Design of Computer-Aided-Diagnosis (CAD) for Self-Assessment Tuberculosis in Indonesia

Faisal Asadi1*, Joko Pebrianto Trinugroho1, and Bens Pardamean2

1Bioinformatics and Data Science Research Center, 11480 Bina Nusantara University, Indonesia
2Computer Science Department BINUS Graduate Program, Master of Computer Science, 11480 Bina Nusantara University, Indonesia

Abstract. Tuberculosis (TB) is one of the highest causes of death in Indonesia. The main reason is lack of the health facilities. Computer-aided diagnosis (CAD) is a tool for early treatment and screening of many diseases, including TB. This paper proposed a design of a CAD system in Indonesia specifically for TB. The design gives the analysis of self-assessment concepts, use-case diagrams, and black-box diagrams. The black box utilizes chest x-ray (CXR) data for the medical image processing (MIP) method, and artificial intelligence (AI) for classification and visualization of the TB. This CAD design of self-assessment of TB has a capability to help the health practitioners read and interpret the diagnosis result more easily.

1 Introduction

In 2021, The Ministry of Health the Republic of Indonesia reported that estimated number of TB cases in Indonesia is 842,000 with 393,323 notification TB, 7,921 cases confirmed TB MDR (multidrug resistance), and 4,590 cases with enrolled TB MDR. From the total cases of TB, the number of death is still high is 13,110 cases, the successful of treatment is only 83% of total cases ([1-3]). Besides, the main problem of the treatment of TB is lack of the health facilities, such as the laboratory for complete inspection is not distributed well in Indonesia. In addition, Indonesia has many islands and rural areas with some regions are only accessible remotely. Based on the epidemiology data and lack of health facilities, we observed that the treatment of TB is still a high priority and needs the government's attention [4-6].

AI is one of the several technologies that have capabilities of answering the challenges and significantly impacting the economic, transportation, government, public policy, and health sector [7-13]. Internet of things (IoT) completed by several fields such as data science (DS), machine learning (ML), and deep learning (DL) have a great potential for serving human tasks [14-16]. The remarkable thing detailed in the application of AI in medicine has been successfully implemented, such as telehealth system, telemedicine system, medical imaging system, genomic sequence system, and (CAD) system [17-21].

The application of AI in medicine has covered many areas such as decision making, expert system, classification of the disease, and the detection system of diseases ([9], [22]). Besides, some research on AI in medicine and digital health (AIDH) has successfully been implemented in public policymaking, for example the public policy during the pandemic of coronavirus disease (COVID-19). Some research successfully built up using tabular data, images data, and medical record data ([23-24]).

Related to the main problem of TB, this research proposes a design of a CAD system for TB for self-assessment in Indonesia. The design utilizes the concepts of AI and MIP techniques. Based on the concept, CAD is the application of information and communication technology (ICT) in medicine, which helps the health practitioners to read and interpret the diagnosis result [25-28]. The design of CAD’s system for TB was elaborated in this research. The novelty of this research is giving solutions in the early treatment and screening of TB in Indonesia.

2 Literature review

2.1 Tuberculosis (TB)

TB is an infectious disease with the highest causes of death worldwide [2, 3]. Referring to WHO, TB cases significantly increase every year [2]. In general, TB treatment must pass gold standards and procedures on diagnosing and judging if the patient is TB or non-TB. This gold standard is from the international standard for tuberculosis care (ISTC2014) [29]. The first procedure is the primary assessment of symptoms, followed by radiography evaluation using CXR data [30]. The last procedure is sputum smear checking, and the specimens were then submitted to the laboratory for checking Xpert MTB/RIF. [29]. Concerning the gold standard of TB diagnosis, treating TB requires coordination with local...
public health services or other agencies. Early treatment is necessary for the successful treatment of TB patients ([3], [31]). Moreover, a manual assessment of TB disease from the practitioners could lead to a false diagnosis ([32-34]), because it is difficult to read and interpret the CXR data to support the diagnosis. This problem is highlighted in this research. Furthermore, the other urgency of this research is the limit of the number of health practitioners. Even general doctors, pulmonologists, are not well distributed in Indonesia ([5], [35-37]).

