International Journal of Engineering Trends and Technology
ISSN: 2231–5381 / https://doi.org/10.14445/22315381/IJETT-V72I12P104
Volume 72 Issue 12, 42-54, December 2024

Original Article

Cancer Detection and Classification Using CNN Model

Percy Okae[1], Theophilus Addo[2], Joseph Boateng Owusu-Afari[3], Gifty Bondzie[4]

[1,2,3,4]*Computer Engineering Department, University of Ghana, Accra, Ghana.*

[1]*Corresponding Author : pokae@ug.edu.gh*

Received: 27 May 2024 Revised: 08 October 2024 Accepted: 22 November 2024 Published: 25 December 2024

Abstract *- The research utilizes the CNN model to develop the machine learning mode due to its image extraction performance. The system was developed to identify and categorize eight (8) different kinds of cancers, namely lymphoma, oral cancer, brain cancer, breast cancer, cervical cancer, kidney cancer, lung and colon cancer, and leukemia. The multi cancer image dataset from Kaggle was utilized to train and test the models. The dataset contained eight (8) types of cancers grouped into different classes. For each class, 2000 images were used for training and 500 for testing. Pre-processing techniques were applied to normalize and standardize the images to ensure the correct format and resolution. Nine (9) CNN models were trained, with eight responsible for classifying each cancer type while the remaining model detects the cancer type. The system was designed to perform two levels of classification for each image. The first level is the detection of the type of cancer, and the second level is the classification of the cancer type. Generally, the manual examination of cancer diagnoses is error-prone, and this work sought to automate the process as best as possible by investigating the performance of the CNN model on selected types of cancer. The results demonstrated the effectiveness of the developed system in accurately detecting and classifying the eight types of cancers and the potential to alleviate the errors faced with the manual examination. All the models obtained accuracies above 90%.*

Keywords *- Cancer detection and classification, Convolutional Neural Network (CNN), Machine learning, Magnetic Resonance Imaging (MRI), Web application, Mobile application.*

1. Introduction

The health challenges cancers of all forms pose to people of all shades globally cannot be over emphasized. [1]. The fact that 9.6 million people were estimated to have died from various forms of cancer in the year 2017 alone indicates the importance of developing solutions that can help in early detection and diagnosis. Now, it has also been established that cancer is the leading cause of death globally – second only to cardiovascular diseases. Thus, timely and accurate cancer detection and classification is crucial in improving patient outcomes, enabling personalized treatment strategies, and ultimately increasing patients' chances of survival. Conventional diagnostic methods often rely on invasive procedures, such as biopsies, which can be time-consuming, expensive, and may pose risks to patients. Therefore, there is a growing need for non-invasive and efficient techniques to aid in early cancer detection and precise cancer classification. In recent years, machine learning has shown remarkable advancements and emerged as a powerful tool for various medical applications, including cancer detection and classification. Over the years, various machine learning techniques such as Support Vector Machine (SVM), Naïve Bayes and Random Forest algorithms have been applied to cancer detection. Several successes have been achieved using these techniques, but they usually experience shortcomings because of their inability to handle the complexity and variability of medical images. Most of these methods rely on shallow learning, limiting their ability to extract deep and detailed features from images, which is essential for accurate cancer classification. As a result, their accuracy and performance have been suboptimal, especially when dealing with multiple types of cancers. This study adopts Convolutional Neural Networks (CNNs), a deep learning algorithm for image recognition tasks and image analysis. CNNs are particularly well-suited for medical imaging because they can automatically learn and extract hierarchical features from images, enabling more effective feature extraction and classification. In this work, a robust system capable of detecting and classifying multiple cancer types, namely leukemia, brain cancer, breast cancer, cervical cancer, kidney cancer, lung and colon cancer, lymphoma, and oral cancer, is developed. If implemented successfully, healthcare providers can make more informed judgments about cancer diagnosis, treatment planning, and patient management. Additionally, it could aid in identifying cancer at an early stage when intervention and treatment are most effective, leading to improved patient outcomes and potentially reducing healthcare costs. This study builds upon existing cancer detection and classification research using machine learning, with a particular emphasis on predicting eight major types of cancers, which was alluded to earlier in this section. The availability of a comprehensive dataset comprising medical images representing these cancer types allowed for a comprehensive evaluation of the proposed system's performance. By addressing the challenges in cancer detection and classification using the CNN technique, this work aims to contribute to the ongoing efforts to improve early diagnosis, treatment outcomes, and personalized medicine for cancer patients.

Even with advances in medical technology, a doctor's examination of microscope slides is still a common method of diagnosing cancer. As such, inexperience, stress, or haste might lead a pathologist to make an error in their analysis. One of the most frequent forms of medical misconception is cancer diagnosis, which arises from the generally wrong perception of the disease. [2]. The current estimate is that 7500–10,000 lives are lost annually in the UK from late diagnosis [3]. According to the data of the World Health Organization, in 2020, there were over 19 million cancer patients in the world [4]. One such trauma that never goes away, even in memory, is cancer.

It is not only extremely painful, but it also occasionally recurs. This is one of the most significant global health issues. Efforts to reduce the misdiagnoses of cancer cases to the barest minimum have precipitated research into this area using machine learning techniques [5]. Machine learning is increasingly being employed in cancer detection and diagnosis. Machine learning, a subset of artificial intelligence, has revolutionized medical diagnostics, offering unprecedented levels of accuracy and efficiency. Cancer, characterized by its diverse array of types such as cervical cancer, brain cancer, breast cancer, kidney cancer, lung and colon cancer, lymphoma, leukemia, or oral cancer, and others, requires nuanced diagnostic approaches. Each type possesses distinct features necessitating tailored strategies for effective identification and classification. This work has been structured into eight major sections. Section one presents an introduction to cancer. Section two reviews research related to this study, the proposed work, and the scope of the study.

Section three talks about comparative analysis to existing models. Section four talks about system design and development. Section five also discusses the developed system's web and mobile application interface. Section six discusses the testing of the systems and the results obtained. Section seven considers the conclusions derived from this work. It captures the challenges faced in the system design and implementation stages and makes recommendations for future work. Finally, section eight talks about the real-world applications of the developed system.

2. Literature Review

Several existing works have all tried in one way or another to propose solutions to the classification of cancers using one machine learning technique. This research is no different in terms of the approach and techniques used. However, this work solely settles on the CNN techniques and applies two extraction levels in its classifications. In the work by [6], six machine learning algorithms were compared by way of performance using the Wisconsin Diagnostic Breast Cancer (WDBC) dataset. These are GRU-SVM, Linear Regression, Multilayer Perceptron (MLP), Nearest Neighbour (NN) search, Softmax Regression, and Support Vector Machine (SVM). The WDBC dataset used in this work comprised digitized images obtained from FNA tests on a breast mass. While 70 % of the datasets were used for training the images, the remaining 30 % were used for the actual testing. All the six algorithms used in this work performed very well to expectations.

Reference [7] proposed a machine learning system to detect and categorize lung cancer from chest CT data. While various modalities are employed in diagnosing and detecting lung cancer, the research concentrated on Computed Tomography (CT) scan images since they offer the most important information on lung infections. The study concentrated on identifying and categorizing several lung cancer forms, including squamous cell carcinoma, large cell carcinoma, and adenocarcinoma. The chest CT scan image collection was assessed for texture feature categorization using machine learning approaches, including Support Vector Machines (SVM) and K-Nearest Neighbors (KNN). With respect to the support vector machine algorithm and the K-nearest neighbors' approach, the suggested technique yielded performance accuracy of 93% and 91%, respectively. Also, [8] studied using machine learning models for texture analysis to identify cervical cancer.

