**“AI-based Image Analysis for Early Disease**

**Detection in Medical Imaging”**

**A Project Work Synopsis**

*Submitted in the partial fulfilment for the award of the degree of*

**BACHELOR OF ENGINEERING**

**IN**

**COMPUTER SCIENCE WITH SPECIALIZATION IN**

**ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING**

**Submitted by:**

| UMA SHANKAR SINGH | 20BCS3789 |
| --- | --- |
|  |  |
|  |  |

**Under the Supervision of:**

**SAKSHI GILL**



**CHANDIGARH UNIVERSITY, GHARUAN, MOHALI - 140413,**

**PUNJAB**

**January -June, 2024**

# Abstract

We develop effective medical image classification techniques, with an emphasis on histopathology and magnetic resonance imaging (MRI). The trainer utilized the curriculum as a starting point for a set of data and a restricted number of samples, and we used it as a starting point for a set of data. As calibrating a machine learning model is difficult, we used alternative methods as unsupervised feature extracts or weight-conditioning factors for identifying pathological histology pictures. As a result, the pretrained models will be trained on 3-channel RGB pictures, while the MRI sample has more slices. To alter the working model using the MRI data, the convolutional neural network (CNN) must be fine-tuned. Pretrained models are placed and then used as feature snippets. However, there is a scarcity of well-done medical photos, making training machine learning models a difficult endeavor to begin with. In any case, data augmentation aids in the generation of sufficient training samples; however, it is unclear if data augmentation aids in the prediction of unknown data samples. As a result, we fine-tuned machine learning models without using any additional data. Furthermore, rather than utilizing a standard machine learning classifier for the MRI data, we created a unique CNN that uses both 3D shear descriptors and deep features as input. This custom network identifies the MRI sample after processing our representation of the characteristics from beginning to end. On the hidden MRI dataset, our bespoke CNN outperforms traditional machine learning. Our CNN model is less prone to overfitting as a result of this. Furthermore, we have given cutting-edge outcomes employing machine learning.

**Keywords:** Artificial intelligence, Alzheimer, Cancer disease, Chronic disease, Heart disease, Tuberculosis

# Table of Contents

| **1.    Introduction** |  |
| --- | --- |
| 1.1 Problem Definition |  |
| 1.2 Project Overview |  |
| 1.3 Hardware Specification |  |
| 1.4 Software Specification |  |
| **2.    Literature Survey** |  |
| 2.1 Existing System |  |
| 2.2 Proposed System |  |
| 2.3 Literature Review Summary |  |
| **3.    Problem Formulation** |  |
| **4.    Research Objective** |  |
| **5.    Methodologies** |  |
| **6.    Experimental Setup** |  |
| **7.    Conclusion** |  |
| **8. Reference** |  |
|  |  |

# 1. INTRODUCTION

In the era of medical healthy recognition of histopathology for cells and tissues under the microscope used in this study, these complement each other in the performance of life functions. Recently, medical image analysis takes a long time to analyze. Therefore, the interpretation of hidden medical experts about the interpretation of images or samples is not rare. In addition, samples can be diagnosed with the help of the computer (CAD), which is considered by medical experts as a high balance that helps in conducting local diagnoses. Therefore, the medical pathologist begins by taking a biopsy of the living tissue. As the biopsy that was examined and then through which the tissue is removed. The tissue removed from the organ is placed in a fixative before being examined under a microscope. To be able to differentiate between the different cellular levels under the microscope, tissues are usually stained for ease of work and precision and tissues are stained with hematoxylin and eosin (H&E).

X-rays are used by specialized devices that help diagnose the disease and in several ways, including magnetic resonance (MRI) to create an image of the particular organ with the technique of generating a magnetic field and thus drawing the outputs of those rays. Here comes the role of the computer to draw an illustration in black and white with multiple levels for each specific time, and as a group of images, they are treated in a complex manner .

Our goal is to develop, apply, and comparatively evaluate techniques that are capable of actively and evidently classifying medical images.

This model was created by developing a multitrainable machine learning technique capable of performing classification tasks using traditional texture-based features .

