MATHEMATICAL MODELING AND COMPUTING, Vol. 12, No. 4, pp. 1295–1304 (2025)

AI-Powered Detection of COVID-19 and Lung Diseases from Chest X-Rays: Boosting Accuracy with CNNs and Top-K Algorithms

El Handri K.[1,2,3], Bouhouch A.[4], Hamal O.[5]

[1] *Medbiotech laboratory, Faculty of Medicine and Pharmacy (FMPR), University Mohammed V in Rabat, Morocco*
[2] *Aivancity School of AI & Data for Business & Society, France*
[3] *IPSS Laboratory Faculty of Sciences, Mohammed V University in Rabat, Morocco*
[4] *I.M.A.G.E. Laboratory, Moulay Ismail University of Meknes, Morocco*
[5] *National School of Architecture of Marrakech (ENAM), Marrakech, Morocco*

(Received 28 March 2025; Revised 6 November 2025; Accepted 10 November 2025)

The COVID-19 epidemic has highlighted the need for easier and more precise diagnoses. Traditional techniques, such as PCR tests, are helpful but can be time-consuming and laborious. In order to further enhance picture quality, this study presents a novel method for identifying COVID-19 and other lung disorders utilizing chest X-rays, convolutional neural networks (CNNs), and histogram equalization. The 1 823 X-ray pictures in the collection were divided into three categories: regular, COVID-19-positive, and additional lung infections. Based on the combination of CNN and Topk algorithms, our proposed approach reached 98.45% accuracy. These promising results suggest that our method may expedite the identification of COVID-19, reducing its consequences on the healthcare system. The dataset will be expanded in the future, along with sophisticated techniques and the use of our created Top-k algorithms to improve decision-making.

Keywords: *COVID-19; pulmonary diseases; convolutional neural network; confusion matrix; Topk algorithms; diagnosis.*

2010 MSC: 68T05, 68P20, 68T45, 92C55, 68U35 **DOI:** 10.23939/mmc2025.04.1295

1. Introduction

Due to the tremendous strain that the COVID-19 epidemic has placed on healthcare systems around the world, effective diagnostic technologies are essential for facilitating early detection and treatment. Rapid and accurate diagnostics are still crucial for limiting the spread of viruses and lowering mortality, as there are approximately 400 million cases and 6 million fatalities globally [1, 2]. Although conventional methods such as PCR testing and symptom analysis are widely used, they can often be slow, resource-heavy, and prone to human error. Medical imaging, particularly chest X-rays, has proven to be important in diagnosing COVID-19 and associated lung conditions [3]. However, manually analyzing these images can be time-consuming and prone to mistakes, especially in the high-pressure environment of clinical settings. However, with recent advances in artificial intelligence (AI) and deep learning, new opportunities have arisen to tackle these challenges [3]. CNNs are especially well-suited for evaluating medical images because they are already effective at identifying conditions like tumors and malignancies. This study suggests a novel method for identifying COVID-19 from chest X-rays by employing a CNN model that has been improved by histogram equalization during the preprocessing phase. By enhancing image contrast and preserving important details, this method makes it possible for CNN to better learn and identify the delicate patterns of COVID-19. An essential feature of this study is the integration of decision-support tools aimed at improving clinical usability. We propose using Multi-Criteria Decision Analysis (MCDA) to assist in interpreting diagnostic results. This will enable a more thorough evaluation by factoring in variables such as patient history, severity of symptoms, and confidence in predictions. Additionally, integrating the Top-k algorithm, which selects the most relevant diagnostic outcomes from the model streamlines the decision-making process for healthcare professionals, allowing them to focus on the most accurate and pertinent predictions while minimizing false positives and negatives. By embedding these decision-support systems, our approach has the potential to not only increase diagnostic accuracy but also provide clinicians with real-time, data-driven insights to make more informed decisions. The following sections will describe the development and testing of our CNN model, its combination with MCDA and Top-k, and the potential impact this approach could have on improving COVID-19 diagnostics.

