Modern Education
and Computer Science
PRESS

Design and Implementation of Diagnosis System for Cardiomegaly from Clinical Chest X-ray Reports

Omolara A. Ogungbe
Department of Computer Science and Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria
E-mail: omolaraogungbe@gmail.com

Abimbola R. Iyanda
Department of Computer Science and Engineering, Obafemi Awolowo University, Ile-Ife, Nigeria
E-mail: abiyanda@oauife.edu.ng

Adeniyi S. Aderibigbe
Department of Radiology, Obafemi Awolowo University Teaching Hospital Complex, Ile-Ife, Nigeria
E-mail: adeniyiribigbe@gmail.com

Abstract: With the increasingly broadening adoption of Electronic Health Record (EHR) worldwide, there is a growing need to widen the use of EHR to support clinical decision making and research particularly in radiology. A number of studies on generation, analysis and presentation of chest x-ray reports from digital images to detect abnormalities have been well documented in the literature but studies on automatic analysis of chest x-ray reports have not been well represented. Interestingly, there is a large amount of unstructured electronic chest x-ray notes that need to be organized and processed in such a way that it can be automated for the purpose of giving urgent attention to abnormal radiographs in clinical findings to allow for quicker report analysis and decision making. This study developed a system to automate this analysis in order to prioritize findings from chest x-rays using support vector machine and Lagrange Multiplier for the constraint optimization. The classification model was implemented using Python programming language and Django framework. The developed system was evaluated based on precision, recall, f1-score, negative predictive value (NPV). Expert's knowledge was also used as gold standard and comparison with the existing system. The result showed a precision of 96.04%, recall of 95.10%, f1-score of 95.57%, specificity of 86.21%, negative predictive value of 83.33% and an accuracy of 93.13%. The study revealed that a limited but important number of relevant attributes provided an effective and efficient model for the detection of cardiomegaly in clinical chest x-ray reports. From the evaluation result, it is evident that this system can help the clinicians to quickly prioritize findings from chest x-ray reports, thereby reducing the delay in attending to patients. Hence, the developed system could be used for the analysis of chest x-ray reports with the purpose of diagnosing the patient for cardiomegaly. Chest X-ray reports are usually textual, therefore, further studies can introduce spell checker to the system to provide higher sensitivity.

Index Terms: Chest X-ray, Electronic Health Record, Clinical, Natural Language Processing, Machine Learning, Cardiothoracic Ratio, Cardiomegaly.

1. Introduction

Chest X-rays is responsible for production of images of the heart, lungs, airways, blood vessels and the bones of the chest and spine. It can also reveal fluid in or around the lungs or air surrounding a lung. Chest X-Rays (CXRs) are widely performed to diagnose and monitor a wide range of abnormalities such as infection, pneumonia, congestive heart failure or different cancers affecting the heart, bones, soft tissues and lung area. Automatically detecting these abnormalities with high accuracy could greatly enhance real world diagnosis processes [1]. Cardiomegaly is a medical condition in which the heart is enlarged and it is commonly referred to as an enlarged heart. Often cases, it results from hypertensive heart disease (HHD) which is a complication of hypertension or high blood pressure. Figs. 1 and 2 show the posterior-anterior chest x-ray of a normal adult male and the posterior-anterior chest x-ray of a patient with congestive cardiac failure (cardiomegaly) respectively. It may result in an ineffective blood pump to the body as a result
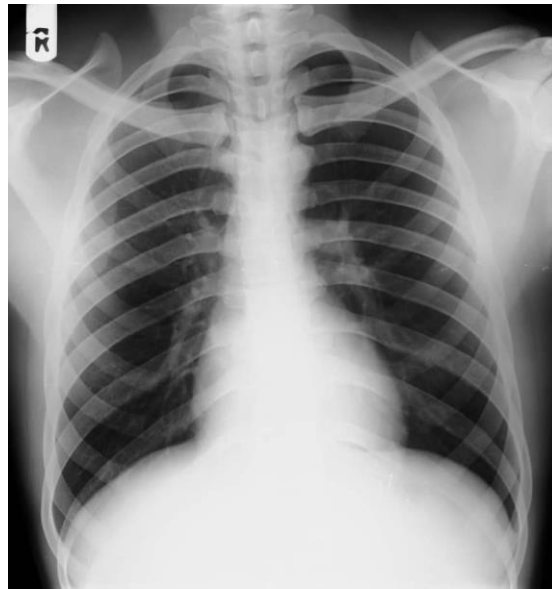
of congestive heart failure.
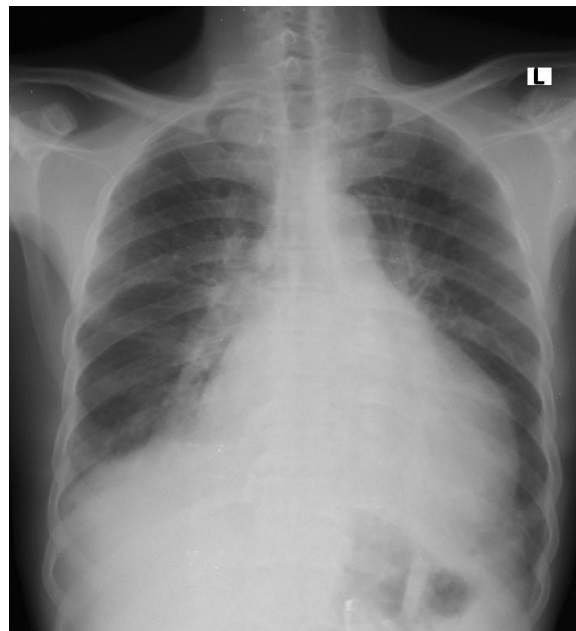

Fig. 1. Chest X-ray for Normal Heart

Fig. 2. Chest X-ray for an Enlarged Heart

Cardiomegaly is not a disease, rather a condition that can result from a host of other diseases such as coronary artery disease or obesity with a risk of sudden cardiac death. Although, a chest x-ray is the most common way to test for cardiomegaly because it is cost effective [2]. Nevertheless, there are other means and part of which are echocardiography and electrocardiography.

Cardiomegaly is an independent prognostic factor for mortality [2]. It is an independent predictor of death that could result in mortality due to some mechanisms such as higher oxygen need by larger ventricular mass, fatal ventricular arrhythmias and endothelial dysfunction. Artificial intelligence system can reduce the time needed to process abnormal chest x-rays and to achieve this, a human language processing algorithm can be developed and validated to read the radiologist report, understand the findings mentioned by the reporting radiologists and automatically infer the priority level of the examination. Automatic analysis and presentation, especially for chest x-ray reports from analog images, remain a challenge in the natural language processing community for healthcare research due to the usage of digital chest x-rays in the generation, analysis, and presentation of chest x-ray abnormalities.

