SCaiN- AI based application to read Medical Images

Introduction:

There is an acute shortage of doctors in Australia and around the world, this is especially true of radiologists who are specialists in reading medical images and diagnosing illnesses that these can notice. With Australia, other countries like the US and UK are also facing an alarming shortage of radiologists. In some cases, the waiting time for the assessment goes up to 30 days. Such a huge amount of waiting time is alarming as it is risky for critical case patients. (Driver M, (n.d.))
With the increase in complex images, assessment becomes difficult. There will be need to train and employ more radiologists to bridge the gap. This solution is not viable because of the hospitals budget concerns. Rural areas with minimal resources don't always have the availability of the required medical professionals. All the developing and underdeveloped countries are facing these issues daily. With increasing number of patients, the demand for the increase in the number of medical professionals is increasing day by day. (The Radiology Labour Shortage, 2020)
This can be avoided by the help of AI. Use of AI in medical industry is starting to prove very vital and essential. AI image recognition techniques have been proved accurate and precise in past few years. The idea is to use AI systems to provide insight and useful information to the experts on time. Using an AI image recognition in medical imaging can result in less wait time for diagnosis of the patients. AI systems are better than humans in recognising abnormalities in provided data. This can help ease pressure off the systems and medical professionals throughout. The major stakeholders in this issue are the hospitals, patients and radiologist. (Long M, 2020)

Problem Definition

Currently, many countries in the world are facing a shortage of radiologists. This is increasing the load on the system and getting the radiologists overworked. A solution to that problem is using an AI image recognition technique to give meaningful insights to medical professionals, resulting in reduced decision-making time and precise results.
Existing AI solutions are costly and inefficient. For resolution of this issue, an unbiased AI should be trained and made available to radiologists as a service. This will be a useful solution for hospitals with little or no access to trained radiologists.

Significance:

Impact of issue:

There is a shortage of radiologists around the globe. Medical imaging and radiology help to diagnose illness such as brain haemorrhage, pulmonary embolism, cancer, and fractures of the bones. Without the knowledge of what exactly is causing a patient to be ill, the treatment of the symptoms of the patient cannot be effectively treated. There are a huge 3.6 billion X-rays each year globally and as such, a vast number of X-rays need a huge workforce to analyse those medical images, one that cannot keep up with demand.
There are only 1.9 radiologists per million people in developing countries while there are 97.9 radiologists per million inhabitants in high-income countries (Hricak et al, 2021). It was found by performing a simulation model of 11 cancers. A significant increase of imaging in poorer areas currently critically underserved by radiology abilities would avert 3 - 2% of 76 million deaths or 2·46

million deaths caused by the modelled cancers worldwide between 2020 and 2030, saving 54 - 92 million years of people's lives, simply from better detection of these cancers.(Hricak et al, 2021). Rural hospitals in poor nations are relying on telehealth and sending scans away into developed countries because they have no one with the expertise to analyse them. They are then relying on volunteer radiologists to look over them who are already strained in their jobs within their country too. These results have exceptionally long wait times for people in developing countries which could determine life or death (Mitchell.C, 2021)

Cost of impact:

In the UK (United Kingdom), it was found just 2% radiology departments fulfilled their reporting requirements within contracted hours resulting in longer wait times for results (Long M, 2020) There are many hospitals in the UK which cannot operate 24/7 radiology labs which means patients miss out on certain medical procedures. The Royal College of Radiologists' (RCR) annual radiology workforce report found: (Our World in Data, 2021)

- Demand for complex CT and MRI scans is growing at three times the speed of the radiologist workforce(The Royal College of Radiologists, n.d).
- In 2018, approximately 200 doctors will qualify as radiology consultants which is not enough to fill even half of ongoing vacancies(The Royal College of Radiologists, n.d).
- The UK radiologist workforce is now 33 per cent short-staffed – without more consultants in training and better staff retention and recruitment, the shortfall will hit 43 per cent by 2020.

