A Systematic Review of Deep Learning Techniques for Tuberculosis Detection From Chest Radiograph

Mustapha Oloko-Oba and Serestina Viriri*

Computer Science Discipline, School of Mathematics, Statistics and Computer Science, University of KwaZulu-Natal, Durban, South Africa

The high mortality rate in Tuberculosis (TB) burden regions has increased significantly in the last decades. Despite the possibility of treatment for TB, high burden regions still suffer inadequate screening tools, which result in diagnostic delay and misdiagnosis. These challenges have led to the development of Computer-Aided Diagnostic (CAD) system to detect TB automatically. There are several ways of screening for TB, but Chest X-Ray (CXR) is more prominent and recommended due to its high sensitivity in detecting lung abnormalities. This paper presents the results of a systematic review based on PRISMA procedures that investigate state-of-the-art Deep Learning techniques for screening pulmonary abnormalities related to TB. The systematic review was conducted using an extensive selection of scientific databases as reference sources that grant access to distinctive articles in the field. Four scientific databases were searched to retrieve related articles. Inclusion and exclusion criteria were defined and applied to each article to determine those included in the study. Out of the 489 articles retrieved, 62 were included. Based on the findings in this review, we conclude that CAD systems are promising in tackling the challenges of the TB epidemic and made recommendations for improvement in future studies.

Keywords: tuberculosis, chest radiograph, computer-aided diagnosis, deep learning, systematic review

INTRODUCTION

Tuberculosis (TB) is ranked among the leading causes of death. About 10 million persons fell ill globally from TB infections in 2019 (1). TB is triggered by the Mycobacterium bacteria that usually affect the lungs (pulmonary) but sometimes affect other parts of the body (extrapulmonary) (2). Many TB patients lose their lives yearly due to diagnostic delay, misdiagnosis, and lack of appropriate treatments (3, 4). Although TB is a global challenge, the mortality rate is more prevalent in low and middle-income nations (5).

TB is certainly treatable if diagnosed early for appropriate treatment. Early diagnosis is essential for successful treatment, preventing further spread, and significantly reducing the mortality rate in line with the World Health Organization (WHO) End TB Strategy (1). The gold standard for TB screening is Sputum culture. However, posterior-anterior chest radiographs (CXR) are an effective technique with low-cost and moderately low radiation doses for screening lung abnormalities to achieve prompt results (6). CXR has been adequately employed in developed countries to analyze individuals exhibiting active TB symptoms. At the same time, its application is limited in developing countries where TB is most prevalent (7, 8). High TB burden regions lack the skilled and radiological expertise required to interpret CXR images adequately (9, 10).
In the last decades, several efforts have been made using Artificial Intelligence (AI) to develop a Computer-Aided Detection (CAD) system to advance automatic object/image recognition tasks and overcome the challenges of a skilled workforce. Machine Learning (ML) and Deep Learning (DL) are the predominant AI techniques employed to develop CAD systems for analyzing CXR images. Both techniques have had a significant impact, but the DL approach, such as Convolutional Neural Network (CNN), has become more prominent for analyzing different pulmonary abnormalities in the medical domain, most importantly in diagnosing TB. The application of an efficient classification tool is vital for improving the quality of diagnosis while reducing the time taken to analyze a large volume of CXRs (11). This endeavor is to achieve the global decline in TB incidence to about 5% annually compared to the current 2% yearly as part of the World Health Organization strategy to end TB (1).

The contribution of this systematic review is to present an extensive summary of the various state-of-the-art CAD system proposed in the literature for the classification of TB. Ultimately, only the CAD system developed using Deep Learning models is considered in this study detailing the diagnostic accuracy between 2017 and 2021. The rest of the paper is structured as follows: section Methodology presents the study methodology. The results are presented in section Results. Section Discussion presents the discussion, while the conclusion is expressed in section Conclusion and Recommendations.

**METHODOLOGY**

This systematic review aims to establish various CAD systems related to Tuberculosis diagnosis from CXR using DL techniques. The study followed the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) procedures (12) to identify the standards for inclusion and exclusion, as shown in Figure 1. These standards were formulated based on the present study objectives and the research questions. All articles that satisfied the following conditions were included and excluded if otherwise:

- Articles that considered only pulmonary tuberculosis disease.
- Employed at least one deep learning technique as a classifier
- CXR is the only medical imagine for screening tuberculosis.
- Articles published between January 2017 and September 2021.
- Articles are entirely written in English.
- Articles must be full text. All others, such as abstract, preprints are excluded.

**Search Strategy**

The search strategy was developed to identify relevant published articles in Scopus (https://www.scopus.com/), IEEEExplore (https://ieeexplore.ieee.org/), Web of Science (www.webofscience.com), and PubMed (https://www.ncbi.nlm.nih.gov/pubmed/). Searches were performed across the four databases using the following search keywords: “tuberculosis,” “chest x-ray,” “classification,” “artificial intelligence,” “computer-aided diagnosis,” “deep learning.” These keywords were used to

![FIGURE 1 | The PRISMA structure for the study selection process.](image-url)
form Boolean search strings according to searching standards on different databases. The configuration of keywords used to retrieve all relevant articles from each database search engine is presented in Table 1.

### Study Selection

A total of 488 articles were retrieved from the initial search on all the databases. The next step then scans the article topics and keywords to identify highly related papers from the irrelevant ones. This step is followed by the overview reading of the abstract, methods, and Conclusion to further screen for relevant papers for full-text reading and better understanding. Thus, 209 articles were considered for the title and abstract screening, out of which 96 papers were found suitable for full-text reading, and a total of 62 were finally included in the analysis after removing 34 duplicates. The detailed structure for the study selection is presented in Figures 1, 2 shows the numbers of articles included per year and categories. It is necessary to note that some relevant articles might have been unintentionally omitted.

