Segmentation of Breast Tissue Lesions Using Explainable Models

Emlőszöveti elváltozások szegmentálása magyarázható modellekkel

Adél Bajcsi, Ditmár Máté and Lehel Csató

Babeș–Bolyai University, Cluj–Napoca
Faculty of Mathematics and Computer Science

2025

Outline

Introduction

- Breast cancer is the most common cancer diagnosed in women
- Early detection through imaging (e.g., mammography) significantly improves outcomes
- Mammography is the most frequently used imaging technique to record the breast tissue

Introduction

- Breast cancer is the most common cancer diagnosed in women
- Early detection through imaging (e.g., mammography) significantly improves outcomes
- Mammography is the most frequently used imaging technique to record the breast tissue

Introduction

- Breast cancer is the most common cancer diagnosed in women
- Early detection through imaging (e.g., mammography) significantly improves outcomes
- Mammography is the most frequently used imaging technique to record the breast tissue

Datasets

- Tissue density has high impact on the lesion detection performance
- However, not all datasets include this information
- There is a growing demand for multi-modal computer-aided diagnosis systems

Datasets

- Tissue density has high impact on the lesion detection performance
- However, not all datasets include this information
- There is a growing demand for multi-modal computer-aided diagnosis systems

Datasets

- Tissue density has high impact on the lesion detection performance
- However, not all datasets include this information
- There is a growing demand for multi-modal computer-aided diagnosis systems

RABMC – Romanian Adipose Breast Mammogram Collection

- The dataset is collected in collaboration with “Prof. Dr. Ion Chiricuță” Oncology Institute
- Contains:
  - Mammograms
  - Tomosinthesis
  - Histopathology
  - Reports from the doctors

RABMC – Romanian Adipose Breast Mammogram Collection

- The dataset is collected in collaboration with “Prof. Dr. Ion Chiricuță” Oncology Institute
- Contains:
  - Mammograms
  - Tomosinthesis
  - Histopathology
  - Reports from the doctors

RABMC – Romanian Adipose Breast Mammogram Collection

- The dataset is collected in collaboration with “Prof. Dr. Ion Chiricuță” Oncology Institute
- Contains:
  - Mammograms
  - Tomosinthesis
  - Histopathology
  - Reports from the doctors

RABMC – Romanian Adipose Breast Mammogram Collection

- The dataset is collected in collaboration with “Prof. Dr. Ion Chiricuță” Oncology Institute
- Contains:
  - Mammograms
  - Tomosinthesis
  - Histopathology
  - Reports from the doctors

RABMC – Romanian Adipose Breast Mammogram Collection

- The dataset is collected in collaboration with “Prof. Dr. Ion Chiricuță” Oncology Institute
- Contains:
  - Mammograms
  - Tomosinthesis
  - Histopathology
  - Reports from the doctors

Metadata

- The type of the tissue
- Presence of a lesion
- Type of the lesion
- Location of the lesion
- The original reports
- An English version of the report (translated with ChatGPT)

Metadata

- The type of the tissue
- Presence of a lesion
- Type of the lesion
- Location of the lesion
- The original reports
- An English version of the report (translated with ChatGPT)

Metadata

- The type of the tissue
- Presence of a lesion
- Type of the lesion
- Location of the lesion
- The original reports
- An English version of the report (translated with ChatGPT)

Metadata

- The type of the tissue
- Presence of a lesion
- Type of the lesion
- Location of the lesion
- The original reports
- An English version of the report (translated with ChatGPT)

Metadata

- The type of the tissue
- Presence of a lesion
- Type of the lesion
- Location of the lesion
- The original reports
- An English version of the report (translated with ChatGPT)

Metadata

- The type of the tissue
- Presence of a lesion
- Type of the lesion
- Location of the lesion
- The original reports
- An English version of the report (translated with ChatGPT)

Distribution of the data

- 289 patients:
  - 167 heterogeneous
  - 131 homogeneous
- 187 healthy scans, lesions: 195 benign and 254 malignant
- Age between 32 and 90

Distribution of the data

- 289 patients:
  - 167 heterogeneous
  - 131 homogeneous
- 187 healthy scans, lesions: 195 benign and 254 malignant
- Age between 32 and 90

Distribution of the data

- 289 patients:
  - 167 heterogeneous
  - 131 homogeneous
- 187 healthy scans, lesions: 195 benign and 254 malignant
- Age between 32 and 90

Distribution of the data

- 289 patients:
  - 167 heterogeneous
  - 131 homogeneous
- 187 healthy scans, lesions: 195 benign and 254 malignant
- Age between 32 and 90

Distribution of the data

- 289 patients:
  - 167 heterogeneous
  - 131 homogeneous
- 187 healthy scans, lesions: 195 benign and 254 malignant
- Age between 32 and 90