Squamous cell carcinoma and adenocarcinoma, the two primary forms of cervical cancer, were the focus of the work. The study concentrated on identifying cancer cells in greyscale photos. A Gabon filter was used to evaluate the texture, and histogram equalization was used to pinpoint high-impact elements. 97% accuracy was achieved in categorizing cancerous and non-cancerous cells using the SVM approach, the suggested strategy. Reference [9] conducted a comparative study on deep learning and machine learning approaches for breast cancer diagnosis. Like well-known machine learning models for breast cancer, the study sought to identify the two types of breast cancer: benign and malignant.

The Convolutional Neural Network (CNN) was used in the study project to extract features for machine learning algorithms. It is a better option for feature extraction than a Support Vector Machine (SVM), Naïve Bayes classifier (NB), K-Nearest Neighbor (KNN), Logistic Regression (LR), and Random Forest (RF). The BreaK-Hist dataset from Kaggle was used to test the framework. CNN achieved the highest accuracy and the lowest loss/error rate in both classes of breast cancer, as demonstrated by the trial, with an F-1 score of 92%. According to [10], an experiment was conducted to identify the various Machine Learning algorithms that aid lung cancer diagnosis and prediction. The research used various medical imagery, including X-rays and CT scan images. The primary goal of the approach was to divide lung cancer into two groups: malignant and benign. Support Vector Machine (SVM), Artificial Neural Networks (ANN), Naïve Bayes (NB), Back Propagation Network (BPN), and Convolutional Neural Network (CNN) were among the machine learning algorithms that were employed. The comparative study showed that the CNN technique was more accurate and had an accuracy rate of 94.37 %.

In [11], the authors proposed a deep learning method that involves creating a web application to run the trained model on the backend to classify kidney images. The suggested study included two models: an Artificial Neural Network (ANN) trained with blood sample test data and a Convolutional Neural Network (CNN) trained with CT scan images. The study used 12,446 distinct CT scan samples from the PACS image dataset on Kaggle. The algorithm was trained to predict kidney CT scan images with tumors and without tumors as part of the research. The comparison analysis yielded validation accuracy values of 96.7 percent for the ANN model and 95.6 percent for the CNN model.

In [12], the work carried out was to create a method for identifying leukemia utilizing machine learning methods, including transfer learning and image processing. The study aimed to classify leukemia into four distinct types: ALL (Acute Lymphoblastic Leukemia), CLL (Chronic Lymphocytic Leukemia), and CML (Chronic Myeloid Leukemia). The study used 260 acute lymphoblastic and healthy images from the Kaggle ALL_IDB dataset. CML, CLL, and AML photos, among other image types, were acquired from the ASH image repository. Data augmentation was done to expand the dataset because there were so few microscopic blood smear images of CML, CLL, and AML in the ASH image collection.

By utilizing the weighted average ensemble learning approach, the proposed model got an accuracy of 91.71%. In [13], a model was proposed to classify brain MRIs into tumorous or non-tumorous. Convolutional Neural Networks (CNN) and other deep learning and machine learning techniques were used in the study. Approximately 258 MRI scans were used in the study to create a small dataset, which was then expanded to approximately 2000 images through data augmentation techniques. 15% of the photos were used for testing and validation, and the remaining 70% were used for training. Under its best performance, the developed model had an accuracy of 88.79% and achieved an F1-score of 0.88. Table 1 summarizes the results of works achieved using various machine learning models.

3. Comparative Analysis

Despite advancements in medical technology, traditional diagnostic methods often suffer from limitations such as high false-positive rates and reliance on subjective interpretation by pathologists. This section critically analyzes existing cancer detection and classification methods, highlighting their strengths and weaknesses compared to the proposed CNN-based approach.

3.1. Traditional Diagnostic Methods

Histopathology is the gold standard for cancer diagnosis, providing highly accurate results when performed by experienced pathologists. These methods allow for detailed examination of cellular morphology, enabling the identification of a wide range of cancer types. Traditional methods are widely accepted in clinical practice, and physicians are trained to interpret these results.

- Histopathology**:** Involves examining tissue samples under a microscope. While it is the gold standard for cancer diagnosis, it is time-consuming and subject to human error.
- Cytology**:** Techniques like fine-needle aspiration provide insights into cellular abnormalities but are prone to false negatives due to sampling errors.
- Imaging Techniques: X-rays, CT scans, and MRIs are essential for initial screenings but require expert analysis and can miss subtle cancer signs.

Procedures like biopsies and histological analysis can take days or weeks, delaying diagnosis and treatment. The interpretation of results is highly subjective and can vary significantly between pathologists, leading to inconsistencies and potential misdiagnoses. Many traditional methods require invasive procedures, posing risks and discomfort to patients. As patient volumes increase, the reliance on human interpretation becomes a bottleneck, limiting the ability to diagnose more patients quickly.

Table 1. Comparison of existing systems

Algorithms	Cancer Type	Accuracy	Ref
MLP	Breast Cancer	99.7%	[6] (Abien Fred, 2018)
SVM	Lung Cancer	93%	[7] (Amjad Rehman et al., 2021)
SVM	Cervical Cancer	97%	[8] (G. Ramkumar et al., 2022)
CNN	Breast Cancer	92%	[9] (Simran Bepari et al., 2022)
CNN	Lung Cancer	94.37%	[10] (Amit Singh et al., 2023)
CNN	Kidney Cancer	99.6%	[11] (K Rajkumar et al., 2023)
Ensemble Learning	Leukemia	91.71%	[12] (S. Rajeswari et al., 2022)
CNN	Brain Tumor	88.79%	[13] (B. Ramya Sree et al., 2022)

3.2. Machine Learning Techniques

Machine learning techniques can automate the detection process and subsequent diagnosis of cancer images, thus speeding up the entire process. These methods can analyze vast amounts of data, identifying patterns that human experts might miss. Machine learning algorithms can be retrained with new data, improving detection accuracy continuously.

- Support Vector Machines (SVM): Effective for binary classification but can struggle with multi-class problems and require careful parameter tuning.
- Decision Trees: Intuitive and easy to interpret but prone to overfitting with small datasets.
- Random Forest: An ensemble method that reduces overfitting but can be computationally expensive.
- K-Nearest Neighbors (KNN): Simple and effective for small datasets but become slow with large datasets.
- Naïve Bayes: Fast and efficient but relies on strong independence assumptions, which may not always hold true.

The quality and quantity of training data generally determine the effectiveness of machine learning models. Poor or biased datasets can lead to inaccurate predictions. Many machine learning algorithms require careful parameter tuning, which can be time-consuming and require expert knowledge. Some machine learning models, particularly those complex (like ensemble methods), can be challenging to interpret, making it difficult for clinicians to understand the decision-making process.

3.3. Deep Learning Approaches

CNNs have demonstrated superior performance in image classification tasks, achieving higher accuracy rates than traditional and other machine learning methods. Unlike traditional methods, CNNs are known to have the capability to automatically learn and extract relevant features from images, reducing the need for manual feature engineering. CNNs are generally robust to variations in input data, making them effective across different imaging modalities and conditions. The CNN algorithms can quickly process the trained images, making them suitable for high-throughput environments such as hospitals and clinics. Recent studies have shown that CNNs outperform traditional machine learning methods in terms of accuracy and efficiency.

CNNs require large amounts of labeled data for training, which can be a significant barrier in some clinical settings where data may be limited. Training CNNs can be resource-intensive, requiring powerful hardware and potentially significant time investments. CNNs can overfit the training data without proper regularisation techniques, leading to poor generalization of unseen data. The "black box" nature of deep learning can make it difficult to explain the reasoning behind specific predictions, which can hinder clinical acceptance and trust. Figure 1 shows the structure of the proposed solution.