### Data Extraction

The details extracted from each article are presented in Tables 2–5. Each table represents the search results from the Scopus, IEEE Xplore, PubMed, and Web of Science databases. Data extracted include study aim and scope, computational techniques (DL models), CXR datasets, evaluation metrics/results achieved, and publication year. The database search shows that most IEEE Xplore articles are conference proceedings and included in this systematic review, along with a few conference articles from springer and ACM. This is due to the articles’ quality and contribution to the subject matter.

### RESULTS

This study reviewed computer-aided diagnosis systems articles to detect pulmonary TB from January 2017 until September 2021. It was observed from the articles that the development of CAD follows a standard framework involving four steps as follows: “pre-processing” is the first step which deals with cleaning up the CXR images by eliminating noise and enhancing for clarity. The second step is “segmentation” of a region of interest from the entire image, which is the lung field region in the case of CXR. The third step is the “Feature extraction,” where discriminative features are identified and selected for further analysis in the “classification” step, where the various images are categorized as normal or abnormal (infected) with TB. Several techniques such as handcraft, machine learning, and deep learning have been employed to diagnose TB, but DL has recorded more success in this regard; hence our interest was to analyze the CAD system based on one or more DL techniques as the classifier for TB detection.

The descriptive analysis of the results is presented in Tables 2–5. These tables show the computational technique, study scope, datasets, evaluation criteria, and results. Articles that utilized CXR as the only imaging modality and employed DL as the only computational technique for developing CAD are considered.

### Imaging Modalities

Different screening procedures are used to confirm the presence of TB. Still, Chest Radiograph, otherwise referred to as Chest X-ray (CXR), is a radiograph tool used to detect abnormalities in the lungs and nearby structures. In recent years, CXR has remained a vital method for screening TB and other lung diseases and hence recommended by WHO due to its high sensitivity, wide availability, and relatively less expensive (70, 71).

### Deep Learning Techniques

Many DL techniques have been used for screening, predicting, and diagnosing TB. In most classification algorithms, a dataset is required for training and testing with many samples of inputs and outputs to learn from. Model is developed using the training set to calculate how best to map examples of input data to a specific class label; then, the model validation is accomplished using the test set (72).

In Tables 2–5, only the results obtained from the test set are extracted. Some studies employed more than one DL technique to find the optimal results for diagnosing TB diseases. Only the best results are documented in this study if more than one accuracy is reported in an article.

### Datasets

Data is crucial for developing CAD required to solve life-threatening diseases, including TB, as one of the leading causes of worldwide death. The various popular datasets that have been used in developing TB detection algorithms contain de-identified CXR images to protect the privacy of the patients. In other words, the identity of the patients is not disclosed. Most of the datasets are accompanied by radiological interpretation of the observed manifestation that can serve as groundtruth. These datasets are made available to the research communities to foster state-of-the-art research into finding lasting solutions to the early diagnosis of TB manifestations. Some of these public datasets include Montgomery County (73), Shenzhen (73), Peruvian (49), KIT (74), MIMIC-CXR (75), Belarus, NIH (chest x-ray 8, 14) (76), JSRT. Figure 3 shows the datasets frequency of use.
**Evaluation Metrics**

Once a model is trained using some training images, the test dataset is then employed to assess the quality of the model. It is evident from the data extraction process that different evaluation metrics exist for assessing model performance. Most of the popular evaluation metrics applied to the development of CAD system includes accuracy, sensitivity, specificity, and AUC. These evaluation metrics are briefly explained as follows:

Accuracy is the rate of the correct samples from the total number of samples examined (77). The equation gives the accuracy:

\[
\text{Accuracy} = \frac{TP + TN}{TP + FP + TN + FN}
\]

Sensitivity is the proportion of the actual positive samples that are correctly identified as positive (78). This metric is given as:

\[
\text{Sensitivity} = \frac{TP}{TP + FN}
\]

Specificity is the ratio of the actual negative samples that are correctly identified as negative (78). This metric is given as:

\[
\text{Specificity} = \frac{TN}{TN + FP}
\]

Precision, otherwise known as a positive predictive value, is the ratio of positive samples that are accurately predicted (77). This is given as:

\[
\text{Precision} = \frac{TP}{TP + FP}
\]

AUC-ROC Curve: sometimes written as AUROC, specify the rate at which a model distinguishes between classes [84]. It tells the probability of separating negative samples from positive samples. Thresholds are set to determine the ROC curve, which is the plot of True Positive Rate (TPR) against False Positive Rate (FPR) given as:

\[
\text{TPR} = \text{Sensitivity} = \frac{TP}{TP + FN}
\]

\[
\text{FPR} = 1 - \text{Specificity} = \frac{FP}{FP + TN}
\]

Where TP, TN, FP, FN, TPR, and FPR are true positive, true negative, false positive, false negative, true positive rate, and false positive rate.