Therefore, in this paper, we use the basis of shearlet positive classification techniques for Alzheimer’s patients. The Alzheimer’s Disease Neuroimaging Initiative (ADNI) database and the Open Access Imaging Studies Series (OASIS) are used to attain this purpose.

## 1.1 Problem Definition

Early detection of disease can be life-saving, but analyzing medical images is a time-consuming and error-prone process for radiologists facing immense workloads. AI-based image analysis offers a promising solution.

Currently, manual image analysis is subjective and prone to fatigue-related errors, potentially missing subtle disease signs. Additionally, the sheer amount of data is overwhelming, making efficient analysis difficult. To address these challenges, AI algorithms can be trained to accurately identify and characterize subtle disease signatures in various imaging modalities, like X-rays, CT scans, and MRIs. This automation leads to faster and more objective analysis, freeing up radiologists for complex cases.

However, ensuring privacy and security of patient data is crucial. By developing robust data anonymization and access control mechanisms, we can ethically utilize AI for personalized medicine approaches. Collaboration between researchers, healthcare professionals, and patients is vital for responsible development and implementation. Furthermore, clinical trials and real-world studies are essential to evaluate the impact of AI-based image analysis on patient care.

Ultimately, AI-based image analysis has the potential to revolutionize early disease detection, leading to better patient outcomes, improved healthcare efficiency, and a healthier population.

## 1.2 Problem Overview

This research is to develop a machine learning system for early detection of diseases in medical images, enhancing the accuracy and speed of diagnosis through the analysis of X-rays, MRIs, or other medical imaging modalities.

To design an AI models for analyzing multimodal data to improve diagnostic accuracy and reliability and Integration of multiple imaging modalities such as MRI, X-rays , and ultrasound to provide a comprehensive view for early disease detection.

We enhance the user interface of software and applications to respond appropriately to users' disease detection in Medical Imaging.

We Develop of AI-based tools for longitudinal monitoring of disease progression and treatment response using sequential medical imaging data.

## 1.3 Hardware Specification

* Windows 10
* webcam
* Minimum RAM 2GB
* Intel core i5 processor

## 1.4 Software Specification

● Microsoft Visual Studio is an integrated development environment from Microsoft. It is used to develop computer programs, as well as websites, web apps, web services and mobile app.

● Operating system (such as window 7, 8,10, XP, Linux etc.)

**Used Language/library:**

● Python

● OpenCV

● Deep learning

● CNN

● Keras

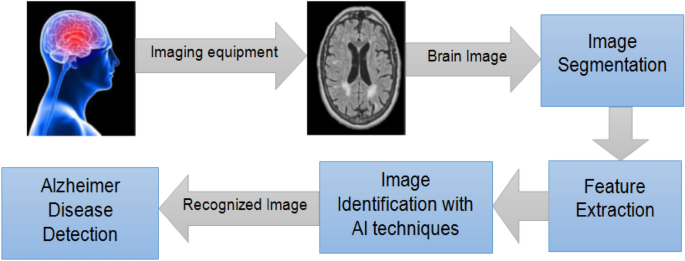
● Image processing

# 2. LITERATURE SURVEY

## 2.1 Existing System

Medical healthy recognition of histopathology for cells and tissues under the microscope used in this study, these complement each other in the performance of life functions. Recently, medical image analysis takes a long time to analyze. Currently, manual image analysis is subjective and prone to fatigue-related errors, potentially missing subtle disease signs. Additionally, the sheer amount of data is overwhelming, making efficient analysis difficult. To address these challenges, AI algorithms can be trained to accurately identify and characterize subtle disease signatures in various imaging modalities, like X-rays, CT scans, and MRIs. This automation leads to faster and more objective analysis, freeing up radiologists for complex cases.