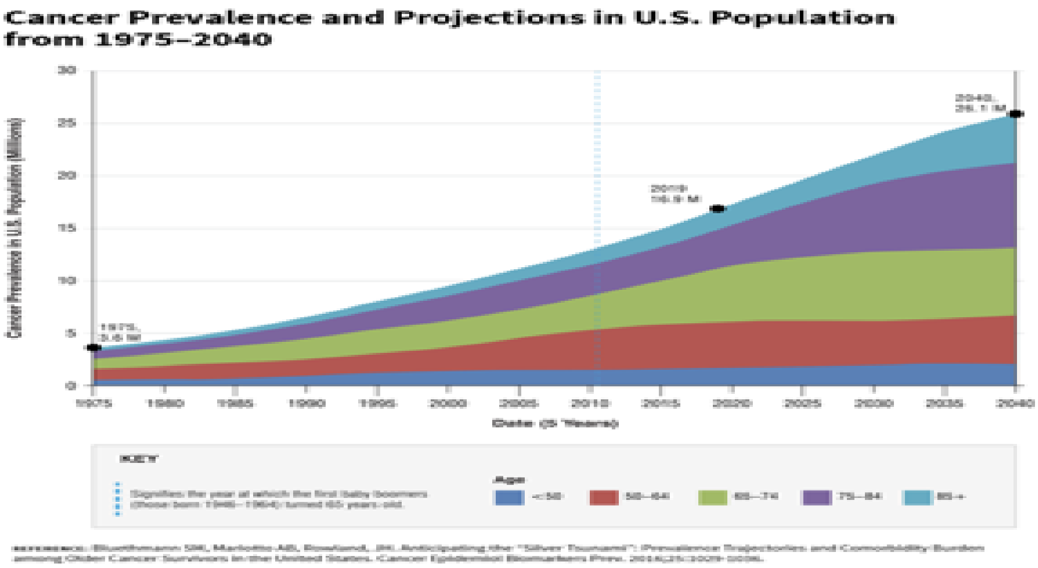

Fig.1 Statistics and Graphs, Division of Cancer Control and Population Sciences

Radiology experts believe radiology services will be in even greater demand as the population ages (The Radiology Labour, 2020).

Proposed data-driven solution:

Solution:

As per traditional existing solution, the scanned images are be analysed by radiology/lab technicians. It usually takes longer times to obtain the results. There are some existing AI based applications. But

they offer image analysis for a limited set of diagnosis and rather expensive. Patients usually tend to be anxious to know the results after a scan, this can’t be done with the existing traditional solution. We developed a solution to cut down the longer waiting times and address the lack of radiologists in healthcare department. Our proposed solution provides faster, efficient and reliable results after analysing the scanned X-rays/MRI/CT-scans. In our solution, we are using various types of datasets to train our AI-models. We are developing our model to work in broad spectrum and analyse different types of scanned images such as X-rays, MRI’s and CT-Scans to provide result for various diagnosis treatments such as Breast cancer, Pneumonia and Lung cancer etc, we implemented 2 different algorithms. A CNN based algorithm and a RADTorch based algorithm (RADTorch, n.d). If both the algorithms return positive, then the patient can consider it as positive. If both the algorithm returns negative, then the patient can consider it as negative and can consult the doctor for further examinations. If there’s a scenario of 1 positive and 1 negative result from the algorithms. Then it is advised to get analysed from the radiologists. Our basic inspiration for this idea derived from RAT tests prescribed for COVID-19. The RAT based results were much quicker and on par with RT-PCR test results.
Our proposed solution is a web/mobile based application and works in simple 3-steps:

- Patient selects which type of scanned image they are going upload.
- Then patient selects which type of diagnosis that they are going for, such as Breast cancer, Pneumonia and Lung cancer etc, to get result.
- Finally, the results will be displayed as positive or negative after analysing the image by both the algorithms.

ML methods and algorithms used:

We used various types of datasets such as DICOME (Digital Imaging and Communications in Medicine) dataset for lung cancer/breast cancer/, CT-scan, X-ray datasets for heart-based diseases and MRI scan datasets (DICOM, n.d.).
To train ML algorithms, we used pneumonia related X-ray datasets, lung cancer related CT scan datasets, Tumour related MRI scan datasets (Guluck M. A, n.d.).
We developed tailor made ML models for each specific scan type and diagnostic type depending on the datasets. We used **CNN algorithm** (Mathur.M, 2020), i.e., CNN architectures with ImageNet pre-trained weights (e.g., vgg16, resnet50) with **RMSProp Optimizer**, a RMSprop optimizer restricts the oscillations in the vertical direction. Therefore, we can increase our learning rate and our algorithm could take larger steps in the horizontal direction converging faster. We also performed **data augmentation** if required, i.e., to overcome bias in the training model we need to expand our dataset artificially. We can make our existing dataset even larger. The idea is to alter the training data with small transformations to reproduce the variations. Approaches that alter the training data in ways that change the array representation while keeping the label the same. For data augmentation, we randomly choose some training images and rotated 10-20%, and in some cases we expanded the image size to 25%. The augmentation is done till we approach a near possible mixture of unbiased data. We perform a grayscale normalization to reduce the effect of illumination's differences. Moreover, the CNN converges faster on [0...1] data than on [0...255].
To improve the results, we used **Binary Cross Entropy**. This compares each of the predicted probabilities to actual class output which can be either 0 or 1. It then calculates the score that penalizes the probabilities based on the distance from the expected value. That means how close or far from the actual value/result (Khadivi.H, 2021).

We also used **ReduceLROnPlateau** class with patience 2, to improve the learning rate of our CNN algorithm. Models often benefit from reducing the learning rate by a factor of 2-10 once learning stagnates. This call-back monitors a quantity and if no improvement is seen for a 'patience' number of epochs, the learning rate is reduced. Also, with this CNN algorithm (Mathur.M, 2020), we are achieving 92.6% of accuracy (Mathur.M, 2020). Also, we are using a pre-trained model for medical image classification to implement 2nd algorithm. The pre-trained model is RADTorch (Elbanan.R, n.d.). Classifier trained on DICOM datasets.

RADTorch provides a framework of higher-level classes and functions that significantly reduce the time needed for implementation of different machine and deep learning algorithms on DICOM and non-DICOM medical images (DICOM, n.d.). RADTorch was built by radiologists for radiologists so they can build, test and implement state-of-the-art machine learning algorithms in minutes. It is User-Friendly as minimal coding required to implement, Comprehensive as it Includes data analysis and visualization, and backed by state-of-the-art machine learning and data visualization frameworks.

We choose these 2 ML models for their performance, reliability, accuracy and the speed that they work. We did not go for an ensemble ML algorithm because both these algorithms are trained on different datasets. By capturing their results individually, we can analyse them well. It is a CNN algorithm, so it gradually improves upon the usage and the RADTorch classifier can act as a baseline algorithm in future.

Benefits of the proposed solution to users and customers:

Our proposed solution will not replace the traditional existing mechanism of scanned images being analysed by radiologists. It acts as a layer in between the patients and radiologists. This algorithm gives the result of various diagnostic cases by analysing different scan types, because of its work rate in broad spectrum it can be a "one-size-fit for all" solution. This reduces the anxiety and longer wait times for patients to get the results from scanned images. It can be an additional help to radiologists to analyse various scanned images quickly. This is a mobile and web-based application, so this can be accessed by any patient/radiologist around the world.

Fairness:

AI should be extensively trained on images that are from all equal genders, races, backgrounds, etc, having even the slightest of bias can influence the accuracy of the returned diagnosis for a particular group or gender of person. This would be very unfair to the group that the AI is inaccurate for as it could potentially deprive them of the time saving diagnoses. There is a need to have an extremely high level of accuracy as the algorithm must work in the medical field. The datasets used to train the model must have all the features checked to ensure there is no bias in the training data which could result in the final model being biased or skewed. Similarly, after the algorithm is trained it must be extensively tested and not rushed out and implemented.

To address Fairness issues in our application, we used data augmentation techniques to get evenly distributed data. Also, we will taking peer-reviewed training datasets and feedback on developed model before implementing it. We are building it as a web and mobile based application. So, it can be used by everyone with an internet access.

Accountability:

In general, a radiologist who analyses the scans in the lab will be held accountable for any final diagnosis presented to the patient.
The results returned by the AI algorithms are to be presented to doctors before proceeding further. In case of any ambiguity the traditional approach of image analyses is to be done by radiologist. The application we developed here acts as a layer between the radiologists and patients. As this application can return faster and accurate results. It still needs to be validated by radiologists. Hence, the accountability is shared between the application developers and the validating radiologists.

