DEEP LEARNING-BASED ANALYSIS OF COVID-19 X-RAY IMAGES: A COMPARATIVE STUDY OF COLOURMAP

Che Daud, M.Z.[1*], Ahmad Zaiki, F.W.[1], Che Azemin, M.Z.[2]

[1]Department of Diagnostic Imaging and Radiotherapy, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia.
[2]Department of Optometry, Kulliyyah of Allied Health Sciences, International Islamic University Malaysia.

*Corresponding author: Mohd Zamzuri Che Daud, Department of Diagnostic Imaging and Radiotherapy, Kulliyah of Allied Health Science, International Islamic University Malaysia, Jalan Sultan Ahmad Shah, 25200 Kuantan, Pahang, MALAYSIA. Email: dzamzuri@iium.edu.my.

ABSTRACT

Background: *With the emergence of the SARS-CoV-2 virus late in 2019, the world's healthcare system has been severely affected by the COVID-19 pandemic, necessitating the need for quick and effective actions to reduce its extensive effects. Chest X-ray (CXR) imaging is critical for accurate assessment, displaying intricate lung structural abnormalities, including ground-glass opacities, consolidation, and bilateral infiltrates in COVID-19 patients. The objective of this study was to examine the comparison between grayscale and 16 colourmap images in terms of their efficacy in COVID-19 detection when used with the DarkNet-53 deep learning architecture.*

Methodology: *We conducted an experiment with a dataset of 9,665 CXRs, consisting of 7,134 normal images and 2,531 COVID-19 images, in order to train deep learning architectures. An additional dataset of 4,143 CXRs, with 3,058 normal and 1,085 COVID-19 images, was used for independent testing. The images underwent pre-processing and were split into grayscale and 16 colourmap images for individual examination. The COVID-19 detection task was fine-tuned on DarkNet-53, a deep learning architecture, with standard data augmentation techniques applied to grayscale and 16 colourmap images.*

Results: *The DarkNet-53 deep learning architecture demonstrated verifying results based on the X-ray image utilised. The bone colourmap achieved the highest accuracy (0.985) and sensitivity (0.952) scores, while the grayscale, pink, and summer colourmaps demonstrated the greatest specificity (0.998).*

Conclusion: *Our study highlights the importance of choosing the right type of X-ray image in association with deep learning architecture for CXR COVID-19 detection. These outcomes have important consequences for automating and upgrading CXR analysis, aiding in the exact detection of COVID-19 and respiratory health issues, and eventually benefiting patient care and outcomes.*

Keywords: *COVID-19, deep learning, convolutional neural network, colourmap*
Manuscript classification: *Research paper*

INTRODUCTION

Since the emergence of the novel coronavirus SARS-CoV-2 in late 2019, the COVID-19 pandemic has had a tremendous impact on global healthcare systems, necessitating immediate and effective measures to contain the virus and mitigate its devastating effects. The virus had rapidly spread across the globe, resulting in millions of infections, and fatalities. The ongoing crisis had placed immense strain on healthcare systems, making accurate and rapid diagnosis challenging. In this regard, chest X-ray (CXR) imaging has become a valuable tool in identifying and managing COVID-19 cases.

In response to this challenge, CXR imaging has emerged as a pivotal tool in the identification and management of COVID-19 cases. By revealing intricate structural abnormalities within the lungs, CXR imaging enables precise assessments of COVID-19 patients (Bernheim et al., 2020). This imaging technique has unveiled a spectrum of radiographic findings in COVID-19 cases, including ground-glass opacities, consolidations, and bilateral infiltrates (Rousan et al., 2020). The attributes of CXR imaging, such as affordability, accessibility, and straightforward implementation, render it especially invaluable in resource-constrained environments. The insightful information provided by CXR imaging aids healthcare professionals in making informed decisions regarding patient prioritising, treatment strategies, and the monitoring of therapeutic responses.

