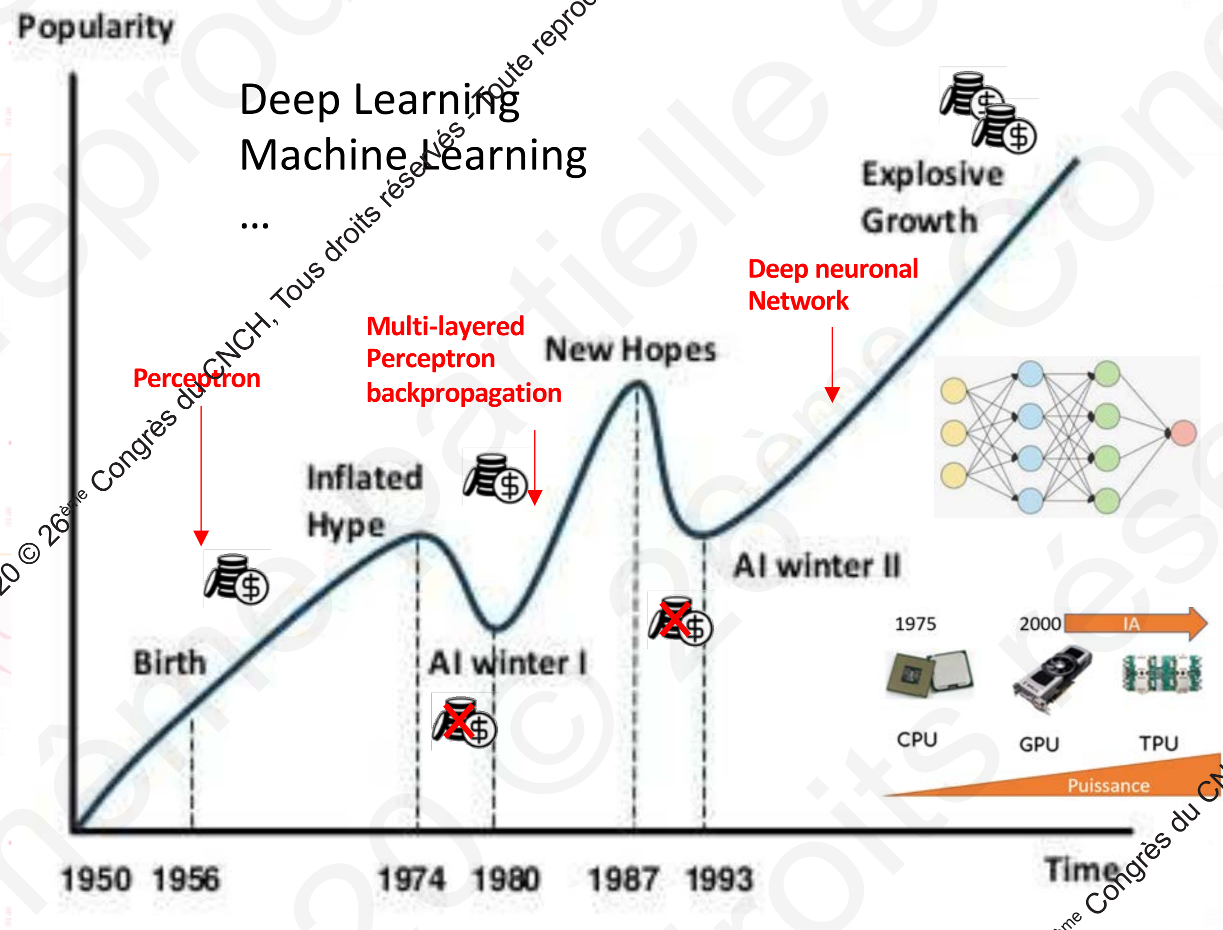
26ème CONGRÈS DU CNCH
Digital & interactif

IA
une chance pour l'expert ?

Dr Eric Saloux
Novembre 2020 CNCH

26ème CONGRÈS DU CNCH
Digital & interactif

IA : est-ce nouveau ?



Timeline of AI Development

- **1950s-1960s**: First AI boom - the age of reasoning, prototype AI developed
- **1970s**: AI winter I
- **1980s-1990s**: Second AI boom: the age of Knowledge representation (appearance of expert systems capable of reproducing human decision-making)
- **1990s**: AI winter II
- **1997**: Deep Blue beats Gary Kasparov
- **2006**: University of Toronto develops Deep Learning
- **2011**: IBM's Watson won Jeopardy
- **2016**: Go software based on Deep Learning beats world's champions