4. System Design and Development

The system design and development phase results in the convergence of advanced technology and healthcare imperatives. The system uses the chosen machine learning algorithm, CNN, for accurate cancer detection and classification. Additionally, the system includes a user-friendly web and mobile application that bridges the gap between machine learning advancements and practical healthcare delivery. The web and mobile applications have an easy-to-use user interface that enables seamless navigation. The whole system is composed of three main components:

- The application or the doctor's portal, which allows for data entry and image upload
- The machine learning model that processes the data and makes the classification based on past data
- The server stores the data generated by the system.

4.1. System Architecture

The physical design setup included the doctor's portal, the machine learning model meant to process or train data and make the classification based on past data, and then the server that stores the data generated by the system, as shown in Figure 2.

4.2. Model Development

Convolutional Neural Network (CNN) machine learning is the method used by our system. CNNs are very good at processing and identifying images among deep learning algorithms. It consists of convolutional, pooling, and fully linked layers, among other layers. The convolution layer is the fundamental component of CNN. It carries the majority of the network's computational load. It needs three things: a feature map, a filter, and input data. The input image is transformed using a convolution layer to extract features. Dimensionality reduction is carried out via the pooling layer, sometimes called the downsampling layer, which lowers the number of parameters in the input. Our models are made up of three convolutional layers and three pooling layers. The number of output layers depends on the sub-types or classes of each type of cancer, as depicted in Figure 3.

4.3. Layer Descriptions

The model was built using a series of layers serving a specific function. From initial data normalization to feature extraction and final classification, these layers work together to process the medical images effectively. Below is a breakdown of the key layers used in the model:

- Rescaling Layer: The first layer is a Rescaling layer, which normalizes the pixel values of the input images. By rescaling the pixel values to a range of [0, 1] (from the original [0, 255]), the model ensures that the input data is standardized and well-suited for training. The input shape is specified as (180, 180, 3) to accommodate RGB images with 180x180 pixels.

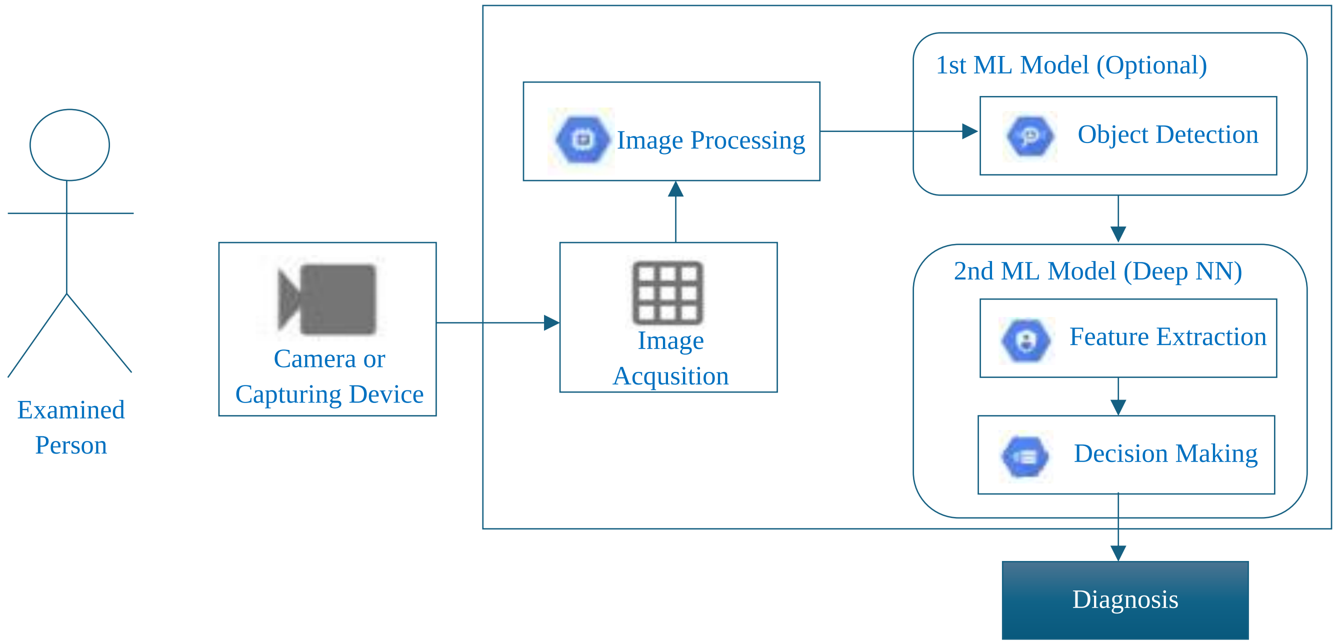
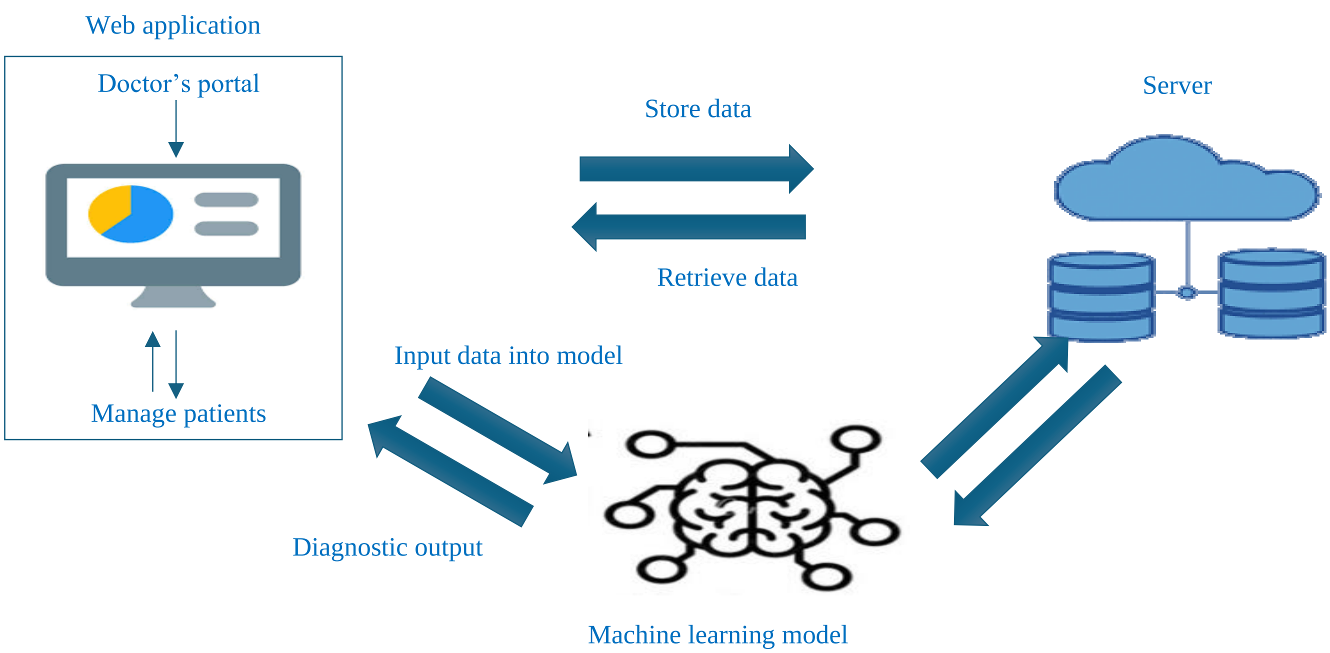


Fig. 1 Structure of the proposed solution [14]



- **Fig. 2 Architectural diagram of the system [15]**

- First Convolutional Layer (Conv2D) & Pooling Layer (MaxPooling2D): The Conv2D layer applies 16 filters of size 3x3 to the input image, allowing the network to learn 16 different feature maps through a ReLU activation function. These filters identify local patterns such as edges or textures. The MaxPooling2D layer then performs a down-sampling operation, reducing the spatial dimensions of the feature maps while preserving the most significant information. This step minimizes computational complexity and helps prevent overfitting.
- Second Convolutional & Pooling Layers: The second Conv2D and MaxPooling2D layers repeat the same operations as the first set. The network adds another 16 filters in this convolutional layer, allowing it to learn more abstract features. As the layers deepen, the model begins to capture more complex patterns from the images.
- Third Convolutional & Pooling Layers: The third Conv2D layer, with 16 filters, further extracts deeper-level features, and the subsequent MaxPooling2D reduces the dimensions once again. The repeated convolution-

pooling structure helps the network focus on hierarchical features of the medical images, gradually learning from local to more global patterns.

