a) random manipulation of image segmentation and b) image rotation and random image segmentations.

The statistical data or the image characteristics selected by Trace4 © (stable and repeatable between groups of images) are the input for multivariate statistical analysis and for the training of predictive systems.

Using Trace4 ©, in fact, different classification/predictive systems based on machine learning can be trained, validated and tested to generate multivariate models with different tasks (e.g. images from group 1 vs images from group 2 / trend of continuous variables) on the basis of supervised training by the medical user, further reducing the number of characteristics to stable characteristics, reproducible to non-redundant, so that the number of degrees of freedom of the problem is appropriate with respect to the number of images / data collected (e.g. approximately 1 characteristic every 10 images / data samples).

As an example, one of the machine-learning systems present in Trace4 © is a set of hundreds of decision trees with the decision rule for the classification of the majority vote; another machine learning system present in Trace4 © is a set of hundreds of Support Vector Machines combined with principal component analysis with the decision rule for majority vote classification.
For each classification system considered, Trace4 © applies the nested cross-validation method to train, validate and test each classification system in a supervised way.
For the minority class, an oversampling technique is applied with the method of adaptive synthetic sampling, in order to balance the training of the system.
The performance of the various Trace4 © classification systems are measured in terms of maximum and average precision, sensitivity, specificity, the area below the Receiver Operating Characteristics (AUC-ROC) curve, confidence interval, standard deviation, statistical significance. The best performing grading system is selected as the best grading system for the task of interest at the single image level (e.g., class 1 vs. class 2), based on supervised training by the medical user. The best image characteristics (more stable, reproducible, non-redundant and accurate in the classification task) are selected and stored.

Image Analytics/Radiomics and AI based tools at the level of single subject

ANALYSIS AND MODELS USING DEEP-LEARNING

The "Analysis and Deep-Learning Modeling" option present in Trace4 © allows the generation of deep-learning classifiers based on quantitative characteristics previously learned from images belonging to different classes of interest and transferable to images or portions of medical images of interest (for example following the segmentation of the image carried out using the "Segmentation" option).
Deep-learning models can be trained, validated and tested in Trace4 © for both binary and multiple classification tasks, based on the supervised training adopted by the medical user.

The deep-learning architectures adopted for image analysis are based on convolutional neural networks (CNNs) composed of many layers whose purpose is to learn a rich representation of the characteristics of the input classes and use this representation to classify the new images as belonging to one of the input classes.

Deep learning architectures can also be obtained from open access CNNs pre-trained and perfected in their latest levels for the activity of interest (Transfer Learning).

In order to increase the diversity of medical images between different training phases (epochs), automatic data augmentation techniques (including image rotations, cuts and reflections) can be applied to images in Trace4 © during training of the classifier.

The performance of Trace4 © deep-learning classifiers can be measured, even through several folds when a cross-validation approach is adopted, in terms of maximum and mean precision, sensitivity, specificity, AUC and standard deviation.

Image Analytics/Radiomics and AI based tools at the level of single subject

USE OF MODELS

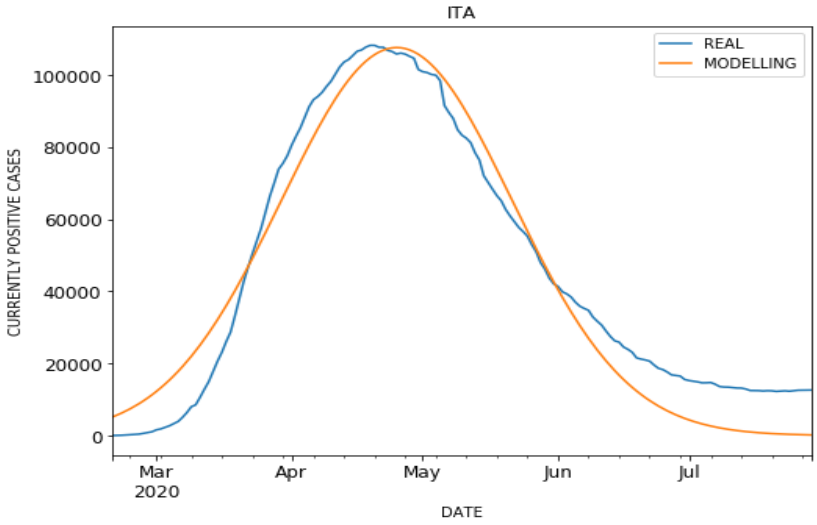
In Trace4© the "Use an existing model" option allows you to perform a multivariate statistical analysis at the single image level on the characteristics of the segmented image using the best classification system selected for the classification task (e.g. membership of the images to group 1 vs. group 2).