https://www.actuaries.digital/2018/09/05/history-of-ai-winters/

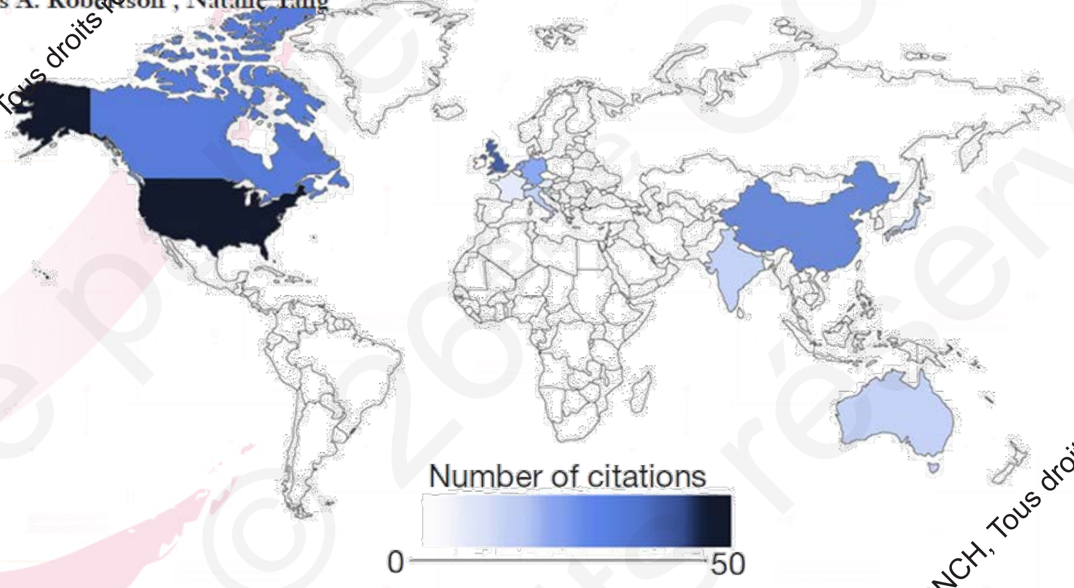
IA: Etat de l'art

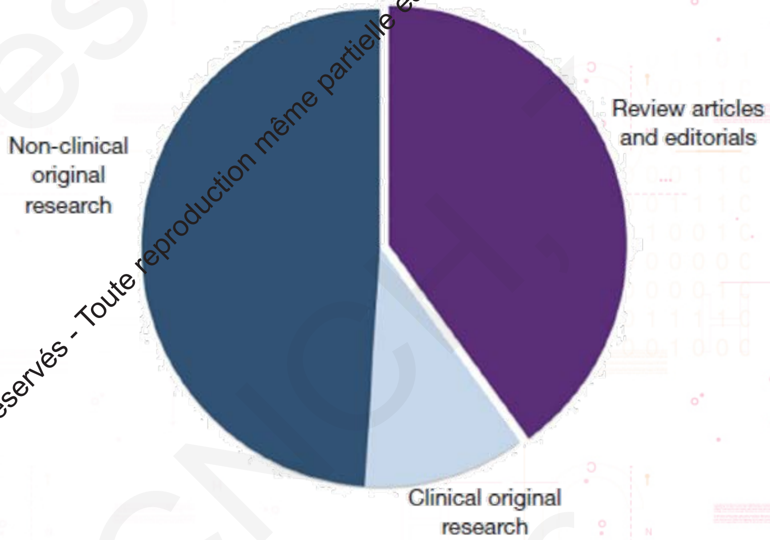
Original Article
Page 1 of 12

The top 100 most cited articles in medical artificial intelligence: a bibliometric analysis

Subhashaan Sreedharan[1], Mustafa Mian[1], Ross A. Robertson[1], Natalie Yang[1,2]

Decade	No of articles
1970–1979	1
1980–1989	2
1990–1999	12
2000–2009	50
2010–2019	35