When these analog diagnostic images are taken, the radiologist further interprets the images and resolves the images into a pros form. This pros form constitutes the chest X-ray report. Usually as practiced in most hospitals in

Nigeria, to determine the cardiomegaly status of a patient, the process starts from the radiologist. A patient is referred to a radiologist for a chest x-ray report to be generated by the radiologist. This radiologist's report is then placed in the patient's case file for the clinician to analyze and make decision on the patient's cardiomegaly status. This status could either be positive, negative, or indeterminate. At this point, most patients would have to wait for the clinician to treat their case file depending on the availability of the clinician. This can take a longer time than necessary thereby leaving the patient unattended to and possibly leading to death in critical cases.

The speedy detection implies that abnormal radiographs with critical findings could be prioritized to receive an urgent experts' clinicians' opinion much sooner than the usual protocol. Therefore, an alternative model of care such as human language processing algorithm could be used to greatly reduce delays in the process of identifying and acting on abnormal chest x-ray [3], particularly for cardiomegaly as reported that methods based on human language processing and machine learning tends to perform better in the extraction task in this domain but more experience is required for its analysis. Cardiovascular diseases are those which affect the structure and the function of the human heart.

The function of the heart sometimes depends on the structure of the heart which could be either normal or enlarged in size. The enlargement of the heart which is also known as cardiomegaly result in sudden cardiac death which emanates from the following as the reasons for this research:

1. shortages of doctors or clinicians particularly during peak times to cause high death rates;
2. patients not being seen within admission time.

These reasons have resulted in the grouping of patients based on their medical record or treatment outcome which in the study is considered to be the outcome of the chest radiograph examination of the patient taken by the Radiographers and interpreted by the Radiologists for the clinicians to act upon.

Therefore, this research seeks to find a way to automatically analyze these reports and further determine the patient's cardiomegaly status from chest x-ray reports written in an unstructured format without having to wait for the clinician's analysis. If such patient is at risk as detected by the system, the patient's case is prioritized and another clinician is to take up the case to save the patient's life.

2. Related Work

Radiography is one of the most frequently performed medical procedures usually carried out to examine changes in human tissue [4]. Clinically, radiology refers to the use of radiation in diagnostic imaging, nuclear medicine and interventional treatment. Patient's examination within the field of radiology are x-rays, computed tomography (CT) scans, magnetic resonance imaging (MRI), positron emission tomography (PET) scans and ultrasounds. Radiology is a means of making visible via imaging modalities abnormalities in patients using x-rays. Radiology has been influential in the diagnosis and treatment of a variety of conditions such as fractured bones, cancers, brain injury, clotted arteries, strokes, tendon and muscle damage, pulmonary conditions, spinal problems, osteoarthritis [5], dental caries [6], heart and lung conditions. X rays are a form of radiation that can penetrate the body and produce an image on an x-ray film.

According to [7], the x-ray is the oldest form of medical imaging and are broadly applicable for diagnosis and treatment. The standard chest x-ray is acquired with the patient standing up and with the x-ray beam passing through the patient from posterior (back) to anterior (front) where the x-ray produced is viewed as if the radiologist is looking at the patient from the front face to face. Sometimes it is not possible for radiographers to acquire a posterior-anterior chest x-ray because the patient is too unwell to stand but the chest x-ray is still viewed as if the radiographer is looking at the patient face to face. However, if the image is not labeled, it is usually fair to assume it is a standard posterior-anterior view as it was considered in our study. Most often, chest x-rays are used to detect abnormalities in the chest which include acute lung injury (ALI), pneumonia, pulmonary edema, broken ribs, lung cancer, and cardiomegaly. Basically, this work informed the proposed research on detecting cardiomegaly from chest x-ray reports.

[8] developed a natural language processing system to recognize device mentions in radiology reports to get an information about the device status. The data used in this research was gotten from the intensive care unit (ICU) of 46 Veteran Affairs Medical Centers from across the country during 1/1/2006 to 1/1/2009. Rule based method was employed as methodology in this work using semantics approach. This system was developed using 300 chest x-ray reports and was evaluated using 90 reports among which none was used in the development of the system. At the evaluation stage, reference standard was used as gold standard with just precision and recall as the evaluation metrics. The system developed here is different from the proposed system as machine learning approach was employed instead of rule-based approach and the evaluation metrics used in proposed system are sensitivity, negative predictive value and accuracy in addition to theirs.

The development and validation of a deep learning system to detect chest x-ray abnormalities was presented by [9] This system was developed using 1.2 million chest x-rays and their corresponding reports to detect abnormalities and a single algorithm was used to detect the following abnormalities: calcification, cardiomegaly, cavity, consolidation, fibrosis, hilar enlargement, opacity, pleural effusion and blunted costophrenic angle. In this work, x-rays from posterior-anterior and anterior-posterior view of position were included while supine position of x-ray were excluded. This work

differs from the proposed study in the sense that the proposed system only made use of chest x-ray reports which are generated from analog images while [9] trained it's model using the digital chest x-rays and its corresponding reports.

[2] emphasized that chest x-ray is the most common imaging examination of the heart. Its focus was to compare the findings from chest x-ray and echocardiography in the detection of cardiomegaly. The study shows that echocardiography is expensive and is not always successful as expected while chest x-ray is more accessible, cost-effective and feasible. However, chest x-ray does not have a high degree of sensitivity despite its specificity but it is recommended. Therefore, the usage of chest x-ray in the proposed research is based on this outcome.

[10] developed automated approaches to identify acute lung injury (ALI) from free-text chest x-ray reports using natural language processing with unigram as the baseline and bigrams and trigrams as the features and supervised text classification method. This work serves as the existing model where maximum entropy was used as the classifier with the precision of 81.70%, recall of 75.59%, f1-score of 78.53%, specificity of 76.80% and negative predictive value (NPV) of 74.61%. This work is used as a benchmark for the proposed research.

[11] developed a novel approach to classifying chest x-rays on the basis of the practical disease condition. It utilized attributes and not features as they are: a subset of data for traits; and subsequent working with attributes offer more granularities in framework configuration just to aid classification order execution. The result of [11] proved that the usage of machine learning algorithm for the analysis of chest x-ray reports can achieve a high accuracy. Therefore, based on the attributes available in the unstructured text, a natural language processing technique was used to formulate the computational model as suggested by [11].