The use of Multivariate Models can support the identification and quantification of the heterogeneity and morphology of the portion of tissue analyzed in the medical image, characteristics that can be associated with screening, diagnosis, prognosis and response to treatment.

Data Analytics and AI based tools at the level of population

INDICATION OF USE

Designed to be used by referring professionals for statistical and epidemiologic studies (e.g., statisticians, clinicians, epidemiologists).
Users can use Trace4© and its derived predictive models to visualize, process, classify, archive, print health data, in particular to extract quantitative characteristics from data, including data acquired from public-private datasets as cumulative and incident positive cases, cumulative death, mortality, estimated actual cases.



TECHNICAL DESCRIPTION

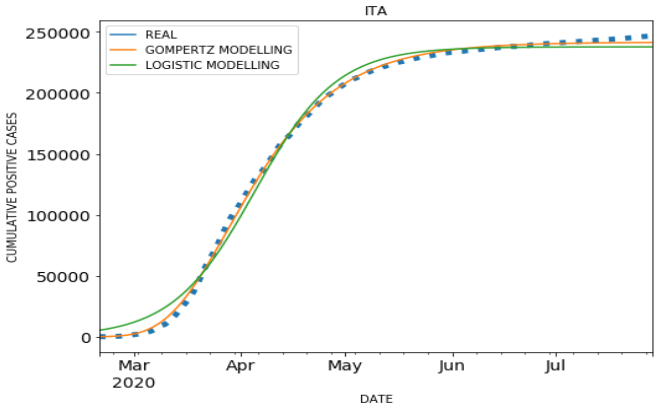
Data acquired by healthcare organizations (such as national and regional health systems), or from government body or statistical data can be analyzed with the Trace4© feature computation, extraction and selection methods.

Data Analytics and AI based tools at the level of population

STATISTICAL ANALYSIS AND MODEL USING DEEP LEARNING

The "Statistical Analysis and Modeling" option of Trace4© allows the statistical analysis of time distribution of specific features of disease spread at the level of population groups or statistical data based on population, and the development of predictive models based on these values.

Specific features selected by Trace4© (e.g. R, R mean over a period, attack rate, mortality) are computed and use as the input for statistical distributions and for the training of predictive systems. As representative examples, the selected features can be modelled by gaussian, logistic or Gompertz distribution.



Using Trace4©, different classification/predictive systems based on machine learning can be trained, validated and tested to generate models with different tasks (e.g. distributions from group 1 vs distributions from group 2 / time trend of selected features) on the basis of supervised training by the user.

As an example, one of the machine-learning systems present in Trace4© is a neural network system known as long-Short Term Memory (STML), a time-series predictive model based on recurrent neural network architectures. Modelling is iteratively updated, thus allowing the proposed system to improve its performance by learning on new data.
For each classification system considered, Trace4© applies the nested cross-validation method to train, validate and test each classification system in a supervised way.
For the minority class, an oversampling technique is applied with the method of adaptive synthetic sampling, in order to balance the training of the system.

The performance of the various Trace4© classification systems are measured in terms of maximum and average precision, sensitivity, specificity, the area below the Receiver Operating Characteristics (AUC-ROC) curve, confidence interval, standard deviation, statistical significance. The best performing grading system is selected as the best grading system for the task of interest at the single image level (e.g., class 1 vs. class 2), based on supervised training by the medical user. The best image characteristics (more stable, reproducible, non-redundant and accurate in the classification task) are selected and stored.

REGULATORY

Trace4 © was developed in accordance with the General Data Protection Regulation, GDPR, 2016/679, considering privacy as the default in the design of the platform.

Trace4 © is not a medical device. It is not CE marked or FDA cleared as a medical device. Any use of this software and associated information is solely for the purpose of research and forecast-statistical analysis.

AGREEMENT ON TERMS OF USE

Acceptance of the Agreement.