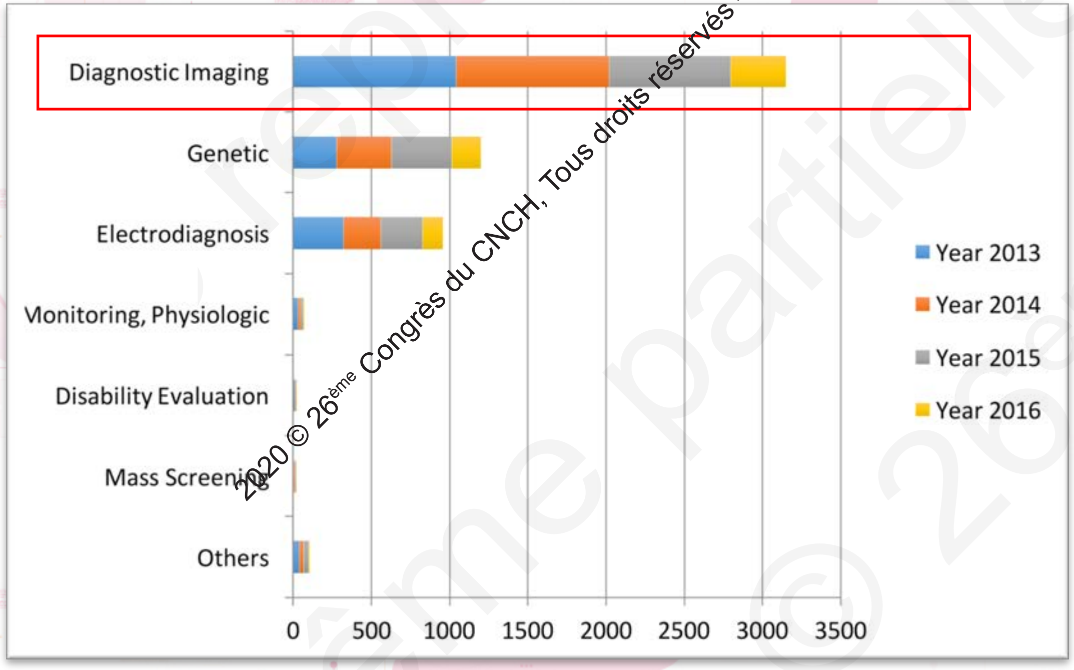


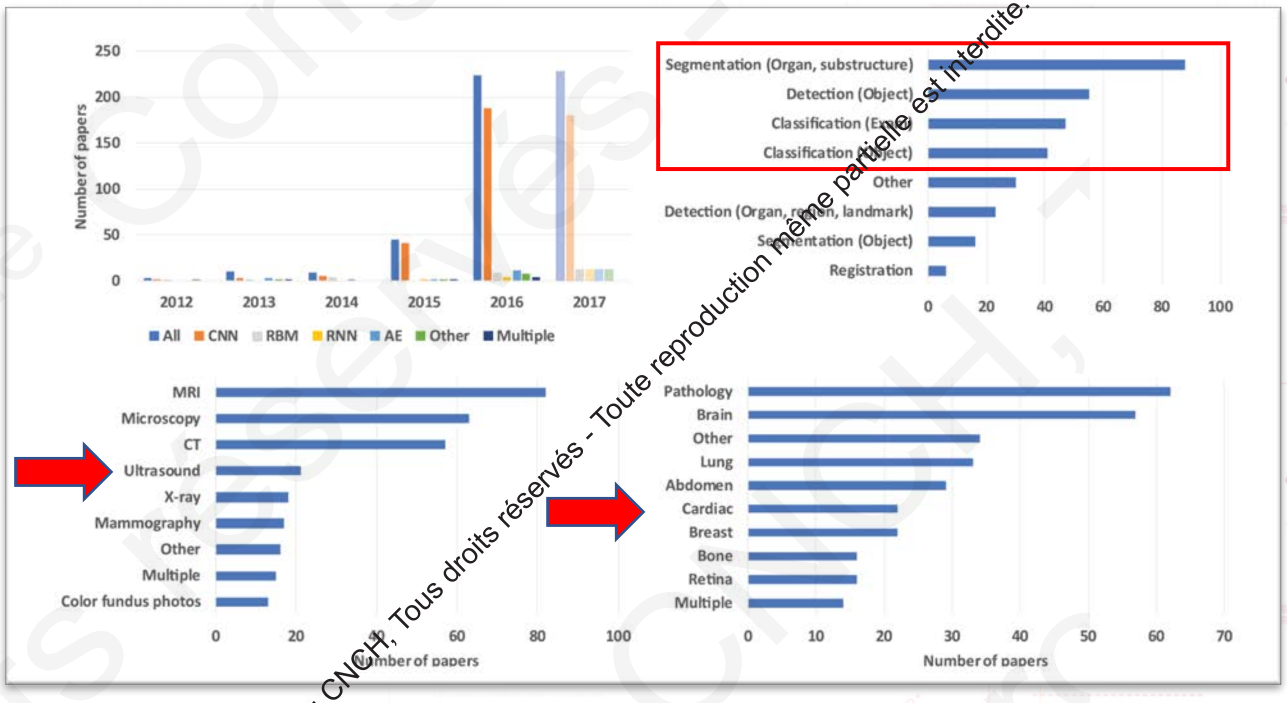
Conclusions: This study highlights that the current citation classics are largely in the non-clinical, experimental phase and have yet to progress to the clinical, integration phase of medical AI.

2020 © 26ème Congrès du CNCH, Tous droits réservés - Toute reproduction même partielle est interdite.

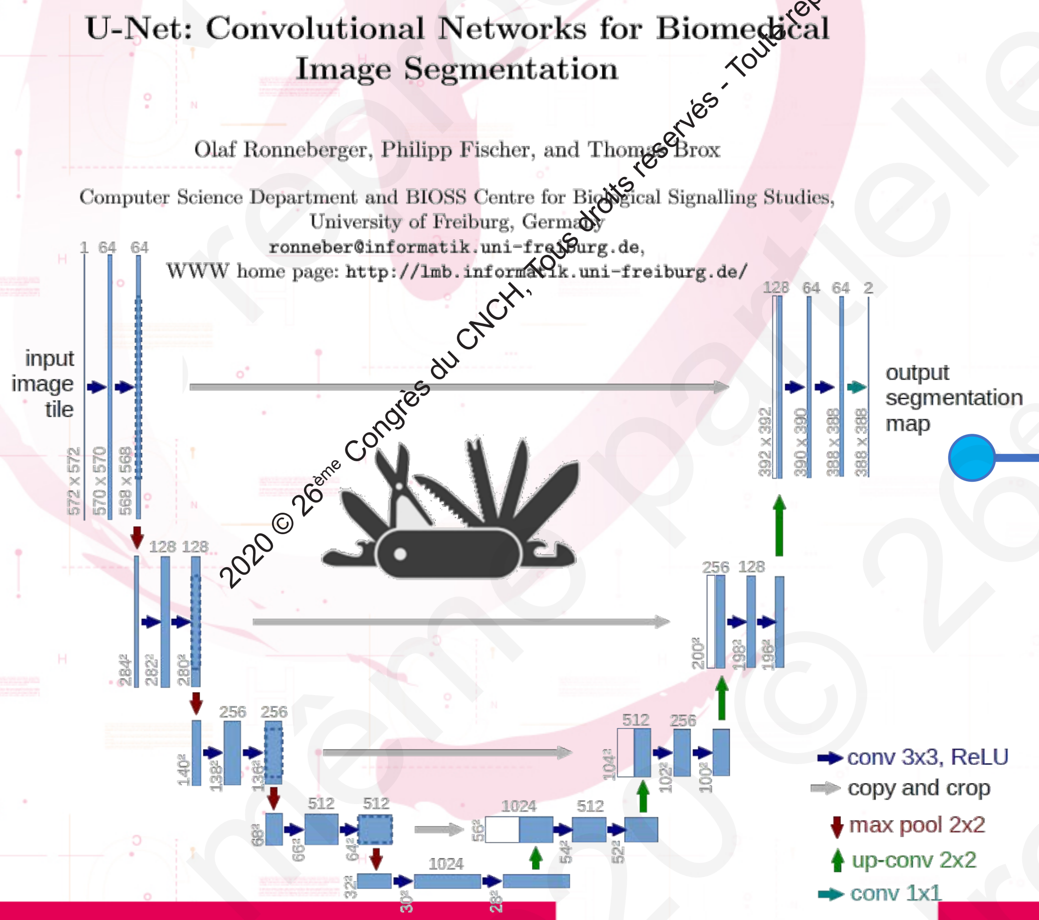
26ème CONGRÈS DU CNCH
Digital & interactif

