HAL open science

Three artificial intelligence data challenges based on CT and ultrasound

Nathalie Lassau, Imad Bousaid, Emilie Chouzenoux, Antoine Verdon, Corinne Balleyguier, François Bidault, Elie Mousseaux, Sana Harguem-Zayani, Loic Gaillandre, Zoubir Bensalah, et al.

To cite this version:

Nathalie Lassau, Imad Bousaid, Emilie Chouzenoux, Antoine Verdon, Corinne Balleyguier, et al.. Three artificial intelligence data challenges based on CT and ultrasound. Diagnostic and Interventional Imaging, 2021, 102 (11), 10.1016/j.diii.2021.06.005 . hal-03797851

HAL Id: hal-03797851

https://hal.science/hal-03797851v1

Submitted on 5 Jan 2024

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial 4.0 International License

Version of Record: https://www.sciencedirect.com/science/article/pii/S2211568421001674
Manuscript_ea070bec7518b3cd2ee8a4f0c9891c97

Three artificial intelligence data challenges based on CT and ultrasound

Short title:

Artificial intelligence based on CT and ultrasound

Authors

Nathalie Lassau [a,b,*]
Imad Bousaid [c]
Emilie Chouzenoux [d]
Antoine Verdon [c]
Corinne Balleyguier [a,b]
François Bidault [a,b]
Elie Mousseaux [e]
Sana Harguem-Zayani [a,b]
Loic Gaillandre [f]
Zoubir Bensalah [g]
Isabelle Doutriaux-Dumoulin [h]
Michèle Monroc [i]
Audrey Haquin [j]
Luc Ceugnart [k]
Florence Bachelle [k]
Mathilde Charlot [l]
Isabelle Thomassin-Naggara [m]
Tiphaine Fourquet [n]
Héloise Dapvril [o]
Joseph Orabona [p]
Foucauld Chamming's [q]
Michael El Haik [a,b]
Jules Zhang-Yin [r]
Marc-Samir Guillot [e]
Mickaël Ohana [s]
Thomas Caramella [t]
Yann Diascorn [t]
Jean-Yves Airaud [u]
Philippe Cuingnet [v]
Umit Gencer [e]
Littisha Lawrance [a]
Alain Luciani [w]
Anne Cotten [x]
Jean-François Meder [y,z]

Affiliations

© 2021 published by Elsevier. This manuscript is made available under the CC BY NC user license
https://creativecommons.org/licenses/by-nc/4.0/

[a] Laboratoire d'Imagerie Biomédicale Multimodale Paris-Saclay. BIOMAPS, UMR 1281. Université Paris-Saclay, Inserm, CNRS, CEA, 94800 Villejuif, France

[b] Department of Imaging, Institut Gustave Roussy, 94800 Villejuif, France

[c] Direction de la Transformation Numérique et des Systèmes d'Information, Institut Gustave Roussy, 94800 Villejuif, France

Villejuif, France

[d] CVN, Inria Saclay, 91190 Gif-sur-Yvette, France

[e] Unité Fonctionnelle d'Imagerie Cardiovasculaire Non Invasive, Hôpital Européen Georges Pompidou, AP-HP, 75015 Paris, France

[f] Centre Libéral d'Imagerie Médicale Agglomération Lille, 59800 Lille, France, France

[g] Department of Radiology, Centre Hospitalier St Jean, 66000 Perpignan, France

[h] Department of Radiology, Institut de Cancérologie de l'Ouest, 44800 Saint-Herblain, France

[i] Department of Radiology, Clinique Saint Antoine, 76230 Bois-Guillaume, France

[j] Department of Radiology, Hôpital de la Croix-Rousse - HCL, 69004 Lyon, France

[k] Department of Radiology, Centre Oscar Lambret, 59000 Lille, France

[l] Department of Radiology, Hôpital Lyon Sud - HCL, 69310 Pierre-Bénite, France

[m] Department of Radiology, Centre Hospitalier Intercommunal de Créteil, 94000 Créteil, France

[n] Department of Radiology, Centre Hospitalier Universitaire de Lille, 59000 Lille, France