- Flatten Layer: After the final pooling layer, the Flatten layer converts the 3D feature maps into a 1D vector, preparing the data for the fully connected layers. This step is essential for feeding the learned features into the subsequent dense layers for classification.
- First Dense Layer: The Dense layer with 128 neurons and a ReLU activation function is fully connected, where all the extracted features are combined. This layer helps learn complex representations from the data, aiding the differentiation between cancerous and non-cancerous images.
- Output Dense Layer: The final Dense layer consists of num_classes neurons, each representing a specific class or subtype of cancer (e.g., breast cancer, lung cancer, etc.). This layer outputs raw scores for each class, which are later passed through a softmax function to generate probabilities during classification.In Figure 4, the structure of the CNN model is depicted showing the layers.

4.4. Preprocessing Techniques

Before training the model, the data was preprocessed to ensure that it was in the optimal form for the machine learning algorithm to work effectively. The image data was normalized to scale pixel values between 0 and 1. This was done using the Rescaling layer in TensorFlow, where each pixel value was divided by 255. This step ensures that all the pixel values are within a consistent range, making it easier for the model to process the data and converge during training. After normalizing the data, the dataset was processed using a mapping function that applies the normalization layer to each image in the dataset. This ensures that every image passed into the model is scaled appropriately. Once the dataset is normalized, a batch of images is fetched for verification. This step allows the user to inspect the transformed data and verify that the pixel values are within the expected range (between 0 and 1). Normalization ensures the model is better equipped to handle image data, reducing training time and improving the system's overall accuracy. Figure 5 is a snapshot of the pre-processing script of the pre-processing techniques to normalize the dataset.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 rescaling_1 (Rescaling)     (None, 180, 180, 3)       0

 conv2d (Conv2D)             (None, 180, 180, 16)      448

 max_pooling2d (MaxPooling2D  (None, 90, 90, 16)       0
 )

 conv2d_1 (Conv2D)           (None, 90, 90, 16)        2320

 max_pooling2d_1 (MaxPooling  (None, 45, 45, 16)       0
 2D)

 conv2d_2 (Conv2D)           (None, 45, 45, 16)        2320

 max_pooling2d_2 (MaxPooling  (None, 22, 22, 16)       0
 2D)

 flatten (Flatten)           (None, 7744)              0

 dense (Dense)               (None, 128)               991360

 dense_1 (Dense)             (None, 8)                 1032

=================================================================
Total params: 997,480
Trainable params: 997,480
Non-trainable params: 0
```

Fig. 3 Summary of model

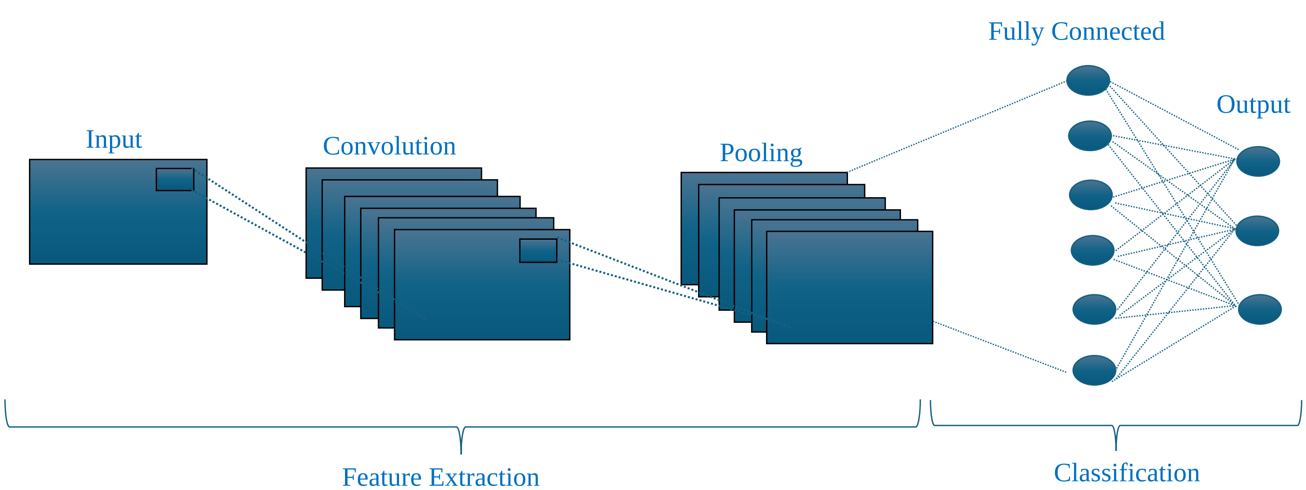


Fig. 4 Structure of CNN model [17]