IA: Etat de l’art





IA: l’univers des outils



Instance Segmentation

Object Detection

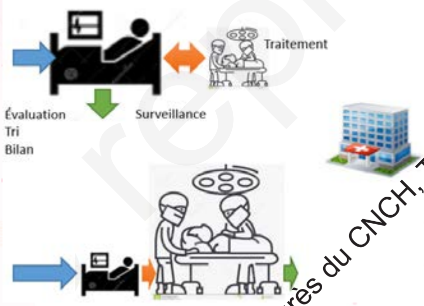
Springer, LNCS, Vol.9351: 234--241, 2015,

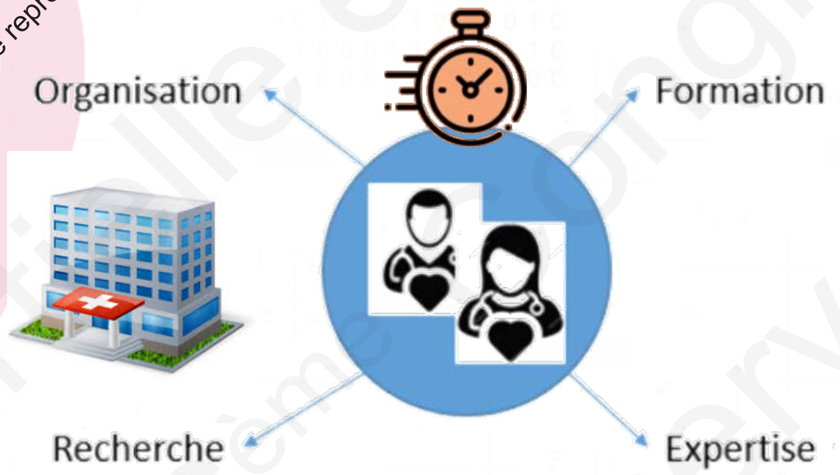
2020 © 26ème Congrès du CNCH, Tous droits réservés - Toute reproduction même partielle est interdite.

26ème CONGRÈS DU CNCH
Digital & interactif

IA: un terreau fertile ?

Contraintes





Rôle propre

2020 © 26ème Congrès du CNCH, Tous droits réservés - Toute reproduction même partielle est interdite.

26ème CONGRÈS DU CNCH
Digital & interactif

IA: cela marche-t-il ?

Intra-observer and overall agreement in the radiological assessment of tuberculosis

J-P. Zellweger,* R. Heinzer,* M. Touray,* B. Vidondo,† E. Altpeter†

* University Medical Outpatient Department, University Hospital, Lausanne, † Swiss Federal Office of Public Health, Berne, Switzerland

SUMMARY

OBJECTIVE: To assess the intra-observer and overall agreement in the interpretation of chest X-rays (CXRs) performed for detecting tuberculosis (TB) among immigrants in Switzerland.

METHOD: Four hundred digitalised CXRs from the files of immigrant registration centres were selected and read twice in random order by three readers. The readers had to assess 1) if the picture was normal or abnormal; 2) if an abnormality was suggestive of TB; and 3) if the suspicion of TB needed an immediate examination (potentially smear-positive TB). The intra-observer and overall agreements were expressed as kappa with standard error.

RESULTS: Due to losses for technical reasons, 377 of the 400 pictures were analysed. The intra-observer agreement was 0.39–0.90 for any abnormality, and 0.60–0.82 for TB needing an urgent examination. The overall agreements were: 0.55 (all three readers) and 0.84 (two best readers) for any abnormality, and 0.64 (all three readers) and 0.80 (two best readers) for active TB.

CONCLUSIONS: The intra-observer and overall agreements for the detection of abnormalities on digitalised CXRs and for the presence of possible active TB depend on the reader's experience. It was good between experienced readers and fair between and within the inexperienced reader.

KEY WORDS: tuberculosis; screening; immigrants; radiology; observer agreement

-The intra-observer agreement was 0.39–0.90 for any abnormality, and 0.60-0.82 for TB needing an urgent examination.
-The intra-observer and overall agreements for the detection of abnormalities on digitalised CXRs and for the presence of possible active TB <u>depend on the reader's experience.</u>

INT J TUBERC LUNG DIS 10(10):1123–1126 2006

Chest x-ray analysis with deep learning-based software as a triage test for pulmonary tuberculosis: a prospective study of diagnostic accuracy for culture-confirmed disease

Faiz Ahmad Khan, Arman Majidulla, Gamuchirai Tavaziva, Ahsana Nazish, Syed Kumail Abidi, Andrea Benedetti, Dick Menzies, James C Johnston, Aamir Javed Khan, Saima Saeed

Summary

Background Deep learning-based radiological image analysis could facilitate use of chest x-rays as triage tests for pulmonary tuberculosis in resource-limited settings. We sought to determine whether commercially available chest x-ray analysis software meet WHO recommendations for minimal sensitivity and specificity as pulmonary tuberculosis triage tests.

Methods We recruited symptomatic adults at the Indus Hospital, Karachi, Pakistan. We compared two software, qXR version 2.0 (qXRv2) and CAD4TB version 6.0 (CAD4TBv6), with a reference of mycobacterial culture of two sputa. We assessed qXRv2 using its manufacturer prespecified threshold score for chest x-ray classification as tuberculosis present versus not present. For CAD4TBv6, we used a data-derived threshold, because it does not have a prespecified one. We tested for non-inferiority to preset WHO recommendations (0·90 for sensitivity, 0·70 for specificity) using a non-inferiority limit of 0·05. We identified factors associated with accuracy by stratification and logistic regression.