**DISCUSSION**

This systematic review extensively searched the various Deep Learning classifiers for diagnosing TB from CXR. Automatic detection of TB has received mammoth attention in the last decades resulting in many publications with state-of-the-art techniques. Figure 4 presents a hierarchical chart of computational techniques categorized according to the frequency of usage in CAD systems for TB based on the included articles. Despite some good accuracies reported in some studies, we found some limitations in the existing studies concerning methods and reported accuracy, which should be a point of consideration for developing CAD systems in the future. Many studies measured diagnostic accuracy without evaluating the risk of bias emerging from the datasets that were used. It is essential that the accuracy...
### TABLE 2 | Scopus search results.

| References | Computational techniques | Study aim/scope | Datasets | Results | Year |
|------------|--------------------------|-----------------|----------|---------|------|
| Duong et al. (13) | EfficientNet, Vision Transformer | Classification of CXR as normal, pneumonia and TB | Montgomery, Shenzhen, Belarus, RSNA, A COVID-19 CXR | Acc = 97.72%, Auc = 100% | 2021 |
| Norval et al. (14) | AlexNet, VGG16, and VGG19 | Improve accuracy and Classification of CXR as Has TB and No TB | Montgomery, Shenzhen, NIH | Acc = 89.99% | 2021 |
| Rahman et al. (15) | ResNet101, VGG19, and DenseNet201 XGBoost | The study utilizes three Deep CNN models as features extractor then classify using xExtreme Gradient Boosting | Montgomery, Shenzhen, Belarus, RSNA, Private | Acc = 99.92%, Auc = 99.93%, Pre = 99.85%, Sen = 100%, Spe = 99.85%, F1 = 99.92% | 2021 |
| Govindarajan and Swaminathan (16) | ELM, OSELM | Identify and classify TB conditions from healthy subjects in chest radiographs using integrated local feature descriptors and variants of extreme learning machine | Montgomery | Acc = 99.2%, Sen = 99.3%, Spe = 99.3%, Pre = 99.0% | 2021 |
| Alawi et al. (17) | CNN | Study proposed automated technique to diagnose TB from CXR | NLM, Belarus, NIAID TB and RSNA | Acc = 98.71%, Sen = 98.86%, Spe = 98.57% | 2021 |
| Khatibi et al. (18) | Complex networks and stacked ensemble (CNISE), CNN | A multi-instance classification model to detect TB from CXR is proposed | Shenzhen, Montgomery | Acc = 99.26%, Auc = 99.00% | 2021 |
| Ayaz et al. (19) | Deep CNN | Present a hybrid method for TB detection | Montgomery, Shenzhen | Auc = 0.99% | 2021 |
| Priya and Virdma (20) | VGG-19, ResNet50, DenseNet121, InceptionV3 | The study employs transfer learning for TB diagnosis. | Montgomery, Shenzhen | Auc = 0.95% | 2021 |
| Dasanayaka and Dissanayake (21) | DC-GAN, VGG16 and InceptionV3 | generate, segment, and classify CXR for TB using three deep architectures | Montgomery, Shenzhen | Acc = 97.10%, Sen = 97.90%, Spe = 96.20% | 2021 |
| Msonda et al. (22) | AlexNet, GoogLeNet, ResNet50 and Spatial Pyramid Pooling (SPP) | Integrate SPP with Deep CNN to improve performance for TB detection | Montgomery, Shenzhen, Private (KERH) | Acc = 0.98%, Pre = 1.00%, Spe = 1.00%, F1 = 1.00% | 2020 |
| Owais et al. (23) | Fusion-based deep classification network | The study proposes a CAD system for the effective diagnosis of TB and provides visual with descriptive information that is useful to the radiologist | Montgomery, Shenzhen | Acc = 0.928%, Auc = 0.965%, Pre = 0.937%, Recall = 0.921%, F1 = 0.929% | 2020 |
| Yoo et al. (24) | ResNet18 | Classify CXR into Normal, TB and Non-TB | Montgomery, Shenzhen, NIH | Acc = 0.98%, Sen = 0.98%, Spe = 0.97%, Auc = 0.965%, Pre = 0.97% | 2020 |
| Sathiratanacheewin et al. (25) | DCNN | To examine the generalization of deep CNN models for classification of CXR as normal or abnormal with different manifestations | Shenzhen, NIH (ChestX-ray8) | Acc = 0.965%, Sen = 72%, Spe = 92% | 2020 |
| Sahiol et al. (26) | MobileNet Artificial Ecosystem-based Optimization (AEO) | Classification of CXR to detect TB | Shenzhen, Private | Acc = 94.1% | 2020 |
| Das et al. (27) | InceptionV3 | Screening CXR for TB abnormalities | Montgomery, Shenzhen | Acc = 91.7%, Auc = 0.96%, Sen = 0.89%, Spe = 0.93%, Pre = 0.93% | 2020 |
| Rahman et al. (28) | ResNet, ChexNet, InceptionV3, Vgg19, DenseNet201, SqueezeNet, MobileNet, and Ensemble | Automatic detection of TB from the CXR | NLM, Belarus, NIAID TB, and RSNA | Acc = 98.6%, F1 = 98.56%, Sen = 98.56%, Spe = 98.54%, Pre = 98.57% | 2020 |

(Continued)
of CAD systems is evaluated using a different set of datasets (CXR images) from the set used for training. In other words, avoid

- Using the same set of CXR images for training and testing.
- Testing with CXR images that were not used for training but originated from the same image subset.

Otherwise, the diagnostic accuracy evaluation is likely to be exaggerated and could impact the overall generalization of the system. In general, about 80% of the studies used the