## 2.2 Proposed System



## 

## 2.3 Literature Review Summary

Alzheimer’s disease (AD) is critical for a pathologist to suggest appropriate treatment for patients. Different methods are used with the pathological tissue classification process. Therefore, we use techniques for visual, gray-level texture-based repetition matrix, variable-scale feature transformation, local binary pattern, and gradient histogram [13]. Besides that, features that depend on texture are extracted from a whole image or at least the specific section in the image like gray matter, brain spinal, or white matter. Since GLCM comes, four descriptors can be obtained: variance, correlation, homogeneity, and energy. As a result, just a portion of some slides is run in the training, while the vector of features will not be averaged if all segments are used [14].

The author suggested many steps for a new diagnosis of the acute lymphoblastic leukemia approach. This method entails acquiring blood images using the public dataset (ALL-IDB1), segmenting blast cells using marker-based segmentation (MBS), and then extracting features from segmented blast cells using the gray level cooccurrence matrix (GLCM) with reduced and selected specific features using probabilistic main component analysis. Furthermore, Random Forest (RF) was introduced to categorize the segmented cell into the normal or abnormal group at the classification level [15].

# 3. PROBLEM FORMULATION

The critical problem addressed by AI-based image analysis in early disease detection lies in the limitations of traditional, manual approaches. While crucial for diagnosis, current manual analysis suffers from several key shortcomings:

Subjectivity and Error: Human interpretation inherently carries subjective biases and can be susceptible to fatigue-related errors, potentially leading to missed diagnoses or misinterpretations of subtle disease indications.

Time-Consuming and Inefficient: The vast amount of medical imaging data creates a significant workload for radiologists, resulting in lengthy analysis times and potentially delaying diagnoses or interventions.

Limited Detection Capabilities: The human eye may struggle to identify faint or nuanced disease signatures, hindering early detection of certain conditions.

These limitations highlight the urgent need for a more objective, efficient, and sensitive approach to analyzing medical images, paving the way for early intervention, improved prognosis, and ultimately, better patient outcomes.

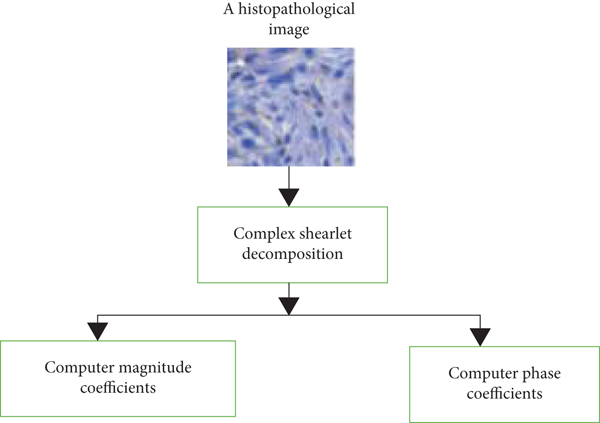
# 4. OBJECTIVES

The overarching objective of AI-based image analysis for early disease detection in medical imaging is to leverage the power of artificial intelligence to dramatically improve the accuracy, efficiency, and objectivity of disease identification within medical scans. By automating tedious analysis tasks, mitigating human error, and uncovering subtle disease signatures invisible to the naked eye, this technology aims to propel early detection rates, ultimately leading to better patient outcomes, reduced healthcare costs, and a healthier population. This objective necessitates a multi-pronged approach, encompassing the development of high-performance AI algorithms, robust data privacy and security measures, and collaborative efforts between researchers, healthcare professionals, and patients. Through successful implementation, AI-based image analysis can become a transformative force in the fight against disease.

Artificial intelligence, Alzheimer, Cancer disease, Chronic disease, Heart disease, Tuberculosis

