ScienceDirect

Expert Systems with Applications

Volume 239, 1 April 2024, 122317

Localization model and rank-based features selection approach for the classification of GGO and consolidation stages of COVID-19

Javaria Amin [a], Muhammad Almas Anjum [b], Nadia Gul [c], Muhammad Imran Sharif [d], Muhammad Irfan Sharif [e], Seifedine Kadry [f g h]

Show more

https://doi.org/10.1016/j.eswa.2023.122317

Get rights and content

Abstract

Diseases like Middle East respiratory syndrome (MERS) and severe acute respiratory syndrome (SARS) have been caused by a large family virus known as Coronavirus. High-resolution computed tomography (HRCT) chest is used as a useful tool for the triage and stratification of COVID-19-suspected patients having respiratory symptoms until they are confirmed by polymerase chain reaction (PCR) test. To further expedite the reporting of HRCT chest in a short time, computer-aided detection algorithms can be more useful. In this work, a method is proposed that consists of three phases including COVID-19 classification, localization, and segmentation. The COVID-19-ResNet-50 model is designed for COVID-19 classification using HRCT slices. The features are obtained from a pre-trained ResNet-50 model, and the process of identifying the most pertinent features is carried out using the Wilcoxon Separability Criteria (WSC). The linear discriminant (LD) classifier is used to classify COVID-19 providing an accuracy of 0.99. To localize the COVID-19-infected region, the classified HRCT slices are fed as input to an open neural network exchange (ONNX) that is a backbone of the Tiny-You Look at Once (YOLO)-v2 model. The proposed COVID-19-YOLOv2 model localized the COVID-19 lesions with 0.948 intersections over union (IoU). A three-dimensional semantic neural network comprising 10 selected layers is proposed whose training based on the best-selected hyperparameters is performed, it accurately segments the localized infected regions having an 0.98 IoU. The proposed method is accessed on national and international datasets that provide improved performance as compared to the most recent works published in this domain.

Introduction

COVID-19 is produced by a virus from the coronavirus family (COVID-19 Resources, 2020). COVID-19 virus causes lung infection leading to pneumonia which in severe cases may cause acute respiratory distress

syndrome, or ARDS which may be fatal. RT-PCR is a widely active molecular test considered the gold standard for spotting COVID-19 in blood, it claims a 90% consistency, although it requires up to 6h for results. Early detection of lung involvement and the extent of infection in both lungs is mandatory in COVID-19 suspected cases and it is very important to segregate the patients with mild and severe acute COVID lung disease to provide the required treatment. HRCT chest is the diagnostic modality that can serve the purpose and tells physicians not only the presence of COVID lung infection but also the extent of lung involvement. Keeping in view the importance of HRCT chest as a useful modality we have developed an algorithm under the guidance of a qualified radiologist that can automatically detect ground glass haze and consolidation in acute COVID lung infection and also assesses the extent of lung infection by using a scoring system. The total score of both lungs is given 25, and each lobe of the lung is given a maximum score of 5 according to the percentage of lung infections. If the total score is less than 7%, it is labeled as a mild disease, and if the score is more than greater18% it is labeled as a severe disease. This algorithm will triage the suspected patients with COVID-19 disease immediately into mild and severe lung disease and will help the physicians to give prompt treatment. Medical experts and radiologists use HRCT chest with 90% precision in spotting COVID-19 lung disease (Alom et al., 2020, Arunkumar et al., 2017, Baskar et al., 2013, Bezdan et al., 2022). In the peak scenario of COVID-19 cases, judging HRCT chest scans for a large number of patients posed a challenge. To address this, doctors utilized computer-aided diagnostic systems to assess CT images for COVID-19 diagnosis (Budak et al., 2021, Chen et al., 2020, Ewen and Khan, 2020). This accelerated the reporting process and eased the workload of radiologists during pandemic peaks. Acute COVID-19 lung disease has four stages on HRCT chest scans. Stage 1, the early stage, may show either normal lungs or Ground Glass Opacity (GGO). Stage II, the progressive stage, is measured by a rise in GGO. Stage III, the peak stage, displays progressive consolidation. Stage IV (occurring 14 or more days after symptom onset) displays a decrease in GGO and consolidation. In chronic cases, it may also present radiological signs of fibrosis, with traction bronchiectasis, thick parenchymal bands, and architectural distortion of the lungs (Exchange, 2018, Exchange, 2019, Fatima Bokhari et al., 2018). Our algorithm is more efficient in acute COVID-19 lung disease diagnosis as it is trained and tested for ground glass haze and consolidation. A three-dimensional CNN visual attention model is employed for COVID-19 classification, achieving an F1 score of 0.811 using clinically acquired CT imaging slices (Fisher, 1936). Several deep learning models, such as COVID-Net (Gomes et al., 2020), ResNets-18 (He, Zhang, Ren, & Sun, 2016), CNN (Hemdan, El-Shafai, & Sayed, 2023); and start migration-learning models (Huyut, 2023), are deployed for the COVID-19 investigation. Features are extracted based on the S-transform model and nourished into fuzzy-c-means and decision tree classifiers (Jangam, Barreto, & Annavarapu, 2022). Features resulting from VGG-16 are utilized to differentiate between normal and COVID pneumonia using CT slices, attaining an accuracy of 93.9% (Jin et al., 2021). A bilateral noise removal filter is applied to improve image quality. Subsequently, CNN is applied for COVID-19 detection with 94.5% accuracy (Kasinathan & Uyyala, 2021). Despite significant progress, precise detection of COVID-19 remains challenging due to aspects like image noise, diverse shapes, sizes, and complex textures (Kathamuthu et al., 2023, Khan et al., 2016).