Transparency:

All the data used for training will be made publicly available by changing the labels and renaming attributes. So, we won't be compromising the patient's confidentiality.
Also, the working of the algorithms is made transparent, and the improvements are regularly published. Public feedback and reviews are rigorously collected to improve the AI model. Patients will be informed that their diagnosis is being found by an AI and able to request a radiologist look over their results. Patients will be educated about AI issues in the analysis of scanned images. The hospitals or radiology labs must also inform the patient of the data they have collected on them and make it clear what data they have given to the AI algorithm and what information stays just with them in the hospital. (Farhud, Zokaei, 2021)

Ethics:

It is important that AI does not become the only thing used to read and analyse medical imaging and it is to be ensured that there are many trained professionals around for validating and overseeing the results and performance of the AI applications respectively. No AI algorithm should be relied upon completely without oversight. (Farhud, Zokaei, 2021)**.** There also must be a real person to run through potential diagnosis found by the medical imaging results with the person as these can be distressing and AI cannot empathise with people well and person to person contact is better off in this regard.
Thus, while developing our application we took various important measures to keep up the FATE standards. Also, by implementing the public review system there is a scope for gradual improvement of application.

Information Privacy Concerns:

Hospitals need to upgrade computer security systems to deal with increasing flows of data to and from the database the AI system is located on to ensure there are no data breaches. They must ensure they have information privacy and security protocols to regulate who can access the patient data and they need to ensure treating physicians can only access the information of the patients they are treating. The owner of the algorithm must limit the amount of people who have access to decryption keys to access the stored data of the algorithms data base and have these people go through numerous security checks. Need to ensure any medical images released to independent radiologists and professionals for quality checking do not have any personalised patient data attached to them and that patients consent to their medical images being used for quality checking. There should also be regular checking and auditing of the data and who has had access to ensure any potential data breaches have been found and patients at risk can be notified (Reuters, n.d.).

It important to ban and regulate the commercialisation of the data used in the AI so it cannot be sold to companies who have a financial interest in acquiring the healthcare data.
We implemented certain techniques to protect user data such as the aspects of patient data which identifies a patient but is not used to analyse the medical images, such as name, address is not uploaded to the AI algorithm database and hence never leaves the hospital or radiology lab and hence reduces the amount of people that have access to that information (Reuters, n.d.).
Also, the attribute labels are changed/renamed to protect patient data privacy.

Methodology:

Data Collection

The algorithm will use publicly available medical imaging datasets that have been uploaded from reputable sources such as the dataset released by the US National Institute of Health Clinical Centre which contains over 100,000 chest X-rays from over 30000 patients which will be a good start to train the model. (*NIH Clinical Centre Provides One of the Largest Publicly Available Chest X-ray Datasets to Scientific Community | National Institutes of Health (NIH)*, n.d.)
We are using multiple datasets in this application. Some of them are:

- Pneumonia related X-ray (Mathur M, 2020) datasets (NIH, 2017).
- lung cancer related CT scan datasets (NIH, 2018).
- Tumour related MRI scan datasets (Guluck, M. A, n.d.).
- DICOM datasets (DICOM, n.d.) that need to be trained on RADTorch classifier (RADTorch, n.d.).

Data Analysis:

To analyse the data, the data must be grouped by all possible biases, gender, age, race, etc to ensure there is an equal proportion of all people as well as medical images of all different parts of the body in all proportions to ensure there is no bias for any one group of people.
Also, each group of datasets are analysed individually and then EDA is performed to understand the distribution and any bias causing scanned images.

Prototyping the selected ML methods and algorithms:

We prototyped the ML algorithms in mobile and web-based application. We implemented different models based on the type of scan types and diagnostic types. We have collected the public datasets and DICOM datasets. But to improve the models we plan to collect the real time data from hospitals and medical laboratories. We will keep up the FATE and Information privacy principles while collecting and developing the data and ML models respectively. Before implementing the models and training the data. We plan to get it peer reviewed by some expert physicians/radiologists before being cleared for training the AI system. In order to account for any grading errors, the evaluation set was also checked by a third expert. The feedback is collected from the experts and then the model is improved.
This application can be used by patients, hospitals, laboratories and radiologists. The ease of access and simplicity of applications makes it usable in all the conditions.