```
normalization_layer = layers.Rescaling(1./255)
normalized_ds = train_ds.map(
          lambda x, y: (normalization_layer(x), y)
      )
image_batch, labels_batch = next(iter(normalized_ds))
first_image = image_batch[0]
print(np.min(first_image), np.max(first_image))
```

Fig. 5 Pre-processing script

Table 2. Summary of dataset description [19]

Cancer	Classes	Images
Acute Lymphoblastic Leukaemia	4	20000
Brain Cancer	3	15000
Breast Cancer	2	10000
Cervical Cancer	5	25000
Kidney Cancer	2	10000
Lung and Colon Cancer	5	25000
Lymphoma	3	15000
Oral Cancer	2	10000

4.5. Data Acquisition

The comprehensive dataset of medical images used to develop the model was obtained from Kaggle. It contained eight different types of cancers, each with a variable number of classes. The dataset is described in Table 2. The dataset had to be very well trained on the machine learning models to recognize images that will be fed into it subsequently. Thus, we used 80 of the images for the training and the remaining 20 % for testing purposes. In total, 52,000 images were used for training, and 13,000 were used for testing. Before feeding medical images into the CNN model for the training, preprocessing techniques such as rescaling and standardization were applied, as medical images are often of different sizes and resolutions. Rescaling ensures uniform dimensions, while standardization reduces variations in pixel intensity. Also, feature engineering central to CNN-based cancer detection is embedded in this work [17]. There are several convolution layers in feature extraction, followed by max-pooling and an activation function.

These features are fundamental to training CNN models effectively and enabling accurate cancer detection and classification. CNNs are structured to automatically extract relevant features from images through a series of convolutional and pooling layers. The model employs three convolutional layers to detect features like edges, textures, and patterns in the images. The convolution layer detects image features by sliding small filters across the data. These filters learn to recognize patterns. Weight sharing makes the process efficient, and the result is a feature map that highlights where specific features are in the input. This process allows the CNN to learn relevant features from the data automatically.

The depth of these layers increases progressively, allowing for the abstraction of higher-level features. Three pooling layers were incorporated to reduce spatial dimensions while preserving critical information. One flattened layer was also introduced to reshape the feature maps extracted by the convolution and pooling layers into a one-dimensional vector. This step prepares the data for input into the dense layers. Two dense layers are employed for classification. These layers facilitate the mapping of high-level features to specific cancer categories. The convolution and pooling layers inherently extract relevant features from the input images. The feature extraction sequence is shown in Figure 6.

5. Web and Mobile Application User Interface

This section shows the design of the User Interface (UI) for both the web and mobile applications, allowing users to interact with the cancer detection system. The UI is designed to be intuitive and user-friendly, enabling easy image uploads, viewing of classification results, and navigation. It ensures seamless access across devices, making the system accessible to healthcare professionals and patients. Figures 7 and 8 show the web application's home page and a typical query results page, respectively. Likewise, Figures 9, 10, 11, and 12 also display some interfaces that this work's mobile application can produce.

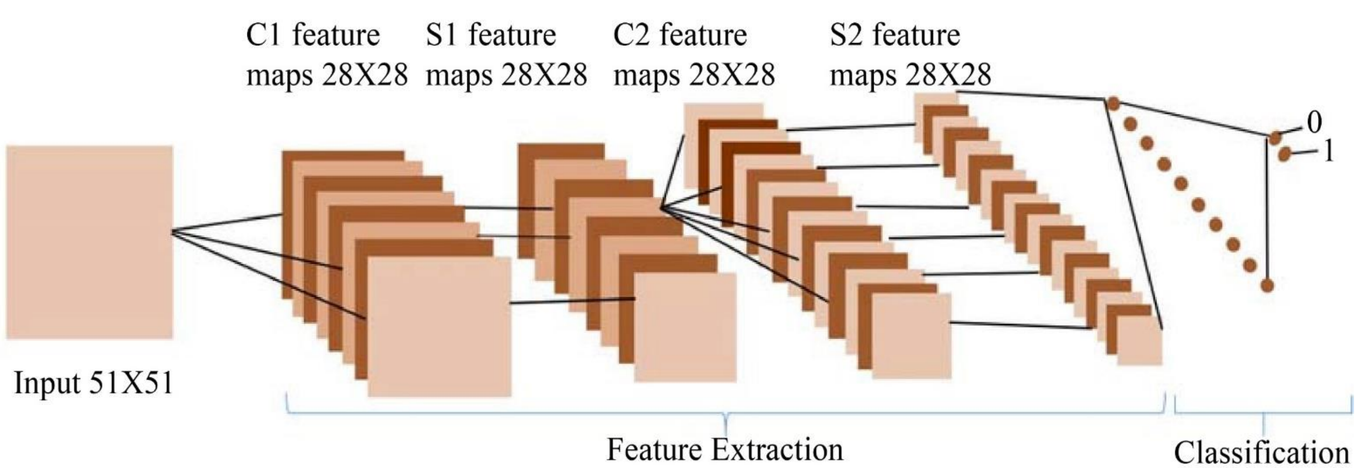


Fig. 6 Feature extraction in CNN [16]

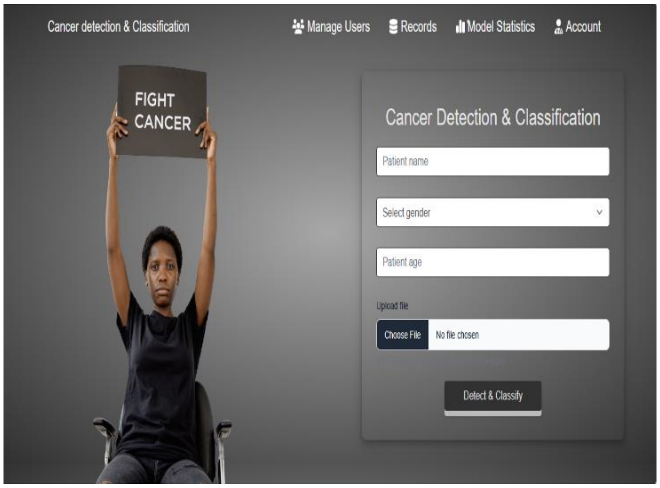


Fig. 7 Home page screen

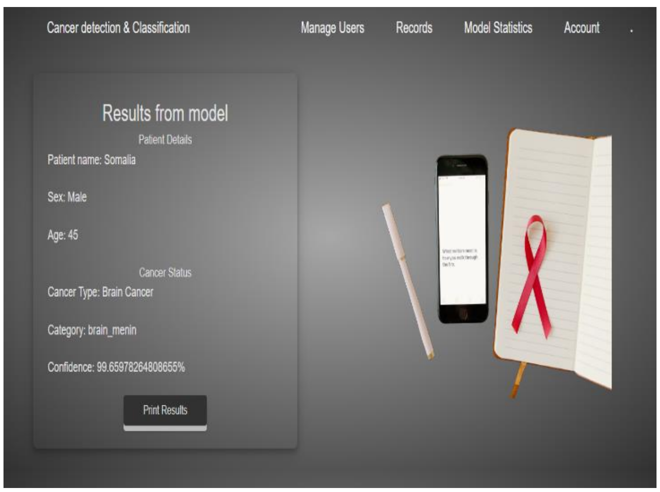


Fig. 8 Results screen

Fig. 9 Sign-in screen

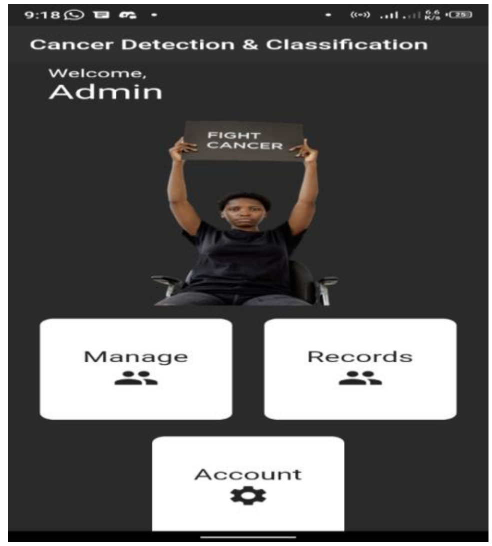


Fig. 10 Home page screen

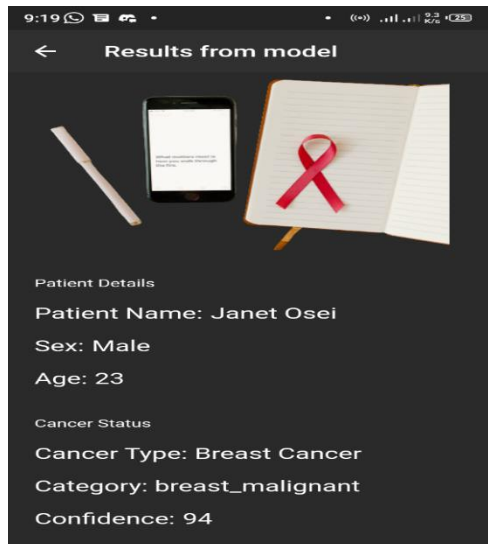


Fig. 11 Results screen

Patient	Date	Delete
Somalia	2023-06-04	
Kofi	2023-08-26	
Jeffrey Nyarko	2023-08-30	
John Doe	2023-06-04	
Fatimah	2023-06-04	
Micheal Addo	2023-08-30	
Kofi	2023-08-26	
Harry	2023-08-26	
Peter	2023-08-28	
Mary	2023-08-28	
Nathaniel	2023-08-30	
Darko Ernest	2023-08-30	
Janet Osei	2023-08-30	

Fig. 12 Records Screen

Table 3. Summary of model results

Model	Accuracy	Precision	Recall	F1-Score
General	0.97	0.98	0.98	0.98
Leukaemia	0.95	0.95	0.95	0.95
Brain Cancer	0.92	0.92	0.91	0.91
Breast Cancer	0.93	0.93	0.93	0.93
Cervical Cancer	0.89	0.90	0.89	0.89
Kidney Cancer	1.00	1.00	1.00	1.00
Lung Cancer	0.91	0.91	0.91	0.91
Lymphoma	0.96	0.96	0.96	0.96
Oral Cancer	0.92	0.93	0.93	0.93

6. Testing and Results

The main aim of this research is to alleviate the errors that occur with the manual examination of cancer by radiologists, leading to accurate and early cancer detection, which can increase patients' chances of survival. Errors in diagnosing cancer can lead to a delay in initiating the appropriate treatment, which reduces the chances of survival. Hyperparameter tuning was done to reduce or minimize the risk of misdiagnosis and ensure the optimum performance for all the models.

The parameters in the CNN model, including the number and arrangement of layers, were repeatedly modified to obtain the best results. After a series of tests, the optimizer and the loss function were chosen carefully to ensure the best performance possible.

The metrics used in this study included accuracy, precision, recall, F1-score, and confusion matrix to evaluate the performances of the nine (9) models. Since the dataset was balanced, accuracy can be used as a performance metric because there are equal probabilities of predicting each class. Table 3 summarizes the performance results obtained for each of the cancer models.

6.1. Discussion of Results and Analysis

