

Breast cancer Detection and Classification

Ced16i013

K V Gnana Vardhani

- Will your work on breast cancer detection and classification be useful? (You don't want to work in something outdated, right?)

❑ Yes, this is useful. Breast cancer is the second leading cause of death for women all over the world. Early detection and diagnosis is the key for breast cancer control, and it can increase the success of treatment, save lives, and reduce cost. [1]

❑ Due to the false-positive mammograms, Five percent to 15 percent of screening mammograms require further testing such as additional mammograms or ultrasound. [2]

If there is an abnormal finding, a follow-up or biopsy may have to be performed. Most of the biopsies confirm that no cancer was present. It is estimated that a woman who has yearly mammograms between ages 40 and 49 has about a 30 percent chance of having a false-positive mammogram at some point in that decade and about a 7 percent to 8 percent chance of having a breast biopsy within the 10-year period.[3]

So **mammography** advancement is significant with the introduction of screen-film **mammography** And automation in the detection of breast cancer.

❑ In order to eliminate the operator dependency and improve the diagnostic accuracy, a computer-aided diagnosis (CAD) system is a valuable and beneficial means for breast cancer detection and classification. [4]

- If useful, who will use it?

Automation of early detection of Breast cancer with better accuracy is crucial.

Cancer is one of the widespread deadly diseases that can be now detected through machine learning and AI-enabled automated machines. [4]

AI is already diagnosing breast cancer but owing to accuracy, the reliability among the doctors was less. It can detect the early signs of breast cancer better than radiologists.

Radiologists, doctors, patients can use them. This will benefit patients.

[google's AI breast cancer screening tool] [5]

Radiologists miss about 20 percent of breast cancers in mammograms.[6]

AI has better eyes than a human [7] subject to it being taught well to do the job better than a doctor.

- If you know who will use it, will they actually use it?
Yes, More than 90% of women diagnosed with breast cancer at the earliest stage survive their disease for at least 5 years compared to around 15% for women diagnosed with the most advanced stage of the disease, which is now possible with AI [4]

References:

- [1]<https://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/>
- [2]<https://ww5.komen.org/BreastCancer/AccuracyofMammograms.html>
- [3]<https://www.radiologyinfo.org/en/info.cfm?pg=mammo>
- [4]<https://www.quora.com/Does-Google-AI-beat-doctors-at-breast-cancer-detection>
- [5]<https://www.mobihealthnews.com/news/study-googles-ai-tech-shows-promise-detecting-breast-cancer>
- [6]<https://www.aljazeera.com/ajimpact/google-ai-system-aid-breast-cancer-detection-study-200101211950569.html>
- [7]<https://www.mobihealthnews.com/news/study-googles-ai-tech-shows-promise-detecting-breast-cancer>

Set0 :

- **What is Google Scholar? (Bonus question: Who created it and why?)**

Google Scholar is a freely accessible web search engine that provides a simple way to broadly search for scholarly literature. One can search across many disciplines and sources: articles, theses, books, abstracts and court opinions, from academic publishers, professional societies, online repositories, universities. Google Scholar helps you find relevant work across the world of scholarly research. [1]

Google Scholar was created by Alex Verstak and Anurag Acharya. Their goal was to "make the world's problem solvers 10% more efficient" by allowing easier and more accurate access to scientific knowledge.

- **What is a bibliographic database? Give at least 5 such databases for Computer Science.**

A **bibliographic database** is a database of bibliographic records, an organized digital collection of references to published literature, including journal and newspaper articles, conference proceedings, reports, government and legal publications, patents, books, etc.[2]

5 research databases with special focus on computer science to help you find research papers, scholarly articles, and conference papers fast.[3]

- ☐ ACM Digital Library
- ☐ IEEE Xplore Digital Library
- ☐ dblp computer science bibliography
- ☐ Microsoft Academic
- ☐ Springer Lecture Notes in Computer Science (LNCS)

- **What does "Stand on the shoulders of giants" mean in Google Scholar?**

Standing on the shoulders of giants is a metaphor which means "Using the understanding gained by major thinkers who have gone before in order to make intellectual progress".[4]

This is a choice to describe the scholarly workflow of discovering, reading, and citing literature.[5]

- **How to cite resources using Google scholar?**

Google scholar provides citations for articles from the search result list (currently MLA, APA, Chicago, Harvard or Vancouver). To grab a citation click on the Cite link below a search result and select from the available citation styles.[6]

- **What are the different styles of citation in a text document? Name 3 citation styles in the text part of a text document.**

Three main types of citation:[7]

- ☐ **Parenthetical citation:** You put the source reference in parentheses directly in your text. This usually includes the author's last name along with the publication date and/or the page number.
- ☐ **Note citation:** You put the source reference in a footnote or endnote.
- ☐ **Numeric citation:** You number each of your sources in the reference list and use the correct number when you want to cite a source.