[12] showed how advanced NLP techniques can aid clinical radiologist to carry out more efficient radiology as extracting information/data is challenging. General technologic approaches to NLP which could be understood by clinical radiologists were reviewed, latest developments in NLP techniques were explained, and up-to-date applications of NLP in radiology of which data was gathered to justify the research were discussed. Some applications of radiology which indicated the advantage of using NLP in terms of quality; language discovery and clinical and clinical research application were pointed out.

Findings show that the average compliance calculated by manual audit using of standardized reports was 91.2% with 89.3% to 92.8% confidence interval while the average rate of compliance calculated by the NLP automated audit was 92.0%, within the confidence interval. Also, in predicting survival for patients with rectal cancer. the data from NLP of radiology reports for CT studies was better than other clinical survival predictors. It is quite obvious that NLP will become the most incorporated feature of clinical activity and workflow in radiology. This study established the focus upon which our research is based.

[13] established the fact that there is a lot of data in the clinical world with majority as unstructured free text. This data requires an expert to read and infer the report, hence it addresses the need to use NLP to reduce time in getting these reports for easy and efficient extraction of meaningful information. The various processes involved in NLP and its applications in radiology reports were discussed. Ontology-driven concept recognition and mapping was suggested as a way of developing some NLP applications.

The result of the approach showed that customizing ontologies has been found to improve NLP algorithm performance from 42% sensitivity to 95%. Several ML based algorithms were explored and it showed a better result when NLP was adopted as highlighted by the authors. Without a doubt, NLP will be useful in furthering and improving research in cancer and aiding in personalized medicine approaches. The recent NLP research tells us the increasing role of NLP in radiology report interpretation, radiology report generation, emergency alert generation, uncertainty detection, data extraction for clinical decision support systems, predictive modeling, and cohort generation for research. The usage of SVM and TF-IDF as used in our study outperformed the ontology-driven approach with a higher sensitivity.

[14] assess and quantify the current literature in NLP used for radiology reports. A computerized exploration of literature yielding 4836 results using automated cleaning, metadata enriching stages and citation exploration in combination with manual review was conducted. The analysis was based on 21 variables comprising but not restricted to radiology features, NLP approach, performance, study, and characteristics of clinical application. A broad analysis of 164 publications extracted with publications in 2019 virtually triple those in 2015 was presented. Individual publication is categorized into one of the 6 categories of clinical application.

They indicated that the though application of deep learning increases in the period, yet conventional machine learning approaches remain widespread and this becomes a challenge when there is limited data and little indication of acceptance into clinical practice. Despite the fact that 17% of studies reported greater than 0.85 F1 scores, it was difficult to quite evaluate these techniques given that most of them use dissimilar datasets. It was reported that only 14 and 15 studies made their data and code available respectively with 10 externally proven results. It is known that automated understanding of clinical descriptions of the radiology reports has the ability to improve the healthcare procedure and it was shown that research in this area continues to grow.

The way models are explained are significant if the field is to transform applications into use clinically. More could be done to share code permitting validation of approaches on diverse institutional data and to lessen heterogeneity in communicating of study properties permitting inter-study assessments. The findings have implication for researchers in the domain giving a methodical synthesis of previous research to build on, detect gaps, prospects for partnership as well as avoid replication. The review study by [14] influenced the usage of conventional machine learning approach in

the proposed study despite the scarcity of dataset.

Several algorithms have been employed to process natural language in radiology but diversity in their performance in constant. This diversity has resulted from a number of factors that has been systematically and interrelatedly evaluated by [15]. A deep learning approach was used to evaluate the impact of dataset size and report complexity on an algorithm using bidirectional encoder representation from transformers, long short-term memory, a fully connected neural network (Dense) and convolutional neural network. Despite having same dataset on four algorithms, the performance of each algorithm was different. Therefore, the need for this proposed algorithm in respect to the radiology report complexity peculiarity in Nigeria.

A study on natural language processing of radiology reports in orthopaedic trauma was carried out by [16]. This study compared different machine learning approach to classify presence of injury(ies) from radiology reports in orthopaedic trauma. Dataset of Dutch radiology reports were used in testing rule based, machine learning and bidirectional encoder representation from transformers classifiers. Although, the study concluded that bidirectional encoder representation from transformers classifier outperformed the traditional machine learning and rule-based classifier on its dataset but it has not been applied on the dataset for the proposed model.

Different approaches have been used in the literature to perform abnormalities detection. The choice of a suitable approach to process human language is based on the specific attributes available for the language domain.

Automatic analysis and presentation of chest x-ray reports from analog images, remain a challenge in the natural language processing community for healthcare research due to the usage of digital chest x-rays in the generation, analysis, and presentation of chest x-ray abnormalities. When automatic analysis and presentation of chest x-ray reports are compared with humans (which is the basis for evaluation), their results are as close as that of humans. This study focuses on how to reduce the delay time of clinicians' response to patients with cardiomegaly of any form of level, hence produce a fast and well reliable system based on the chest x-ray reports generated by the Radiologists.

3. Methodology

This part explains the method used in the development of the system.

3.1. Data Collection and Preprocessing

Data were collected on Radiologists' interpretation written in a pros form of diagnostic imaging of an analog format. The data contains the date the chest x-ray was taken, the x-ray number (usually six digits), the text generated by the Radiologist(s) (chest x-ray interpretation) and the Radiologists name(s). Therefore, to produce a well-structured data from these, identification of attributes was done by interviewing experts (Radiologists) in the domain. The collected data were saved in a .docx form and converted to a usable scale format, with all of the significant features and variables extracted from the collected corpus. The dataset was preprocessed using the NLTK. Stop-words, punctuation marks and any other non-English characters were removed from the datasets. After the preprocessing stage, the dataset became structured, saved in .csv format with appropriate post-processing performed on the data before feeding them into the machine learning algorithm. For the purpose of this research, only the x-ray number and the chest x-ray interpretation were extracted from the report. This is because the x-ray number is unique for each interpretation of the chest x-ray.

3.2. System Design

To have an abstract view of how the system works, flowchart and usecase diagrams were used for the system design as shown in Figs. 3 and 4 respectively. The flowchart represents the step by step procedure used in interacting with the system by the user to detect cardiomegaly while the usecase diagram shows the activities the clinicians and the system administrator can carry out on the system both in the front end and the back end.

The conceptual view for the cardiomegaly detection system is shown in Fig. 5. The techniques applied to the developed system is as follows: First, the corpus were collected electronically as raw chest x-ray reports from the chosen domain (radiology); Second, data understanding for interpreting the structure of the dataset to know the attributes in the corpus was made after which term frequency - inverse document frequency vectorizer (TFIDFvectorizer) was used for the data preprocessing which include punctuation removal, ordered list creation, word count and count weighting.