[o] Service d'Imagerie de la Femme, Centre Hospitalier de Valenciennes, 59300 Valenciennes, France

[p] Department of Radiology, Centre Hospitalier de Bastia, 20600 Bastia, France

[q] Department of Radiology, Institut Bergonié, 33000 Bordeaux, France

[r] Department of Radiology, Hôpital Tenon, AP-HP, 75020 Paris, France

[s] Department of Radiology, Centre Hospitalier Universitaire de Strasbourg, 67200 Strasbourg, France

[t] Department of Radiology, Institut Arnault Tzanck, 06700 Saint-Laurent du Var, France

[u] Groupe Radioniort, 79000 Niort, France,

[v] Department of Radiology, Centre Hospitalier de Douai, 59507 Douai, France

[w] Collège des Enseignants de Radiologie de France, 75013 Paris, France

[x] Musculoskeletal Imaging Department, Lille Regional University Hospital, 59000 Lille, France

[y] Department of Neuroradiology, Centre Hospitalier Sainte-Anne, 75014 Paris, France

[z] Université de Paris, Faculté de Médecine, 75006 Paris, France

*Corresponding author: Nathalie.LASSAU@gustaveroussy.fr

Abstract

The 2020 edition of these Data Challenges was organized by the French Society of Radiology (SFR), from September 28 to September 30, 2020. The goals were to propose innovative artificial intelligence solutions for the current relevant problems in radiology and to build a large database of multimodal medical images of ultrasound and computed tomography (CT) on these subjects from several French radiology centers. This year the attempt was to create data challenge objectives relative to the clinical routine of radiologists, with less preprocessing of data and annotation, leaving a large part of the preprocessing task to the participating teams. The objectives were proposed by the different organizations depending on their core areas of expertise. A dedicated platform was used to upload the medical image data, to automatically anonymize the uploaded data. Three challenges were proposed including classification of benign or malignant breast nodules on ultrasound examinations, detection and contouring of pathological neck lymph nodes from cervical CT examinations and classification of calcium score on coronary calcifications from thoracic CT examinations. A total of 2076 medical examinations were included in the database for the three challenges, in three months, by 18 different centers, of which 12% were excluded. The 39 participants were divided into six multidisciplinary teams among which the coronary calcification score challenge was solved with a concordance index > 95%, and the other two with scores of 67% (breast nodule classification) and 63% (neck lymph node calcifications).

Keywords: Artificial intelligence; Ultrasonography; Computed tomography; Radiology; Data management

Abbreviations

2D: Two dimensional; 3D: Three dimensional; AI: Artificial intelligence; AUC: area under the curve; AUROC: area under the receiver operating characteristic curve; BI-RADS: Breast imaging-reporting and data system; C-index: Concordance index; CT: Computed tomography; DECT: Dual-energy computed tomography; ENT: Ear nose and throat; GDPR: General Data Protection and Regulation; INRIA : Institut National de Recherche en Informatique et en Automatique (National Research Institute in Digital Sciences and Technologies); MRI:

Magnetic resonance imaging; ROC: Receiver operating characteristic; SFR: Société Française de Radiologie (French Society of Radiology);

1. Introduction

In recent years, artificial intelligence (AI) has become increasingly important in medical research [1]. Indeed, the continuous creation of medical data makes machine learning and deep learning more effective, helping radiologists in their work [2,3], either in complex tasks like classification and detection or repetitive segmentation tasks [4,5]. These new technologies represent an important advancement not only in terms of diagnostic capabilities but also in terms of quality of work life for radiologists [3]. This requires a change in work habits in the health community, as it is important to involve patients, radiologists and AI researchers to work together, while abiding by the new regulations of the General data Protection and Regulation (GDPR) [6, 7]. The first two editions of this Data Challenge were held during the Journées Francophones de Radiologie in 2018 and 2019. Its main goals were: (*i*), to help radiologists answer difficult issues in their field; (*ii*), create large medical databases that can be reused in the future; and (*iii*), gather and challenge multidisciplinary teams to raise awareness of new challenges among all the persons involved in AI and radiology research.