The user accepts the terms and conditions outlined in the Terms of Use Agreement ("Agreement") in relation to the "Trace4©" software platform (the "software") created by DeepTrace Technologies S.r.l.

Intellectual property.

Any unauthorized use of the Materials, Content or Tools within this Software may violate the laws and Terms of Use. The user undertakes not to copy, republish, frame, download, transmit, modify, rent, lend, sell, assign, distribute, license, reverse engineer or create derivative works and instruments based on this software, its materials, content or tools or their selection and arrangement, except as expressly authorized in this document by DeepTrace Technologies SRL

The user agrees not to decompile, reverse engineer or otherwise attempt to discover the source code of this software.

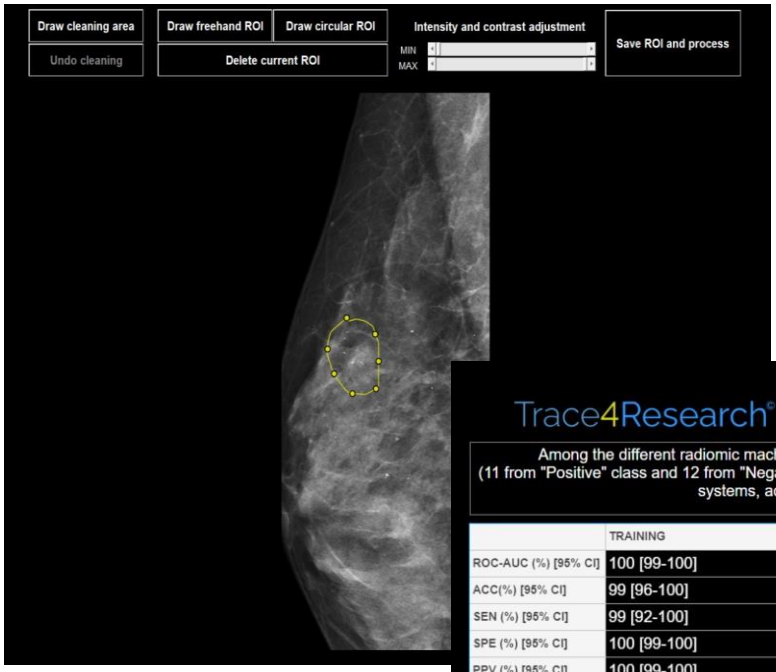
Limited right of use.

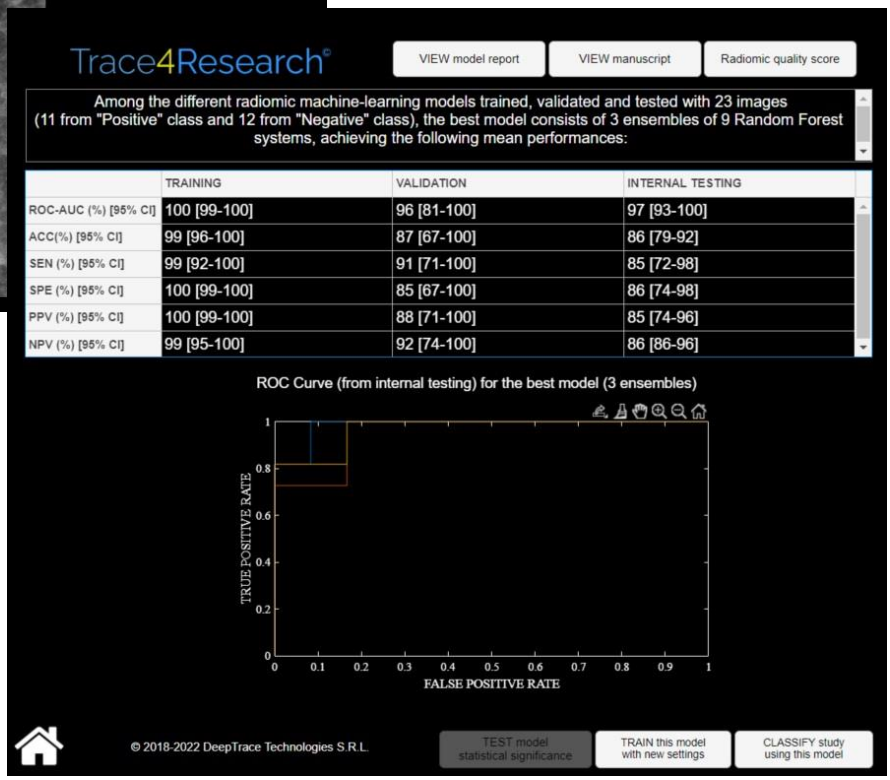
Viewing, using, printing or downloading any Materials, Content, Graphics, Forms or Documents from this software grants the user only a limited, non-exclusive license to use solely for the purpose defined in the Agreement with DeepTrace Technologies SRL, and not for the republication, distribution, assignment, license, sale, preparation of derivative works or other uses not authorized by DeepTrace Technologies S.r.l.

