

Machine learning for breast cancer diagnosis in developing countries

Citation for published version (APA):

Kakileti, S. T. (2020). *Machine learning for breast cancer diagnosis in developing countries*. ProefschriftMaken. <https://doi.org/10.26481/dis.20201109st>

Document status and date:

Published: 01/01/2020

DOI:

[10.26481/dis.20201109st](https://doi.org/10.26481/dis.20201109st)

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.

Appendix I

Summary

In this thesis, we investigated the feasibility of automated breast thermography using computer aided diagnosis for breast cancer in developing countries. We wanted to explore alternatives and think creatively about solving the problems in our setting (developing countries). Existing breast cancer detection technology poses some challenges in our setting due to minimal human expertise, high equipment cost, social and cultural barriers such as privacy and stigma. The technology developed during this thesis attempted to address some of these issues by combining the benefits of breast thermography with the advances in machine learning. A brief summary of different chapters presented in this thesis is described below.

Chapter 2 of this thesis is a review paper that was focused on understanding the challenges associated with the conventional breast imaging modalities and appreciating the potential benefits of infrared thermography as an imaging modality for breast cancer detection. We first explored the biological relation between the abnormal heat patterns and the presence of malignancy and provided a detailed review on the different criteria used for manual interpretation of the breast thermal images and the challenges such as high expertise, cognitive overload and subjectivity associated with the manual interpretation. This was followed by a survey on the existing large-scale studies since 1960, that showed the efficacy of breast thermography for malignancy detection. Finally, we studied the recent advances in the hardware and the software technologies of infrared breast thermography for achieving better accuracy with the combination of automated interpretation and highly sensitive infrared cameras.

To automate the crucial preprocessing steps in the imaging protocol, we discussed novel segmentation and view tagging approaches that are required for automated diagnosis. Due to the limited data available in the initial stages of experimentation, we proposed an image processing approach with heuristics as discussed in **Chapter 3** for tagging and segmentation of the thermal images. However, this technique was highly dependent on the imaging protocol and any deviation from the imaging protocol resulted in erroneous predictions and thus making it unsuitable for clinical practice. Hence, we proposed a deep learning architecture in **Chapter 4** to accurately predict the breast region and the view angle by learning high level semantics from a large training dataset. Interestingly, the obtained dice index of 0.92 (92%) with the deep learning approach was very close to the inter-observer dice correlation between two human experts. Also, the obtained dice index was approximately 25% higher than the prior heuristic based segmentation, demonstrating the strength of the deep learning architectures when there is large training data.

From the detected breast regions, blood vessels were automatically extracted as discussed in **Chapter 5** to study the vascular properties. Vascular changes are one of the early changes that occur inside the body during the onset of cancer. Hence, examination of this vascularity can help in early detection of cancerous regions. Automated

extraction of these vessels was challenging as the heat signature of the vessel boundaries diffused while its transmission to the breast surface. Hence, we proposed a three-level vessel enhancement technique followed by shape and temperature filters to detect potential vessel pixels. We also defined a new diffusivity metric to remove vessels with high diffusivity which allowed for the identification of the prominent vessel structures. The results with the proposed approach were observed to be better and less dependent on the threshold parameters when compared to vessel extraction techniques that were used for other imaging modalities.

Chapter 6 explored different deep learning architectures for extracting thermally active regions (hotspots) from the segmented breast regions using convolutional neural networks. The hotspot extraction helps in localization and understanding the nature and the properties of the abnormal region. We observed that the encoder-decoder architectures are better compared to the point-wise architectures in terms of inference time, dice index and accuracy for hotspot extraction in thermal images. Though the accuracies of the encoder-decoder architectures were high, these results were still inferior compared to an adaptive histogram thresholding approach. This could be due to the small size of the labeled dataset used for training the architectures. Therefore, in **Chapter 7**, we used an adaptive histogram thresholding for the hotspot detection and extracted medically interpretable features with the help of domain knowledge to characterize the detected hotspot regions. These features were then fed to train a random forest (RF) classifier to predict the probability of malignancy based on the hotspot properties.