In recent years, the convergence of deep learning techniques in the analysis of medical images, including CXR, has gained remarkable prominence (Litjens et al., 2017). **Figure 1** demonstrates the relationship between artificial intelligence, machine learning, and deep learning. A subset of machine learning, deep learning employs artificial neural networks to unravel intricate patterns and features from extensive datasets. The advent of deep learning algorithms has unlocked the potential to harness the wealth of medical imaging data, thereby enhancing diagnostic precision, expediting workflows, and elevating patient care. In the realm of COVID-19, the application of deep learning methodologies to CXR images holds tremendous promise in fortifying the capabilities of healthcare practitioners to effectively combat the pandemic (Wang et al., 2020).

This paper introduces a comprehensive and comparative exploration of colour spaces within the context of deep learning-driven analysis of COVID-19 CXR images. Colour spaces are pivotal components in image analysis, significantly influencing the efficacy of deep learning algorithms (Ballester et al., 2022). Thus, it becomes imperative to meticulously investigate and assess the utility if diverse colour spaces in capturing and representing crucial information for precise classification and detection of COVID-19 within CXR images. Our research

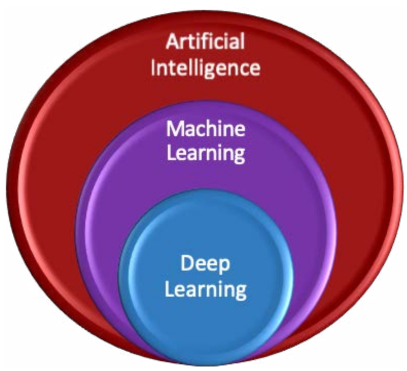

***Figure 1:** Relationship between artificial intelligence, machine learning, and deep learning*

investigates the application of deep learning methodologies in the domain of medical imaging, delivering invaluable insights into the optimised employment of colour spaces, specifically concerning the analysis of COVID-19 CXRs. Our primary aim is to enhance diagnostic methodologies and streamline the management of COVID-19 cases by harnessing the power of CXR imaging.

A critical consideration in the analysis of CXR images lies in the transformation of single-channel grayscale images into the three-channel configuration of the RGB colour space. Although CXR images are inherently monochromatic, the replication of the single-channel format into three channels, red, green, and blue has garnered traction in deep learning-driven analyses (Gui et al., 2022). This replication, transforming the single-channel grayscale images into the three-channel format, facilitates the utilisation of pre-trained models, and algorithms designed for RGB images. This extension of resources enriches the arsenal available for the analysis of COVID-19 CXRs (Nanni et al., 2023). Nonetheless, it is imperative to meticulously evaluate the repercussions of this replication process, as it may introduce superfluous information or artifacts that could potentially impact the performance and interpretability of deep learning models.

This study aims to assess the effects of colour space transformation and channel replication on the precision and reliability of COVID-19 identification and categorisation in CXR images. Our goal is to gain a thorough understanding of the use of colour spaces and channel replication methods so that deep learning-based analysis of COVID-19 CXR images can be more reliable and effective. By examining and examining in detail a range of colour spaces, we seek to gain a deeper insight into their benefits and shortcoming in recognising the unique radiographic characteristics indicative of COVID-19. This assessment will provide a roadmap for further research and development, directing us towards more accurate and reliable deep-learning models for the automated analysis of COVID-19 CXR images.

METHODOLOGY

In this study, we conducted an experiment with a dataset of 9,665 CXRs, consisting of 7,134 normal images and 2,531 COVID-19 images, in order to train learning architectures. An additional dataset of 4,143 CXRs, with 3,058 normal and 1,085 COVID-19 images, was used for independent testing. The CXR images were acquired from publicly available datasets (Chowdhury et al., 2020). We utilised purposive sampling because it was accessible to us. The images underwent pre-processing and were split into grayscale and 16 colourmap images for individual examination. The colourmap includes Jet, Bone, HSV, Cool, Cooper, Pink, Turbo, Flag, Parula, Colourcube, Winter, Summer, Spring, Autumn, Hot, and Lines. As seen in **Figure 2**, several types of colourmap CXR images are depicted. This study uses de-identified publicly available data, so no ethical considerations are needed.