No part of any content, form or document may be reproduced in any form or incorporated into any information retrieval system, electronic or mechanical, other than the use defined in the Agreement (but not for resale or redistribution). The rights granted to the user constitute a license and not a transfer of title.

Trace4Research

Automatic analysis of image characteristics(>3000), training, validation and competitive testing of AI predictive models of outcomes for personalized medicine on research cases (machine learning, deep learning)





	TRAINING	VALIDATION	INTERNAL TESTING
ROC-AUC (%) [95% CI]	100 [99-100]	96 [81-100]	97 [93-100]
ACC(%) [95% CI]	99 [96-100]	87 [67-100]	86 [79-92]
SEN (%) [95% CI]	99 [92-100]	91 [71-100]	85 [72-98]
SPE (%) [95% CI]	100 [99-100]	85 [67-100]	86 [74-98]
PPV (%) [95% CI]	100 [99-100]	88 [71-100]	85 [74-96]
NPV (%) [95% CI]	99 [95-100]	92 [74-100]	86 [86-96]

Automatic creation of technical reports and scientific manuscripts on developed AI models with full transparency on performance and explainability of models and predictors

Manuscript draft created by TRACE4Research© AI platform on October 15, 2021 at 12:42 PM and associated to the machine-learning model 20211015-1236_model created on October 15, 2021 at 12:36 PM by user unknown from unknown.

A Machine Learning Model Based on Breast Mammography Radiomics to Predict Histological Diagnosis From Biopsy And Definite Surgery in Subjects at Risk of Breast Cancer.

ABSTRACT

Aim of this work was to develop a machine learning model based on breast mammography images to stratify the single-subject risk of breast cancer.

For this purpose we collected images of mammography from 30 subjects. Among these subjects, 15 patients (50%) belonged to class "Malignant" and 15 patients (50%) belonged to class "Benign", according to histological diagnosis from biopsy and definite surgery. This image set was used for the training and cross-validation of different machine learning models. A robust radiomic approach was applied, under the hypothesis that radiomic feature could be able to capture the disease heterogeneity among the two groups. Three models consisting of several ensembles of machine-learning classifiers (random forests, support vector machines and k-nearest neighbor classifiers) were developed, for the binary classification task of interest (Malignant vs. Benign), based on supervised learning, using histological diagnosis from biopsy and definite surgery as reference standard. The best model showed ROC-AUC (%) of 81.9** [72.4-91.3], Accuracy (%) of 74.4** [61.8-87.1], Sensitivity (%) of 77.8** [68.2-87.3], Specificity (%) of 71.1* [52-90.2], PPV (%) of 73.1** [58.1-88.1], and NPV (%) of 76.1** [65.3-86.9] (*p<0.05, **p<0.005).

MATERIALS AND METHODS

Image Sets

We collected images of mammography from 30 subjects. Among these subjects, 15 patients (50%) belonged to class "Malignant" and 15 patients (50%) belonged to class "Benign", according to histological diagnosis from biopsy and definite surgery. This image set was used for the training, cross-validation and internal testing of 3 machine-learning models.

Radiomic-based machine-learning modelling

Radiomic methodology was applied to collected images, according to the International Biomarker Standardization Initiative (IBSI) guidelines (https://arxiv.org/abs/1612.07003). For this purpose, the TRACE4© radiomic platform was used (http://www.deeptracetech.com/files/TechnicalSheet_TRACE4.pdf) allowing the whole IBSI-compliant radiomic workflow to be obtained in a fully-automatic way. IBSI radiomic workflow included: i) a segmentation of the ROI from each patient image, ii) a preprocessing of image intensities within the segmented ROI required to measure

Use of AI models and predictor analysis on new research cases