- **Where should you place the reference section in a text document? Name 5 reference section citation styles.**

This **section** goes at the end of your paper. The **references** are to be alphabetized by the first author's last name, or (if no author is listed) the

organization or title. If **you cite** more than **one** paper by the same first author, sort them by year of publication, earliest year first.[7]

Reference section citation styles: [8]

- ☐ American Meteorological Society (AMS) style.
- ☐ Chicago style.
- ☐ Harvard referencing.
- ☐ OSCOLA referencing.
- ☐ Oxford referencing.
- ☐ Vancouver (numeric) referencing

- **Who are research publishers? (Bonus question: Is it a business?) Name 3 famous research publishing organizations in the field of computer science.**

Research Publisher is an academic **publisher** of peer-reviewed open access journals covering a wide range of academic disciplines. ResearchPub aims to develop the highest quality knowledge-based products and services for the academic, scientific, professional, **research**, and student communities worldwide.[9]

Elsevier, Springer, and IEEE are very **popular publishers** for **research** paper **publication** in **computer science/engineering**. [10]

References:

- [1]<https://scholar.google.com/intl/en/scholar/about.html>
- [2]https://en.wikipedia.org/wiki/Bibliographic_database
- [3]<https://paperpile.com/g/research-databases-computer-science/>
- [4]https://en.wikipedia.org/wiki/Standing_on_the_shoulders_of_giants
- [5]<https://blogs.lse.ac.uk/impactofsocialsciences/2015/11/19/standing-on-the-shoulders-of-the-google-giant/>
- [6]<https://libguides.com.edu/c.php?g=649172&p=4554037>
- [7]<https://www.google.com/search?client=ubuntu&channel=fs&q=Where+should+you+place+the+reference+section+in+a+text+document%3F&ie=utf-8&oe=utf-8>
- [8]<https://libguides.reading.ac.uk/citing-references/referencingstyles>
- [9]<http://researchpub.org/>
- [10][https://www.researchgate.net/post/Apart from Elsevier Springer and IEEE which are other players in publishing very good research articles in Computer Science and Engineering](https://www.researchgate.net/post/Apart_from_Elsevier_Springer_and_IEEE_which_are_other_players_in_publishing_very_good_research_articles_in_Computer_Science_and_Engineering)

Proposed solutions:

Histology:

- AggNet- Aggregation layer to Convolution Neural Network[3]
- Deep Convolution Neural Network [5,14]
- Combination of a DCNN model and a boosting trees classifier.[7]
- CNN [1,6,8,9,11]
- CNN deep features [2]
- Watershed segmentation[10]
- Stacked Sparse Autoencoder (SSAE) plus
- Softmax Classifier (SMC) CNN[12]
- CN and Bi-LSTM model[13]
- Deep residual networks[15]
- Inception Recurrent Residual Convolutional Neural Network[4]

Mammography:

- Multi-Regional forecast+Multi-Layer Perceptron+KNN [16]
- CNN[17,19]
- Deep CNN[18]

Image data DataSets:

Histology:

<https://www.kaggle.com/paultimothymooney/breast-histopathology-images>

<https://www.kaggle.com/simjeg/lymphoma-subtype-classification-fl-vs-cll>

<https://rdm.inesctec.pt/dataset/nis-2017-003/resource/df04ea95-36a7-49a8-9b70-605798460c35>

<https://www.kaggle.com/ankur1809/breakhist-dataset>

<https://www.kaggle.com/kritika397/breast-cancer-dataset-from-breakhis>

<https://www.kaggle.com/andrewmvd/breast-cancer-cell-segmentation>

Mammography :

<https://www.kaggle.com/kmader/mias-mammography>

<https://www.kaggle.com/skooch/ddsm-mammography>

- What is a mammogram?

Mammography is a specialized medical imaging that uses a low-dose x-ray system to see inside the breasts.

A mammography exam, called a mammogram, aids in the early detection and diagnosis of breast diseases in women.

An x-ray (radiograph) is a noninvasive medical test that helps physicians diagnose and treat medical conditions. Imaging with x-rays involves exposing a part of the body to a small dose of ionizing

radiation to produce pictures of the inside of the body. X-rays are the oldest and most frequently used form of medical imaging.

- Why is the automatic detection and classification of breast cancer necessary?
- What all public datasets are available online?
- What are the methods and techniques used to detect breast cancer?
 - Mammogram: A mammogram is an X-ray of the breast.
Mammograms are commonly used to screen for breast cancer. If an abnormality is detected on a screening mammogram, your doctor may recommend a diagnostic mammogram to further evaluate that abnormality.
 - Breast ultrasound. Ultrasound uses sound waves to produce images of structures deep within the body. Ultrasound may be used to determine whether a new breast lump is a solid mass or a fluid-filled cyst.
 - Removing a sample of breast cells for testing (biopsy). During a biopsy, your doctor uses a specialized needle device guided by an X-ray or another imaging test to extract a core of tissue from the suspicious area.
Biopsy samples are sent to a laboratory for analysis where experts determine whether the cells are cancerous. A biopsy sample is also analyzed to determine the type of cells involved in breast cancer.

- Breast magnetic resonance imaging (MRI). An MRI machine uses a magnet and radio waves to create pictures of the interior of your breast.

Breast ultrasound is generally not used as a screening tool for breast cancer detection because it does not always detect some early signs of cancer such as microcalcifications, which are tiny calcium deposits.

- How do you identify good and reliable sources to read about the topic?