| References                      | Computational techniques | Study aim/scope                                      | Datasets                  | Results                  | Year  |
|---------------------------------|--------------------------|------------------------------------------------------|---------------------------|--------------------------|-------|
| Munadi et al. (29)              | ResNet and EfficientNet  | Enhances CXR images for improving TB detection accuracy | Shenzhen                  | Acc = 91.7%              | 2020  |
| Oloko-Oba and Viriri (30)       | CNN                      | Detection of TB from CXR and classification as normal and abnormal | Shenzhen                  | Acc = 87.8%              | 2020  |
| Xie et al. (31)                 | Faster RCNN              | Detection of multiple categories of TB lesions in CXR | Montgomery, Shenzhen      | Acc = 0.926%             | 2020  |
| Verma et al. (32)               | InceptionV3, faster RCNN | Classify CXR as pulmonary TB and Pneumonia           | Shenzhen                  | Acc = 99.01%             | 2020  |
| Tasci (33)                      | AlexNet, VGGNet          | Classifying CXR ROI for TB detection                 | Montgomery, Shenzhen      | Acc = 88.32%             | 2020  |
| Rajaraman and Antani (34)       | Inception-V3, ResNet-V2, VGG-16, Xception, DenseNet-121, Ensemble | Improve state-of-the-art architecture for TB detection from CXR | Shenzhen, RSNA, Indiana   | Acc = 0.941% F1 = 0.941% Sen = 0.926% Spe = 1.00% Auc = 0.990% | 2020  |
| Abideen et al. (35)             | B-CNN                    | Identification and classification of CXR as TB and Non-TB | Montgomery, Shenzhen      | Acc = 96.42%             | 2020  |
| Hijazi et al. (36)              | Ensemble of VGG16 InceptionV3 | Detection of TB from CXR                                  | Montgomery, Shenzhen      | Acc = 89.77%              | 2019  |
| Pasa et al. (37)                | CNN                      | Developed a faster TB detection algorithm              | Montgomery, Shenzhen, Belarus | Acc = 84.4%              | 2019  |
| Meraj et al. (38)               | VGGNet, RestNet50, GoogleLeNet | Detection of TB abnormalities from CXR                 | Montgomery, Shenzhen      | Acc = 86.74%              | 2019  |
| Ahsan et al. (39)               | VGG16                    | Screening of CXR to identify the presence of TB       | Montgomery, Shenzhen      | Acc = 81.25%              | 2019  |
| Nguyen et al. (40)              | ResNet-50, VGGNet, DenseNet-121, Inception, ResNet.       | Improving detection rate of TB                        | Montgomery, Shenzhen, NIH-14 | Auc = 0.99%              | 2019  |
| Ho et al. (41)                  | InceptionResNetV2, ResNet150, DenseNet-121 | Classification of CXR as pulmonary TB or healthy. | ChestX-ray14, Shenzhen, Montgomery | Auc = 0.95%              | 2019  |
| Heo et al. (42)                 | VGG19, InceptionV3, ResNet-50, DenseNet-121, InceptionResNetV2. | Detection of TB from CXR                                | Private (Yonsei University) | Auc = 0.9213%             | 2019  |
| Hernández et al. (43)           | Ensemble of VGG19, InceptionV3, ResNet-50 | Automatic classification of CXR for TB detection.     | Private                   | Acc = 0.8642%            | 2019  |
| Hijazi et al. (44)              | Ensemble of InceptionV3, VGG16 | Detection of TB from CXR without segmentation           | Shenzhen, Montgomery      | Acc = 91.0%              | 2019  |
| Abbas and Abdelsamea (45)       | AlexNet                  | Classification of CXR as healthy or having TB manifestation | Montgomery                  | Auc = 0.998%              | 2018  |
| Karnkawinpong and Limpiyakorn (46) | AlexNet, VGG16, CapsNet | CAD for early diagnosis of TB                          | Private (Thai), Shenzhen, Montgomery | Acc = 90.79%              | 2018  |
| Strenko et al. (47)             | DCNN                     | Prediction of the presence of TB from CXR              | Shenzhen                  | ---                      | 2018  |
| Becker et al. (48)              | CNN                      | Detection and classification of different TB pathogens from CXR | Private                   | Auc = 0.98%              | 2018  |
| Liu et al. (49)                 | AlexNet, GoogLeNet       | Detection and classification of TB manifestations in CXR images | Peruvian                   | Acc = 85.68%             | 2017  |
| Hooda et al. (3)                | DCNN                     | Detect and classify TB from CXR as normal and abnormal | Shenzhen, Montgomery      | Acc = 82.09%             | 2018  |
### TABLE 3 | PubMed search results.

| References                        | Computational techniques                                   | Study aim/scope                                      | Datasets               | Results                  | Year |
|-----------------------------------|-----------------------------------------------------------|-----------------------------------------------------|------------------------|--------------------------|------|
| Oloko-Oba and Viriri (50)         | Ensemble of VGG-16, ResNet50, Inception V3               | Automatic detection of TB from CXR                  | Shenzhen, Montgomery   | Acc = 96.14%              | 2021 |
|                                   |                                                           |                                                     |                        | Sen = 90.03%              |      |
|                                   |                                                           |                                                     |                        | Spe = 92.41%              |      |
| Lee et al. (51)                   | DLAD                                                      | Detection of active TB and classification of relevant abnormalities on CXR | Private                | Auc = 0.967%              | 2021 |
|                                   |                                                           |                                                     |                        | Sen = 0.821%              |      |
|                                   |                                                           |                                                     |                        | Spe = 0.997%              |      |
| Zhang et al. (52)                 |Convolutional Block Attention Module (CBAM)               | Classification of TB from CXR                       | Private                | Acc = 87.7%               | 2020 |
|                                   |                                                           |                                                     |                        | Auc = 94.3%               |      |
|                                   |                                                           |                                                     |                        | Recall = 89.7%            |      |
|                                   |                                                           |                                                     |                        | Spe = 85.9.7%             |      |
| Hwang et al. (53)                 | DLAD                                                      | Developed a Deep Learning-based automatic detection algorithm (DLAD) for active Pulmonary TB on CXR and validate its performance using various datasets compared to physicians' results. | Shenzhen, Montgomery, Private (SNUH) | Auc = 0.977%              | 2019 |
|                                   |                                                           |                                                     |                        | Sen = 94.3%               |      |
|                                   |                                                           |                                                     |                        | Spe = 91.1%               |      |
| Rajpurkar et al. (54)             | CNN (CheXNeXt)                                           | To evaluate the effectiveness of CheXNeXt in detecting TB and other abnormalities from CXR | ChestX-ray14           | Auc = 0.862%              | 2018 |
|                                   |                                                           |                                                     |                        | Sen = 0.594%              |      |
|                                   |                                                           |                                                     |                        | Spe = 0.927%              |      |
| Lakhani and Sundaram (55)         | Ensemble of AlexNet, GoogLeNet                           | Evaluates the efficacy of deep models for detecting TB on CXR | Belarus, Shenzhen, Montgomery | Auc = 0.99%              | 2017 |
|                                   |                                                           |                                                     |                        | Sen = 97.3%               |      |
|                                   |                                                           |                                                     |                        | Spe = 94.7%               |      |