Several publications have resulted from prior Data Challenges [8,9], and also on related subjects for which the work was initiated from these challenges, such as the works on meniscus tear detection on two-dimensional (2D) magnetic resonance imaging (MRI) [10,11], thyroid cartilage classification on 2D computed tomography (CT) [12], breast lesion characterization on 2D MRI [13], liver lesion classification on 2D ultrasound [14], kidney cortex segmentation using 2D CT [15], sarcopenia assessment on 2D CT [16], prediction of clinical disability in patients with multiple sclerosis using fluid attenuated inversion recovery MRI [17] and classification of pulmonary CT [18] are a few examples. The third edition of these challenges took place in September 2020. The objective was to answer the issues faced in clinical routine with limited preprocessing performed on the data by the radiologists. This made the objectives more difficult to reach, but made the AI more efficient. The data protection was optimized to be more compliant with the GDPR standards, by automatically anonymizing while uploading the medical examinations.

2. Material and methods

2.1 Clinical objectives

Radiologists from several organizations of the French Society of Radiology (SFR) proposed the objectives, in order to cover all the different domains, related to the current reality of medical diagnosis. These objectives were discussed with data scientists of the National Research Institute in Digital Sciences and Technologies (INRIA, Saclay, France) in order to verify the comprehensiveness of the objectives by AI, and the relevance of using the AI instead of another technology. This first step was essential in order to determine the important issues in both the medical and AI fields. A review of the scientific literature was done to verify if the proposed AI projects were not already answered in the publications, on Kaggle and on Grand-Challenge websites. Finally, a feasibility assessment was done to verify the collection of a sufficient number of medical examinations for the different challenges. The final dataset contained about 600 medical image series for each challenge.

2.2 Security and data protection

All the stages of the project, from data uploading to the challenge phase, were designed to be aligned with the GDPR. The Commission Nationale Informatique et Libertés was consulted, and the SFR assumed the role of Data Protection Officer. Automatic anonymization of the examinations was ensured by the platform developed by the startup Cleverdoc©. This process ensured the protection of patients' data, before the participants had access to it. The DICOM metadata on the age, gender, image size, modality, and pixel spacing was retained post anonymization. In addition, every radiologist involved was asked to send an information letter to patients about the use of their medical examinations, with the option of refusal of consent. A data chart was also sent to each radiologist to help them abide by the GDPR rules, as well as guidelines on the terms of use of the data for the participants to the challenge.

2.3 Communication and uploading

The data uploading phase began on June 19, 2020, however the first step was to contact as many radiologists as possible. The SFR contacted its network of radiologists. The interested radiologists had to register on the platform, and they were welcomed by email with a user guide, the technical and clinical specifications of the question and the format of the medical data to be uploaded was sent to them. Our aim was to have an all-in-one interface through the

platform for the radiologists, to upload and annotate with a uniform format throughout the dataset. They were trained through webinars and provided with a support contact email and phone numbers. A follow-up of their progress was also done. In addition, a weekly update was sent to all radiologists that included an overall report on the project's progress. The purpose of this newsletter was to motivate radiologists to participate actively in the project and for them to add the required medical examinations. The uploaded data was monitored on a daily basis to check the conformity of the medical image examinations and an expert checked the whole datasets before sending it to the participants.

2.4 Team gathering and challenge phases

Each team had to have multidisciplinary member background: at least one radiologist, one engineer/data-scientist and an engineering/PhD student. The team could be a startup, a big company or a research laboratory. Each team would have to register on the platform with all the members' details. Three datasets were sent to them for each challenge: the first dataset batch was sent on September 3, the second on September 28, for the teams to train their algorithms. Finally, the validation dataset was sent during the data challenge on September 30 at 1PM, and teams had until 2 PM to submit their results. The jury calculated the scores of each team, with the help of an INRIA researcher, and the winners were announced on the next day, on October 1. A prize of €3000 was awarded to the winner of each challenge.