2.2 Computer-aided diagnosis (CAD) for TB

The previous research related to AI and CAD system for TB was conducted by [38], which used AI in CXR’s data for the early screening of TB. The accuracy of the DL algorithm has achieved the area under curve (AUC) is 0.92 with confidential interval (CI) from 0.91-0.94 for abnormal scans detection. For detection of blunted costophrenic angle and cardiomegaly got the AUC 0.96 with CI = 0.94-0.98. Cavity the AUC 0.95 with CI = 0.87-1, consolidation the AUC 0.95 with CI = 0.92-0.98, fibrosis the AUC 0.93 with CI = 0.90-0.96, hilar enlargement the AUC 0.89 with CI = 0.83-0.94, nodule the AUC 0.91 with CI = 0.87-0.96, opacity the AUC 0.94 with CI = 0.93-0.96, pleural effusion the AUC 0.98 with CI = 0.97-1. The research using the total data of 2000 data sets. The same research, from [39]. Evaluated twelve CAD systems, and six were used by xpert readers such as Qure.ai software, DeepTek software, Delft Imaging software, JF Healthcare software, OXIPIT software, and the last is Lunit Insight software. The research showed that the AUC of those systems had given the best result and performance to obtaining the TB scores. Table 1 provides a comparison of CADs software with the AUC and, precision-recall (ROC).

Table 1. Receiver operating characteristic (ROC) - area under curve (AUC), and precision-recall (RC) with CI = 95% [39].

<table>
<thead>
<tr>
<th>Vendor/CAD Software</th>
<th>ROC-AUC (CI = 95%)</th>
<th>PR-AUC (CI = 95%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qure.ai software</td>
<td>0.82 (0.76-0.86)</td>
<td>0.41 (0.33-0.50)</td>
</tr>
<tr>
<td>Delft Imaging software</td>
<td>0.82 (0.78-0.85)</td>
<td>0.39 (0.31-0.47)</td>
</tr>
<tr>
<td>DeepTek software</td>
<td>0.78 (0.73-0.82)</td>
<td>0.28 (0.22-0.34)</td>
</tr>
</tbody>
</table>

The other research related to the CAD for TB is from [40]. This research compared three DL systems of CAD TB, which are CAD4TB from the Delft-Netherland, Lunit INSIGHT from South Korea, and qure.ai (qXR) from India. The AUC of the three systems was similar, those are for Lunit INSIGHT getting the AUC = 0.94 and CI = 0.93-0.96, for qXR getting the AUC = 0.94, and CI = 0.92-0.97, and the last is CAD4TB getting the AUC = 0.92 and CI = 0.90-0.95. Utilizing DL systems to read the CXR data has capability to reduce the number of xpert while tests maintaining sensitivity at 95% or above [40]. Similar research for treating TB in Indonesia was conducted by [41], which utilize the CAD for reading the TB screening for people with comorbid disease diabetes mellitus (DM). The research found the confirmation of CAD4TB with the radiologist and found the AUC for CAD4TB was 0.89 with CI 0.73-1.00.

3 Methodology

The methodology of this research started from the literature review. This part is the literature review, which describe the data of TB in Indonesia, lack of health facilities problems, and the other research related to the design of CAD for TB. The second part of this research is to define the problem of the TB and relate it to the CAD for TB using CXR data, and then will study and design the self-assessment of the CAD for TB as a solution for treating TB in Indonesia. The third part is the study and design, elaborating the use case diagram, block box diagram, and the detail of the black box CAD system. Moreover, the methodology part focuses on designing self-assessment CAD for TB using the CXR data as a modal for the input to the system. The methodology will be shown in Figure 1.

![Fig. 1. Research methodology.](image)

4 Results and discussion

CAD system is the medical device and application of AI for treating patients, including reading and interpreting the disease [19, 26, 28, 42]. The CAD’s system can solve medical problems such as interpreting the CXR data.
image for diagnosis of lung disease, mammography for breast abnormalities diagnosis [43, 44], and analysis and detection of brain cancer [45-47]. WHO has legalized and recommended the CAD system for early treatment and diagnosis of TB. Figure 2. compares the CAD system and manual reading interpretation for TB based on threshold [29, 31].