Rationale for choosing these methods:

This is a web/mobile application, because this application should be accessible to everyone around the world. This application is useful in underdeveloped and remote countries. We implemented CNN algorithm, which improves itself after using it more and more. So, the improvement in efficiency and performance is guaranteed. The process and application are simple. So, it can be understood and used by any person with minimal technical knowledge. Also, we used 2 different algorithms.

Evaluating the solutions:

The applications return the overall result of the diagnosis as positive or negative based on the results from both the algorithms.

- If the result from both the algorithms are positives, then we can assume that the patient is diagnosed with the selected form of illness.
- If the result from both algorithms are negatives, then we can assume that the patient is not diagnosed with the selected form of illness.
- If the result from one algorithm is positive and another algorithm is negative. It raises ambiguity in the result. In these cases, the patients need to wait and hear their diagnostic analysis from the radiologists.

The CNN based algorithm with X-ray datasets in related to pneumonia gave 92.6% accuracy levels (Mathur M, 2020). In a similar manner other ML models that we trained with various datasets gave over 85% of accuracy. Also, the RADTorch classifier trained by using DICOM dataset covers broader scale and achieved an accuracy of 84%.

Design prototype (no-code):

The below attached snapshot is the design prototype of our application. User can select the type of scan and type of diagnosis from the drop-down menu. User can upload the scanned image from the upload scanned image button. The results from the Algorithm-1 and algorithm-2 are displayed as positive/negative.

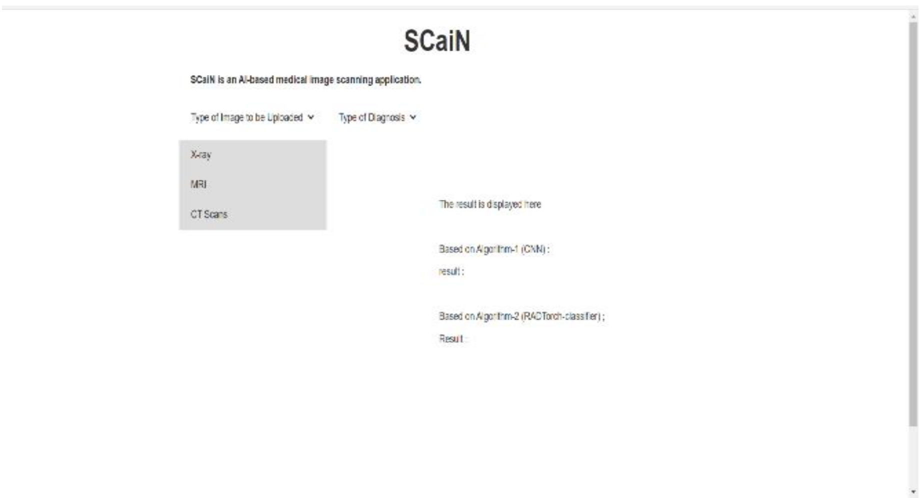

Fig.2 Application interface displaying type of scan

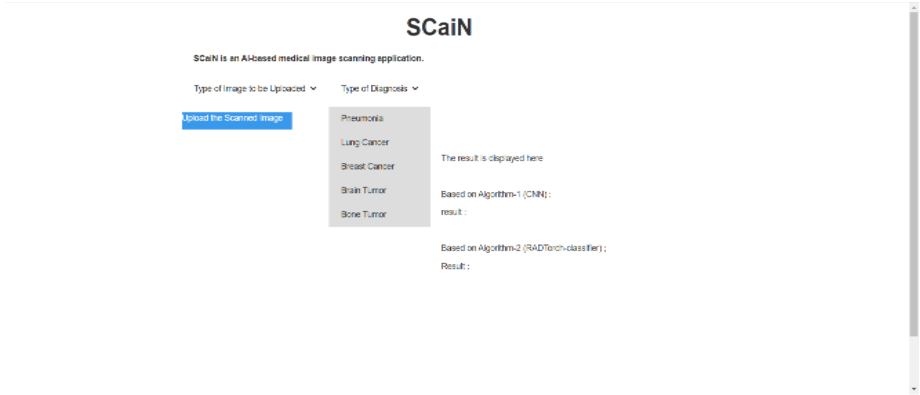

Fig.3 Application interface displaying type of illness

Working:

As an example, we consider the X-ray type and Pneumonia for diagnosis. We passed scanned image.