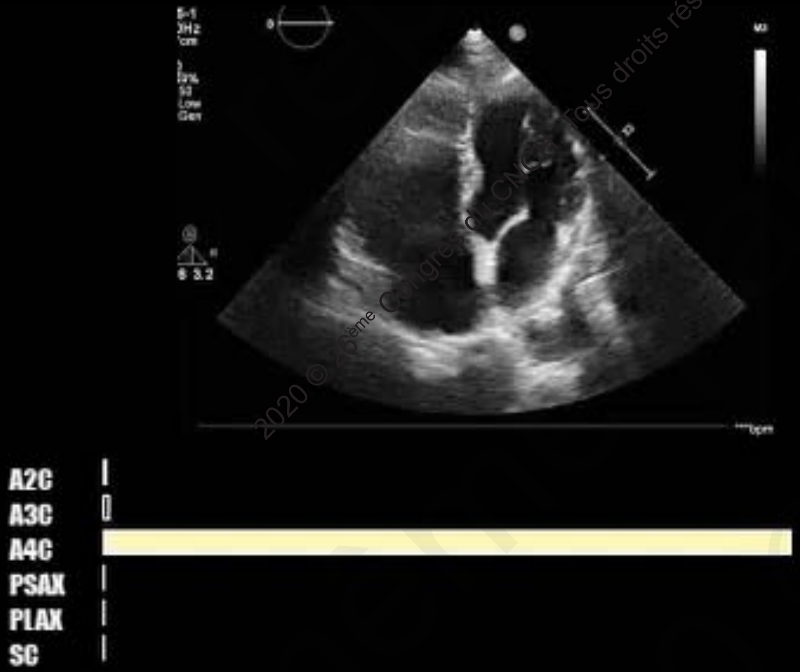
Findings We included 2198 (92·8%) of 2370 enrolled participants. 2187 (99·5%) of 2198 were HIV-negative, and 272 (12·4%) had culture-confirmed pulmonary tuberculosis. For both software, accuracy was non-inferior to WHO-recommended minimum values (qXRv2 sensitivity 0·93 [95% CI 0·89–0·95], non-inferiority p=0·0002; CAD4TBv6 sensitivity 0·93 [0·90–0·96], p<0·0001; qXRv2 specificity 0·75 [0·73–0·77], p<0·0001; CAD4TBv6 specificity 0·69 [0·67–0·71], p=0·0003). Sensitivity was lower in smear-negative pulmonary tuberculosis for both software, and in women for CAD4TBv6. Specificity was lower in men and in those with previous tuberculosis, and reduced with increasing age and decreasing body mass index. Smoking and diabetes did not affect accuracy.

Interpretation In an HIV-negative population, these software met WHO-recommended minimal accuracy for pulmonary tuberculosis triage tests. Sensitivity will be lower when smear-negative pulmonary tuberculosis is more prevalent.

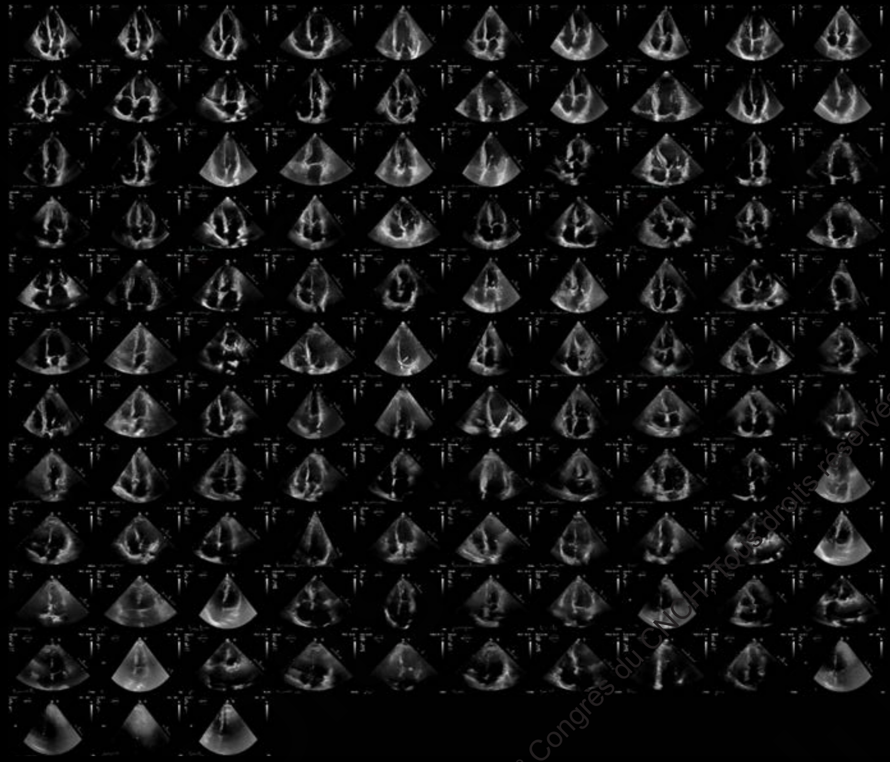
Funding Canadian Institutes of Health Research.