The preprocessing result was then taken to be the clean chest x-ray reports upon which the linear support vector classifier was applied to extract and analyse the attributes of cardiomegaly. Linear support vector classifier provides decision function method which predicts confidence score for the three classes (positive, negative, indeterminate) of the dataset. The confidence score for each class is the signed distance of the class to the hyperplane.

Therefore, performing classification with its associated probability in this study gives some kind of confidence on the prediction. Hence, the usage of calibrated classifierCV to get the probabilities for all classes serves as step four in the conceptual model for a desirable post-processing. Finally, the outcome of this phase was used to determine the cardiomegaly status of the concerned patient whether it is positive (present of cardiomegaly), negative (absence of cardiomegaly) or indeterminate. Summarily, the tested value on the model serves as the result of the model used to develop the system and this study makes use of linear support vector classifier for the data training and calibrated

classifierCV for the probability calibration to provide high level of confidence in the predicted value.

3.3. System Implementation

The system was implemented using python programming language and the graphical user interface was designed using hypertext markup language (HTML) and cascading stylesheet (CSS) in collaboration with Django frame work. For proper navigation, responsiveness of the user interface was achieved using javascript. Although, there are several code editors for all the programming and scripting languages used in this study but visual studio was preferred due to its flexibilty and responsiveness. The file views.py is the application programming interface (API) for the cardiomegaly detection system and it was tested using postman API tester for handling all request as experienced in this work to be POST and GET requests. The three functions in the API are explained as follows:

• Function 'login' is a POST request that allows the user to enter the username and password in to the system while the system communicate with the database (postgreSQL) as used in this work for validity.
• Function 'dashboard' is that which gives access to the user to enter the patient's xray number and the chest x-ray report via the interface into the system to detect cardiomegaly. It is also a POST request. This function makes use of the formulated model to make classification of the chest x-ray report inputed to the system.
• Function 'logout' allows the logged in user to end session of usage after interacting with the system.
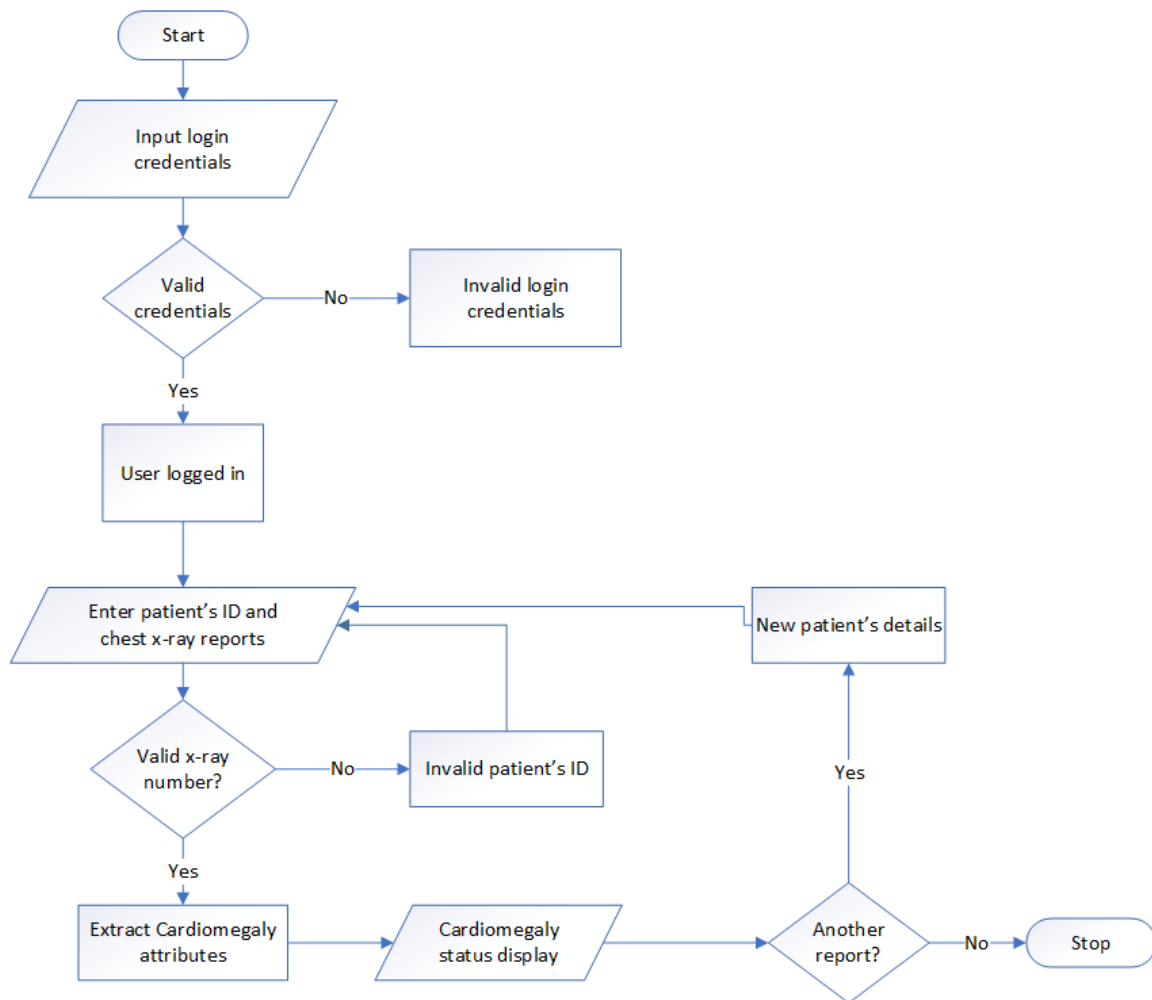