3. Results

There were three main challenges for the Data Challenge 2020: The classification of benign or malignant breast nodules present from ultrasound examinations proposed by the Société d'Imagerie de la Femme (SIFEM; French Women's Imaging Society), the detection and contouring of pathological neck lymph nodes from cervical CT examinations, proposed by the College d'Imagerie pour la Recherche et l'Enseignement en Otorhino Larygologie (CIREOL, Francophone Society of Head and Neck Imaging), and the classification of calcium score on coronary calcifications from thoracic CT examinations proposed by the Société Française d'Imagerie Cardiaque et Vasculaire (SFICV, French Society of Cardiac and Vascular Imaging).

3.1 Breast nodule classification on ultrasound

The radiologists use the Breast Imaging-Reporting and Data System (BI-RADS) classification for mammography, ultrasound and MRI examinations to define the abnormalities seen and to indicate clinical management, (*i.e.*, return to screening, close follow-up or biopsy). The purpose of this question was to determine the risk of malignancy of the findings to establish appropriate diagnostic strategy for the patient. It was challenging for even the human eye to categorize some. Thus, there was a need to have a more powerful algorithm for the particular abnormalities in question. The test dataset composed of ultrasound images of the breast with abnormal findings, was annotated as “benin” or “malignant” depending on the corresponding histological proof.

3.2 Pathological lymph nodes classification on cervical CT examinations

The presence of abnormal lymph nodes is an important finding in the interpretation of ear nose and throat (ENT) CT examinations in patients with infection or cancer. The presence of cervical lymph nodes has an initial prognostic impact in ENT cancer. CT is widely used for a variety of clinical and therapeutic situations involving ENT adenopathy. The challenge was therefore to detect and assess the pathological lymph nodes on cervical dual energy (DE) CT examinations. The test dataset composed of cervical three-dimensional (3D) DECT examinations with normal and/or pathological lymph nodes. The locations of some specific lymph nodes were annotated on the images, with pathological or normal assignments. All nodes were not annotated as there were more than 70 lymph nodes present per patient, in which the amount of normal lymph nodes is much more than the pathological ones.

3.3 Coronary calcifications classification on thoracic CT examinations

Coronary calcification is responsible for many cardiovascular diseases and its early detection is an important prevention tool. When calcium score is low, the patient is followed up to limit his or her risk of disease. The challenge was to determine the calcification class, based on Agatston's calcium score for a score of 0 as Class A, score between 1 and 9 as Class B, between 10 and 99 as Class C, between 100 and 399 as Class D, for a score greater than 400, classified as class E. The test dataset was composed of thoracic 3D CT examinations with or without coronary calcification. The location of the calcifications was annotated on the images.

3.4 Score computation

3.4.1 Breast nodule classification

The final score was calculated using area under the receiver operating characteristic curve (AUROC) method. The final score ratio of the AUROC between benign / malignant was defined using the following equation:

$$Final\ score = AUROCC\ (Benign/Malignant)$$

3.4.2 Pathological lymph nodes

For each patient's examination, a different score was considered depending on the patient's condition. If the individual was sick the score was the Dice score of the area of the pathological nodes, else if the individual was healthy, the score was defined using the following equation:

$$Score = \frac{M}{A+M},$$

where M was the mean of the surface area of pathological nodes on the examinations of sick patients and A was the surface area of pathological nodes erroneously detected by the team. There was a need to define an alternative to the Dice score due to some examinations with no pathological nodes. For these images, the Dice sore was not representative. The teams were evaluated by penalizing them if they detected pathological nodes on these examinations. The bias arising due to this evaluation method was minimized by including the variable M in the calculations.

3.4.3 Coronary calcifications

The final score was defined as the concordance index (C-index) score that compared the classification of examinations based on the Agatston's calcification score. It was basically similar to the area under the curve (AUC) analysis, but allowed a score to be established for multi-class predictions by establishing an order among classes on a scale (class A to class E). At the beginning, the mean AUC on the five classes was opted, but the C-index score allowed an order of the classes, which was more relevant for this challenge because classes A to E were defined as the successive stages of coronary calcification.