The COVID-19 detection task was fine-tuned on the DarkNet-53 model, a deep learning architecture, with standard data augmentation techniques applied to grayscale and 16 colourmap images. The method used in this study comprised image pre-processing, transfer learning for classifying X-ray images (training and testing), and performance metrics evaluation. The system's entire working process is depicted in **Figure 3**.

All the image pre-processing steps analysis was executed on a computer having an operating system of Windows 11, NVDIA GeForce RTX 3060, Intel Core i7 12700 GPU 3.6 GHz, 16.0 GB RAM. The rest of the operations, including training and testing of the classification model, were executed on Matlab version R2022a.

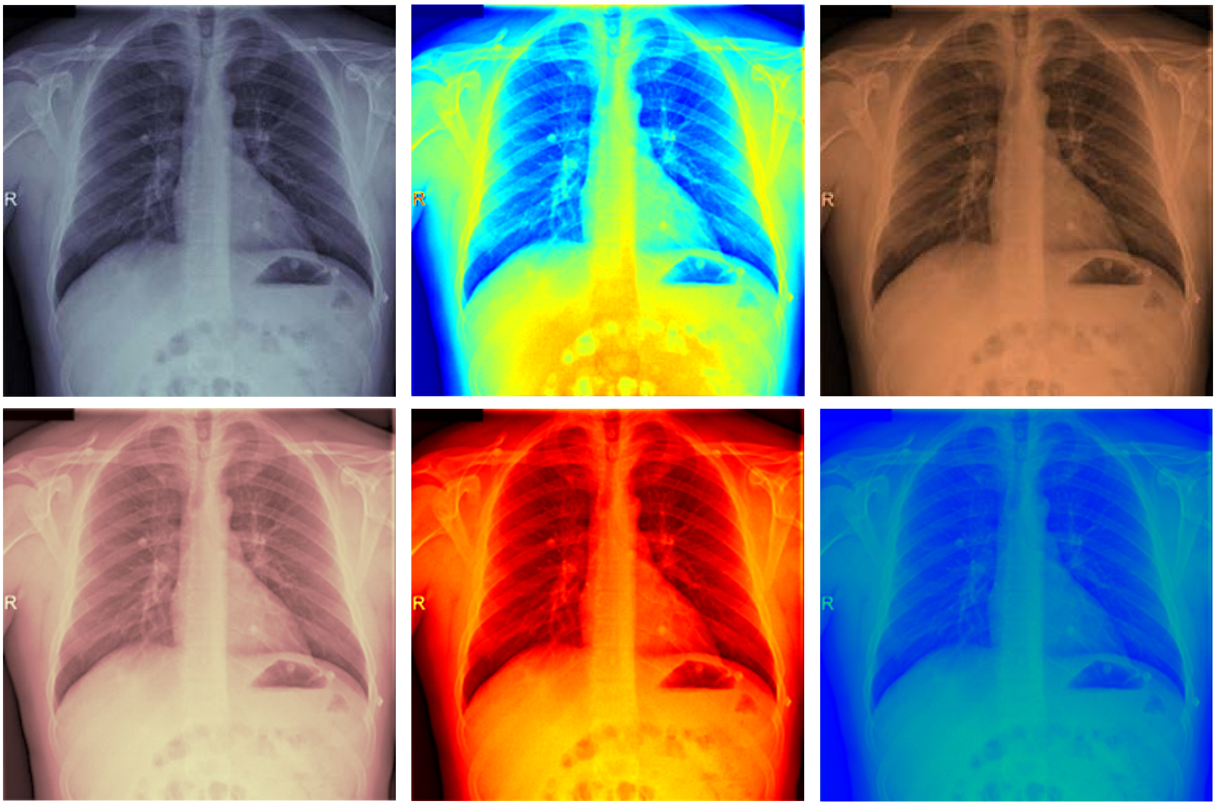

Figure 2: *(A) Bone Colourmap; (B) Parula Colourmap; (C) Copper Colourmap; (D) Pink Colourmap; (E) Hot Colourmap, and (F) Winter Colourmap (from left to right)*

In this study, the Stochastic Gradient Descent Momentum (sgdm) optimiser was applied to compile the classification model. Sgdm is a highly effective optimising algorithm. It has been utilised to update the learnable parameters. The learning rate was set to 0.0003. A batch size of 16 and six epochs was used when fitting the training data on the model.