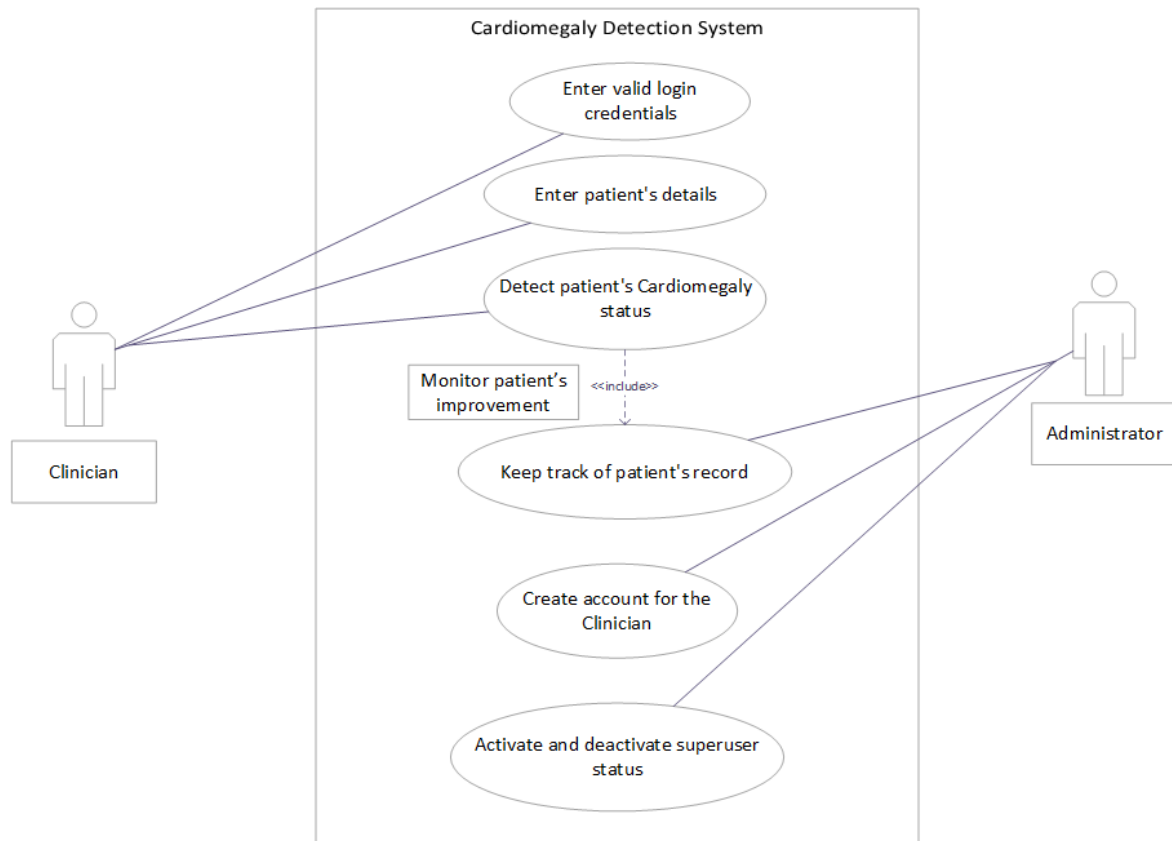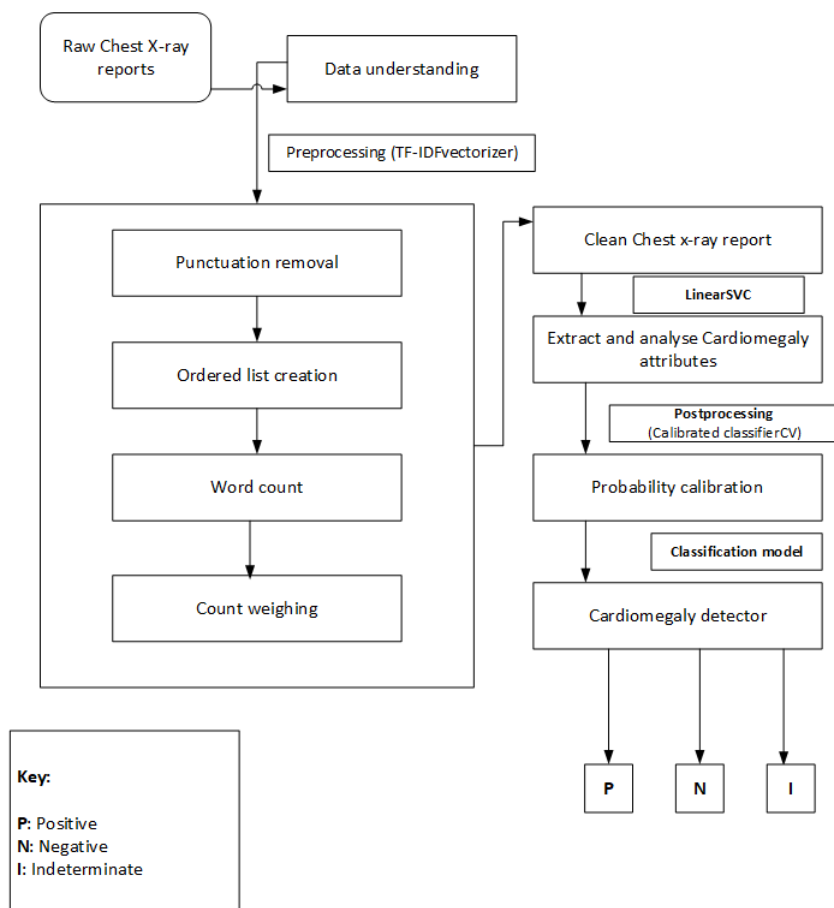Fig. 3. Flowchart for the Cardiomegaly Detection System

Fig. 4. Use Case Diagram

Fig. 5. Conceptual view of the system

4. Results and Discussion

The implementation and evaluation of the system are discussed in this section.