The foremost contributions of this research are:

- The COVID-19-CNN model is proposed for classification based on the selection of a pre-trained ResNet-50 model and best features selection method after extensive experimentation performed on different models and feature selection methods. The best features selection method is selected after analysis of *t*-test, Bhattacharya, entropy, and Wilcoxon based on separability criteria. These methods are tested on different percentages of features such as 5%, 10%, 15%, 20%, 25%, and 30%. Out of these, Wilcoxon has outperformed having a minimum classification error rate on 20% features and selected for further experimentation.

- After classification, the COVID-19-YOLOv2 model is proposed for localization. In this model, features are extracted from the LeakyReLU (Activation7) layer of the pre-trained ONNX model and passed to the tinyYOLOv2 detector. The model is finally trained on the optimum hyperparameters that are selected after extensive experimentation.
- After localization, a 10-layer CNN-based lightweight model is proposed for semantic segmentation of COVID-19 and named COVID-19-CNN. This model is capable of detecting small COVID-19 lesions more accurately as compared to the latest existing works.

The article is structured as follows: Section II covers a discussion of related work, Section III explains the proposed models, Section IV, the results are presented and elaborated upon, and Finally, Section V discusses the conclusion of the article.

Section snippets

Related work

The method is proposed in which the UNet++ model is used as the backbone of pre-trained ResNet-50 that is trained on the 46,096 CT images of 106 patients and detects only COVID-19. This method has limitations because it cannot detect the mild, moderate, and severe COVID-19 stages (Chen et al., 2020). U-net with deep supervision, wide U-net, and U-net are used as segmentation models, and U-net with deep supervision outperformed the others. However, this technique still needs development to…

Proposed method

The proposed architecture contains three deep-learning models such as COVID-19-ResNet-50 for classification, COVID-19-YOLOv2 for localization, and COVID-19-CNN for segmentation.

In the proposed COVID-19-ResNet-50 model features are derived from a fully connected FC-1000 layer. For the best feature selection, the statistical analysis is performed on *t*-test, Bhattacharya, Wilcoxon, and entropy with different percentages of the features. The statistical analysis provided less classification error…

Experimental discussion

In this article, three separate experiments have been performed for classification, localization, and segmentation. The proposed technique is evaluated on a total of 3893 lung CT images which have been collected from national and international hospitals. The names given to the collected images from different Hospitals such as datasets I-V. The dataset-I is UCSD-AI4H/COVID-CT which contains 349 positive and 405 negative slices of 216 COVID-19 patients (Zhao, Zhang, He, & Xie, 2020). Dataset-II…

Discussion research findings (radiologists' point of view)

The proposed innovative method can be applied for real-time application for accurate classification. This model helps to identify infected cohorts more efficiently and speedily, which will aid in early triage and stratification. This model will provide an easy automatic diagnosis system, especially in developing countries with limited health care resources, insufficient diagnostic equipment, and trained radiologists. This AI tool in CT chest reporting will save time as well as provide…

Importance of this research work in medical applications

Radiologically, pure ground-glass opacity with up to 3 abnormalities<3cm in max diameter is considered a CT severity score of mild disease. If GGO represents more than 3 abnormalities>3cm in max diameter, then the CT severity score of the disease is moderate to severe. GGO mixed with consolidation and architectural distortion also indicates a CT severity score of moderate to severe disease. Different medicines are being used in Covid patients in many clinical trials worldwide to see the…

Conclusion

The method proposed consists of three models such as classification, localization, and segmentation that are integrated. The classified images are input to the localization model, and the localized images are provided to the segmentation model that finally supports segmenting the lesion regions more accurately. The images that have been classified are then fed into the localization model. Afterward, the localized images are fed into the segmentation model as input, which in turn enhances the…

Future work

In the future, there is potential to expand this research to encompass the grading of COVID-19 based on the combined percentages of infection present in all five lobes of both lungs. This provides help in the segregation of patients with healthy, mild, and severe diseases more accurately and precisely as it involves information on all five lobes.…

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.…

Recommended articles

References (51)

J. Chen *et al.*
Deep learning-based model for detecting 2019 novel coronavirus pneumonia on high-resolution computed tomography
Scientific reports (2020)

M. Huyut
Automatic detection of severely and mildly infected COVID-19 patients with supervised machine learning models
IRBM (2023)

Q. Jin *et al.*
Domain adaptation based self-correction model for COVID-19 infection segmentation in CT images
Expert Systems with Applications (2021)

N.D. Kathamuthu *et al.*
A deep transfer learning-based convolution neural network model for COVID-19 detection using computed tomography scan images for medical applications
Advances in Engineering Software (2023)

M.H. Nadimi-Shahraki *et al.*
Enhanced whale optimization algorithm for medical feature selection: A COVID-19 case study
Computers in biology and medicine (2022)

N. Prakash *et al.*
Deep transfer learning for COVID-19 detection and infection localization with superpixel based segmentation
Sustainable Cities and Society (2021)

J.H. Shah *et al.*
Facial expressions classification and false label reduction using LDA and threefold SVM
Pattern Recognition Letters (2020)

A.M. Tahir *et al.*
COVID-19 infection localization and severity grading from chest X-ray images
Computers in biology and medicine (2021)

Alom, M. Z., Rahman, M., Nasrin, M. S., Taha, T. M., & Asari, V. K. (2020). COVID_MTNet: COVID-19 detection with...

N. Arunkumar *et al.*
Classification of focal and non focal EEG using entropies
Pattern Recognition Letters (2017)

View more references

Cited by (0)

View full text