This study evaluated various colourmaps for extracting discriminative features from COVID-19 chest radiographs using the DarkNet-53 neural architecture. Four key performance metrics were assessed - true positive rate (sensitivity), true negative rate (specificity), false positive rate, and false negative rate. Overall accuracy was also evaluated.

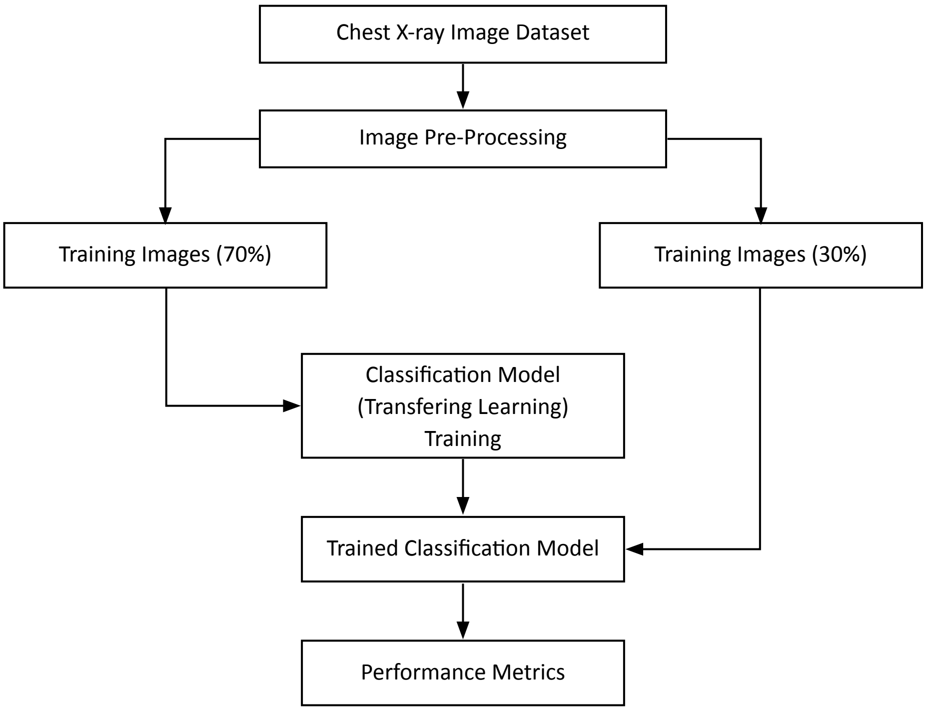

Figure 3: *The flow of data collection*

RESULTS

This study evaluated the impact of various colourmap transformations on COVID-19 classification performance using the DarkNet-53 deep convolutional neural network. The evaluation was conducted on a dataset of CXRs including conformed COVID-19 viral pneumonia cases and normal CXRs. Performance metrics of overall accuracy, sensitivity, and specificity were computed (refer to **Table 1**).

Table 1: *Accuracy, Sensitivity, and Specificity of DarkNet-53 on COVID-19 images*

Colourmap	Accuracy	Sensitivity	Specificity
Grayscale	0.984	0.945	0.998
Jet	0.974	0.912	0.996
Bone	0.985	0.952	0.996
HSV	0.960	0.854	0.997
Cool	0.971	0.906	0.994
Copper	0.983	0.949	0.995
Pink	0.969	0.888	0.998
Turbo	0.977	0.929	0.993
Flag	0.889	0.638	0.978
Parula	0.963	0.867	0.997
Colourcube	0.940	0.927	0.944
Winter	0.966	0.881	0.997
Summer	0.981	0.934	0.998

cont.