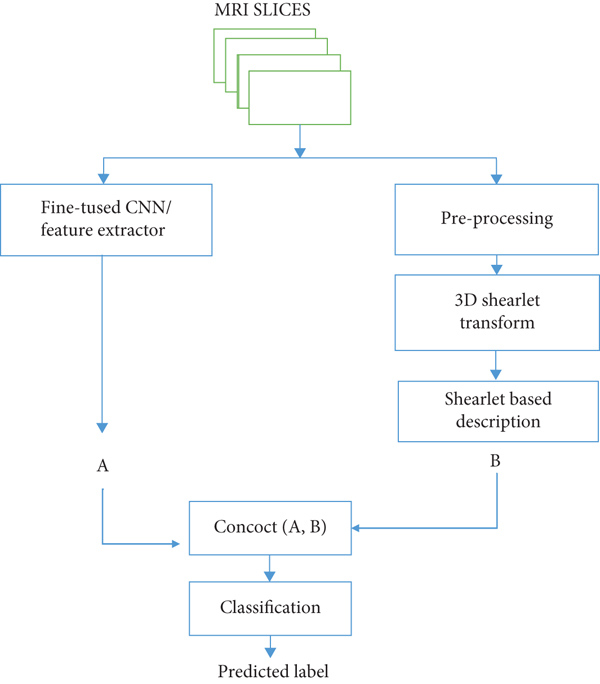
# 5. METHODOLOGY

In order to implement the histopathological image as used in AD, we apply training classifiers to the complex shearlet transform and with the complex coefficients compute the magnitude and relative phase (RP), as shown in Figure 3.



# 6.EXPERIMENTAL SETUP

Figure 4 depicts the model for our proposed project. The fusion of both deep features and shearlet descriptors is worthy here in the proposed study. The proposed system comprising two MRI samples and pipelines is preprocessed in the first step of our study, as indicated in the picture. The shearlet transform coins are then summed up using some important procedures, as each summarization differs from one way to another in supporting a useful set in descriptors of MRI samples so that the gain of high-end performance is achieved.



# 7.CONCLUSION

The major purpose of this study is to develop a model for medical image classification that is both efficient and accurate. To that end, we have presented two pipelines, one that uses manual procedures based on clipping descriptors and the other that uses machine training methods. MRI scans, on the other hand, can be used to carefully design prescriptions and extract relevant information and features from the content of medical images. In addition to the difficulties in histopathology pictures, which employ magnetic resonance imaging, there are other issues to consider. Each sample contains a large number of 2D image slices. As a result, we will have to adapt our methods to deal with class cation MRI samples. As a result, using a three-dimensional shearlet transform, calculate a shearlet-based feature representation. We do not employ pretrained machine learning (CNN) models as unsupervised feature extractors, on the other hand, because we would have to feed each slice separately to generate a trait vector for each slice. As a result, we use the MRI samples to tweak a pretrained model, which we then employ as a feature extractor for the entire MRI sample.

**REFERENCES**

1. J. Yanase and E. Triantaphyllou, “A systematic survey of computer-aided diagnosis in medicine: past and present developments,” *Expert Systems with Applications*, vol. 138, article 112821, 2019.

View at: [Publisher Site](https://doi.org/10.1016/j.eswa.2019.112821) | [Google Scholar](https://scholar.google.com/scholar_lookup?title=A%20systematic%20survey%20of%20computer-aided%20diagnosis%20in%20medicine%3A%20past%20and%20present%20developments&author=J.%20Yanase&author=E.%20Triantaphyllou&publication_year=2019)

1. E. Arvaniti, K. S. Fricker, M. Moret et al., “Automated Gleason grading of prostate cancer tissue microarrays via deep learning,” *Scientific Reports*, vol. 8, no. 1, 2018.

View at: [Publisher Site](https://doi.org/10.1038/s41598-018-30535-1) | [Google Scholar](https://scholar.google.com/scholar_lookup?title=Automated%20Gleason%20grading%20of%20prostate%20cancer%20tissue%20microarrays%20via%20deep%20learning&author=E.%20Arvaniti&author=K.%20S.%20Fricker&author=M.%20Moret&publication_year=2018)

1. A. M. Scholz, L. Bünger, J. Kongsro, U. Baulain, and A. D. Mitchell, “Non-invasive methods for the determination of body and carcass composition in livestock: dual-energy X-ray absorptiometry, computed tomography, magnetic resonance imaging and ultrasound: invited review,” *Animal*, vol. 9, no. 7, pp. 1250–1264, 2015.