2. Related works

Shibly et al. proposed a hybrid model combining CNN [4] architecture with long short-term memory (LSTM) networks to automatically diagnose COVID-19 from CXR images [2]. The combination above highlights the value of integrating complementary deep learning models to increase the diagnostic's accuracy by utilizing both the CNNs' feature extraction capabilities and the LSTM networks' temporal dependencies. Similarly, other research efforts have focused on the role of image preprocessing techniques in enhancing model performance showed in papers [5, 6] in which the authors explored various image enhancement methods, such as contrast adjustment and noise reduction, to improve the quality of CXR images prior to analysis. These enhancements have been shown to greatly boost the accuracy of COVID-19 detection by permitting CNNs to capture finer details in lung images [7]. In addition to these advancements, some studies have also emphasized the use of decision-support mechanisms to enhance diagnostic workflows [7, 8]. Multi-Criteria Decision Analysis (MCDA) and Top-k algorithms presented in our previous papers [9–12], although not widely adopted yet in COVID-19 diagnostic models. It should be noted that the proposed approaches represent a new direction for further research. Moreover, the findings shows that it can help refine and prioritize outputs from CNN models, ensuring that the focus on the most relevant and accurate results will be easy for healthcare providers. For instance, the authors Wang et al. proposed a new ensemble learning approach that combines multiple CNN models with different architectures to improve the robustness and generalization ability of COVID-19 detection [13, 14]. According to their findings, the ensemble model performed better in sensitivity and accuracy than individual CNN models. Furthermore, another work presented by the authors in [8] studies the augmentation techniques' impact on data on the performance of CNN models for COVID-19 detection. Their finding shows the augmenting the training dataset with rotated, flipped, and cropped images can ameliorate and improve model accuracy and prevent overfitting. Finally, several studies have explored the use of transfer learning to accelerate the training of CNN models for COVID-19 detection. Researchers can improve performance and shorten training times by using pre-trained models that have been developed on large-scale image datasets to fine-tune the models on smaller COVID-19 datasets. In addition, the collective findings from these studies [15] illustrate the growing importance of deep learning in medical imaging, particularly in the context of pandemic response. By building on these foundations, our work aims to further optimize COVID-19 detection using advanced CNN models enhanced with preprocessing techniques and decision-support systems.

3. Methodology

Chest X-ray image database. A subset of the COVID-QU database, a large collection of chest X-ray pictures created to aid in the creation of machine-learning algorithms for COVID-19 detection, was used to train our model. The chest X-ray image database used in this study is a subset of the COVID-QU Dataset (Kaggle). We selected 1 823 images [16, 17], which are distributed across three key categories as presented in Table 1.

Table 1. Image categories and counts.

Category	Number of Images
COVID-19 positive	536
Normal	668
Other lung viruses	619

As illustrated in Figure 1, the dataset includes examples of COVID-19 positive chest X-ray images. Figure 1 shows normal lung X-rays, while Figure 2 depicts the distribution of images across the three diagnostic categories. This specific distribution in

Fig. 1. Sample images from the COVID-QU dataset: Positive COVID-19 case.

Fig. 2. Sample images from the COVID-QU dataset: Normal lung case.

Figure 3 allows the model to differentiate between normal and infected lungs while also learning the subtle features of COVID-19 in contrast to other viral lung infections. The kind of data mentioned above has to be used. Thus, the COVID-19 positive X-Rays images were selected to help the model recognize unique patterns associated with the disease, such as lung opacities and ground-glass appearances. Standard images serve as a reference for comparison, enabling the model to distinguish between healthy and diseased lungs. Lastly, the inclusion of lung virus images from other viral infections ensures that the model does not overfit COVID-19-specific features. Consequently, this aspect still maintains broader diagnostic applicability. This balanced dataset aims to provide CNN with a well-rounded understanding of different lung conditions, ensuring it is capable of making accurate distinctions not only between COVID-19 cases and healthy lungs but also between COVID-19 and other viral infections.

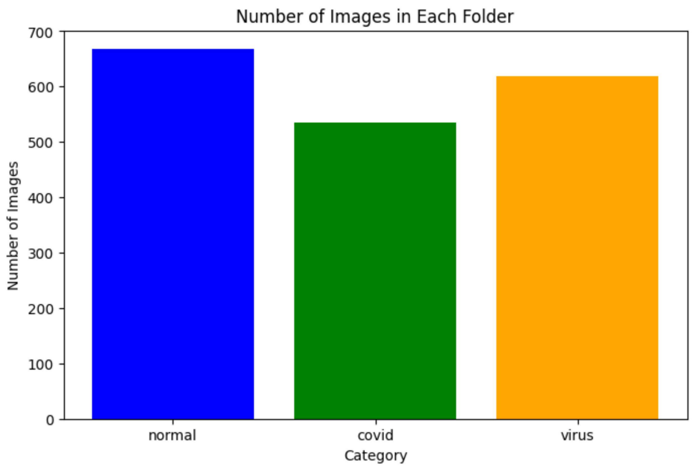


Fig. 3. Distribution of images in the COVID-QU database.

By carefully curating this selection of images, a robust and diverse dataset was aimed to be created to allow the model to generalize well across various pulmonary conditions, while ensuring optimal performance in detecting COVID-19.

4. Project stages

4.1. Image size standardization

To ensure consistency and compatibility with the CNN model, all images were resized to a uniform dimension. This preprocessing step is crucial for reducing computational complexity and ensuring consistent feature extraction across the dataset [18, 19].