4. Communication and team gathering

Six among the eight registered teams participated in the data challenge (BrAIn, CS/EURECOM; GAMC, Owkin, Philips, Radioadvisor). Thirty-nine team members

participated in this challenge, including six researchers, 13 students, 12 employees, and eight radiologists, in six different teams. The upload phase began on June 19, and the first images were uploaded by mid-July. A total of 2,085 images were uploaded from 18 radiology services (15 public hospitals, three profit private clinics, and three cancer centers) with 681 for the breast nodules, 737 for the pathological lymph nodes and 667 for the coronary calcifications (Table 1).

5. Data processing

Only a few preprocessing operations were performed on the data, such as automatic anonymization of the examinations and a pixel slope. The aim was to preserve the data as close to the ones used in reality by not filtering the image noise, as these AI models were intended to be used in routine in medical imaging without being pre-processed. Therefore, this strategy was adopted from the point of choosing the clinical objectives till the end of the challenge.

An important issue faced was in the gathering of data from multiple centers with the same characteristics, for the AI to train on a homogeneous format. Each challenge had its own characteristic format. For the Breast Nodules Challenge, data consisted of a B-mode ultrasound image in JPEG format. For the Pathological Lymph Nodes Challenge, data were extracted from 3D DECT examinations obtained after intravenous administration of iodinated contrast material with the patient in supine position in DICOM format with a soft part filter. For Coronary Calcification Challenge, data were extracted from 3D CT examinations with a slice thickness of 2.5- or 3 mm without gaps and iteratively reconstructed without filters.

When imposing these characteristics on radiologists, some poor quality data were excluded such as cervical CT examinations containing metal artifacts. Some images were also not suitable due to features that would have disturbed the training of the AI models, such as calipers within the lesion. However, some noisy images were not removed if the noise was not in the area of interest. The teams had to avoid these details to focus on the area of interest. Therefore, the medical images were sorted to have relevant datasets, with the same format with homogeneous distribution between classes and quality (Tables 2, 3 and 4). On October 1st, 2020, the three winners of the data challenges were announced based on the results calculated by all the teams. Philips team obtained the best results for Breast Nodule Classification Challenge (AUROC = 0.666 compared to 0.624 for Owkin and 0.643 for

Radioadvisor) and GAMC team for Pathological Lymph Node Challenge (score = 0.631) and Coronary Calcification Challenge (C-index = 0.951 compared to 0.909 for Owkin).

6. Discussion

The Data Challenge 2020 was designed to resemble routine clinical examinations of the patients. The data challenge in itself was an innovation because the AI models produced by the teams were useful for the radiologists. The aim was, for the AI to be effective on real-time medical examinations. In addition, the objectives concerned the human capacities of the radiologists, to classify the benign/malignant breast nodules and also the tedious task of contouring pathological neck lymph nodes on 3D examinations represented as a very relevant task for radiologists and as a complex task for an AI. The technical difficulty was therefore present on both detection and classification objectives. The optimization of uploading also presented an improvement in the method to organize the challenge. Gathering data from different medical centers across France was a complex task. A dedicated platform for the recuperation of data was created the previous year, but this year it was outsourced it to a startup Cleverdoc©, which specializes in medical data collection and AI models, to propose a more optimized solution to facilitate uploading and annotation. Thus, it was simpler and faster for radiologists, and made it possible to automate their registration, uploading and anonymization of examinations. Such a platform was a good solution for the future as it ensured simplicity and being completely compliant with the GDPR rules.

The results of the Data Challenge 2020 were quite disappointing for the Breast Nodules Challenge and the Pathological Lymph Nodes challenge. The scoring system was also a matter of debate. For the Breast Nodules Challenge, AUC is a commonly used score for binary prediction and was relevant. However, for the Coronary Calcification Challenge, which was a non-binary classification problem, an initial option was to take a mean AUC over the five classes. Since these classes are ordered by Agatston's score ($A_A < A_B < A_C < A_D < A_E$), it seemed preferable to use a score based on the C-index, as the values on which they were ordered was more important than identifying the right class. Finally, for the Pathological Lymph Nodes challenge, which combined detection and segmentation, selection of a suitable score was difficult considering that the Dice score did not take into account clinical findings for patients with no pathological lymph nodes. Thus, an alternative score for those examinations had to be determined. Having two different scores for two different types of examinations implied a bias in the final score, which was mitigated by using a factor related

to the mean of the pathological lymph node surfaces, such that the two scores were close. Nevertheless, the score used presented a certain bias, and it would have been better to have an alternative solution in further Data Challenges, to limit this bias and to have an objective score.