Anonymization

- Ensuring compliance with privacy and clinical data standards
- All personal identifiers are removed from medical images and metadata
- Conversion and standardization of imaging formats

Anonymization

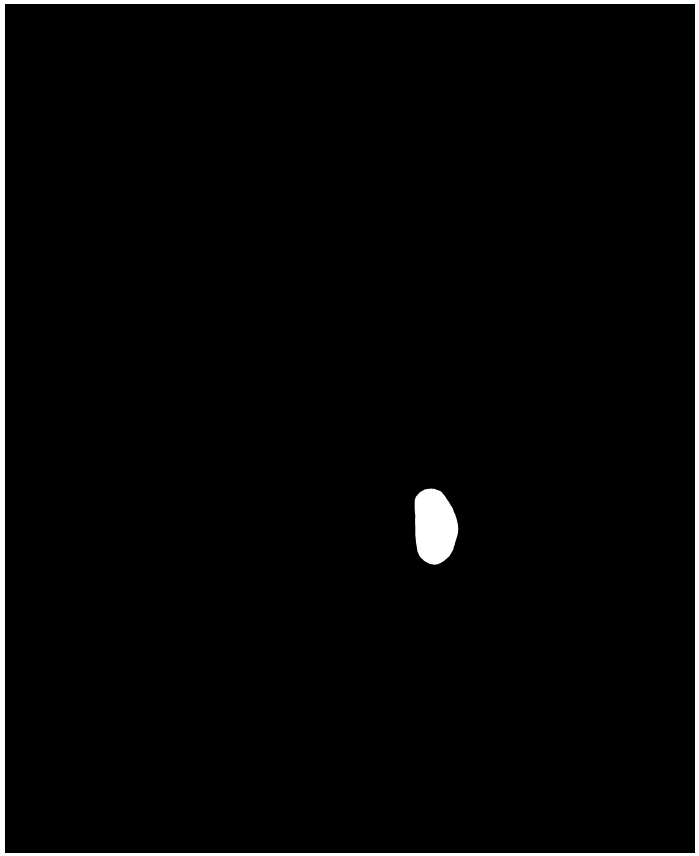
- Ensuring compliance with privacy and clinical data standards
- All personal identifiers are removed from medical images and metadata
- Conversion and standardization of imaging formats

Anonymization

- Ensuring compliance with privacy and clinical data standards
- All personal identifiers are removed from medical images and metadata
- Conversion and standardization of imaging formats