4.2. Grayscale conversion

Converting images to grayscale allows the model to focus on essential texture and intensity variations, reducing overall data complexity. By eliminating color information, the model emphasizes critical patterns in lung opacity. The grayscale transformation is mathematically represented as:

$$f(x) = \max(0, x). \tag{1}$$

As shown in Equation (1), this transformation enhances the model's ability to detect abnormalities [20].

4.3. Data splitting

The dataset was divided into training and validation sets to assess model performance on unseen data:

$$\text{Training Set} = 80 \tag{2}$$

Equation (2) ensures that the model generalizes well to new data [18].

4.4. CNN architecture design

The CNN architecture was designed with multiple convolutional and pooling layers to extract meaningful features from images. The convolution operation is defined as:

$$O(i, j) = \sum_{m=0}^{k-1} \sum_{n=0}^{k-1} I(i + m, j + n) \cdot K(m, n). \tag{3}$$

As shown in Equation (3), convolution aids in the identification of patterns at various spatial locations. Moreover, feature maps are downsampled using max pooling:

$$P(i, j) = \max_{m,n \in R} O(i + m, j + n). \tag{4}$$

Equation (4) reduces spatial dimensions while retaining essential features [19]. The CNN architecture used in this study is illustrated in Figure 4.

4.5. Model hyperparameter tuning

To maximize the model's performance, key hyperparameters such as the learning rate, batch size, and the number of training epochs were fine-tuned. This process involved several iterations to achieve an optimal balance between training time and accuracy [18].

4.6. Model training

The model was trained using cross-entropy loss:

$$L = -\sum_{c=1}^{C} y_c \log(\hat{y}_c), \tag{5}$$

where y_c is the true label and $\hat{y}_c$ is the predicted probability. The weights were updated using gradient descent:

$$W_{\text{new}} = W_{\text{old}} - \eta \frac{\partial L}{\partial W}. \tag{6}$$

Equation (6) ensures that the model learns by minimizing classification errors [19].

4.7. Evaluation metrics

The performance of the model was evaluated using standard metrics derived from the confusion matrix:

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}, \tag{7}$$

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}, \tag{8}$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \tag{9}$$

$$\text{F1-Score} = 2 \cdot \frac{\text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}}. \tag{10}$$

Equations (7)–(10) provide insight into the model's predictive performance [18].

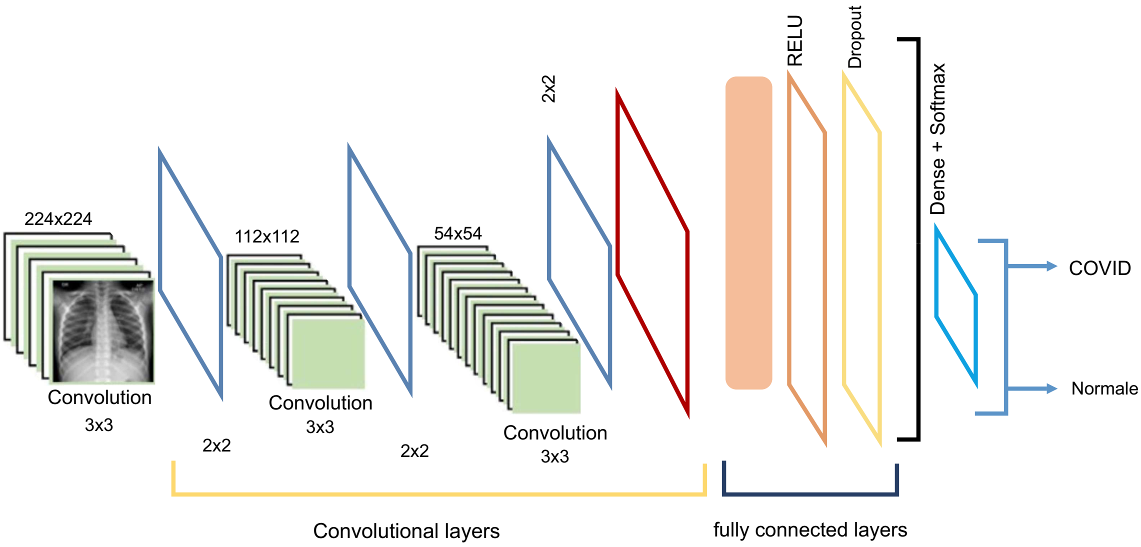


Fig. 4. The CNN architecture used in this study.

4.8. ROC curve

The Receiver Operating Characteristic (ROC) curve was used to evaluate the classification performance across various threshold values. The true positive rate (TPR) and false positive rate (FPR) are given by:

$$\text{TPR} = \frac{\text{TP}}{\text{TP} + \text{FN}}, \quad \text{FPR} = \frac{\text{FP}}{\text{FP} + \text{TN}}. \tag{11}$$

Equation (11) helps visualize the trade-off between sensitivity and specificity in classification performance [19].