The BI-RADS classification of breast nodules gives a probability of malignancy of a nodule and thus helps radiologists diagnose benign or malignant nodules. In particular, nodule category helps decide to perform a biopsy and reduce the rate of useless biopsies of benign nodules. This classification was initially made by clinical observations on mammography, breast ultrasound, and breast MRI [19], but no study has been able to determine precisely the malignancy of these nodules [20]. Through the creation of this classification, numerous shape parameters are used on different image modalities such as "skin", "vascular", "coarse or popcorn-like", "large rod-like" [21]. The Breast Nodule Classification Challenge was focused on ultrasound, one single image, and some patient data (age, gender), which represented a major challenge and also a significant step forward for clinicians [22]. Three machine learning algorithms were used for this challenge and the results showed that AI remains significantly less effective alone than with the help of a radiologist. Therefore, as anticipated, results were quite disappointing. Nevertheless, the development of deep learning algorithms for breast nodule classification should be encouraged as computer-aided diagnosis tools [23].

Presence of pathological lymph nodes is an important finding in patients with ENT infection or cancer. The presence of cervical lymph nodes has an initial prognostic impact in ENT cancers taking into account by the Tumor Node Metastasis classification [24]. In addition, lymph node is a frequent form of recurrence of ENT cancers. A meta-analysis gives a precise definition of the radiological semiology of ENT adenopathy as well as the performances of the semiological signs and imaging techniques for the diagnosis [25]. There are only a few articles on the detection or contouring of lymph nodes whereas studies have used texture analysis from DECT data to classify lymph nodes into pathological or benign ones [26]. For the Pathological Lymph Nodes challenge, machine learning algorithms were used on DECT data, and on a rather limited database (50 patients with 412 lymph nodes). The objective was to detect and contour lymph nodes from 3D DECT examinations, which was challenging. However, GAMC succeeded to build a solution yielding a Dice score of 0.631.

In conclusion, three Data Challenges with over 2000 GDPR-compliant, multi-centric, 3D CT and ultrasound databases were organized for 6 multidisciplinary teams. For the future challenges, a single platform to manage the whole data challenge workflow is recommended.

It could be also stimulating to work on one or two objectives but to increase the number of data for each of them.

Acknowledgements

All authors thank the Société Française de Radiologie for the opportunity to organize this challenge and its support and Cleverdoc© for the platform and for the data hosting.

Human and animal rights

The authors declare that the work described has been carried out in accordance with the Declaration of Helsinki of the World Medical Association revised in 2013 for experiments involving humans as well as in accordance with the EU Directive 2010/63/EU for animal experiments.

Informed consent and patient details

The authors declare that they obtained a written informed consent from the patients and/or volunteers included in the article. The authors also confirm that the personal details of the patients and/or volunteers have been removed.

All the stages of the project, from data uploading to the challenge phase, was designed to be aligned with the GDPR. The Commission Nationale Informatique et Libertés was consulted, and the SFR assumed the role of Data Protection Officer. Automatic anonymization of the examinations was ensured by the platform developed by the startup Cleverdoc©. Therefore, the image data was anonymized during the upload, if not anonymized before such that the participants received anonymized medical data. This process ensured the protection of patients' data. In addition, every radiologist involved was asked to send an information letter to patients about the use of their medical examinations, with the option of refusal of consent. A data chart was also sent to each radiologist to abide by the GDPR laws, as well as guidelines on the terms of use of the data for the participants to the challenge.

Disclosure of interest

The authors declare that they have no competing interest.

Funding

This work did not receive any grant from funding agencies in the public, commercial, or not-for-profit sectors.

Author contribution

All authors attest that they meet the current International Committee of Medical Journal Editors (ICMJE) criteria for Authorship.