Chapter 8 extended the scope of thermal imaging into prognosis by proposing an automated technique for estimating the hormonal status of the detected malignant regions from the breast thermal images. This hormonal status biomarker is an important prognostic factor for treatment planning and survival prediction. Our proposed approach is a preliminary study that showed an accuracy of 80% for estimating the hormonal status using a combination of machine learning and thermography. If these results are replicable on a large-scale population, this could eliminate the need for an additional invasive procedure to obtain these parameters, thereby reducing the cost and providing prompt treatment initiation.

Furthermore, we investigated the robustness of different machine learning classifiers for predicting the risk of breast cancer from the common breast cancer risk factors as discussed in **Chapter 9**. This risk estimation could help in identifying high-risk women, thereby providing personalized care and diagnosis to these women for either reducing the risk of breast cancer or for early detection of breast cancer. Incomplete or inaccurate information of the risk factors is a major problem that occurs due to fear of social stigmatization or patient's lack of knowledge about the risk factors and could lead to inaccurate risk prediction. Hence, we evaluated three prominent machine learning

classifiers for their robustness to the incomplete and inaccurate information. The proposed custom neural network architecture was found to be superior compared to logistic regression and random forests especially when less than 50% of the data was incorrect or missing.

Though the neural network architecture produced superior results compared to other classifiers even in the presence of incomplete or inaccurate information, the risk estimate was not accurate due to the weak correlation of the risk factors to the actual screening outcome. Therefore, we proposed a novel combination of machine learning with breast thermography and non-imaging features such as lump and nipple discharge for breast cancer pre-screening through a personalized risk assessment, as discussed in **Chapter 10**. The use of breast thermography made the risk prediction personalized, as it considered the heat emitted from the subject for risk estimation. We further used the predicted risk score to assign the individuals into four risk cohorts, where the likelihood of malignancy monotonically increased with the risk grouping level. These initial results are encouraging and could help in creating a tailored screening regime for the individuals based on their risk to detect malignancies in their early stages.

Chapter 11 summarized the overall research, discussed the potential implications of this thesis in the breast cancer continuum of care and shared our thoughts on further research needed to improve early detection of breast cancer in developing and underdeveloped countries.

Appendix II

Samenvatting

Machine learning voor borstkankerdiagnose in ontwikkelingslanden

In dit proefschrift onderzochten hebben we de haalbaarheid van geautomatiseerde borstthermografie met behulp van computerondersteunde diagnose voor borstkanker in ontwikkelingslanden. We wilden alternatieven verkennen en creatief denken over het oplossen van de problemen in onze omgeving (ontwikkelingslanden). Bestaande borstkankerdetectie-technologie levert in onze setting een aantal uitdagende problemen op vanwege minimale menselijke expertise, hoge apparatuurkosten, sociale en culturele barrières zoals privacy en stigma. De technologie die tijdens dit proefschrift werd ontwikkeld, probeerde een aantal van de problemen aan te pakken door de voordelen van de borstthermografie te combineren met de vooruitgang in machine learning. Een korte samenvatting van de verschillende hoofdstukken die in dit proefschrift worden gepresenteerd, wordt hieronder beschreven.

Hoofdstuk 2 van dit proefschrift is een overzichtsartikel dat was gericht op het begrijpen van de uitdagingen die gepaard gaan met de conventionele borstbeeldvormingsmodaliteiten en het waarderen van de potentiële voordelen van infraroodthermografie als beeldvormingsmodaliteit voor de detectie van borstkanker. We verkenden eerst de biologische relatie tussen de abnormale warmtepatronen en de aanwezigheid van kwaadaardigheid en gaven een gedetailleerd overzicht van de verschillende criteria die worden gebruikt voor handmatige interpretatie van de borstwarmtebeelden en de uitdagingen zoals hoge expertise, cognitieve overbelasting en subjectiviteit geassocieerd met de handmatige interpretatie. Dit werd gevolgd door een onderzoek naar de bestaande grootschalige onderzoeken sinds 1960, die de werkzaamheid van borstthermografie voor maligniteitsdetectie aantoonde. Ten slotte hebben we de recente vooruitgang in de hardware en de softwaretechnologieën van infrarood-borstthermografie bestudeerd om betere nauwkeurigheid te bereiken met de combinatie van geautomatiseerde interpretatie en zeer gevoelige infraroodcamera's.