### TABLE 4 | IEEE Xplore search results.

| References                        | Computational techniques                                   | Study aim/scope                                      | Datasets               | Results                  | Year |
|-----------------------------------|-----------------------------------------------------------|-----------------------------------------------------|------------------------|--------------------------|------|
| Cao et al. (56)                   | DenseNet121, VGGNet16, VGGNet19, ResNet152                | Evaluates the performance of deep learning for classification of CXR for TB | Shenzhen, Montgomery   | Acc = 90.38%              | 2021 |
|                                   |                                                           |                                                     |                        | F1 = 90.36%               |      |
|                                   |                                                           |                                                     |                        | Pre = 90.33%              |      |
|                                   |                                                           |                                                     |                        | Recall = 90.53%           |      |
| Karaca et al. (57)                | VGG16, VGG19, DenseNet121, MobileNet, InceptionV3         | Development of a TB detection system                | Montgomery             | Acc = 98.8%               | 2021 |
|                                   |                                                           |                                                     |                        | Auc = 1.00%               |      |
| Saif et al. (58)                  | DenseNet169, ResNet-50, InceptionV3                       | Detection of TB from CXR                            | Shenzhen, Montgomery   | Acc = 99.7%               | 2021 |
|                                   |                                                           |                                                     |                        | Sen = 97.5%               |      |
|                                   |                                                           |                                                     |                        | Spe = 98.4%               |      |
| Das et al. (27)                   | InceptionV3                                              | Screening TB from CXR to eliminate patients diagnosis delay | Shenzhen, Montgomery   | Acc = 91.7%               | 2021 |
|                                   |                                                           |                                                     |                        | Auc = 0.96%               |      |
| Imam et al. (59)                  | Modified Inception                                       | They analyzed patients’ CXR to determine those infected with TB or not. | Shenzhen, Montgomery   | Acc = 91%                 | 2020 |
| Griffin et al. (60)               | R-CNN                                                     | Location of TB manifestations on CXR                | Peruvian               | Auc = 0.753%              | 2020 |
|                                   |                                                           |                                                     |                        | Sen = 0.922%              |      |
|                                   |                                                           |                                                     |                        | Spe = 0.666%              |      |
| Rashid et al. (61)                | Ensemble of ResNet, Inception-ResNet, DenseNet            | Development of a CAD system to classify CXR as normal and infected with TB | Shenzhen              | Acc = 90.5%               | 2018 |
| Abbas and Abdelsamea (45)         | AlexNet                                                   | Classification of CXR as healthy and unhealthy with TB manifestation | Shenzhen, Montgomery   | Auc = 0.998%              | 2018 |
|                                   |                                                           |                                                     |                        | Sen = 0.999%              |      |
|                                   |                                                           |                                                     |                        | Spe = 0.997%              |      |

same CXR images for training and testing or did not report on it.

**CONCLUSION AND RECOMMENDATIONS**

This systematic review intends to inform researchers of the existing Deep Learning classifiers and assist them in developing a CAD system for the efficient diagnosis of TB. Different state-of-the-art Deep Learning techniques that have been used to detect TB have been explored and presented in Tables 2–5. Generally, it is challenging to compare the methods with each other extensively. Factors like the types of datasets used for evaluation, number of image samples used, evaluation metrics approach, and model parameters tunning make the comparison complex.
### TABLE 5 | Web of Science search results.

| References          | Computational techniques | Study aim/scope                                                                 | Datasets                  | Results                  | Year |
|---------------------|--------------------------|---------------------------------------------------------------------------------|---------------------------|--------------------------|------|
| Fehr et al. (62)    | CAD4TBv5                 | Implement CAD4TB to screen for TB on CXR                                         | Private                   | Sen = 82.8%              | 2021 |
| Vats et al. (63)    | CNN (iDoc-X)             | Diagnosis of TB manifestations from CXR and classify images as TB and Non-TB     | Private                   | Acc = 91.10%             | 2021 |
| Khatibi et al. (18) | CNNs, complex networks and stacked ensemble | TB recognition from CXR images                                                   | Shenzhen, Montgomery     | Acc = 99.26%              | 2021 |
| Rajpurkar et al. (64) | DenseNet121             | Development of CheXaid for diagnosing TB                                          | Private                   | Acc = 0.78%               | 2020 |
| Grivkov and Smirnov (65) | InceptionV3             | Screening of CXR to detect TB pathologist                                        | Shenzhen, Montgomery     | Acc = 0.868%              | 2020 |
| Msonda et al. (22)  | AlexNet, GooLeNet, ResNet50, SPP | Integrate SPP with DCNN for the diagnosis of TB on CXR.                            | Shenzhen, Montgomery, Private (KERH) | Auc = 0.98% | 2020 |
| Gozes and Greenspan (66) | DenseNet121             | Learned specific features from CXR to detect TB                                  | Chest X-ray14             | Auc = 0.965%              | 2019 |
| Karnkawinpong and Limpiyakorn (67) | AlexNet, VGG-16, CapsNet | Classification of TB from CXR                                                    | NLM, Private             | Acc = 94.56%              | 2019 |
| Sivaramakrishnan et al. (68) | AlexNet, VGG16, VGG19, Xception, ResNet-50 | evaluate the performance of Deep models toward improving the accuracy of TB screening from CXR | Shenzhen, Montgomery, Private | Auc = 0.855% | 2018 |
| Vajda et al. (69)    | CNN                      | Screening CXR to determine which CXR images are normal or abnormal with TB.     | Shenzhen, Montgomery     | Acc = 97.03%              | 2018 |