5. Results and discussions

Our proposed approach was trained and validated using the database, which contains X-ray images categorized as COVID-19-positive patients, normal X-ray images, and other lung virus cases. This diverse dataset allowed the model to learn the distinctive characteristics of COVID-19 while establishing a baseline for comparison with other lung disorders. The training and validation sets guaranteed the model's strong generalization to unknown data, essential for clinical applications. Analyzing the ROC curve, we notice that the CNN model has enhancing the ability to differentiate between regular, other pneumonia, and COVID-19 patients. Therefore, the combination with the TopK algorithm enhances this ability. The model continuously demonstrates great accuracy across all categories with an outstanding Area Under the Curve (AUC) of 0.96 for standard and pneumonia cases and 0.97 for COVID-19. Consequently, each curve approaches the graph's upper-left corner, suggesting strong sensitivity and specificity. These AUC values illustrate the model's dependability in identifying the conditions for which it was trained. In practical terms, the model's high AUC for COVID-19 reflects its ability to accurately identify true cases while keeping false alarms low. This little advantage in identifying COVID-19 implies that the model is particularly good at identifying this illness. Because of this, it may be a useful tool in clinical situations where prompt and accurate diagnosis is crucial. Finally, all things considered, the CNN precision in identifying certain lung disorders gives medical professionals an extra degree of assistance. This little advantage in identifying COVID-19 implies that the model is particularly good at identifying this illness. Because of this, it may be a useful tool in clinical situations where a prompt and accurate diagnosis is crucial. Despite the precision

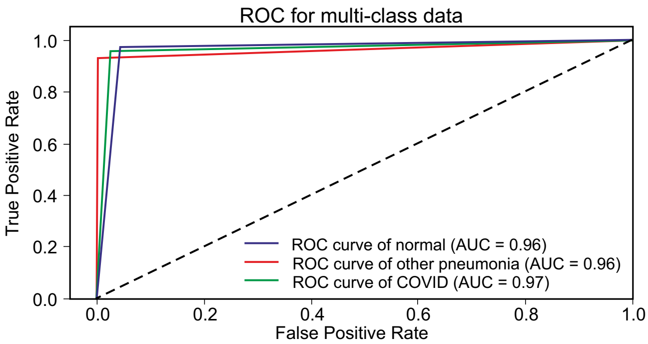
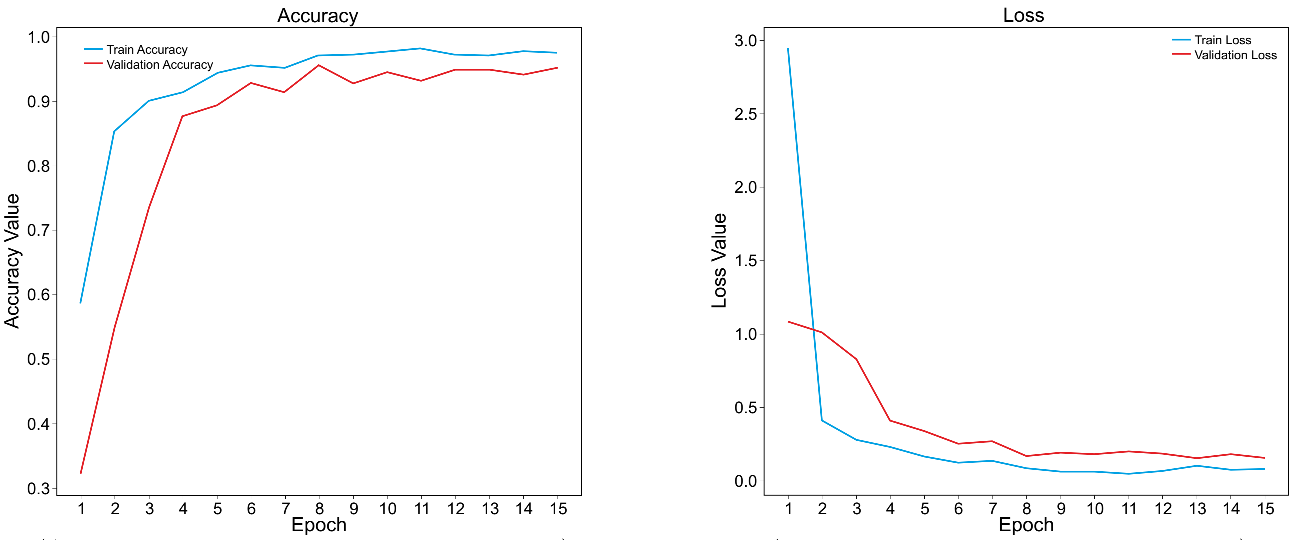


Fig. 5. ROC curve of the CNN model.

of the CNN model in identifying certain lung disorders, medical professionals receive additional help when diagnosing chest X-rays. Finally, this methodology has the potential to improve patient care and resource allocation in healthcare systems by increasing diagnostic confidence and aiding in the prioritization of urgent cases. Figures 6 and 5 illustrate the accuracy and the ROC curve, respectively.