4.1. Software Environment

The Cardiomegaly detector environment contains a graphical user interface (GUI) that is made up of two text areas and a button as shown in Fig. 6. The text areas accept the login credentials (username and password) of the clinician. After correctly entering the username and password the clinician (user) can click on the 'login' button to navigate to the next page. If the user enters a wrong login credentials, then it displays 'invalid credentials' else, it displays the page in Fig. 7.
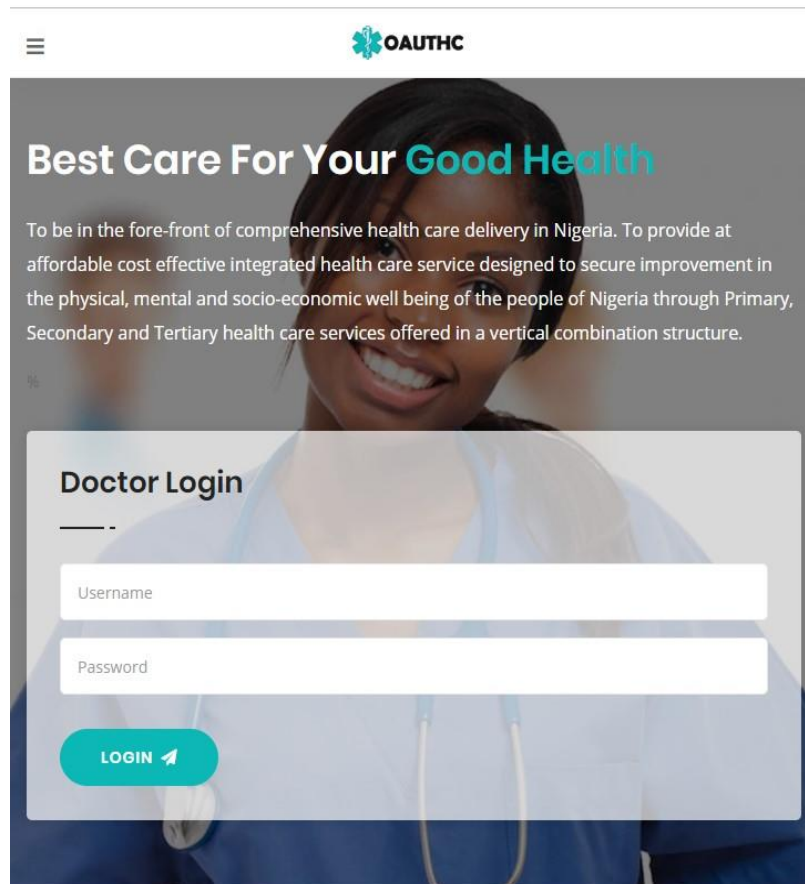

Fig. 6. Login page for all users.

4.2. Software Demonstration

The system detects cardiomegaly in chest x-ray reports written in human language which is usually unstructured and in pros form by taking in an x-ray number together with its corresponding chest x-ray report from the user (clinician) (features and variables are extracted automatically by the system) and the cardiomegaly status of the patient is predicted by the system. Peradventure the user enters a chest x-ray report without an x-ray number or an incorrect x-ray number, the system returns an error "Patient id invalid" else, if correct x-ray number has been entered (e.g. 202992) and the chest x-ray report is entered appropriately (e.g. Rotated radiograph. In spite of this, the cardiac silhouette is enlarged with a left ventricular prependence and a CTR of 0.52. There is fullness of both hila. The aorta is however not unfolded. Both pleural recesses are preserved. Normal bony thorax and overlying soft tissue) as shown in Fig. 8, the patient's information provided serve as a test to the model generated by the classification algorithm to detect the cardiomegaly status of the patient (The result shows positive which implies the sentence the clinician entered has attributes of presence of cardiomegaly) as in Fig. 9.
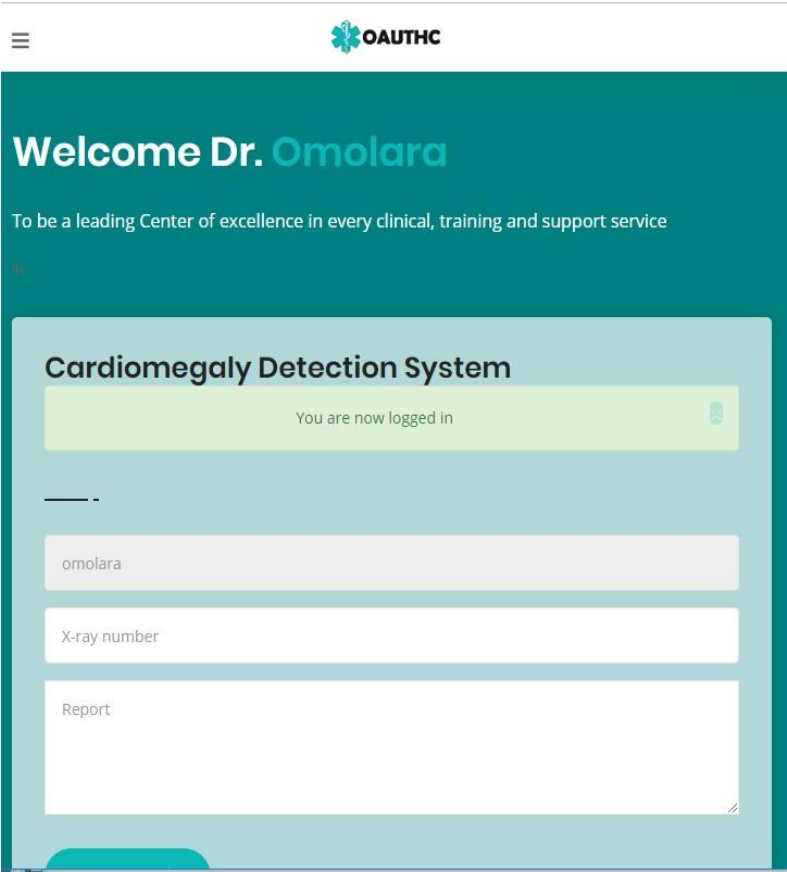
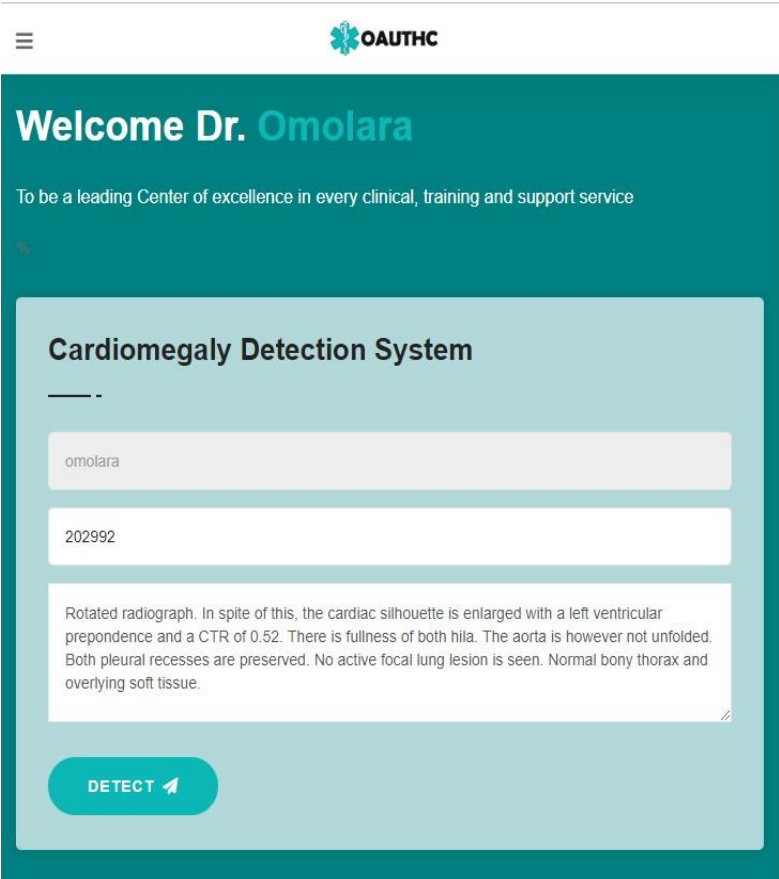Fig. 7. Display page for successful login credentials

Fig. 8. Interface with user input

Fig. 9. Output from user's input data

4.3. Evaluation of the System

The metrics used in evaluating the system are precision, recall, f1-score, specificity, negative predictive value (NPV) and accuracy using gold standard as a benchmark. During the system evaluation, only the positive and the negative responses were considered in the identification of the cardiomegaly status of patients.