As evident from the literature, these pre-trained models (VggNet, ResNet, AlexNet, DenseNet, and Inception) are the most popular and have been extensively explored for the classification of TB, as shown in Figure 4. Despite the effectiveness of Deep Learning models in detection and classification tasks, CAD systems for clinical diagnosis are still challenging in a real-world scenario. The physicians and radiologists see CAD intervention as a threat to their jobs rather...
than a supporting system to improve physicians’ performance in terms of time, effort, efficiency, and affordability, especially in developing countries. This review found that most existing works focused on development studies rather than clinical studies.

One of the likely weaknesses of this review is that it is limited to only the studies written in English because there might be other high-quality studies written in other languages. Also, we restricted the computational techniques to only Deep Learning classifiers, which could increase the risk of classifiers bias where studies that employed a hybrid of both Machine Learning and Deep Learning could have achieved better performance. Furthermore, this study did not undertake meta-analyses due to variations of algorithms used. Also, the raw data required to meta-analyze the diagnostic accuracy for most studies were unavailable.

However, it is recommended that studies be carried out using standardized public datasets that contain additional masks of the images that can be used as groundtruth to detect the infected aspect of the pulmonary images. It is highly recommended that models are trained with a set of images and evaluated on a different set of images. For instance, training a model with the Shenzhen datasets and evaluating it on the Montgomery dataset will validate better generalization. It is also recommended that future CAD systems focus more on clinical evaluation and should be able to identify foreign objects such as buttons and rings that look like a nodule on the CXR images, which may lead to misclassification.

### DATA AVAILABILITY STATEMENT

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

### AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct, and intellectual contribution to the work and approved it for publication.

### REFERENCES

1. World Health Organization. Global Tuberculosis Report 2020: Executive Summary. Available online at: https://apps.who.int/iris/handle/10665/337538 (accessed November 29, 2021). License: CC BY-NC-SA 3.0 IGO.
2. Baral SC, Karki DK, Newell JN. Causes of stigma and discrimination associated with tuberculosis in Nepal: a qualitative study. BMC Public Health. (2007) 7:1–6. doi: 10.1186/1471-2458-7-211
3. Hooda R, Sofat S, Kaur S, Mittal A, Meriaudeau F. Deep-learning: a potential method for tuberculosis detection using chest radiography. In: 2017 IEEE International Conference on Signal and Image Processing Applications (ICSIPA) (IEEE) (2017), 497–502.
4. Zumla A, George A, Sharma V, Herbert RH, Oxley A, Oliver M. The WHO 2014 global tuberculosis report—further to go. Lancet Global Health. (2015) 3:e10–2. doi: 10.1016/S2214-109X(14)70361-4
5. Sathitratanacheewin S, Pongpirul K. Deep learning for automated classification of tuberculosis-related chest x-ray: dataset specificity limits diagnostic performance generalizability. arXiv[Preprint].arXiv:1811.07985. (2018).
6. World Health Organization. Chest Radiography in Tuberculosis Detection: Summary of Current WHO Recommendations and Guidance on Programmatic Approaches. (2016). World Health Organization. Available online at: https://apps.who.int/iris/handle/10665/252424 (accessed November 29, 2021).
7. Williams FH. The use of X-ray examinations in pulmonary tuberculosis. Boston Med Surg J. (1907) 157:850–3. doi: 10.1056/NEJM190712261572602
8. Pande T, Pai M, Khan FA, Denkinger CM. Use of chest radiography in the 22 highest tuberculosis burden countries. Eur Respir J. (2015) 46:1816–9. doi: 10.1183/13993003.01064-2015
9. Noor NM, Rijal OM, Yunus A, Mahayiddin AA, Peng GC, Abu-Bakar SA. A statistical interpretation of the chest radiograph for the detection of pulmonary tuberculosis. In: 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES) (IEEE). (2010), 47–51.
10. Pedrazzoli D, Lalli M, Boccia D, Houben R, Kranzer K. Can tuberculosis patients in resource-constrained settings afford chest radiography? Eur Respir J. (2017) 49:1601877. doi: 10.1183/13993003.01877-2016
11. Schmidhuber J. Deep learning in neural networks: an overview. Neuro. Netw. (2015) 61:85–117. doi: 10.1016/j.neunet.2014.09.003
12. Moher D, Liberati A, Tetzlaff J, Altman DG, Prisma Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS Med. (2009) 6:e1000097. doi: 10.1371/journal.pmed.1000097
13. Duong LT, Le NH, Tran TB, Ngo VM, Nguyen PT. Detection of tuberculosis from chest X-ray images: boosting the performance with vision transformer and transfer learning. Expert Syst Appl. (2021) 184:115519. doi: 10.1016/j.eswa.2021.115519
14. Norval M, Wang Z, Sun Y. Evaluation of image processing technologies for pulmonary tuberculosis detection based on deep learning convolutional neural networks. J Adv Inform Technol. (2019) 12:253–9. doi: 10.12720/aitj.12.3.253-259
15. Rahman M, Cao Y, Sun X, Li B, Hao Y. Deep pre-trained networks as a feature extractor with XGBoost to detect tuberculosis from chest X-ray. Comput Elect Eng. (2021) 93:107252. doi: 10.1016/j.compeleceng.2021.107252
16. Govindarajan S, Swaminathan R. Extreme learning machine based differentiation of pulmonary tuberculosis in chest radiographs using integrated local feature descriptors. Comput Methods Prog Biomed. (2021) 204:106058. doi: 10.1016/j.cmpb.2021.106058
17. Alawi AE, Al-basser A, Sallam A, Al-sabaei A, Al-khateeb H. Convolutional neural networks model for screening tuberculosis disease. In: 2021 International Conference of Technology, Science and Administration (ICTSA) (IEEE). (2021), 1–5.