From the results obtained from the testing phase, it was observed that most models had an accuracy above 90%, with kidney cancer having a perfect accuracy of 100%. Accurate cancer detection is critical for timely treatment initiation and improved patient outcomes. The metrics obtained proved that the model is very accurate and less susceptible to errors, which is the aim of this work. The metrics were obtained after continuously changing and testing the model's parameters to achieve optimum performance. The most accurate model was found to be kidney cancer, which had an accuracy of 100% on the testing dataset. The model correctly detected and classified all the images in the testing set. Kidney cancer had two output classes: kidney typical and kidney tumour, each 500 images for testing, making a total of 1000 images. The model correctly predicted each picture with 100% accuracy and high confidence in the individual prediction. Figures 13 through 20 give a confusion matrix for each type of cancer.

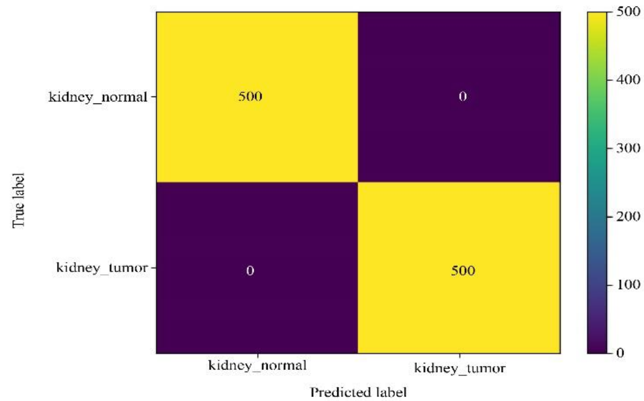


Fig. 13 Confusion matrix for kidney cancer

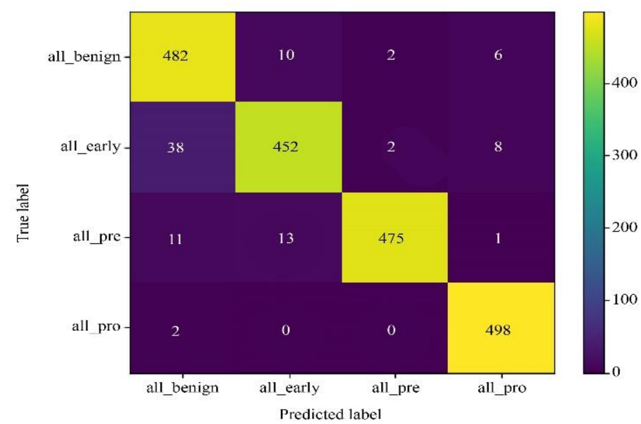


Fig. 14 Confusion matrix for Leukaemia

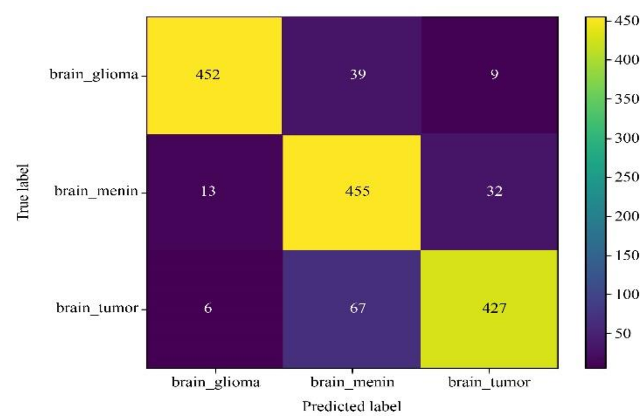


Fig. 15 Confusion matrix for brain cancer

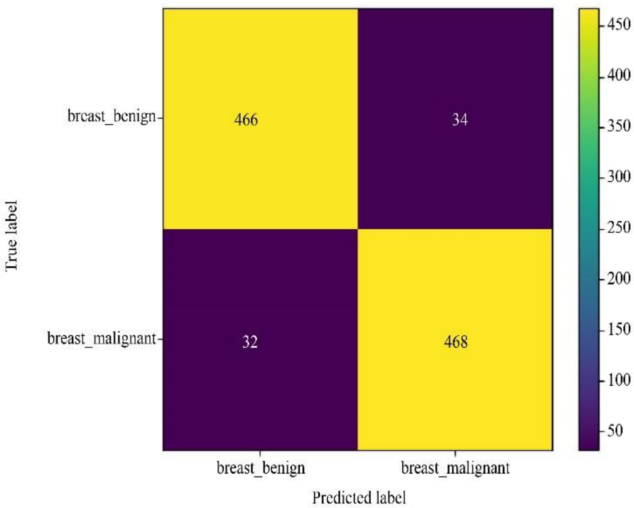


Fig. 16 Confusion matrix for breast cancer

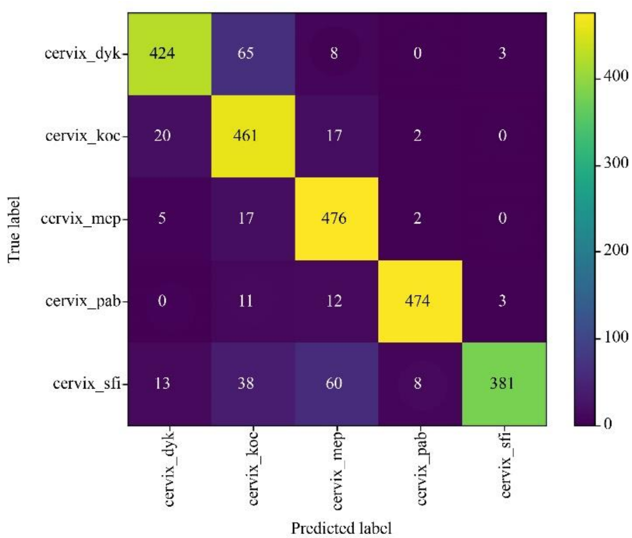


Fig. 17 Confusion matrix for cervical cancer

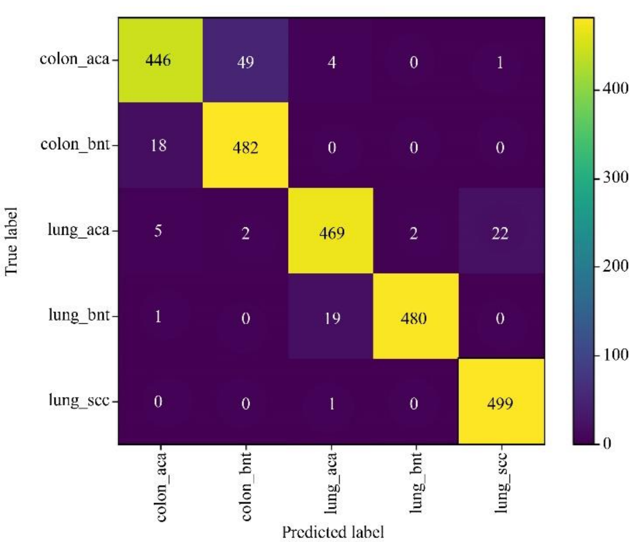


Fig. 18 Confusion matrix for lung and colon cancer

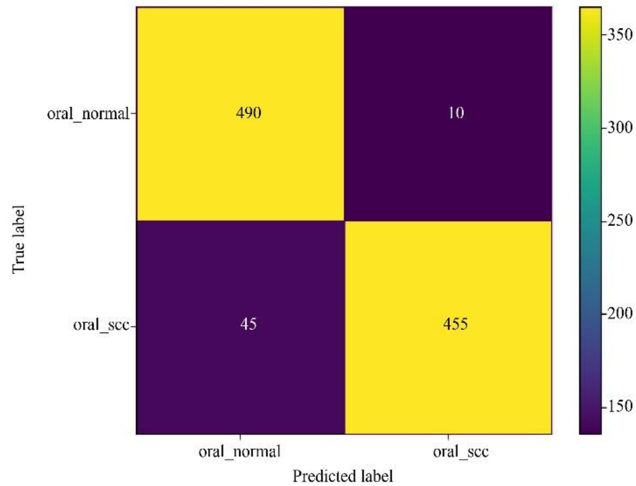


Fig. 19 Confusion matrix for oral cancer

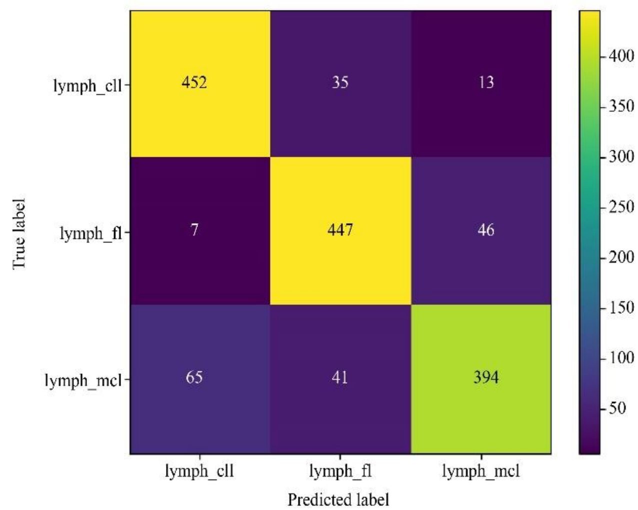


Fig. 20 Confusion matrix for Lymphoma

7. Conclusion

This work reflects an improvement over existing systems through the performance metrics obtained compared to existing works on cancer detection and classification. As outlined in the related works section, earlier studies predominantly focused on single cancer types using machine learning techniques like Support Vector Machine (SVM), Naïve Bayes, and Random Forest for medical image analysis. While these techniques are effective, they have limitations when applied to complex image datasets, especially in medical imaging, where features are subtle and require deep, hierarchical analysis. In contrast, this study's use of Convolutional Neural Networks (CNN) significantly enhances performance, particularly in image classification tasks. CNN's layered architecture allows it to automatically learn and extract intricate features and patterns from images through multiple convolutional layers, which are then pooled and passed through fully connected layers for final classification. This multi-layered feature extraction process gives CNNs an advantage over traditional machine learning algorithms, which may only capture superficial patterns and lack the depth required to classify complex images accurately. Additionally, the ability of CNNs to handle large volumes of data and learn increasingly abstract features as they progress through each layer contributes to their higher accuracy. This detailed hierarchical learning allows the system to differentiate between even the subtlest variations in medical images, which is critical for detecting different types of cancers. Other algorithms, while capable of pattern extraction, operate at a lower abstraction level, making them less suited for the complex task of medical image classification. The proposed system demonstrated that deep learning techniques are more effective for medical image analysis in cancer detection.

7.1. Major Findings of this Work

Through this work, several revelations have come to the fore. The key thing is that it significantly improves cancer diagnosis accuracy. It contributes to the existing body of knowledge in medical AI. Thus, the major findings are:

- Accurate Cancer Classification: The CNN-based model demonstrated high accuracy in classifying cancer types, holding promise for improving diagnostic accuracy in clinical settings.
- Efficient Image Processing: The system's efficiency in processing medical images and providing timely predictions contributes to its practicality in real-world healthcare scenarios.
- Usability and User Acceptance: Usability testing with healthcare practitioners revealed positive feedback regarding the user interface and overall user experience, underscoring the system's potential for adoption.
- Ethical and Privacy Compliance: The system adheres to ethical and privacy considerations, safeguards patient data, ensures informed consent, and makes the system HIPAA compliant.

7.2. Major Findings in the Work

Accessing a diverse and extensive dataset of high-quality medical images proved a significant challenge. The limited availability of such datasets hindered the ability to train the model comprehensively. Developing and deploying the system, particularly training the CNN model, required substantial computational resources. This resource intensiveness may pose constraints for healthcare facilities with limited access to powerful hardware. The reliance on a stable internet connection could be a limitation in regions with unreliable connectivity. Ensuring offline functionality is a potential solution. Integrating the system into existing clinical workflows and gaining acceptance from healthcare institutions is a complex process that requires ongoing efforts and collaboration. Maintaining compliance with evolving healthcare regulations and addressing legal and ethical concerns, particularly concerning liability, was a challenge that had to be solved. The system can only detect eight (8) types of cancers, creating the need for a continuously expanding dataset to account for rare and diverse cancer cases, necessitating ongoing data collection efforts. Also, strategies for optimizing resource usage, including cloud computing and data compression, are being explored to make the system more accessible.

8. Real-World Applications