The ROC curve, presented in Figure 5, demonstrates the model's ability to balance sensitivity and specificity across different threshold settings. Better performance in differentiating between positive and negative examples is shown by a curve that is closer to the top-left corner. According to our findings, the CNN is a trustworthy and accurate tool for clinical decision-making for diagnosing COVID-19.



a (Accuracy of the CNN model during training) ***b*** (Loss of the CNN model during training)

Fig. 6. Fast-CovNet performance.

The training and validation loss curves depicted in Figure 6*b* further support the model's performance. The decreasing loss values over epochs indicate that the model is learning effectively from the training data and generalizing well to the validation set. The model's dependability in practical applications is ensured by the convergence of the training and validation loss curves, which indicates that the model is not overfitting. These findings underscore the potential of integrating deep learning techniques like CNNs in the healthcare sector, especially for rapidly and accurately detecting respiratory diseases. Future work will explore enhancing model robustness through the integration of Multi-Criteria Decision Analysis (MCDA) and the Top-k algorithm to further support decision-making processes by selecting the most relevant cases for diagnosis. In addition, if the dataset is imbalanced, techniques like class weighting or oversampling might be necessary to address potential biases. The choice of model architecture and hyperparameters can also impact performance. Experimenting with different models and tuning hyperparameters could lead to further improvements.

6. Our approach, future directions, and potential for clinical trials

To ensure the practical applicability of this AI model within the Moroccan healthcare system, future clinical trials must leverage a more extensive and carefully curated dataset that reflects local demographics and specific clinical needs. This initiative will be supported by a Big Data infrastructure, enhanced by the SPTopKws mechanism a newly developed Top-K algorithm optimized for large-scale data processing. While PySpark was used in this study, deploying the model within a full-scale Big Data environment will be crucial for validating its scalability and performance on larger datasets.

Additionally, a recommender system powered by SPTopKws [9–11] will be integrated as a decision-support tool to rank potential diagnoses based on the model's outputs. By utilizing Top-K query processing and ranked lists, this approach aims to assist healthcare providers in prioritizing cases and efficiently allocating resources. The design also emphasizes interoperability [21, 22], ensuring seamless integration into existing clinical workflows. While initially developed for COVID-19 detection [11], the model can be extended to other respiratory diseases, further enhancing clinical decision-making across diverse healthcare settings.

By combining Top-K algorithms with CNNs, our study brings a powerful new tool to the table to improve medical diagnostics. CNNs are incredibly good at detecting subtle patterns in chesradiographs such as those linked to COVID diseases, but they can produce so many difficult for physicians to focus on the most critical cases. That is where Top-K algorithms come in. They assist in making sure that critical cases receive attention by going through the projections and emphasizing the most crucial ones. Furthermore, this decreases the possibility of overlooking serious illnesses while also expediting the diagnostic process, as shown in earlier research by [23] and Chen et al. [24]. In addition, Top-K algorithms make model decisions easier to understand, giving doctors more precise information on why certain predictions are made, as highlighted by Zheng et al. [8, 25, 26] in their work on medical image analysis, additional work that relies on traditional Topk combined with the DL models is in [27] and [26]. Therefore, compared to the state-of-the-art, our key contribution is bringing these two technologies together: CNNs for detecting patterns and Top-K algorithms for prioritizing decisions to create a more intelligent and reliable diagnostic system. Based on the conceived algorithm, Topkws and Sptopkws, we achieve a higher accuracy. This approach delivers fast and precise results and helps healthcare professionals make better, more informed decisions. Ultimately, this innovation could ease the load on healthcare systems and improve patient outcomes.

Algorithm 1 CNN-Based Image Processing and Top-K Recommendation.

Input: Dataset $\mathcal{D} = \{I_1, I_2, \ldots, I_n\}$.
Output: Top-K recommendations $R_{\text{Top-K}}$.
Step 1: Image Preprocessing and Feature Extraction
Initialize SparkSession.
for each image $I_i \in \mathcal{D}$
  Convert I_i to grayscale and normalize: $I_i^{\text{norm}} = \frac{\text{Grayscale}(I_i)}{255}$.
Define CNN architecture $\mathcal{C}, \mathcal{P}, \mathcal{F}$.
Train CNN using cross-entropy loss $\mathcal{L} = -\sum_{c=1}^{M} y_c \log(\hat{y}_c)$.
Extract features $F = \{f_1, f_2, \ldots, f_n\}$ from CNN.
Step 2: Query Processing and Ranking
Initialize User's Query Processing Agent (UQPA).
Apply Pre-Skyline Processing Agent (PSPA) to filter queries: $Q_{\text{filtered}} = \text{PSPA}(F)$.
Apply ELECTRE IS Agent (EISA) to rank queries: $Q_{\text{ranked}} = \text{EISA}(Q_{\text{filtered}})$.
Step 3: Top-K Query Optimization
Apply Top-K Query Processing Agent (TopKQPA): $Q_{\text{Top-K}} = \text{TopKQPA}(Q_{\text{ranked}}, K)$.
Apply Weighted Sum Agent (WSA): $\text{Score}(q) = \sum_{i=1}^{m} w_i \cdot f_i(q)$.
Step 4: Recommendation Generation
Predict ratings using SVD: $R_{\text{predicted}} = U \cdot \Sigma \cdot V^T$.
Store results in Spark RDD.
Return $R_{\text{Top-K}}$ if $|R_{\text{Top-K}}| > 0$; otherwise, return $\varnothing$.