18. Khatibi T, Shalasvari A, Farahani A. Proposing a novel multi-instance learning model for tuberculosis recognition from chest X-ray images based on CNNs, complex networks and stacked ensemble. Phys Eng Sci Med. (2021) 44:291–311. doi: 10.1016/j.s1324-021-00980-w

19. Ayaz M, Shaukat F, Raja G. Ensemble learning based automatic detection of tuberculosis in chest X-ray images using hybrid feature descriptors. Phys Eng Sci Med. (2021) 44:183–94. doi: 10.1016/j.s1324-020-00966-0

20. Priya PA, Vimina ER. Tuberculosis detection from CXR: an approach using transfer learning with various CNN architectures. In: International Conference on Communication, Computing and Electronics Systems: Proceedings of ICCCCES 2020. Vol. 733. (Springer Nature). (2021), 407p.

21. Dasanayaka C, Dissanayake MB. Deep learning methods for screening pulmonary tuberculosis using chest X-rays. Comput Methods Biomech Biomed Eng Imaging Vis. (2021) 9:39–49. doi: 10.1080/21681163.2020.1808532

22. Owais M, Arsalan M, Mahmood T, Kim YH, Park KR. Comprehensive computer-aided decision support framework to diagnose tuberculosis from chest X-ray images: data mining study. JMIR Med Inform. (2020) 8:e21790. doi: 10.2196/21790

23. Yoo SH, Geng H, Chiu TL, Yu SK, Cho DC, Heo J, et al. Study on the TB and non-TB diagnosis using two-step deep learning-based binary classifier. J Instrument. (2020) 15:P10011. doi: 10.18088/ijeat.v15i10.10101

24. Rahman T, Khandakar A, Kadir MA, Islam KR, Islam KF, Mazhar R, et al. Reliable tuberculosis detection using chest X-ray with deep learning, segmentation and visualization. IEEE Access. (2020) 8:191586–191591. doi: 10.1109/ACCESS.2020.3031384

25. Munadi K, Muchtar K, Maulina N, Pradhan B. Image enhancement for tuberculosis detection. IEEE Access. (2020) 8:217897–907. doi: 10.1109/ACCESS.2020.3041867

26. Das D, Santosh KC, Pal U. Inception-based deep learning architecture for tuberculosis screening using chest X-rays. In: 2020 25th International Conference on Pattern Recognition (ICPR). Vol. 10 (IEEE). (2021), 3612–9.

27. Rajaraman S, Antani SK. Modality-specific deep learning model ensembles toward improving TB detection in chest radiographs. IEEE Access. (2020) 8:27318–26. doi: 10.1109/ACCESS.2020.2971257

28. Abdeen ZU, Ghafoor M, Munir K, Saqib M, Ullah A, Zia T, et al. Uncertainty assisted robust tuberculosis identification with bayesian convolutional neural networks. IEEE Access. (2020) 8:22812–25. doi: 10.1109/ACCESS.2020.2970023

29. Hijazi MH, Hwa SK, Bade A, Yaakob R, Jeffree MS. Ensemble deep learning for tuberculosis detection using chest X-ray and candy edge detected images. IAES Int J Artif Intell. (2019) 8:429. doi: 10.11591/iaes.v8i4.p429-435

30. Pasa F, Gollov V, Pfeiffer F, Cremers D, Pfeiffer D. Efficient deep network architectures for fast chest X-ray tuberculosis screening and visualization. Sci Rep. (2019) 9:1–9. doi: 10.1038/s41598-019-42557-4

31. Meraj SS, Yaakob R, Azman A, Rum SN, Shahril A, Nazri A, et al. Detection of pulmonary tuberculosis manifestation in chest X-rays using different convolutional neural network (CNN) models. Int J Eng Technol. (2019) 9:2270–5. doi: 10.35940/iet.a623.109119