**Fig. 2.** CAD reading and interpretation (based on a threshold).

For the general procedure of CAD in reading and interpretations, 1-2 sputum samples should be collected from prospectively selected participants for the testing, performed according to standard local operating protocols. The specified reference standard test results were done at the first consultation and should be collected from the selected data source for retrospective case-control studies [31].

Besides, several CAD systems show the same result in terms of accuracy compared to human expert readers, but the CAD system is considerably better than the beginner reader. Table 2. compares expert readers and pieces of CAD systems for TB. The pieces of the CAD systems were compared Qure.ai software, Delf Imaging CAD4TB software, Lunit INSIGHT software, DeepTek software [39].

**Table 2.** Comparison CAD systems for tuberculosis disease [39].

<table>
<thead>
<tr>
<th>CAD System</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Accuracy (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Expert Interpretation</td>
<td>95.5% (90.4-98.3)</td>
<td>42.2% (38.8-45.5)</td>
<td>49.0% (43.5-52.1)</td>
</tr>
</tbody>
</table>

Abnormality scores obtained by FIT

<table>
<thead>
<tr>
<th>CAD System</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Accuracy (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Qure.ai software</td>
<td>95.5% (90.4-98.3)</td>
<td>48.7% (45.1-52.0)</td>
<td>54.7% (51.7-57.8)</td>
</tr>
<tr>
<td>DeepTek software</td>
<td>95.5% (90.4-98.3)</td>
<td>46.8% (43.6-49.6)</td>
<td>52.6% (49.5-55.7)</td>
</tr>
<tr>
<td>Delf Imaging software</td>
<td>95.5% (90.4-98.3)</td>
<td>45.3% (42.6-48.6)</td>
<td>51.7% (48.1-54.8)</td>
</tr>
</tbody>
</table>

Abnormality scores provided by CAD company

<table>
<thead>
<tr>
<th>CAD System</th>
<th>Sensitivity (95% CI)</th>
<th>Specificity (95% CI)</th>
<th>Accuracy (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>JF Health care software</td>
<td>95.5% (90.4-98.3)</td>
<td>41.6% (37.1-44.9)</td>
<td>48.1% (43.6-51.2)</td>
</tr>
<tr>
<td>OSEFT software</td>
<td>95.5% (90.4-98.3)</td>
<td>40.5% (37.6-44.1)</td>
<td>47.9% (44.8-51.0)</td>
</tr>
<tr>
<td>Lunit</td>
<td>95.5% (90.4-98.3)</td>
<td>38.7% (35.1-42.0)</td>
<td>46.0% (43.0-49.1)</td>
</tr>
<tr>
<td>InferVision</td>
<td>95.5% (90.4-98.3)</td>
<td>26.5% (23.6-29.3)</td>
<td>35.4% (32.3-38.4)</td>
</tr>
</tbody>
</table>

**4.1 Use-case diagram**

The CAD system is designed for medical practitioners such as general doctors, radiologists, pulmonologists, and the patients to assess TB’s disease. The users of the CAD system are the medical practitioners and the patient with the modal in CXR’s data. The CAD system will process data in several formats, such as digital imaging and communication in medicine (DICOM), the joint photographic expert group (JPEG), portable network graphics (PNG). The flow of the CAD’s system can be used by the medical practitioners starting from accessing the CAD’s system, registering the patients, and saving the patient’s data. After the patient’s data is registered, the next step is to input the CXR patient’s data to the CAD’s system with the data have been annotated and positive of TB. In this case, the data is inputted manually by the expert. The next step of the workflow is self-assessment with the CAD system by the medical practitioners, the patient is then assessed by the radiologist observing the self-assessment result. In this step, the result of the system shows the lesion area of TB from CXR data. After finishing the self-assessment by the radiologist, the last step is saving the patient’s data into the CAD system. The design CAD system for self-assessment of TB is shown in Figure 4.

**Fig. 3.** Use-case diagram of CAD system.

**Fig. 4.** The design of CAD for self-assessment of TB.
4.2 Black box of CAD for TB

The black box is the core of the CAD for TB system that does the whole data processing starting from the medical practitioners inputting the image to the system and processing the image data. Moreover, the black box concept of the CAD system for TB shown in Figure 5.