Extracting manual segmentation

(a) Manual segmentation



(b) Extracted lesion mask

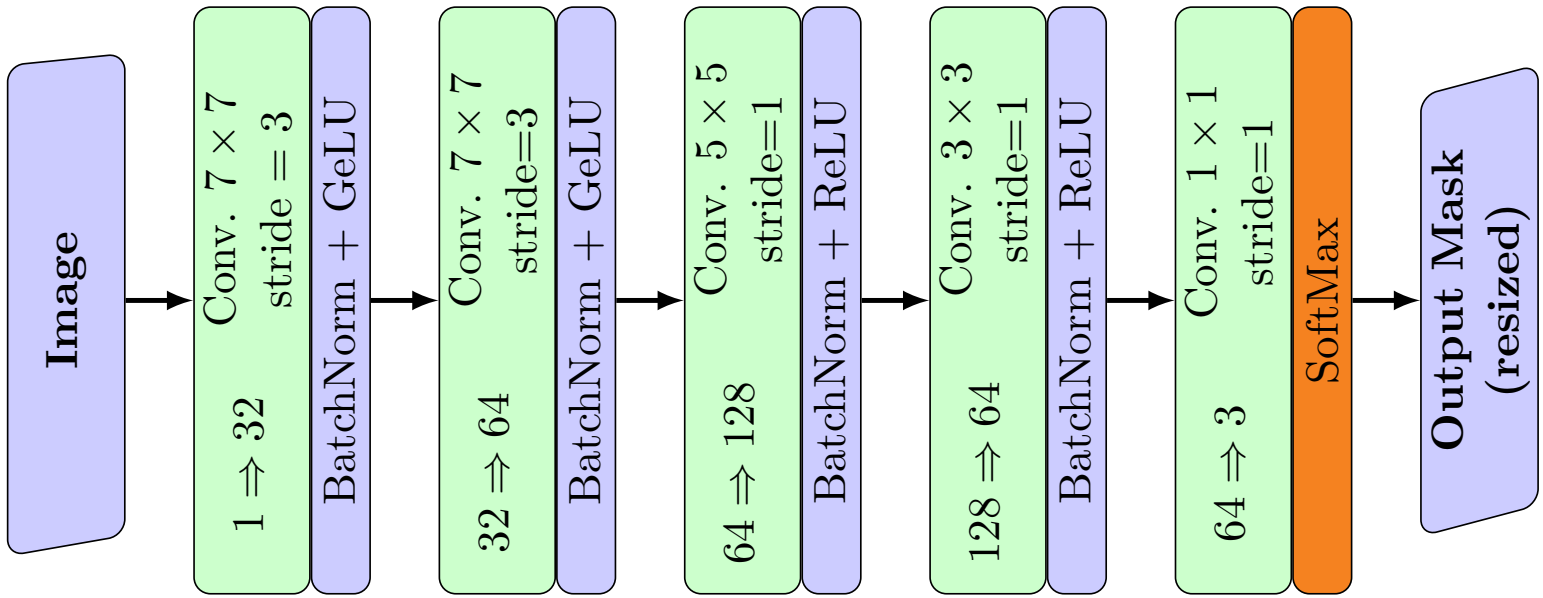
Segmentation of Breast Tissue

- **Objective**: Distinguish lesion regions (foreground) from healthy breast tissue (background)
- Enhances the performance and reliability of automated breast cancer detection systems
  $\Rightarrow$ we propose the use of a convolutional neural network

Segmentation of Breast Tissue

- **Objective**: Distinguish lesion regions (foreground) from healthy breast tissue (background)
- Enhances the performance and reliability of automated breast cancer detection systems
  $\Rightarrow$ we propose the use of a convolutional neural network

Lightweight Segmentation Model



Ditmár Máté and Lehel Csató. "Salivary ferning segmentation with a lightweight CNN" (2025) *2025 27th International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (SYNASC)*

Lightweight Segmentation Model

Receptive field

- The region of the input space that contributes to a neuron's activation
- In convolutional networks, this receptive field expands progressively with each added layer
- Visualization of the receptive field produced by two consecutive convolutional layers using 3×3 kernels with a stride of 1

Lightweight Segmentation Model

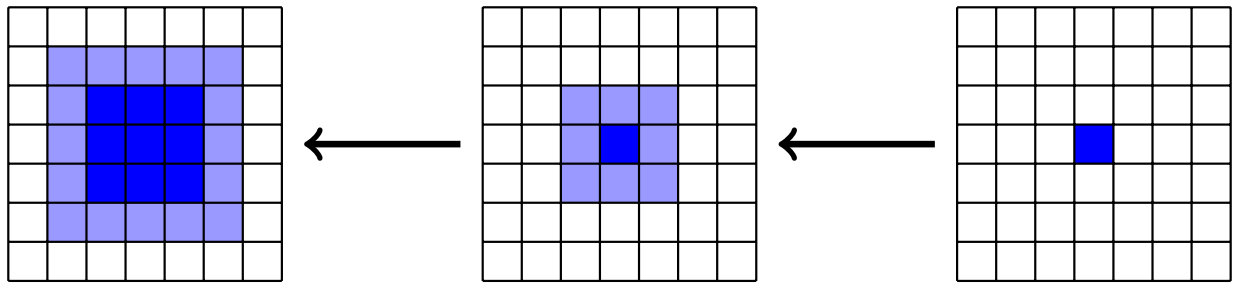
Receptive field

- The region of the input space that contributes to a neuron's activation
- In convolutional networks, this receptive field expands progressively with each added layer
- Visualization of the receptive field produced by two consecutive convolutional layers using 3×3 kernels with a stride of 1