thelancet.com/digital-health Vol 2 November 2020

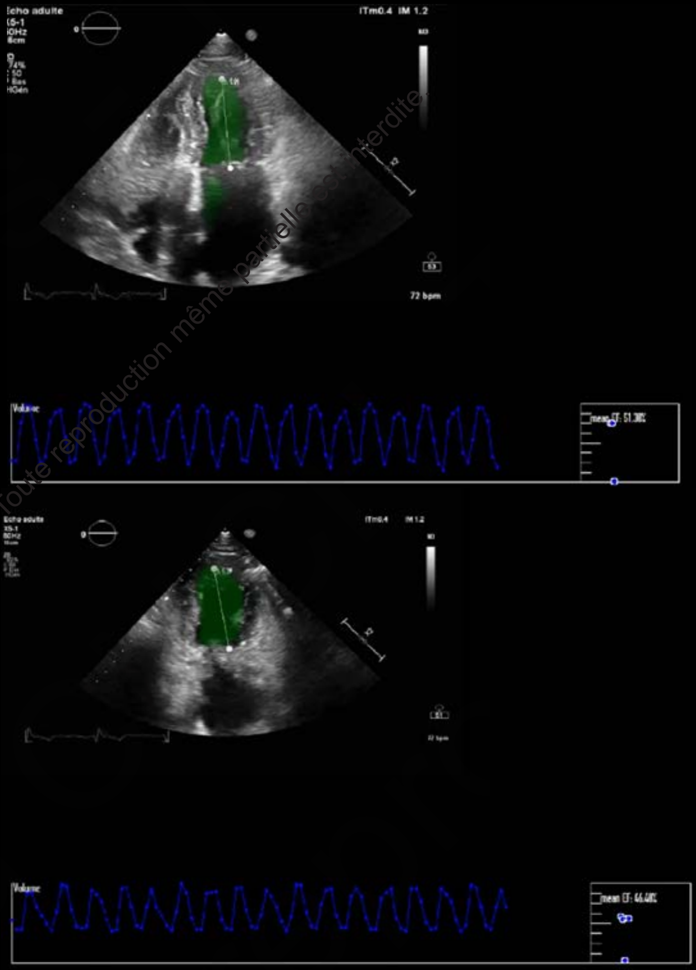
IA: cela marche-t-il ?



Reconnaissance de vue



Tri qualité



Segmentation

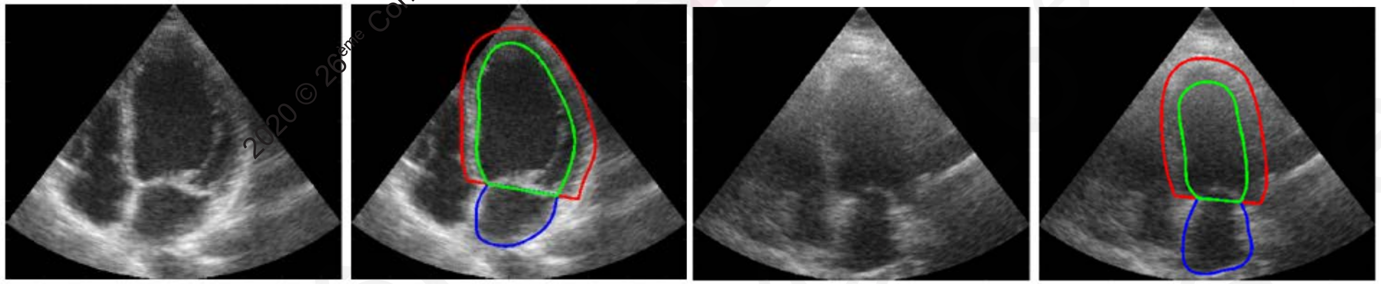
26ème CONGRÈS DU CNCH
Digital & interactif

IA: mieux qu'avant ?

> IEEE Trans Med Imaging. 2019 Sep;38(9):2198-2210. doi: 10.1109/TMI.2019.2900516.
Epub 2019 Feb 22.

Deep Learning for Segmentation Using an Open Large-Scale Dataset in 2D Echocardiography

Sarah Leclerc, Erik Smistad, Joao Pedrosa, Andreas Ostvik, Frederic Cervenansky, Florian Espinosa, Torvald Espeland, Erik Andreas Rye Berg, Pierre-Marc Jodoin, Thomas Grenier, Carole Lartizien, Jan Dhooge, Lasse Lovstakken, Olivier Bernard