*cont. **table 1***

Colourmap	Accuracy	Sensitivity	Specificity
Spring	0.975	0.921	0.994
Autumn	0.967	0.883	0.997
Hot	0.981	0.940	0.996
Lines	0.911	0.764	0.963

The bone colourmap yielded the optimal accuracy of 98.5%, with a sensitivity of 95.2% and a specificity of 99.6%. This indicates the bone colourmap enabled robust discrimination of COVID-19 pneumonia from normal CXRs. The grayscale (accuracy 98.4%, sensitivity 94.5%, specificity 99.8%), copper (accuracy 98.3%, sensitivity 94.9%, specificity 99.5%), hot (accuracy 98.1%, sensitivity 94.0%, specificity 99.6%), and summer (accuracy 98.1%, sensitivity 93.4%, specificity 99.8%) colourmaps also achieved accuracy over 98.0%.

In contrast, the flag and lines colourmaps performed significantly worse, with accuracy of 88.9% and 91.1%, respectively. Sensitivity was low at 63.8% and 76.4% for these colourmaps, suggesting poorer identification of true COVID-19 cases. Specificity also dropped below 98.0% for both flag and lines colourmaps, indicating higher false positive rates.

Our findings demonstrate the critical importance of colourmap selection in developing reliable deep-learning systems for COVID-19 detection. Particular colourmaps like flags and lines lead to suboptimal model performance and should be avoided. Careful colourmap evaluation on target datasets is essential to maximise discrimination of COVID-19 viral pneumonia from normal CXRs.

DISCUSSIONS

The study's findings reveal significant strengths and limitations in the application of various colourmaps for COVID-19 detection from medical images. The Bone colourmap emerged as the most robust performer, achieving an impressive classification accuracy of 98.5%. Notably, it demonstrated a remarkable true positive rate of 95.2% and negligible false negative rate of 4.8%, highlighting its capability to accurately identify both positive and negative COVID-19 cases. This was complemented by a true negative rate of 99.6%, maintaining a low false positive rate of 0.4%. Such exceptional performance underscores the Bone colourmap's efficacy in maximising correct classifications.

Similarly, the grayscale images exhibited strong results, with notable true positive (94.5%) and true negative (99.8%) rates. However, its marginally higher false negative rate compared to Bone suggests the latter's superior sensitivity. Across all metrics, Bone outperformed Grayscale, boasting a 0.1% higher accuracy and a 0.7% higher sensitivity, despite a 0.2% lower specificity. Conversely, the study identified clear limitations in the Flag and Lines colourmaps, which displayed reduced accuracy and elevated false classification rates. The Flag colourmap's sensitivity was notably low at 63.8%, indicating a failure to detect a considerable number of positive COVID-19 cases. Similarly, Lines suffered from a high false negative rate of 23.6%. Both colourmaps exhibited lower true negatives and higher false positives compared to Bone and Grayscale. The study attributes the poorer performance of Flag and Lines to the absence of pseudo-colouring, which may result in less distinct feature representations. The advantageous colour encoding of Bone likely contributes to more discerning COVID-19 indicators.

In light of these findings, the study suggests the need for further research to elucidate the specific mechanisms driving performance. Additionally, the study acknowledges that its analysis was confined to a specific set of colourmaps.

Table 2: *Comparison of the accuracy performance for COVID-19 detection studies*

References	Dataset	Methodology	Accuracy (%)
This research	X-ray image	Transfer learning	98.4
Chakraborty et al., (2022)	X-ray image	Transfer learning	97.1
Kumar Sethy et al., (2020)	X-ray image	Deep learning and SVM	95.3
Heidari et al., (2020)	X-ray image	Transfer learning	94.5
Pathak et al., (2022)	X-ray image	Transfer learning	93.0
Oh et al., (2020)	X-ray image	Statistical approach	88.9
Wang et al., (2021)	CT scan	Transfer learning	85.2

Table 2 compared previous studies of COVID-19 detection that employed grayscale medical images (Chakraborty et al., 2022; Heidari et al., 2020; Kumar Sethy et al., 2020; Oh et al., 2020; Pathak et al., 2022; Wang et al., 2021). This table also outlines the methodology for all the tasks. This research obtained the highest accuracy of 98.4% among the referenced studies by utilising the grayscale images dataset.