![Black box for CAD system.](image)

Fig. 5. Black box for CAD system.

The design of CAD for TB system using the algorithm of MIP combined with AI. The most of algorithm of the black box of CAD for TB systems is ML and DL. Several algorithms of the black box of the CAD system that exists that have been approved by food and drug administration (FDA) and CE mark certification shown in Table 3.

<table>
<thead>
<tr>
<th>CAD for TB</th>
<th>Classification Algorithm</th>
<th>Certification</th>
<th>Commercially Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>LUNIT Insight</td>
<td>Artificial Neural Network, Deep Learning</td>
<td>CE</td>
<td>Yes</td>
</tr>
<tr>
<td>CAD4TB</td>
<td>Artificial Neural Network, Deep Learning</td>
<td>CE</td>
<td>Yes</td>
</tr>
<tr>
<td>Qure.ai</td>
<td>Artificial Neural Network, Deep Learning</td>
<td>CE</td>
<td>Yes</td>
</tr>
<tr>
<td>CAD4TB Delft Imaging</td>
<td>Artificial Neural Network, Deep Learning</td>
<td>CE</td>
<td>Yes</td>
</tr>
</tbody>
</table>

Table 3. CAD for TB (FDA and CE mark certification).

Figure 6. is the detail of the black box design for the CAD system, this part will be detailed for the MIP concept using several techniques and algorithms for heatmap color and classification using ML or DL algorithms. The black box CAD design is the core processor of the machine for executing the data (48-49). The research data is CXR image. Previous research shows that most methods are used for the CXR images extraction and the result classification [40, 50].

![Detail of the design of black box CAD for TB.](image)

Fig. 6. Detail of the design of black box CAD for TB.

5 Conclusion

Due to the increasing number of TB cases in Indonesia throughout the year and lack of the health facilities, this paper analyzes CAD design for TB to treat and early diagnose TB in Indonesia. Several CAD systems for TB have been approved by FDA and CE mark certification, such as LUNIT Insight, Qure.ai, CAD4TB Delft Imaging, and DeepTek. When we look at other low-middle income countries (LMICs) countries such as Africa, Zimbabwe, Mongolia, and Ghana, the CAD system for TB has been implemented. Furthermore, especially in Indonesia, as a part of LMICs, the CAD system is appropriate for Indonesia as an early treatment for screening and diagnosis tool for TB. Moreover, the possibility of future research is developing the CAD design system as the pioneer of CAD for tuberculosis in Indonesia with the coordination and collaboration of the ministry of health, the CAD system for TB is a possible solution for treating and handling TB cases in Indonesia.

References

4. G. Fund, Current status of integrated community based TB service delivery and the global fund work plan to find missing TB cases, Indonesia National TB Program (n.d.)
6. A. Noviyani, T. Nopsonop, K. P. Id, Variation of tuberculosis prevalence across diagnostic approaches and geographical areas of Indonesia, PLoS ONE 9, pp. 1–12 (2021)
importance in classification models for colorectal cancer cases phenotype in Indonesia, Procedia Computer Science 157, pp. 313–320 https://doi.org/10.1016/j.procs.2019.08.172 (2019)


12. J. D. Morgenstern, L. C. Rosella, M. J. Daley, V. Goel, H. J. Schünemann, T. Pigott, “AI’s gonna have an impact on everything in society, so it has to have an impact on public health”: a fundamental qualitative descriptive study of the implications of artificial intelligence for public health, BMC Public Health 21, 1, pp. 1–14 (2021)


18. N. Dominic, Daniel, T. W. Cenggoro, A. Budiarto, B. Pardamean, Transfer learning using inception-resnet-v2 model to the augmented neuroimages data for autism spectrum disorder classification, Communications in Mathematical Biology and Neuroscience, pp. 1–21 (2021)


22. Y. Zhang, Y. Weng, J. Lund, Applications of explainable artificial intelligence in diagnosis and surgery, MDPI-Diagnostics 12, 2, pp. 1–18 https://doi.org/10.3390/diagnostics12020237 (2022)


31. World Health Organization (WHO), A toolkit to support the effective use of CAD for TB screening (World Health Organization, Geneva, 2021c)