This section focuses on how the proposed Convolutional Neural Network (CNN)-based cancer detection and classification system can be applied in clinical settings. The goal is to integrate machine learning technology into real-world healthcare environments, exploring its practical benefits, challenges, and pathways to overcome barriers to widespread adoption.

- Integration into Clinical Workflows: Given that the developed system can analyze medical images such as MRI, CT scans, and biopsies more quickly than traditional methods, it will allow for faster diagnosis and timely treatment decisions. This can be especially useful in high-pressure environments such as emergency rooms or oncology departments.
- Decision Support Systems (DSS): Implementing the CNN model as part of a Decision Support System enables healthcare professionals to have a second opinion in real-time. This could reduce errors in diagnosis caused by human factors like fatigue, bias, or inexperience.
- Telemedicine and Remote Diagnostics: CNN systems can be deployed remotely in areas with limited access to specialized healthcare. This enables radiologists or oncologists to remotely interpret images and make diagnoses, providing critical healthcare access to underserved regions.

8.1. Challenges in Real-World Applications

This section discusses the key challenges faced when implementing cancer detection systems in real-world healthcare settings.

- Obtaining Approval: Regulatory bodies such as the Food and Drug Administration (FDA) or the European Medicine Agency (EMA) ensure that AI systems for cancer detection undergo rigorous mandatory clinical trials before they can be deployed in clinical practice. This process can be lengthy and requires thorough safety, accuracy, and clinical efficacy testing.
- Integration with Existing Systems: Most hospitals use various systems (e.g., Electronic Health Records, PACS systems), so seamlessly integrating the CNN model into existing clinical workflows will require substantial effort. This includes ensuring interoperability between the AI tools and existing hospital information systems.
- Data Privacy and Compliance: Protecting patient data is paramount. Compliance with regulations such as HIPAA

(for the US) or GDPR (for Europe) is essential when implementing AI in healthcare. Ensuring secure data storage and processing is crucial to avoid breaches.
- Trust and Acceptance: Despite technological potential, clinicians may be reluctant to rely solely on AI for critical decisions like cancer diagnosis. Building trust in the system through clinical trials, validations, and user-friendly platforms is crucial for widespread adoption.
- Resistance to Change: Some clinicians may resist adopting new technologies, particularly those that change their workflow. Addressing this resistance through training, awareness programs, and evidence of the model's accuracy and reliability is vital for success.

8.2. Overcoming Real-World Implementation Challenges

Several strategies must be employed to mitigate potential challenges to integrate AI-based systems in clinical practice successfully.

- Pilot Programs: Before full-scale deployment, the system should be tested in a pilot program within a hospital or clinical setting. Feedback from clinicians during the pilot phase will help refine the model and the user interface.
- Gradual Integration: AI systems should be integrated gradually to complement existing methods rather than replacing current diagnostic workflows. This hybrid approach will build trust among healthcare professionals while allowing them to familiarize themselves with the new technology.
- Collaborative Learning: A model incorporating data from multiple institutions can improve generalizability and robustness, leading to better outcomes when deployed in new clinical environments.

8.3. Limitations

The dataset used in this study for both training and testing was cloned from websites such as Kaggle. This is a limitation because acquiring a large and diverse dataset of medical images is challenging in many cancer types. The study may have relied on limited publicly available datasets, which might not fully represent the variability seen in clinical practice. Some cancer types may be underrepresented in the dataset, making a model biased towards the more common classes. This imbalance could affect the model's ability to detect and classify rare cancer types accurately. Our system could perform well with the training dataset but may struggle to generalize unseen data, especially when tested on images from different sources, modalities, or patient populations. The system may have been tested on specific data sets and conditions that do not fully represent real-world scenarios. This could mean the model might not perform as expected in clinical environments. Also, significant computational resources such as very powerful GPUs and high memory capacity are required to train deep learning models, especially CNNs. This limitation makes it difficult for the model to be deployed in low-resource settings or smaller medical facilities with limited infrastructure. CNN models, while accurate, are often considered “black boxes” because they do not provide clear explanations for their decisions. This lack of interpretability can hinder their acceptance in the medical field, where clinicians must understand how the system arrives at its predictions.

9. Future Work