CONCLUSION

In conclusion, the Bone colourmap and grayscale proved optimal, whilst Flag and Lines were significantly inferior. The colour encoding enhanced discrimination of COVID-19 morphological patterns. The usage of optimal colourmaps will enhance diagnostic accuracy and expedite clinical workflows. By selecting the right colourmaps, the DarkNet-53 framework can be optimised to better COVID-19 detection. These findings hold significant implications for automating and improving CXR analysis, aiding in the detection of COVID-19, and ultimately benefiting patient care and outcomes.

ACKNOWLEDGEMENT

This research was supported by the Ministry of Higher Education – SLAB and the International Islamic University Malaysia scholarship.

CONFLICT OF INTEREST

The authors declare that there is no conflict of interest in this research.

REFERENCES

Ballester, C., Bugeau, A., Carrillo, H., Clément, M., Giraud, R., Raad, L., Vitoria, P., 2022. Influence of Color Spaces for Deep Learning Image Colorization. arXi.

Bernheim, A., Mei, X., Huang, M., Yang, Y., Fayad, Z.A., Zhang, N., Diao, K., Lin, B., Zhu, X., Li, K., Li, S., Shan, H., Jacobi, A., Chung, M., 2020. Chest CT findings in coronavirus disease 2019 (COVID-19): Relationship to duration of infection. Radiology 295(3), 200463. doi: 10.1148/radiol.2020200463.

Chakraborty, S., Paul, S., Hasan, K.M.A., 2022. A Transfer Learning-Based Approach with Deep CNN for COVID-19- and Pneumonia-Affected Chest X-ray Image Classification. SN Comput Sci 3, 17.

Chowdhury, M.E.H., Rahman, T., Khandakar, A., Mazhar, R., Kadir, M.A., Mahbub, Z. Bin, Islam, K.R., Khan, M.S., Iqbal, A., Emadi, N. Al, Reaz, M.B.I., Islam, M.T., 2020. Can AI Help in Screening Viral and COVID-19 Pneumonia? IEEE Access 8, 132665–132676.

Gui, T., Zhan, W., Gu, X., Hu, J., 2022. Adversarial Image Colorization Method Based on Semantic Optimization and Edge Preservation. Electronics (Switzerland) 11, 3006.

Heidari, M., Mirniaharikandehei, S., Khuzani, A.Z., Danala, G., Qiu, Y., Zheng, B., 2020. Improving the performance of CNN to predict the likelihood of COVID-19 using chest X-ray images with preprocessing algorithms. Int J Med Inform 144, 104284.

Kumar Sethy, P., Kumari Behera, S., Kumar Ratha, P., Biswas, P., 2020. Detection of coronavirus Disease (COVID-19) based on Deep Features and Support Vector Machine. Preprints (Basel) 2020030300.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A.W.M., van Ginneken, B., Sánchez, C.I., 2017. A survey on deep learning in medical image analysis. Med Image Anal 42, 60-88. doi: 10.1016/j.media.2017.07.005

Nanni, L., Faldani, G., Brahnam, S., Bravin, R., Feltrin, E., 2023. Improving foraminifera classification using Convolutional Neu-ral Networks with Ensemble Learning. Preprints (Basel) 2023020396.

Oh, Y., Park, S., Ye, J.C., 2020. Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets. IEEE Trans Med Imaging 39, 2688–2700.

Pathak, Y., Shukla, P.K., Tiwari, A., Stalin, S., Singh, S., 2022. Deep Transfer Learning Based Classification Model for COVID-19 Disease. Ingenierie et Recherche Biomedicale 43, 87.

Rousan, L.A., Elobeid, E., Karrar, M., Khader, Y., 2020. Chest x-ray findings and temporal lung changes in patients with COVID-19 pneumonia. BMC Pulm Med 20.

Wang, L., Lin, Z.Q., Wong, A., 2020. COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. Sci Rep 10, 19549.

Wang, S., Kang, B., Ma, J., Zeng, X., Xiao, M., Guo, J., Cai, M., Yang, J., Li, Y., Meng, X., Xu, B., 2021. A deep learning algorithm using CT images to screen for Corona virus disease (COVID-19). Eur Radiol 31, 6096.