View at: [Publisher Site](https://doi.org/10.1017/S1751731115000336) | [Google Scholar](https://scholar.google.com/scholar_lookup?title=Non-invasive%20methods%20for%20the%20determination%20of%20body%20and%20carcass%20composition%20in%20livestock%3A%20dual-energy%20X-ray%20absorptiometry%2C%20computed%20tomography%2C%20magnetic%20resonance%20imaging%20and%20ultrasound%3A%20invited%20review&author=A.%20M.%20Scholz&author=L.%20B%C3%BCnger&author=J.%20Kongsro&author=U.%20Baulain&author=A.%20D.%20Mitchell&publication_year=2015)

1. A. Nowogrodzki, “The world's strongest MRI machines are pushing human imaging to new limits,” *Nature*, vol. 563, no. 7729, pp. 24–26, 2018.

View at: [Publisher Site](https://doi.org/10.1038/d41586-018-07182-7) | [Google Scholar](https://scholar.google.com/scholar_lookup?title=The%20world%27s%20strongest%20MRI%20machines%20are%20pushing%20human%20imaging%20to%20new%20limits&author=A.%20Nowogrodzki&publication_year=2018)

1. S. Pathan, K. G. Prabhu, and P. C. Siddalingaswamy, “Techniques and algorithms for computer aided diagnosis of pigmented skin lesions--a review,” *Biomedical Signal Processing and Control*, vol. 39, pp. 237–262, 2018.

View at: [Publisher Site](https://doi.org/10.1016/j.bspc.2017.07.010) | [Google Scholar](https://scholar.google.com/scholar_lookup?title=Techniques%20and%20algorithms%20for%20computer%20aided%20diagnosis%20of%20pigmented%20skin%20lesions--a%20review&author=S.%20Pathan&author=K.%20G.%20Prabhu&author=P.%20C.%20Siddalingaswamy&publication_year=2018)

1. M. Hosseinzadeh, O. H. Ahmed, M. Y. Ghafour et al., “A multiple multilayer perceptron neural network with an adaptive learning algorithm for thyroid disease diagnosis in the internet of medical things,” *The Journal of Supercomputing*, vol. 77, no. 4, pp. 3616–3637, 2021.

View at: [Publisher Site](https://doi.org/10.1007/s11227-020-03404-w) | [Google Scholar](https://scholar.google.com/scholar_lookup?title=A%20multiple%20multilayer%20perceptron%20neural%20network%20with%20an%20adaptive%20learning%20algorithm%20for%20thyroid%20disease%20diagnosis%20in%20the%20internet%20of%20medical%20things&author=M.%20Hosseinzadeh&author=O.%20H.%20Ahmed&author=M.%20Y.%20Ghafour&publication_year=2021)

1. S. Savalia and V. Emamian, “Cardiac arrhythmia classification by multi-layer perceptron and convolution neural networks,” *Bioengineering*, vol. 5, no. 2, p. 35, 2018.

View at: [Publisher Site](https://doi.org/10.3390/bioengineering5020035) | [Google Scholar](https://scholar.google.com/scholar_lookup?title=Cardiac%20arrhythmia%20classification%20by%20multi-layer%20perceptron%20and%20convolution%20neural%20networks&author=S.%20Savalia&author=V.%20Emamian&publication_year=2018)

1. S. Ciucci, Y. Ge, C. Durán et al., “Enlightening discriminative network functional modules behind principal component analysis separation in differential-omic science studies,” *Scientific Reports*, vol. 7, no. 1, article 43946, 2017.

View at: [Publisher Site](https://doi.org/10.1038/srep43946) | [Google Scholar](https://scholar.google.com/scholar_lookup?title=Enlightening%20discriminative%20network%20functional%20modules%20behind%20principal%20component%20analysis%20separation%20in%20differential-omic%20science%20studies&author=S.%20Ciucci&author=Y.%20Ge&author=C.%20Dur%C3%A1n&publication_year=2017)

1. X. Sui, Y. Zheng, B. Wei et al., “Choroid segmentation from optical coherence tomography with graph-edge weights learned from deep convolutional neural networks,” *Neurocomputing*, vol. 237, pp. 332–341, 2017.