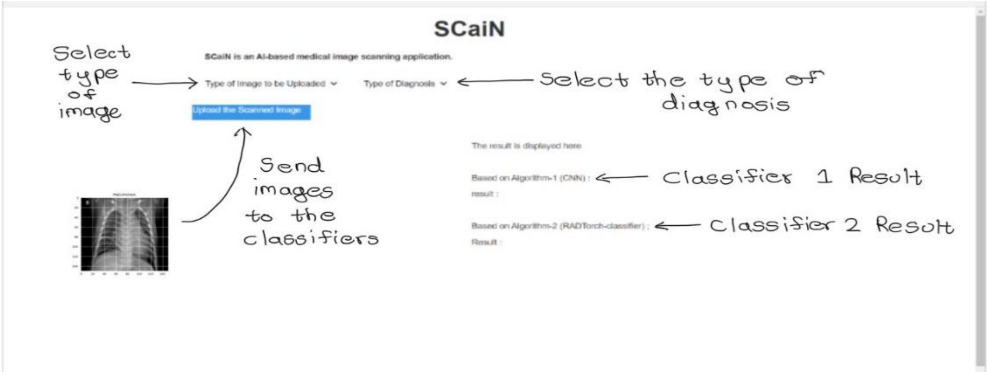

Fig.4 application work example

After uploading the above image, both the algorithms, i.e., CNN (Khadivi H, 2021) and RADtorch classifier (Elbanan.R, n.d.) analyses the image for some time and then returns the result message.

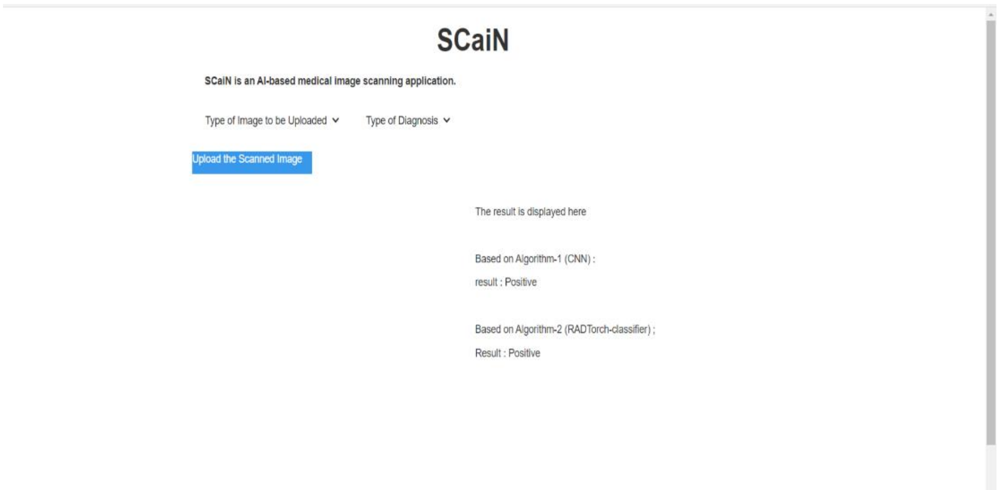

Fig.5 result displayed by the application

As we passed a positive pneumonia scanned image, we got the result as positive in both the algorithms. So, the predicted and actual diagnostic result of the scanned image is same.

Conclusion:

Autonomous AI system can never replace the experienced professionals. This is especially true in medical domain. But an AI system can always improve the existing system. Our application SCaiN acts a layer in between patients and radiologists. It analyses the scanned images quickly and user gets the results instantly. It also reduces burden over radiologist to analyse all scanned images in a short time. In this application, we used 2 different AI algorithms to compensate the various important aspects of both algorithms. This shows that our application valued accuracy over the cost of computation. As we used CNN algorithm in this application, the improvement in performance of algorithm is expected.