We enhance previous work through a combination that starts by following a similar approach and then adds CNN. In fact, the recommender system begins with user interaction data collection and profile creation, followed by query processing and utility matrix extraction. A K-dominating query

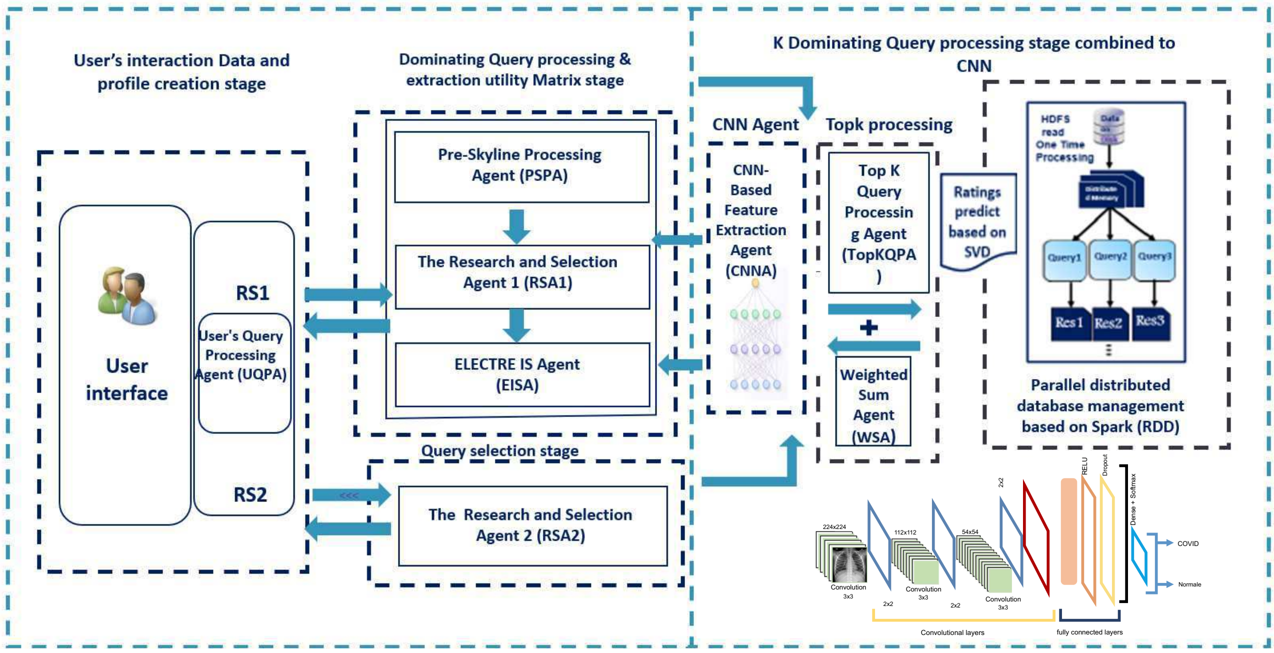


Fig. 7. Recommender System based on SPTopkws algorithm [10].

processing and prediction model, based on Singular Value Decomposition (SVD), was then employed to predict user preferences. The system incorporated multiple agents, including the User's Query Processing Agent (UQPA), Research and Selection Agents (RSA1 and RSA2), the ELECTRE IS Agent (EISA), the TopK Query Processing Agent (TopKQPA), and the Weighted Sum Agent (WSA), all working collaboratively to refine query selection and ranking. The integration of Spark RDD Database ensured efficient data handling, while a CNN-based image processing and feature extraction module contributed to generating Top-K recommendations. As demonstrated in our previous studies, this multiagent framework significantly improved recommendation accuracy and system efficiency [21]. Building upon this foundation, our current approach incorporates:

Users Interaction Data and Prole Creation Stage
↓
Dominating Query Processing & Extraction Utility Matrix
↓
K Dominating QueryProcessing & PredictionModel (SVD)
↓
User Interface
↓
Users Query ProcessingAgent (UQPA)
↓
Research and SelectionAgent 1 (RSA1)
↓
ELECTRE IS Agent (EISA)
↓
Query Selection Stage
↓
Research and SelectionAgent 2 (RSA2)
↓
Top K Query ProcessingAgent (TopKQPA)
↓
Weighted Sum Agent (WSA)
↓
Spark RDD Database
↓
Image Processing & CNN Feature Extraction
↓
Provide TopK Recommendations

Fig. 8. AI-Enhanced Query Processing and Recommendation System.