View at: [Publisher Site](https://doi.org/10.1016/j.neucom.2017.01.023) | [Google Scholar](https://scholar.google.com/scholar_lookup?title=Choroid%20segmentation%20from%20optical%20coherence%20tomography%20with%20graph-edge%20weights%20learned%20from%20deep%20convolutional%20neural%20networks&author=X.%20Sui&author=Y.%20Zheng&author=B.%20Wei&publication_year=2017)

1. K. R. Kruthika, Rajeswari, and H. D. Maheshappa, “Multistage classifier-based approach for Alzheimer's disease prediction and retrieval,” *Informatics in Medicine Unlocked*, vol. 14, pp. 34–42, 2019.

View at: [Publisher Site](https://doi.org/10.1016/j.imu.2018.12.003) | [Google Scholar](https://scholar.google.com/scholar_lookup?title=Multistage%20classifier-based%20approach%20for%20Alzheimer%27s%20disease%20prediction%20and%20retrieval&author=K.%20R.%20Kruthika&author=Rajeswari&author=H.%20D.%20Maheshappa&publication_year=2019)

1. N. Wang, M. Chen, and K. P. Subbalakshmi, “Explainable CNN-attention networks (C-attention network) for automated detection of Alzheimer's disease,” 2020, <https://arxiv.org/abs/2006.14135>.

View at: [Google Scholar](https://scholar.google.com/scholar_lookup?title=Explainable%20CNN-attention%20networks%20(C-attention%20network)%20for%20automated%20detection%20of%20Alzheimer%27s%20disease&author=N.%20Wang&author=M.%20Chen&author=K.%20P.%20Subbalakshmi&publication_year=2020)

1. N. Patel, *Towards Robust and Secure Perception for Autonomous Robotic Systems*, Doctoral dissertation, New York University Tandon School of Engineering, 2021.
2. A. A. Farag, M. Farag, J. Graham, S. Elshazly, M. al Mogy, and A. Farag, “Modeling of lung nodules from LDCT of the human chest: algorithms and evaluation for CAD systems,” in *Shape Analysis in Medical Image Analysis*, pp. 259–290, Springer, Champions, 2014.

View at: [Google Scholar](https://scholar.google.com/scholar_lookup?title=Modeling%20of%20lung%20nodules%20from%20LDCT%20of%20the%20human%20chest%3A%20algorithms%20and%20evaluation%20for%20CAD%20systems&author=A.%20A.%20Farag&author=M.%20Farag&author=J.%20Graham&author=S.%20Elshazly&author=M.%20al%20Mogy&author=A.%20Farag&publication_year=2014)

1. A. Chaddad, C. Desrosiers, and M. Toews, “Multi-scale radiomic analysis of sub-cortical regions in MRI related to autism, gender and age,” *Scientific Reports*, vol. 7, no. 1, article 45639, 2017.

View at: [Publisher Site](https://doi.org/10.1038/srep45639) | [Google Scholar](https://scholar.google.com/scholar_lookup?title=Multi-scale%20radiomic%20analysis%20of%20sub-cortical%20regions%20in%20MRI%20related%20to%20autism%2C%20gender%20and%20age&author=A.%20Chaddad&author=C.%20Desrosiers&author=M.%20Toews&publication_year=2017)

1. A. Shah, S. S. Naqvi, K. Naveed, N. Salem, M. A. U. Khan, and K. S. Alimgeer, “Automated diagnosis of leukemia: a comprehensive review,” *IEEE Access*, vol. 9, pp. 132097–132124, 2021.

View at: [Publisher Site](https://doi.org/10.1109/ACCESS.2021.3114059) | [Google Scholar](https://scholar.google.com/scholar_lookup?title=Automated%20diagnosis%20of%20leukemia%3A%20a%20comprehensive%20review&author=A.%20Shah&author=S.%20S.%20Naqvi&author=K.%20Naveed&author=N.%20Salem&author=M.%20A.%20U.%20Khan&author=K.%20S.%20Alimgeer&publication_year=2021)