The dataset used for evaluating the detection system were one hundred and fifty (150) chest x-ray reports with their respective x-ray number and were in .docx format which makes them digitized text. The x-ray number is a 6-digit value while the chest x-ray report is in form of unstructured text written in human language, generated by an experienced Radiologist with over four (4)-year experience in reporting chest x-rays. The Radiologist was blinded to the result of the evaluating and detection system while the researcher (operating the detection system) was also blinded to the Radiologist report.

Using the gold standard which is the expert's labeling of the dataset in comparison to the proposed system performance, the following outcome was obtained: ninety seven (97) out of one hundred and fifty (150) chest x-ray reports were labeled to be negative by the expert and the system accordingly (true negative value), twenty five (25) out of one hundred and fifty (150) chest x-ray reports were labeled to be positive by the expert and the system accordingly (true positive value), four (4) out of one hundred and fifty (150) chest x-ray reports were labeled to be positive by the expert but labeled to be negative by the system (false negative value) while five (5) out of one hundred and fifty (150) chest x- ray reports were labeled to be negative by the expert but labeled to the positive by the system (false positive value).

Therefore, the cardiomegaly diagnostic system had a precision of 96.04% which implies that 96.04% of the chest x-ray reports were rightly identified. It also has a recall of 95.10% which implies that 95.10% of the chest x-ray reports in the test data-set were identified.

Furthermore, it has a specificity of 86.21%, negative predictive value of 83.33%, f1-score of 95.57% and an accuracy of 93.13%. The performance of the system was further extended to be compared with the existing model [10] which was made use of to detect Acute Lung Injury (ALI) where the data classification was done using maximum entropy. This is represented in Table 1.

Table 1. Results comparison with the existing system

Model (%)	Precision (%)	Recall (%)	F1-Score (%)	Specificity (%)	NPV (%)	Accuracy (%)
Yetisgen-Yildiz et al. (2013)	81.70	75.59	78.53	76.80	74.61	-
Proposed system	97.14	85.00	90.67	98.59	92.11	93.69

4.4. Discussion of Results

During the evaluation of the implemented system, the following observations were made:

1. Most of the false positive results (reports which indicate presence of cardiomegaly that are classified by the system as absent cardiomegaly) gotten while evaluating the system performance based on the aforementioned metrics was because the chest x-ray report was generated from either a rotated radiograph or a penetrated radiograph which sometimes makes it difficult to be specific on whether the patient's heart size is normal or enlarged.

A sample chest x-ray report in this category is as follows: *Underpenetrated rotated radiograph. Within this limit, the cardiac shadow is enlarged with left ventricular preponderance (CTR= 0.54). Both the ascending and descending aorta are unfolded. No active focal lung lesion is seen, but there is plate atelectasis in both lung bases. Both hila appear full (L greater than R). The costophrenic recess are preserved. There is however a well define rounded opacity of soft tissue density seen in the right lower lung field? artefactual? significance. It measures 0.9cm in diameter. The other portions of the lung field are preserved bilaterally. The rib cage and overlying soft tissues are preserved. No cortical discontinuity is seen in the ribs.* The total number of reports that falls in this category is four (4) out one hundred and fifty (150).

2. Most of the false negative results (reports which indicated absence of Cardiomegaly but are classified by the system as having cardiomegaly) during the system evaluation was based on the following facts:

(a) Any cardio thoracic ratio (CTR) greater than 0.5 is considered to be as a result of an enlarged heart which in some cases might be wrong due to the x-ray projection of the patient which could either be posterior anterior (which is the standard projection) or anterior-posterior projection. These projections depend on the level of wellness of the patient such that anterior-posterior projection might not give the true size of the heart

(b) Cardiomegaly is a heart condition that could emanate from other heart diseases such as hypertension. Therefore, the chest x-ray reports in this category considered presence of the word "hypertension" in the corpus as an indication for likelihood of enlarged heart which may not be so in some cases. The total number of reports in this category is five (5) and a sample chest x-ray report is as follows: *The apparent size is preserve. There is unfolding of the aorta. This may be due to aging/or hypertension. No active focal lung lesion is seen. Although the hemodiaphragms appear low set. Both cos tophrenic sulci are intact. Normal overlying ribcage and soft tissue shadows.*

From Table 1, the performance of the algorithm used by [10] was less effective in comparison to the proposed algorithm. The proposed algorithm has resulted in a higher recall, precision, f1-score, negative predictive value and specificity. With these metrics and accuracy inclusive, the proposed study has been able to introduce a better classifier for abnormalities from clinical chest x-ray reports.

Summarily, all chest x-ray reports are dependent and subject to the interpreter which in this research is the Radiologist. A detailed report often produces positive system evaluation than a less detailed Radiologist report. Therefore, the higher the level of experience of the Radiologist, the more exact the reports are. Hence, increases the efficiency of the developed system in detecting Cardiomegaly.

5. Conclusions

A number of chest x-rays analysis and diagnosis systems have been developed for digital images with various levels of precision, recall, specificity and f1-score. However, researches on automatic detection of abnormalities in analogue x-rays is scarce and yet to be materialized into standard software application. This research focused on the development of a software system for the detection of Cardiomegaly from chest x-ray reports which are generated from analogue x-rays. The model was formulated using support vector machine and Lagrange Multiplier for the constraint optimization.

Result showed a precision of 96.04%, recall of 95.10%, f1-score of 95.57%, specificity of 86.21%, negative predictive value of 83.33% and an accuracy of 93.13%. This result indicates that the Cardiomegaly detection process for analogue x-rays can be automated to a great extent even though there were some false positive and false negative results. Based on this, it is reasonable to conclude that no theory of intelligence can perfectly stand on its own to process human language to mimic human intelligence. In determining the most effective computational model to capture and implement the automation of abnormalities detection from chest x-ray reports, the results of an existing model (used to detect Acute Lung Injury (ALI)) and the proposed system results were compared.