— Image preprocessing using histogram equalization for better intensity distribution.
— A lightweight CNN architecture for fast and accurate COVID-19 detection.
— Evaluation through the confusion matrix and the ROC curve, yielding a 98.45% accuracy rate.

As shown in Figure 8, the AI-enhanced query processing system integrates Top-K algorithms and CNN-based feature extraction to improve recommendation accuracy, also presented in Figure 7.

7. Conclusion and prospects

This study presents a novel method for detecting COVID-19 Virus using chest X-rays. The key contributions of this study include a combination of CNN and Topkws algorithms. Further improvements include using a more extensive image database and Big Data infrastructure using GPUs and additional Spark library as in our previous work, Algorithm 1 show the combination of the studied method with our previous approach based on the Recommender system showing in the Figure 7 [10]. To further improve accuracy and resilience, we aim to use more advanced X-rays image segmentation techniques will be integrated to further enhance accuracy and robustness. Furthermore, depending on the model's output, a TopK algorithm more precisely, the SPTopkws [10] mechanism may be used to rank the most likely diagnoses, potentially acting as a useful decision-support tool. Finally, this would enable healthcare providers to prioritize cases and allocate resources more effectively.

[1] Mathieu E., Ritchie H., Rodés-Guirao L., et al. Coronavirus (COVID-19) Cases. Our World in Data (2020). https://ourworldindata.org/covid-cases.

[2] Gupta V., Jain N., Sachdeva J., et al. Improved COVID-19 detection with chest x-ray images using deep learning. Multimedia Tools and Applications. **81** (26), 37657–37680 (2022).

[3] Careful.Online. The Future of Healthcare: 2024 AI Impact Analysis. Careful Online Publications (2024). https://careful.online/future-healthcare-ai-2024/.

[4] Aya T., Handri K. E., Fouzia O. Automated Detection of Malicious Software in Medical Environments Using Deep Learning. International Conference on the Global Summit of Digital Health. 100–106 (2025).

[5] Gupta V., Jain N., Sachdeva J., et al. Improved COVID-19 detection with chest x-ray images using deep learning. Multimedia Tools and Applications. **81** (1), 37657–37680 (2020).

[6] Rahman T., Khandakar A., Qiblawey Y., et al. Exploring the effect of image enhancement techniques on COVID-19 detection using chest X-ray images. Computers in Biology and Medicine. **132**, 104319 (2021).

[7] Ozturk T., Talo M., Yildirim E. A., et al. Automated detection of COVID-19 cases using deep neural networks with X-ray images. Computers in Biology and Medicine. **121**, 103792 (2020).

[8] Beddiar D. R., Oussalah M., Muhammad U., Seppänen T. A Deep learning based data augmentation method to improve COVID-19 detection from medical imaging. Knowledge-Based Systems. **280**, 110985 (2023).

[9] El Handri K., El Ahrach S., Touil M. A new synesthesia-based approach for AI in healthcare. International Journal of Research and Ethics. **5** (1), (2022).

[10] El Handri K., Idrissi A. Parallelization of Top_k Algorithm Through a New Hybrid Recommendation System for Big Data in Spark Cloud Computing Framework. IEEE Systems Journal. **15** (4), 4876–4886 (2020).

[11] El Handri K., Idrissi A. Correlations and Hierarchical Clustering Investigation Between Weather and SARS-CoV-2. Recent Advances in Computer Science and Communications (Formerly: Recent Patents on Computer Science). **15** (6), 859–867 (2022).

[12] Handri K. E., Idrissi A. Comparative study of Topk based on Fagin's algorithm using correlation metrics in cloud computing QoS. International Journal of Internet Technology and Secured Transactions. **10** (1/2), 143 (2020).

[13] Tang S., Wang C., Nie J., Kumar N., Zhang Y., Xiong Z. EDL-COVID: Ensemble Deep Learning for COVID-19 Case Detection From Chest X-Ray Images. IEEE Transactions on Industrial Informatics. **17** (9), 6539–6549 (2021).

[14] Abdullah M., Abrha F. B., Kedir B., Tamirat Tagesse T. A Hybrid Deep Learning CNN model for COVID-19 detection from chest X-rays. Heliyon. **10** (5), e26938 (2024).