References:

1. *A DICOM dataset for evaluation of medical image de-identification (Pseudo-PHI-DICOM-Data) - The Cancer Imaging Archive (TCIA) Public Access - Cancer Imaging Archive Wiki*. (n.d.). DICOM. https://wiki.cancerimagingarchive.net/pages/viewpage.action?pageId=80969777
2. Driver, M. (n.d.). *AI has Huge Potential to Address the Crisis in Medical Imaging* -. Journal of Mhealth. https://thejournalofmhealth.com/ai-has-huge-potential-to-address-the-crisis-in-medical-imaging/
3. Elbanan, R. (n.d.). *GitHub - radtorch/radtorch: Official repository for RADTorch - The Medical Imaging Machine Learning Framework*. GitHub. https://github.com/radtorch/radtorch
4. Farhud, D. D., & Zokaei, S. (2021). Ethical Issues of Artificial Intelligence in Medicine and Healthcare. *Iranian Journal of Public Health*. https://doi.org/10.18502/ijph.v50i11.7600
5. Guluck, M. A. (n.d.). *Classification learning*. Openfmri. https://openfmri.org/dataset/ds000002/
6. *Home - RADTorch*. (n.d.). RADTorch. https://www.radtorch.com/
7. Hricak, H., Abdel-Wahab, M., Atun, R., Lette, M. M., Paez, D., Brink, J. A., Donoso-Bach, L., Frija, G., Hierath, M., Holmberg, O., Khong, P. L., Lewis, J. S., McGinty, G., Oyen, W. J. G., Shulman, L. N., Ward, Z. J., & Scott, A. M. (2021). Medical imaging and nuclear medicine: a Lancet Oncology Commission. *The Lancet Oncology*, *22*(4), e136–e172. https://doi.org/10.1016/s1470-2045(20)30751-8
8. Khadivi, H. (2021, October 1). *Medical Diagnosis with CNN& Transfer Learning*. Kaggle. https://www.kaggle.com/code/homayoonkhadivi/medical-diagnosis-with-cnn-transfer-learning/data
9. Long, M. (2020, April 19). *Radiologist Shortage - How will AI alleviate this growing challenge?* Aidoc. https://www.aidoc.com/blog/is-radiologist-shortage-real/
10. Malek, L. A., Jain, P., Johnson, J., Malek, L. A., Jain, P., & Johnson, J. (2022, March 17). *Data privacy and artificial intelligence in health care*. Reuters. https://www.reuters.com/legal/litigation/data-privacy-artificial-intelligence-health-care-2022-03-17/
11. Mathur, M. (2020, July 3). *Pneumonia Detection using CNN (92.6% Accuracy)*. Kaggle. https://www.kaggle.com/code/madz2000/pneumonia-detection-using-cnn-92-6-accuracy#Data-Visualization-&-Preprocessing
12. Mitchell, C. (2021). *World Radiography Day: Two-Thirds of the World's Population has no Access to Diagnostic Imaging*. Pan American Health Organization / World Health Organization. https://www3.paho.org/hq/index.php?option=com_content&view=article&id=7410:2012-dia-radiografia-dos-tercios-poblacion-mundial-no-tiene-acceso-diagnostico-imagen&Itemid=1926&lang=en

13. *NIH Clinical Center provides one of the largest publicly available*. (2017, September 27). National Institutes of Health (NIH). https://www.nih.gov/news-events/news-releases/nih-clinical-center-provides-one-largest-publicly-available-chest-x-ray-datasets-scientific-community
14. *NIH Clinical Center releases dataset of 32,000 CT images*. (2018, July 20). National Institutes of Health (NIH). https://www.nih.gov/news-events/news-releases/nih-clinical-center-releases-dataset-32000-ct-images
15. *Population by broad age group projected to 2100*. (2021). Our World in Data. https://ourworldindata.org/grapher/population-by-age-group-to-2100?country=%7EOWID_WRL
16. *Radiologist census underlines ongoing toll of workforce shortages | The Royal College of Radiologists*. (n.d.). Rcr. https://www.rcr.ac.uk/posts/radiologist-census-underlines-ongoing-toll-workforce-shortages-0
17. *The Radiology Labor Shortage*. (n.d.). American College of Radiology. https://www.acr.org/Practice-Management-Quality-Informatics/ACR-Bulletin/Articles/March-2022/The-Radiology-Labor-Shortage
18. *Statistics and Graphs | Division of Cancer Control and Population Sciences (DCCPS)*. (n.d.). Cancercontrol. https://cancercontrol.cancer.gov/ocs/statistics#graphs.