

AI applications in routine clinical imaging

Citation for published version (APA):

Vaidyanathan, A. (2023). *AI applications in routine clinical imaging: detection, segmentation, diagnosis, and prognosis*. [Doctoral Thesis, Maastricht University]. Maastricht University. <https://doi.org/10.26481/dis.20230228av>

Document status and date:

Published: 01/01/2023

DOI:

[10.26481/dis.20230228av](https://doi.org/10.26481/dis.20230228av)

Document Version:

Publisher's PDF, also known as Version of record

Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

www.umlib.nl/taverne-license

Take down policy

If you believe that this document breaches copyright please contact us at:

repository@maastrichtuniversity.nl

providing details and we will investigate your claim.

SUMMARY

PART 1 – AI based diagnostic models for respiratory diseases

In **chapter 2**, we developed and externally validated a machine learning model that is able to discriminate between COVID-19 positive and negative patients, and which has been trained and validated using a regularized logistic regression model. The model showed an AUC of 0.882 (95% CI: 0.851–0.913) in the independent test dataset (641 patients). The optimal decision threshold, considering the cost of false negatives twice as high as the cost of false positives, resulted in an accuracy of 85.18%, a sensitivity of 69.52%, a specificity of 91.63%, a negative predictive value (NPV) of 94.46% and a positive predictive value (PPV) of 59.44%.

In **chapter 3**, We developed and externally validated a deep learning AI model for the classification of no-infection, COVID-19, or Influenza/CAP cases based upon CT imaging. The model showed a performance in the external validation set with an AUC of 0.90, 0.92 and 0.92 for COVID-19, Influenza/CAP and No infection respectively. The selection of the input slices based on automatic segmentation of the abnormalities in the lung reduces analysis time (56 second per scan) and computational burden of the model.

In **chapter 4**, We have developed and externally validated an AI model for classification of pulmonary embolism in CTPA images. The model showed an area under the curve (AUC) of 0.86 [0.800-0.919], a sensitivity of 82.68 % [75.16 - 88.27] and a specificity of 79.31 % [61.61 - 90.15] on the external validation set. The activation maps of the slices rightly predicted positive by the PE classifier showed good visual correspondence with areas of PE. This was also quantitatively confirmed as 79.2% of PE regions in the GT were highlighted in the activation maps and the percentage of activated regions corresponding to GT PE is 80.3%.

PART 2 – AI based models for treatment outcome prediction and detection of disease in oncological use cases

In **chapter 5**, we have externally validated the prognostic value of Signature-0 in a prospective cohort using a manual and an automated segmentation method for survival prediction. The results of this study showed that the original radiomics signature, Signature-0, developed in 2012, outperforms the contemporaneous standard of care (TNM 6th edition) producing superior stratification between survivors and non-survivors.

In **chapter 6**, we have developed and externally validated a DL based algorithm that is able to detect metastatic bone disease on Bone Scintigraphy images. The developed DL based algorithm is able to detect MBD on BSs, with high specificity and sensitivity (0.80 and 0.82 respectively on the external test set), in a shorter time compared to the nuclear medicine physicians (2.5 minutes for AI and 30 minutes for nuclear medicine physicians to classify 134 BSs), that could be applied to any BS regardless of the patient's gender and history of cancer.

PART 3 – AI based model for diagnosis of a disorder in inner ear

In **chapter 7**, we have validated a machine learning model trained on radiomics features extracted from inner ear region on MRI for diagnosis of Menière's disease. The classification accuracy of the model on the validation set was 82%, with a sensitivity of 83% and a specificity of 82%. The positive and negative predictive values were 71% and 90%, respectively.

In **chapter 8**, we have trained and externally validated an AI based model for auto-segmentation of inner ear on MRI images. The model showed precise Dice Similarity Coefficient scores (mean DSC- 0.8790) with a high True Positive Rate (91.5%) and low False Discovery Rate and False Negative Rates (14.8% and 8.49% respectively) across images from three different centers. The model proved to perform well with a DSC of 0.8768 on the clinical validation dataset.

SAMENVATTING

DEEL 1 – Op AI gebaseerde diagnostische modellen voor luchtwegaandoeningen

In **hoofdstuk 2** hebben we een machine learning-model ontwikkeld en extern gevalideerd dat onderscheid kan maken tussen COVID-19-positieve en negatieve patiënten, en dat is getraind en gevalideerd met behulp van een geregulariseerd logistisch regressiemodel. Het model vertoonde een AUC van 0,882 (95% BI: 0,851-0,913) in de onafhankelijke testdataset (641 patiënten). De optimale beslissingsdrempel, gezien de kosten van valse negatieven die twee keer zo hoog zijn als de kosten van valse positieven, resulteerde in een nauwkeurigheid van 85,18%, een gevoeligheid van 69,52%, een specificiteit van 91,63%, een negatief voorspellende waarde (NPV) van 94,46 % en een positief voorspellende waarde (PPV) van 59,44%.