(a) Good image quality (c) Poor image quality

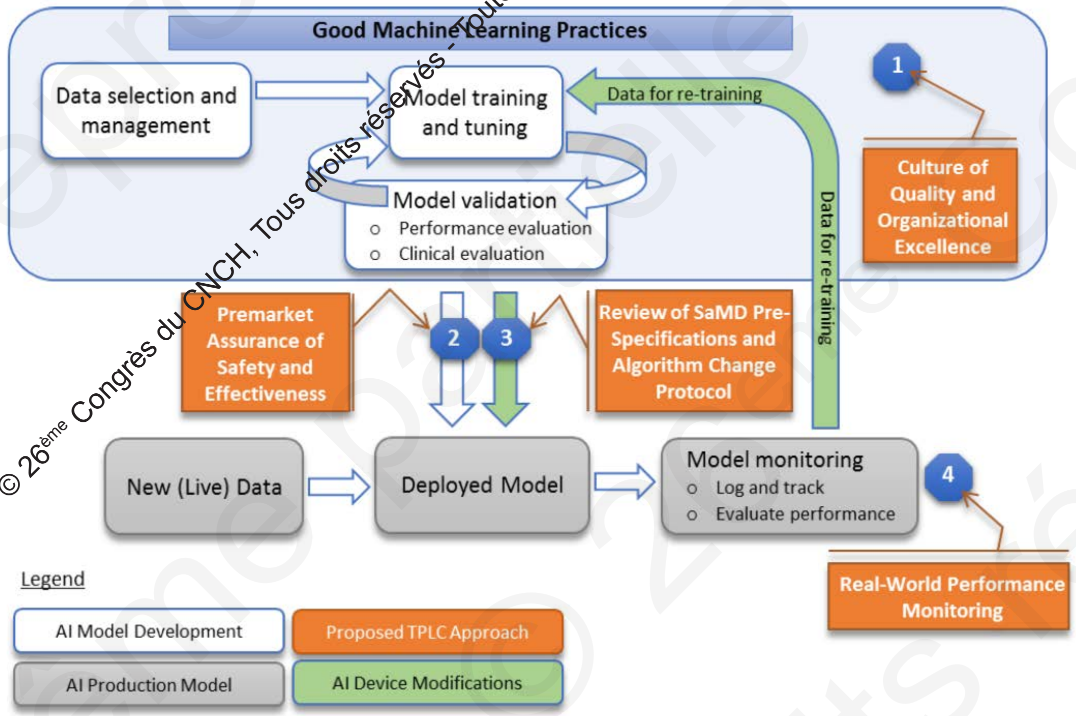
Abstract—Delineation of the cardiac structures from 2D echocardiographic images is a common clinical task to establish a diagnosis. Over the past decades, the automation of this task has been the subject of intense research. In this paper, we evaluate how far the state-of-the-art encoder-decoder deep convolutional neural network methods can go at assessing 2D echocardiographic images, *i.e.* segmenting cardiac structures as well as estimating clinical indices, on a dataset especially designed to answer this objective. We therefore introduce the Cardiac Acquisitions for Multi-structure Ultrasound Segmentation (CAMUS) dataset, the largest publicly-available and fully-annotated dataset for the purpose of echocardiographic assessment. The dataset contains two and four-chamber acquisitions from 500 patients with reference measurements from one cardiologist on the full dataset and from three cardiologists on a fold of 50 patients. Results show that encoder-decoder based architectures outperform state-of-the-art non-deep learning methods and faithfully reproduce the expert analysis for the end-diastolic and end-systolic left ventricular volumes, with a mean correlation of 0.95 and an absolute mean error of 9.5 ml. Concerning the ejection fraction of the left ventricle, results are more contrasted with a mean correlation coefficient of 0.80 and an absolute mean error of 5.6%. Although these results are below the inter-observer scores, they remain slightly worse than the intra-observer's ones. Based on this observation, areas for improvement are defined, which open the door for accurate and fully-automatic analysis of 2D echocardiographic images.

CNN, Dataset, Comparaison à expert, Critères de comparaison

26ème CONGRÈS DU CNCH
Digital & interactif

IA: est ce validé ?

FDA (2016 – 2017)



HAS (2020)

HAS HAUTE AUTORITÉ DE SANTÉ

Consultation publique

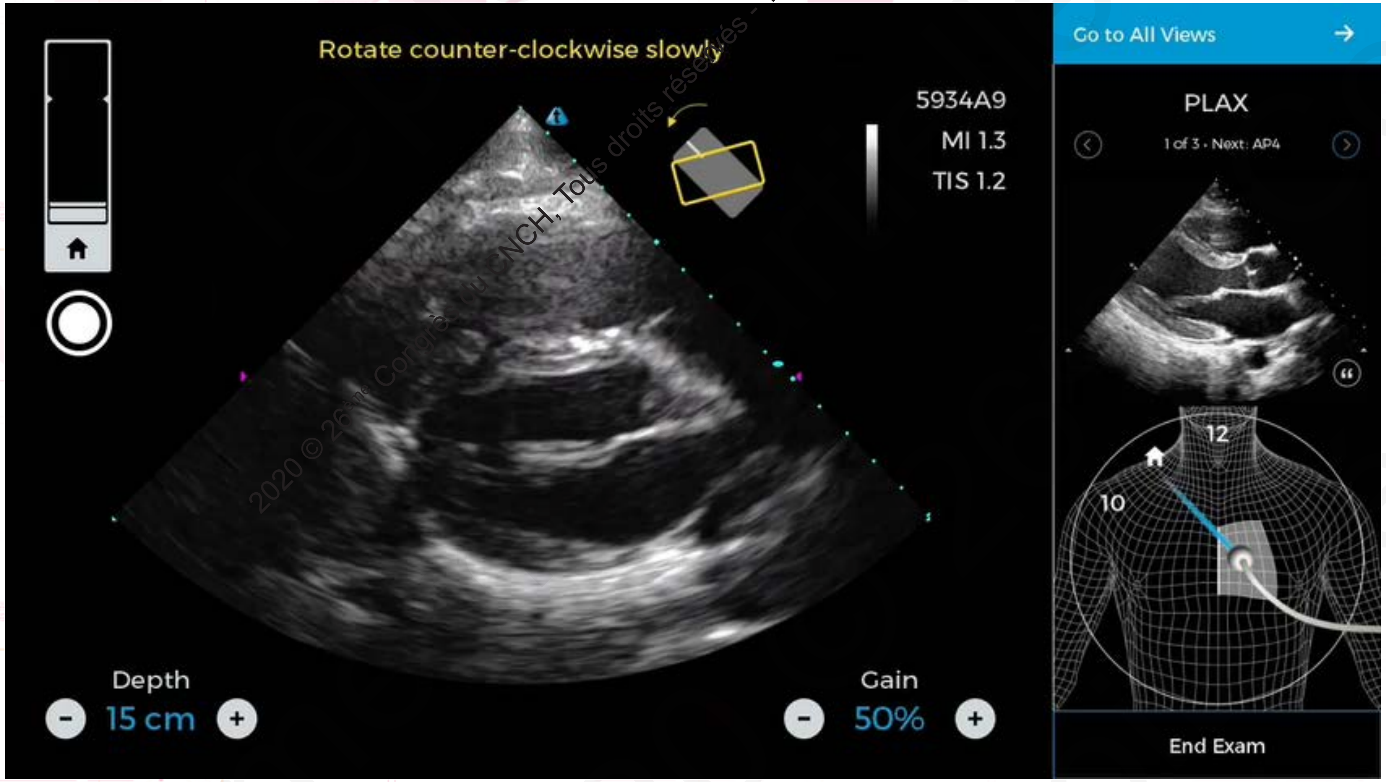
sur le projet de grille d'analyse destinée à être utilisée par la CNEDiMTS pour contribuer à son évaluation de **dispositifs médicaux embarquant des systèmes décisionnels s'appuyant sur des procédés d'apprentissage automatique (« Intelligence artificielle »)**