32. Abideen ZU, Ghafoor M, Munir K, Saqib M, Ullah A, Zia T, et al. Uncertainty assisted robust tuberculosis identification with bayesian convolutional neural networks. IEEE Access. (2020) 8:22812–25. doi: 10.1109/ACCESS.2020.2970023
active pulmonary tuberculosis on chest radiographs. Clin Infect Dis. (2019) 69:739–47. doi: 10.1093/cid/ciy967
54. Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: a retrospective comparison of the CheXNeXt algorithm to practicing radiologists. PLoS Med. (2018) 15:e1002686. doi: 10.1371/journal.pmed.1002686
55. Lakhan P, Sundaram B. Deep learning at chest radiography: automated classification of pulmonary tuberculosis by using convolutional neural networks. Radiology. (2017) 4:574–82. doi: 10.1148/radiol.2017162326
56. Cao K, Zhang J, Huang M, Deng T. X-ray classification of tuberculosis based on convolutional networks. In: 2021 IEEE International Conference on Artificial Intelligence and Industrial Design (AIID) (IEEE). (2021), 125–9. doi: 10.1109/AIID51893.2021.9456476
57. Karaca BK, Güney S, Dengiz B, Agăldere M. Comparative study for tuberculosis detection by using deep learning. In: 2021 44th International Conference on Telecommunications and Signal Processing (TSP) (IEEE) (2021), 88–91.
58. Saif AF, Imtiaz T, Shahnaz C, Zhu WP, Ahmad MO. Exploiting cascaded ensemble of features for the detection of tuberculosis using chest radiographs. IEEE Access. (2021) 9:112388–99. doi: 10.1109/ACCESS.2021.3102077
59. Imam OT, Haque M, Shahnaz C, Imran SA, Islam MT, Islam MT. Detection of tuberculosis from chest X-ray images based on modified inception deep neural network model. In: 2020 IEEE International Women in Engineering (WIE) Conference on Electrical and Computer Engineering (WIECON-EC) (IEEE). (2020), 360–3.
60. Griffin T, Cao Y, Liu B, Brunette MJ. Object detection and segmentation in chest X-rays for tuberculosis screening. In: 2020 Second International Conference on Transdisciplinary AI (TranSAI) (IEEE) (2020), 34–42.
61. Rashid R, Khawaja SG, Akram MU, Khan AM. Hybrid RID network for efficient diagnosis of tuberculosis from chest X-rays. In: 2018 9th Cairo International Biomedical Engineering Conference (CIBEC) (IEEE) (2018), 167–70.
62. Fehr J, Konigorski S, Olivier S, Gunda R, Surujdeen A, Gareta D, et al. Computer-aided interpretation of chest radiography reveals the spectrum of tuberculosis in rural South Africa. NPJ Dig Med. (2021) 4:1–10. doi: 10.1016/S2397-9889(20)30009-3
63. Vats S, Singh S, Kala G, Tarar R, Dhawan S. iDoc-X: an artificial intelligence model for tuberculosis diagnosis and localization. J Discrete Math Sci Cryptogr. (2021) 24:1257–72. doi: 10.1080/09720529.2021.1932910
64. Rajpurkar P, O’Connell C, Schechter A, Asnani N, Li J, Kiani A, et al. CheXNet: deep learning assistance for physician diagnosis of tuberculosis using chest x-rays in patients with HIV. NPJ Dig Med. (2020) 3:1–8. doi: 10.1038/s41746-020-00322-2
65. Grivkov AV, Smirnov AA. Application of convolutional neural networks for diagnostics of tuberculosis. AIP Conf Proc. (2020) 2313:080011.
66. Gozes O, Greenspan H. Deep feature learning from a hospital-scale chest X-ray dataset with application to TB detection on a small-scale dataset. In: 2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC) (IEEE) (2019), 4076–9.
67. Karnkawinpong T, Limpiyakorn Y. Classification of pulmonary tuberculosis lesion with convolutional neural networks. J Phys Conf Ser. (2019) 1195:012007. doi: 10.1088/1742-6596/1195/1/012007
68. Sivaramakrishnan R, Antani S, Candemir S, Xue Z, Abuja J, Kohli M, et al. Comparing deep learning models for population screening using chest radiography. Med Imaging 2018 Comput Aided Diag. (2018) 10575:10575E. doi: 10.1117/12.2293410
69. Vajda S, Karagiris A, Jaeger S, Santosh KC, Candemir S, Xue Z, et al. Feature selection for automatic tuberculosis screening in frontal chest radiographs. J Med Syst. (2018) 42:1–11. doi: 10.1007/s10916-018-0991-9
70. Puddy E, Hill C. Interpretation of the chest radiograph. Continuing education in anaesthesia. Crit Care Pain. (2007) 7:71–5. doi: 10.1093/bjaecap/mkn014
71. Gilpin C, Korobitsyn A, Migliori GB, Raviglione MC, Weyer K. The world health organization standards for tuberculosis care and management. Eur Respir J. (2018) 51:1800098. doi: 10.1183/13993003.00098-2018
72. Brownlee J. 4 Types of Classification Tasks in Machine Learning. Machine Learning Mastery. 4p. Available online at: https://machinelearningmastery.com/types-of-classification-in-machine-learning (accessed November 25, 2021).
73. Jaeger S, Candemir S, Antani S, Wang YX, Lu PX, Thoma G. Two public chest X-ray datasets for computer-aided screening of pulmonary diseases. Quant Imaging Med Surg. (2014) 4:475. doi: 10.3978/j.issn.2223-4292.2014.11.20
74. Irvin J, Rajpurkar P, Ko M, Yu Y, Ciurea-Ilcus S, Chute C, et al. CheXpert: a large chest radiograph dataset with uncertainty labels and expert comparison. Proc AAAI conf Artif Intell. (2019) 33:590–7. doi: 10.1609/aaai.v33i30.3301590
75. Johnson AE, Pollard TJ, Greenbaum NR, Lungren MP, Deng CY, Peng Y, et al. MIMIC-CXR-JPG, a large publicly available database of labeled chest radiographs. arXiv[Preprint].arXiv:1901.07042. (2019).
76. Hwang S, Kim HE, Jeong J, Kim HI. A novel approach for tuberculosis screening based on deep convolutional neural networks. Med Imaging Comput Aided Diag. (2016) 9785:9785W2. doi: 10.1117/12.2216198
77. Lalkhen AG, McCluskey A. Clinical tests: sensitivity and specificity. Continuing Educa Anesth Crit Care Pain. (2008) 8:221–3. doi: 10.1093/bjaecap/mkn041
78. Davis J, Goadrich M. The relationship between precision-recall and ROC curves. In: Proceedings of the 23rd International Conference on Machine Learning 2006 Jun 22, 233–40.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher’s Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.