Om de cruciale voorverwerkingsstappen in het beeldvormingsprotocol te automatiseren, hebben we nieuwe segmentatie en beeldmarkering benaderingen besproken die nodig zijn voor geautomatiseerde diagnose. Vanwege de beperkte gegevens die beschikbaar waren in de beginfase van het experiment, stelden we een beeldverwerkingsbenadering voor met heuristische zoals besproken in **Hoofdstuk 3** voor het labelen en segmenteren van de warmtebeelden. Deze techniek was echter sterk afhankelijk van het beeldvormingsprotocol en elke afwijking van het beeldvormingsprotocol resulteerde in foutieve voorspellingen, waardoor het ongeschikt werd voor de klinische praktijk. Daarom hebben we in **Hoofdstuk 4** een deep learning-architectuur voorgesteld om het borstgebied en de kijkhoek nauwkeurig te voorspellen door semantiek op hoog niveau van een grote trainingsdataset te leren. Interessant is dat de verkregen dice index van 0,92 (92%) met de deep learning-benadering heel dicht

bij de inter-observeer dice correlatie tussen twee menselijke experts lag. Ook was de verkregen dice index ongeveer 25% hoger dan de eerdere heuristische segmentatie, wat de kracht van de deep learning-architecturen aantoont wanneer er een grote hoeveelheid trainingsgegevens is.

Uit de gedetecteerde borstgebieden werden automatisch bloedvaten geëxtraheerd zoals besproken in **Hoofdstuk 5** om de vasculaire eigenschappen te bestuderen. Vasculaire veranderingen zijn een van de eerste veranderingen die in het lichaam optreden tijdens het ontstaan van kanker. Daarom kan onderzoek van deze vasculariteit helpen bij het vroeg opsporen van kankergebieden. Geautomatiseerde extractie van deze bloedvaten was een uitdaging omdat de hitesignatuur van de vaatgrenzen diffundeerde tijdens de transmissie naar het borstoppervlak. Daarom hebben we een vatverbeteringstechniek met drie niveaus voorgesteld, gevolgd door vorm- en temperatuurfilters om potentiële bloedvatpixels te detecteren. We hebben ook een nieuwe diffusiviteitsmetriek gedefinieerd om vaten met een hoge diffusiviteit te verwijderen, waardoor de prominente vaatstructuren konden worden geïdentificeerd. De resultaten met de voorgestelde aanpak bleken beter en minder afhankelijk van de drempelparameters in vergelijking met vaatextractietechnieken die werden gebruikt voor andere beeldvormende modaliteiten.

Hoofdstuk 6 onderzocht verschillende deep learning-architecturen voor het extraheren van thermisch actieve regio's (hotspots) uit de gesegmenteerde borstgebieden met behulp van convolutionele neurale netwerken. De hotspot-extractie helpt bij het lokaliseren en begrijpen van de aard en de eigenschappen van de abnormale regio. We hebben waargenomen dat de encoder-decoderarchitecturen beter zijn in vergelijking met de puntsgewijze architecturen in termen van inferentietijd, dice index en nauwkeurigheid voor hotspot-extractie in warmtebeelden. Hoewel de nauwkeurigheid van de encoder-decoderarchitecturen hoog was, waren deze resultaten nog steeds inferieur in vergelijking met een adaptieve histogram drempelanalyse. Dit kan te wijten zijn aan de kleine omvang van de gelabelde dataset die wordt gebruikt voor het trainen van de architecturen. Daarom gebruikten we in **Hoofdstuk 7** een adaptief histogram drempelwaarde voor de hotspot-detectie en extraheerden we medisch interpreteerbare functies met behulp van domeinkennis om de gedetecteerde hotspot-regio's te karakteriseren. Deze kenmerken werden vervolgens ingevoerd om een random forest (RF) -classifier te trainen om de kans op maligniteit te voorspellen op basis van de hotspot-eigenschappen.