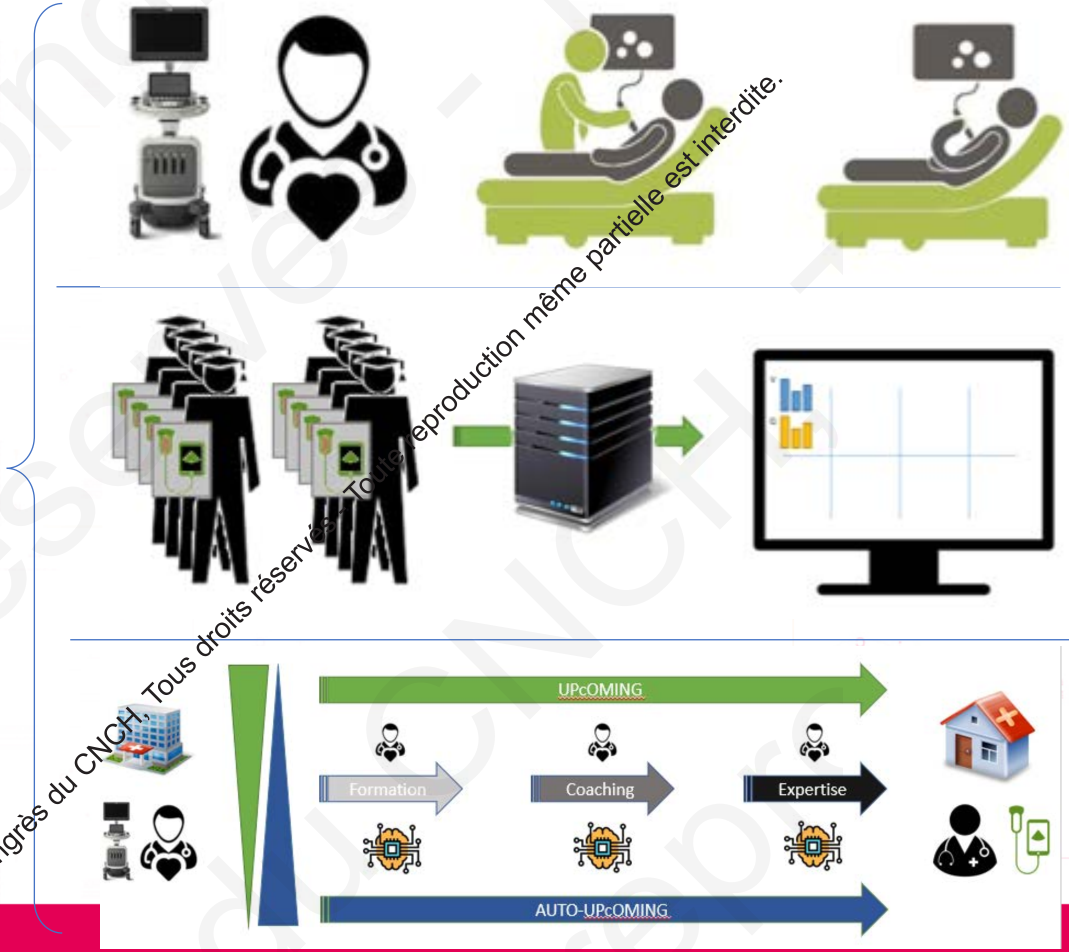
Cette grille comporte 8 catégories de critères qui concernent :

- **la finalité d'usage**
- **l'apprentissage**
- **les données d'entrée d'apprentissage initial ou de réapprentissage**
- **les données d'entrée impliquées dans la décision**
- **la performance**
- **la validation**
- **la résilience du système**
- **l'explicabilité / interprétabilité.**

IA en cardiologie

Caption Health





26ème CONGRÈS DU CNCH
Digital & interactif

En conclusion

- Oui l'IA est une chance, un espoir pour l'expert
  - Leurs performances explosent celles des algorithmes classiques
    - Capables de réaliser des taches impossibles (ou presque) avec les algo conventionnels
    - Capables de travailler en temps réel
  - Certains outils sont validés FDA au même titre qu'un appareil biomédical
  - Gain de temps précieux pour l'expert
  - L'informatique métier est et sera transformée par ces outils
- Cependant…
  - Le domaine d'applicabilité est actuellement restreint (imagerie >>> clinique)
    - Problème de la structure des DB médicales (PACS plus facile à exploiter que les données éparses du dossier médical).
  - L'évaluation en terme de santé publique reste à prouver
    - Critères cliniques d'évaluation != critères algorithmiques
    - Populations != datasets
  - Les traitements personnalisés et l'intelligence augmentée nécessitent des efforts consubstantiels

26ème CONGRÈS DU CNCH
Digital & interactif

Collège National des Cardiologues des Hôpitaux

Suivez le CNCH sur le Social Média !

#CNCHcongres

@CNCHcollege

@CNCHcollege

2020 © 26ème Congrès du CNCH, Tous droits réservés - Toute reproduction même partielle est interdite.

Si vous voulez devenir Ambassadeur social média CNCH adressez-nous un email à cnch@sfcardio.fr