The current research focuses on a limited set of cancer types. Expanding the system to cover additional cancer types, including rare cancers, will broaden its utility in the medical field. By incorporating new datasets representing other forms of cancer, the model can be retrained to detect and classify a wider range of cancer subtypes. This will enhance the system's applicability in different areas of oncology. While the current model primarily uses medical images for cancer detection and classification, there is potential for integrating other types of medical data, such as genomic data, blood tests, or patient history. Future research could explore multimodal learning approaches, where the model processes images, lab results, and clinical notes to improve diagnostic accuracy and provide a more comprehensive view of the patient's condition. There is a growing need for AI-powered diagnostic systems that can be deployed in low-resource settings where healthcare infrastructure is limited. Future work could optimize the CNN model for mobile platforms or cloud-based solutions to make cancer detection tools available in remote or underserved regions. Research can also explore reducing computational and data storage requirements to ensure the system's functionality in low-resource environments. The current model may be limited in learning from new data as it becomes available. Continuous learning systems that evolve over time can enhance performance and maintain relevance in changing medical landscapes. Implementing online learning techniques that allow the model to update itself with new patient data, medical images, or advances in treatment can make the system adaptive and self-improve. Longitudinal studies that track the model's performance over time in clinical settings would also provide valuable insights for further refinement.

Acknowledgements

The authors are grateful to all the participants involved in the study.

Consent to Publish

All the authors have consented to the contents of this paper.

Data availability

Data used in this study is publicly available online.

References

[1] Muhammed Coşkun Irmak et al., "Comparative Breast Cancer Detection with Artificial Neural Networks and Machine Learning Methods," *29th Signal Processing and Communications Applications Conference (SIU)*, Istanbul, Turkey, pp. 1-4, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[2] William Hamilton, "Cancer Diagnosis in Primary Care," *British Journal of General Practice*, vol. 60, no. 571, pp. 121-128, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[3] Mohammed Odeh et al., "i.LLL.CancerCare: Towards an Intelligent Life Long Learning Framework for Cancer Care," *1st International Conference on Cancer Care Informatics (CCI)*, Amman, Jordan, pp. 244-246, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[4] Saloni Dattani et al., Cancer, Cancers are one of the leading causes of death globally. Are we making progress against them? Our World in Data, 2015. [Online]. Available: https://ourworldindata.org/cancer

[5] Ganta Sruthi et al., "Cancer Prediction Using Machine Learning," *2nd International Conference on Innovative Practices in Technology and Management (ICIPTM)*, Gautam Buddha Nagar, India, pp. 217-221, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[6] Abien Fred M. Agarap, "On Breast Cancer Detection: An Application of Machine Learning Algorithms on the Wisconsin Diagnostic Dataset," *Proceedings of the 2nd International Conference on Machine Learning and Soft Computing*, New York, NY, USA, pp. 5-9, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[7] Amjad Rehman et al., "Lung Cancer Detection and Classification from Chest CT Scans Using Machine Learning Techniques," *1st International Conference on Artificial Intelligence and Data Analytics (CAIDA)*, Riyadh, Saudi Arabia, pp. 101-104, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[8] T. J. Nagalakshmi et al., "Detection of Cervical Cancer with Texture Analysis using Machine Learning Models," *International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*, Chennai, India, pp. 1-6, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[9] Sweta Bhise et al., "Detection of Breast Cancer Using Machine Learning and Deep Learning Methods," *3rd International Conference on Intelligent Engineering and Management (ICIEM)*, London, United Kingdom, pp. 1-6, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[10] Amit Singh, Rakesh Kumar, and Rajul Rastogi, "Study of Machine Learning Models for the Prediction and Detection of Lungs Cancer," *11th International Conference on System Modeling & Advancement in Research Trends (SMART)*, Moradabad, India, pp. 1243-1248, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] K. Rajkumar et al., "Kidney Cancer Detection using Deep Learning Models," *7th International Conference on Trends in Electronics and Informatics (ICOEI)*, Tirunelveli, India, pp. 1197-1203, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[12] S. Rajeswari et al., "Detection and Classification of Various Types of Leukemia Using Image Processing, Transfer Learning and Ensemble Averaging Techniques," *2nd Asian Conference on Innovation in Technology (ASIANCON)*, Ravet, India, pp. 1-6, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[13] B. Ramya Sree et al., "Brain Tumor Detection and Classification using Magnetic Resonance Imaging and Machine Learning Approaches," *6th International Conference on Computing Methodologies and Communication (ICCMC)*, Erode, India, pp. 1729-1734, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[14] Theodore V. Maliamanis, and George A. Papakostas, "Chapter 3 - Machine Learning Vulnerability in Medical Imaging," *Machine Learning, Big Data, and IoT for Medical Informatics*, pp. 53-70, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[15] N.M. Saravana kumar et al., "Predictive Methodology for Diabetic Data Analysis in Big Data," *Procedia Computer Science*, vol. 50, pp. 203-208, 2015. [CrossRef] [Google Scholar] [Publisher Link]

[16] Mina Khoshdeli, Richard Cong, and Bahram Parvin, "Detection of Nuclei in H&E-Stained Sections Using Convolutional Neural Networks," *IEEE EMBS International Conference on Biomedical & Health Informatics (BHI)*, Orlando, FL, USA, pp. 105-108, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[17] Sai Balaji, Binary Image Classifier CNN Using TensorFlow, Medium, 2020. [Online]. Available: https://medium.com/techiepedia/binary-image-classifier-cnn-using-tensorflow-a3f5d6746697.

[18] Prashant Gurav, Flutter VS React Native: A Comparison Based On Criteria, Cuelogic, 2020. [Online]. Available: https://www.cuelogic.com/blog/flutter-vs-react-native-a-comparison-based-on-criteria.

[19] Obuli Sai Naren, Multi Cancer Dataset, Kaggle, 2022. [Online]. Available: https://www.kaggle.com/datasets/obulisainaren/multi-cancer.