References

[1] Syeda-Mahmood T. Role of big data and machine learning in diagnostic decision support in radiology. J Am Coll Radiol 2018;15:569-576.

[2] Recht M, Bryan RN. Artificial intelligence: threat or boon to radiologists? J Am Coll Radiol 2017;14:1476-1480.

[3] Thrall JH, Li X, Li Q, Cruz C, Do S, Dreyer K, et al. Artificial intelligence and machine learning in radiology: opportunities, challenges, pitfalls, and criteria for success. J Am Coll Radiol 2018;15:504-508.

[4] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Med Image Anal 2017;42:60-88.

[5] European Society of Radiology (ESR). What the radiologist should know about artificial intelligence – an ESR white paper. Insights Imaging 2019;10:44.

[6] Liew C. The future of radiology augmented with artificial intelligence: a strategy for success. Eur J Radiol 2018;102:152-156.

[7] Pesapane F, Volonte C, Codari M, Sardanelli F. Artificial intelligence as a medical device in radiology: ethical and regulatory issues in Europe and the United States. Insights Imaging 2018;9:745-753.

[8] Lassau N, Estienne T, de Vomecourt P, Azoulay M, Cagnol J, Garcia G, et al. Five simultaneous artificial intelligence data challenges on ultrasound, CT, and MRI. Diagn Interv Imaging 2019;100:199-209.

[9] Lassau N, Bousaid I, Chouzenoux E, Lamarque JP, Charmettant B, Azoulay M et al. Three artificial intelligence data challenges based on CT and MRI. Diagn Interv Imaging 2020;101:783-788.

[10] Roblot V, Giret Y, Bou Antoun M, Morillot C, Chassin X, Cotten A, et al. Artificial intelligence to diagnose meniscus tears on MRI. Diagn Interv Imaging. 2019 Apr;100(4):243-249.

[11] Couteaux V, Si-Mohamed S, Nempont O, Lefevre T, Popoff A, Pizaine G, et al. Automatic knee meniscus tear detection and orientation classification with Mask-RCNN. Diagn Interv Imaging 2019;100:235-242.

[12] Santin M, Brama C, Théro H, Ketheeswaran E, El-Karoui I, Bidault F, et al. Detecting abnormal thyroid cartilages on CT using deep learning. Diagn Interv Imaging 2019;100:251-257.

[13] Herent P, Schmauch B, Jehanno P, Dehaene O, Saillard C, Balleyguier C, et al. Detection and characterization of MRI breast lesions using deep learning. Diagn Interv Imaging 2019;100:219-225.

[14] Schmauch B, Herent P, Jehanno P, Dehaene O, Saillard C, Aubé C, et al. Diagnosis of focal liver lesions from ultrasound using deep learning. Diagn Interv Imaging 2019;100:227-233.

[15] Couteaux V, Si-Mohamed S, Renard-Penna R, Nempont O, Lefevre T, Popoff A, et al. Kidney cortex segmentation in 2D CT with U-Nets ensemble aggregation. Diagn Interv Imaging 2019;100:211-217.

[16] Blanc-Durand P, Schiratti JB, Schutte K, Jehanno P, Herent P, Pigneur F, et al. Abdominal musculature segmentation and surface prediction from CT using deep learning for sarcopenia assessment. Diagn Interv Imaging 2020;101:789-794.

[17] Roca P, Attye A, Colas L, Tucholka A, Rubini P, Cackowski S, et al; OFSEP Investigators; Steering Committee; Investigators; Imaging group. Artificial intelligence to predict clinical disability in patients with multiple sclerosis using FLAIR MRI. Diagn Interv Imaging 2020;101:795-802.

[18] Blanc D, Racine V, Khalil A, Deloche M, Broyelle JA, Hammouamri I, et al. Artificial intelligence solution to classify pulmonary nodules on CT. Diagn Interv Imaging 2020;101:803-810.

[19] Lee KA, Talati N, Oudsema R, Steinberger S, Margolies LR. BIRADS 3: Current and future use of probably benign. Curr Radiol Rep 2018;6:5.