In **hoofdstuk 3** hebben we een deep learning AI-model ontwikkeld en extern gevalideerd voor de classificatie van gevallen zonder infectie, COVID-19 of Influenza/CAP op basis van CT-beeldvorming. Het model liet een prestatie zien in de externe validatieset met een AUC van 0,90, 0,92 en 0,92 voor respectievelijk COVID-19, Influenza/CAP en Geen infectie. De selectie van de invoerschijfjes op basis van automatische segmentatie van de afwijkingen in de long vermindert de analysetijd (56 seconden per scan) en de rekenbelasting van het model.

In **hoofdstuk 4** hebben we een AI-model ontwikkeld en extern gevalideerd voor classificatie van longembolie in CTPA-beelden. Het model toonde een oppervlakte onder de curve (AUC) van 0,86 [0,800-0,919], een sensitiviteit van 82,68 % [75,16 - 88,27] en een specificiteit van 79,31% [61,61 - 90,15] op de externe validatieset. De activeringskaarten van de plakjes die terecht positief waren voorspeld door de PE-classificator, vertoonden een goede visuele overeenkomst met gebieden van PE. Dit werd ook kwantitatief bevestigd, aangezien 79,2% van de PE-regio's in de GT werden gemarkeerd in de activeringskaarten en het percentage geactiveerde regio's dat overeenkomt met GT PE is 80,3%.

DEEL 2 – Op AI gebaseerde modellen voor voorspelling van behandelresultaten en detectie van ziekte in oncologische gebruiksgevallen

In **hoofdstuk 5** hebben we de prognostische waarde van Signature-0 extern gevalideerd in een prospectief cohort met behulp van een handmatige en een geautomatiseerde segmentatiemethode voor overlevingsvoorspelling. De resultaten van deze studie toonden aan dat de originele radiomic-signatuur, Signature-0, ontwikkeld in 2012, beter presteert dan de huidige standaard van zorg (TNM 6e editie) en een superieure stratificatie produceert tussen overlevenden en niet-overlevenden.

In **hoofdstuk 6** hebben we een op DL gebaseerd algoritme ontwikkeld en extern gevalideerd dat metastatische botziekte kan detecteren op botsintigrafiebeelden. Het ontwikkelde op DL gebaseerde algoritme kan MBD detecteren op BS'en, met een hoge specificiteit en gevoeligheid (respectievelijk 0,80 en 0,82 op de externe testset), in een kortere tijd vergeleken met de nucleair geneeskundigen (2,5 minuten voor AI en 30 minuten voor nucleaire geneeskundigen om 134 BS'en te classificeren), die op elke BS kan worden toegepast, ongeacht het geslacht van de patiënt en de voorgeschiedenis van kanker.

DEEL 3 – AI-gebaseerd model voor diagnose van een stoornis in het binnenoor

In **hoofdstuk 7** hebben we een machinaal leermodel gevalideerd dat is getraind op radiomic-kenmerken die zijn geëxtraheerd uit het binnenoorgebied op MRI voor de diagnose van de ziekte van Menière. De classificatienauwkeurigheid van het model op de validatieset was 82%, met een

sensitiviteit van 83% en een specificiteit van 82%. De positief en negatief voorspellende waarden waren respectievelijk 71% en 90%.

In **hoofdstuk 8** hebben we een op AI gebaseerd model voor autosegmentatie van het binnenoer op MRI-beelden getraind en extern gevalideerd. Het model toonde nauwkeurige Dice-overeenkomstcoëfficiëntscores (gemiddelde DSC-0,8790) met een hoge True Positive Rate (91,5%) en een lage False Discovery Rate en False Negative Rates (14,8% en 8,49% respectievelijk) over afbeeldingen van drie verschillende centra. Het model bleek goed te presteren met een DSC van 0.8768 op de dataset voor klinische validatie.

3 REFERENTIES

- [1] W. Samek and K.-R. Müller, "Towards Explainable Artificial Intelligence," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 11700 LNCS, pp. 5–22, Sep. 2019, doi: 10.1007/978-3-030-28954-6_1.