The evaluation showed that the proposed system achieved higher precision, recall, f1score, specificity and negative predictive value while accuracy was not considered as part of their metrics. Hence, it is reasonable to recommend the developed system as a Cardiomegaly detector in hospitals. This has the potential to reduce delay in attending to patients and also help the clinicians prioritize findings from chest x-ray reports. Although this system meets its aim but there is still a need to incorporate spellchecker since the spelling of each word in the report sentence has to be correct. Spell checking is almost a solved problem in biomedical intelligence with the usage of unified medical language system (UMLS).

Further study should therefore be carried out to integrate this system with a spell checker. This system can be upgraded to helps future researchers detect other abnormalities in the chest region from chest x-ray reports.

References

[1] Islam, M. T., Aowal, M. A., Minhaz, A. T., & Ashraf, K. (2017). Abnormality detection and localization in chest x-rays using deep convolutional neural networks. *arXiv preprint arXiv:1705.09850*, *3*(1), 1–16.

[2] Monfared, A. B., Farajollah, S. A., Sabour, F., Farzanegan, R., & Taghdisi, S. (2015). Comparison of radiological findings of chest x-ray with echocardiography in determination of the heart size. *Iranian Red Crescent Medical Journal*, *17*(1), 1–6.

[3] Sridevi, M., & Arunkumar, B. (2016). Information extraction from clinical text using nlp and machine learning: Issues and opportunities. *International Journal of Computer Applications*, *975*(8887), 11–16.

[4] Ilovar, M., & Sajn, L. (2011). Analysis of radiograph and detection of cardiomegaly. In *2011 Proceedings of the 34th International Convention MIPRO*, pp. 859–863. IEEE.

[5] Gornale, S. S., Patravali, P. U., Marathe, K. S., & Hiremath, P. S. (2017). Determination of osteoarthritis using histogram of oriented gradients and multiclass SVM. International Journal of Image, Graphics and Signal Processing, 9(12), 41

[6] Geetha, V., & Aprameya, K. S. (2019). Textural analysis based classification of digital X-ray images for dental caries diagnosis. Int J Eng Manuf (IJEM), 9(3), 44-5.

[7] van Gelderen, F. (2004). A brief history of radiology. In *Understanding X-Rays*, pp. 597–602. Springer.

[8] Rubin, D., Wang, D., Chambers, D. A., Chambers, J. G., South, B. R., & Goldstein, M. K. (2010). Natural language processing for lines and devices in portable chest x-rays. In *AMIA Annual Symposium Proceedings*, Vol. 2010, p. 692. American Medical Informatics Association.

[9] Putha, P., Tadepalli, M., Reddy, B., Raj, T., Chiramal, J. A., Govil, S., Sinha, N., KS, M., Reddivari, S., Rao, P., et al. (2018). Can artificial intelligence reliably report chest x-rays?: Radiologist validation of an algorithm trained on 1.2 million x-rays. *arXiv preprint arXiv:1807.07455*, *1*, 1–13.

[10] Yetisgen-Yildiz, M., Bejan, C., & Wurfel, M. (2013). Identification of patients with acute lung injury from free-text chest x-ray reports. In *Proceedings of the 2013 Workshop on Biomedical Natural Language Processing*, pp. 10–17.

[11] Savitha, S., & Naveen, N. (2018). Comprehensive classification model for diagnosing multiple disease condition from chest x-ray. *International journal of advanced computer science and applications*, *9*(9), 326–337.

[12] Donnelly, L. F., Grzeszczuk, R., & Guimaraes, C. V. (2022). Use of natural language processing (nlp) in evaluation of radiology reports: An update on applications and technology advances. In *Seminars in Ultrasound, CT and MRI*. Elsevier.

[13] Mithun-Nair, S., Jha, A., Rangarajan, V., Wee, L., & Dekker, A. (2021). Natural language processing in radiology reports. In *XIX Annual Conference on Evidence Based Management of Cancers in India: Technology and Cancer Care-Promise and Reality of the Brave New World*, pp. 461–472. Tata Memorial Centre.

[14] Casey, A., Davidson, E., Poon, M., Dong, H., Duma, D., Grivas, A., Grover, C., Su árezPaniagua, V., Tobin, R., Whiteley, W., et al. (2021). A systematic review of natural language processing applied to radiology reports. *BMC medical informatics and decision making*, *21*(1), 1–18.

[15] Olthof, A. W., van Ooijen, P. M. A., & Cornelissen, L. J. (2021). Deep Learning-Based Natural Language Processing in Radiology: The Impact of Report Complexity, Disease Prevalence, Dataset Size, and Algorithm Type on Model Performance. Journal of medical systems, 45(10), 1-16.

[16] Olthof, A. W., Shouche, P., Fennema, E. M., IJpma, F. F. A., Koolstra, R. H. C., Stirler, V. M. A. Stirler, P. M. A. van Ooijen & Cornelissen, L. J. (2021). Machine learning based natural language processing of radiology reports in orthopaedic trauma. Computer methods and programs in biomedicine, 208, 106304

Authors' Profiles

Omolara Ogungbe is a researcher with a keen interest in machine learning, image processing and natural language processing. She has a Master's degree in Intelligent Systems Engineering from Obafemi Awolowo University, Nigeria. She is currently a PhD student.

Abimbola R. Iyanda holds a B.Sc. degree in Computer Engineering, an M.Sc. and Ph. D. degrees in Computer Science from Obafemi Awolowo University, Ile-Ife, Nigeria. The thrust of her research is in the area of Computing and Intelligent Systems Engineering with focus on Speech and Language Engineering research aiming at domesticating computer technology and the computational rendering of indigenous ideas. She is a Member of the Nigerian Society of Engineers, Association of Professional Women Engineer in Nigeria, Council for the Regulation of Engineering in Nigeria, Association for Women in Science for the Developing World (OWSD) and Nigeria Computer Society. Her present employment is with the Computer Science and Engineering Department, Obafemi Awolowo University, Ile-Ife, Nigeria.

Dr Aderibigbe Adeniyi Sunday is an academic Radiologist and Honorary Consultant Radiologist at the Obafemi Awolowo University/ Obafemi Awolowo University Teaching Hospitals Complex Ile-Ife. He has special interests in musculoskeletal and Vascular imaging as well as artificial intelligence applications in radio diagnosis..