Hoofdstuk 8 breidde de reikwijdte van warmtebeeldtechnologie uit naar prognose door een geautomatiseerde techniek voor te stellen voor het schatten van de hormonale status van de gedetecteerde kwaadaardige gebieden vanuit de warmtebeelden van de borst. Deze hormonale biomarker is een belangrijke prognostische factor voor het behandelplanning en de overlevingsvoorspelling. Onze voorgestelde aanpak is een

voorstudie die een nauwkeurigheid van 80% aantoonde voor het schatten van de hormonale status met een combinatie van machine learning en thermografie. Als deze resultaten replicerbaar zijn op een grootschalige populatie, kan dit de noodzaak van een extra invasieve procedure om deze parameters te verkrijgen elimineren, waardoor de kosten worden verlaagd en een snelle start van de behandeling wordt geboden.

Verder onderzochten we de robuustheid van verschillende classificatoren voor machine learning voor het voorspellen van het risico op borstkanker door de veelvoorkomende risicofactoren voor borstkanker, zoals besproken in **Hoofdstuk 9**. Deze risicoschatting zou kunnen helpen bij het identificeren van vrouwen met een hoog risico, en zou zo persoonlijke zorg en diagnose kunnen bieden aan deze vrouwen voor het verminderen van het risico op borstkanker of voor het vroeg opsporen van borstkanker. Onvolledige of onnauwkeurige informatie over de risicofactoren is een groot probleem dat optreedt als gevolg van angst voor sociale stigmatisering of het gebrek aan kennis van de patiënt over de risicofactoren en kan leiden tot een onnauwkeurige risicovoorspelling. Daarom hebben we drie prominente classificatoren voor machine learning beoordeeld op hun robuustheid voor de onvolledige en onnauwkeurige informatie. De voorgestelde aangepaste neurale netwerkarchitectuur bleek superieur te zijn in vergelijking met logistische regressie en random forests, vooral wanneer minder dan 50% van de gegevens onjuist was of ontbrak.

Hoewel de architectuur van het neurale netwerk superieure resultaten opleverde in vergelijking met andere classificatoren, zelfs in de aanwezigheid van onvolledige of onnauwkeurige informatie, was de risicoschatting niet nauwkeurig vanwege de zwakke correlatie van de risicofactoren met het daadwerkelijke screeningsresultaat. Daarom stelden we een nieuwe combinatie voor van machine learning met borstthermografie en niet-beeldvormende kenmerken zoals klonter- en tepelafscheiding voor pre-screening van borstkanker door middel van een gepersonaliseerde risicobeoordeling, zoals besproken in **Hoofdstuk 10**. Het gebruik van borstthermografie maakte de risico voorspelling persoonlijk, aangezien het de warmte die door het onderwerp werd uitgestoten voor risicoschatting in aanmerking nam. We gebruikten de voorspelde risicoscore verder om de individuen in vier risicocohorten toe te wijzen, waarbij de kans op maligniteit monotoon toenam met het risicogroeperingsniveau. Deze eerste resultaten zijn bemoedigend en kunnen helpen bij het creëren van een op maat gemaakt screeningsregime voor de individuen op basis van hun risico om maligniteiten in hun vroege stadia op te sporen.

Hoofdstuk 11 vatte het algehele onderzoek samen, besprak de mogelijke implicaties van dit proefschrift in het continuüm van de zorg voor borstkanker en deelde onze gedachten over verder onderzoek dat nodig is om de vroege opsporing van borstkanker in ontwikkelings- en onderontwikkelde landen te verbeteren.