Lightweight Segmentation Model

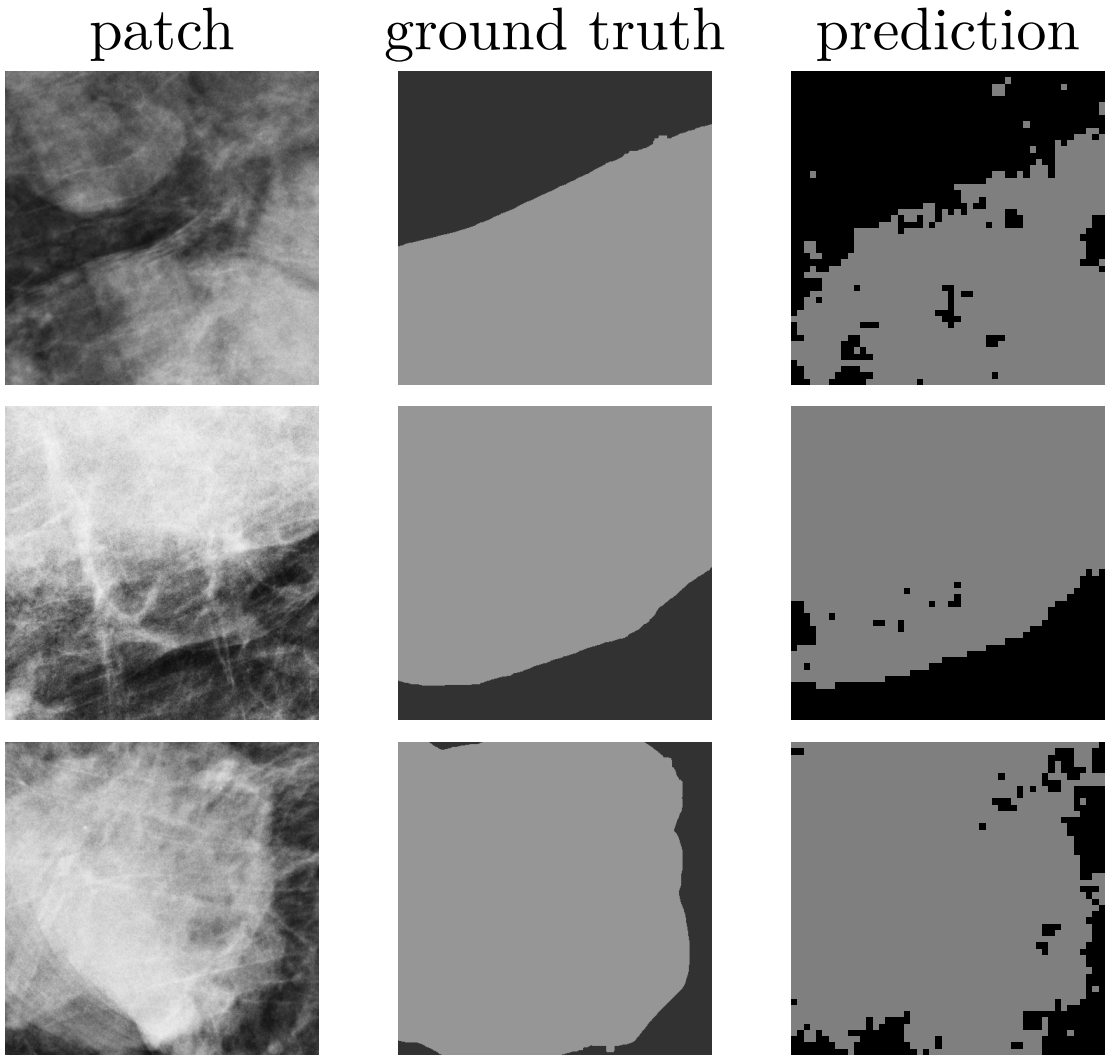
Receptive field

- The region of the input space that contributes to a neuron's activation
- In convolutional networks, this receptive field expands progressively with each added layer
- Visualization of the receptive field produced by two consecutive convolutional layers using 3×3 kernels with a stride of 1



Lightweight Segmentation Model

Results



Conclusions and Future Work

- The manual segmentation must be revised
- The preliminary results of the lightweight CNN model are promising

- Fine-tuning the proposed lightweight model
- Training more complex self-explanatory models like ProtoSeg
- Training a multi-modal system, considering the histopathological reports

Conclusions and Future Work

- The manual segmentation must be revised
- The preliminary results of the lightweight CNN model are promising

- Fine-tuning the proposed lightweight model
- Training more complex self-explanatory models like ProtoSeg
- Training a multi-modal system, considering the histopathological reports

Conclusions and Future Work

- The manual segmentation must be revised
- The preliminary results of the lightweight CNN model are promising

- Fine-tuning the proposed lightweight model
- Training more complex self-explanatory models like ProtoSeg
- Training a multi-modal system, considering the histopathological reports

Conclusions and Future Work

- The manual segmentation must be revised
- The preliminary results of the lightweight CNN model are promising

- Fine-tuning the proposed lightweight model
- Training more complex self-explanatory models like ProtoSeg
- Training a multi-modal system, considering the histopathological reports

Conclusions and Future Work

- The manual segmentation must be revised
- The preliminary results of the lightweight CNN model are promising

- Fine-tuning the proposed lightweight model
- Training more complex self-explanatory models like ProtoSeg
- Training a multi-modal system, considering the histopathological reports

[1] Máté, D. & Csató, L. (2025) Salivary ferning segmentation with a lightweight CNN, International Symposium on Symbolic and Numeric Algorithms for Scientific Computing (Accepted)

[2] Sacha, M., Rymarczyk, D., Struski, Ł., Tabor, J., & Zieliński, B. (2023). ProtoSeg: Interpretable Semantic Segmentation with Prototypical Parts (Version 2). arXiv. https://doi.org/10.48550/ARXIV.2301.12276