[15] Luo X., Chen X., Yao Y., Yao Y. MVMSGAT: Integrating Multiview, Multi-Scale Graph Convolutional Networks with Biological Prior Knowledge for Predicting Bladder Cancer Response to Neoadjuvant Therapy. Applied Sciences. **14** (2), 669 (2024).

[16] Tahir A. M., Chowdhury M. E. H., Khandakar A., et al. COVID-19 infection localization and severity grading from chest X-ray images. Computers in Biology and Medicine. **139**, 105002 (2021).

[17] Tahir A. M., Chowdhury M. E. H., Qiblawey Y., et al. COVID-QU-Ex Dataset. Kaggle (2021).

[18] An J., Zhang N., Mahmoud W., Denis M. A COVID-19 CXR Image Recognition Method Based on Deep Transfer Learning. 2024 International Joint Conference on Neural Networks (IJCNN). 1–6 (2024).

[19] Goodfellow I., Bengio Y., Courville A. Deep Learning. MIT Press (2016).

[20] François-Lavet V., Henderson P., Islam R., Bellemare M. G., Pineau J., et al. An introduction to deep reinforcement learning. Foundations and Trends in Machine Learning. **11** (3–4), 219–354 (2018).

[21] Handri K. E., Benedress L. G., Wakrim L., Lamzabi S., Idrissi A. AI-Based QoS Optimization in Post-Disaster Ad Hoc Networks: Top-K Path Enumeration with Multicriteria Decision Analysis. 2024 IEEE 4th International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA). 561–566 (2024).

[22] Lamzabi S., Madi B. A. W., Benyoussef M., et al. Minimizing the Viruses Propagation in Interconnected Complex Networks. 2024 IEEE 4th International Maghreb Meeting of the Conference on Sciences and Techniques of Automatic Control and Computer Engineering (MI-STA). 544–549 (2024).

[23] Wang W., Xu Y., Gao R., Lu R., Han K., Wu G., Tan W. Detection of SARS-CoV-2 in different types of clinical specimens. JAMA. **323** (18), 1843–1844 (2020).

[24] Chen J., Wu L., Zhang J., Zhang L., Gong D., Zhao Y., Chen Q., Huang S., Yang M., Yang X., et al. Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography. Scientific Reports. **10** (1), 19196 (2020).

[25] Zhang L., Luo M., Liu J., Chang X., Yang Y., Hauptmann A. G. Deep Top-k Ranking for Image–Sentence Matching. IEEE Transactions on Multimedia. **22** (3), 775–785 (2019).

[26] Zheng Y., Zheng Y., Suehiro D., Uchida S. Top-rank convolutional neural network and its application to medical image-based diagnosis. Pattern Recognition. **120**, 108138 (2021).

[27] Petersen F., Kuehne H., Borgelt C., Deussen O. Differentiable Top-k Classification Learning. Proceedings of the 39th International Conference on Machine Learning. **162**, 17656–17668 (2022).

Виявлення COVID-19 та захворювань легень за допомогою штучного інтелекту та рентгенівських знімків грудної клітки: підвищення точності за допомогою CNN та алгоритмів Top-K

Ель Хандрі К.[1,2,3], Бууш А.[4], Хамал О.[5]

[1] *Медбіотехнологічна лабораторія, Медичний та фармацевтичний факультет (FMPR), Університет Мохаммеда V у Рабаті, Марокко*

[2] *Школа штучного інтелекту та даних для бізнесу та суспільства Айвансіті, Франція*

[3] *Лабораторія IPSS факультету природничих наук, Університет Мохаммеда V у Рабаті, Марокко*

[4] *Лабораторія I.M.A.G.E., Університет Мулая Ісмаїла в Мекнесі, Марокко*

[5] *Національна школа архітектури Марракеша (ENAM), Марракеш, Марокко*

Епідемія COVID-19 висвітлила потребу в простішій та точнішій діагностиці. Традиційні методи, такі як ПЛР-тести, корисні, але можуть бути трудомісткими та трудомісткими. Для подальшого покращення якості зображень у цьому дослідженні представлено новий метод виявлення COVID-19 та інших захворювань легень з використанням рентгенографії грудної клітки, згорткових нейронних мереж (CNN) та вирівнювання гістограм. 1 823 рентгенівські знімки в колекції були розділені на три категорії: звичайні, COVID-19-позитивні та додаткові інфекції легень. Завдяки поєднанню алгоритмів CNN та Topk, запропонований нами підхід досяг точності 98.45%. Ці багатообіцяючі результати свідчать про те, що запропонований метод може пришвидшити ідентифікацію COVID-19, зменшуючи його наслідки для системи охорони здоров'я. Набір даних буде розширено в майбутньому разом із вдосконаленими методами та використанням створених нами алгоритмів Top-k для покращення прийняття рішень.

Ключові слова: *COVID-19; легеневі захворювання; згорткова нейронна мережа; матриця плутанини; алгоритми Topk; діагностика.*