[20] Melnikow J, Fenton JJ, Whitlock EP, Miglioretti DL, Weyrich MS, Thompson JH, et al. Supplemental screening for breast cancer in women with dense breasts: a systematic review for the U.S. Preventive Service Task Force. Ann Intern Med 2016;164:268-78.

[21] Spak DA, Plaxco JS, Santiago L, Dryden MJ, Dogan BE. BI-RADS fifth edition: A summary of changes. Diagn Interv Imaging 2017;98:179-190

[22] Ghosh A. Artificial intelligence using open source BI-RADS data exemplifying potential future use. J Am Coll Radiol 2019;16:64-72

[23] Morgan MB, Mates JL. Applications of artificial intelligence in breast imaging. Radiol Clin North Am 2021;59:139-148.

[24] James D. Brierley, Mary K. Gospodarowicz and Chritian Wittekind (Eds.).TNM classification of malignant tumours, Eighth Edition (2017). John Wiley & Sons, Ltd.

[25] De Bondt RBJ, Nelemans PJ, Hofman PAM. Detection of lymph node metastases in head and neck cancer: a meta-analysis comparing US, USgFNAC, CT and MR imaging. Eur J Radiol 2007;64:266-72.

[26] Seidler M, Forghani B, Reinhold C, Pérez-Lara A. Dual-energy CT texture analysis with machine learning for the evaluation and characterization of cervical lymphadenopathy. Comput Struct Biotechnol J 2019;17:1009-1015

Figures & tables

Table 1. Participating centers that provided imaging data for the three Data Challenges.

Table 2. Examinations received and included for each challenge.

Table 3. Training, validation, and testing dataset for each challenge.

Table 4. Classification of data for each challenge.

Breast nodules	Pathological lymph nodes	Coronary calcification
Ultrasound (n = 672)	3D DECT (n = 737)	3D CT (n = 667)
Gustave Roussy* (n = 309) Centre Hospitalier de Valenciennes* (n = 101) Centre Oscar Lambret† (n = 76) Clinique St Antoine‡ (n = 67) Institut Bergonié† (n = 58) Institut de Cancérologie de l'Ouest† (n = 20) Centre intercommunal de Créteil*(n = 20) Hôpital de la Croix-Rousse* (n = 11) Hôpital Lyon Sud* (n = 7) Centre Hôspitalier de Bastia* (n = 2) Centre Hospitalier Universitaire de Lille* (n = 1)	Centre Libéral Imagerie Médicale Agglomération Lille‡ (n = 495) Gustave Roussy* (n = 229) Hôpital Tenon* (n = 10) Hôpital La conception* (n = 3)	Hôpital de Perpignan* (n = 244) Hôpital Européen Georges Pompidou* (n = 169) Centre Hospitalier Universitaire de Strasbourg* (n = 121) Groupe Vidi‡ (n = 95) Centre Hospitalier de Douais* (n = 38)

CT indicates computed tomography; DECT indicates double-energy computed tomography

*Public hospital; †Non-profit private clinic; ‡ Private clinic

Challenge	Number of medical examinations received	Number of medical examinations kept in the dataset
Breast nodules	672	598 (89%)
Pathological lymph nodes	737	645 (88%)
Coronary calcification	667	594 (89%)
Total	**2076**	**1837 (88%)**

Challenge	Breast nodules	Pathological lymph nodes	Coronary calcification
1: Training set	100 (17%)	108 (17%)	100 (17%)
2: Training set + Validation set	361 (60%)	395 (61%)	396 (67%)
3: Test set	137 (23%)	142 (22%)	98 (16%)
Total (# examinations)	**598**	**645**	**594**

	Breast nodules		Pathological lymph nodes		Coronary calcification	
Number of examinations received	**598**		**645**		**594**	
Classifications achieved as per the data challenge	Benign	385 (64.4%)	Normal	229 (35.2%)	Class A	129 (21.5%)
	Malignant	213 (35.6%)	Pathological	416 (64.8%)	Class B	72 (12%)
					Class C	126 (21%)
					Class D	102 (17%)
					Class E	119 (20%)
					Unclassified	46 (7.5%)