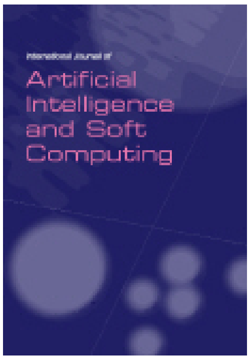


International Journal of Artificial Intelligence and Soft Computing

ISSN online: 1755-4969 - ISSN print: 1755-4950
https://www.inderscience.com/ijaisc

Optimisation of spatial-exploitation CNN models through hyperparameter-tuning and human-in-the-loop combination

Luke Beveridge, Keshav Dahal

DOI: 10.1504/IJAISC.2024.10064253

Article History:

Received:	17 December 2022
Last revised:	19 December 2023
Accepted:	04 January 2024
Published online:	04 July 2024

Optimisation of spatial-exploitation CNN models through hyperparameter-tuning and human-in-the-loop combination

Luke Beveridge* and Keshav Dahal

School of Computing Engineering and Physical Sciences,
University of the West of Scotland,
Paisley, PA1 2BE, UK
Email: luke.beveridge@uws.ac.uk
Email: keshav.dahal@uws.ac.uk
*Corresponding author

Abstract: Spatial-exploitation convolutional neural networks (CNNs) have a simplified architecture compared to other CNN models. However, devices with limited computational resources could struggle with processing spatial-exploitation CNNs. To address this, we investigate two methods to optimise spatial-exploitation CNN models for time efficiency and classification accuracy: hyperparameter-tuning, and human-in-the-loop (HITL). We apply grid-search to optimise the hyperparameter space, whilst HITL is used to identify whether the time-to-accuracy relationship of the optimised model can be improved. To show the versatility of combining the two methods, CIFAR-10, MNIST, and Imagenette are used as model input. This paper contributes to spatial-exploitation CNN optimisation by combining hyperparameter-tuning and HITL. Results show that this combination improves classification accuracy by 1.47–2.34% and reduces the time taken to conduct this task by 27–28%, depending on dataset. We conclude that combining hyperparameter-tuning and HITL are a viable approach to optimise spatial-exploitation CNNs for devices with limited computational resources.

Keywords: deep learning; convolutional neural network; CNN; image classification; hyperparameter-tuning; human-in-the-loop; HITL.

Reference to this paper should be made as follows: Beveridge, L. and Dahal, K. (2024) 'Optimisation of spatial-exploitation CNN models through hyperparameter-tuning and human-in-the-loop combination', *Int. J. Artificial Intelligence and Soft Computing*, Vol. 8, No. 2, pp.147–158.

Biographical notes: Luke Beveridge holds an MRes in AI from the University of the West of Scotland (UWS) and is a PhD candidate in the UWS Creative Computing Technologies Research Group. His research focuses on design processes for extended reality applications and how these are applied to industrial training, with an interest in serious games and AI to support learning and content generation.

Keshav Dahal is a Professor of Intelligent Systems and the Director of the Artificial Intelligence, Visual Communication and Network (AVCN) Research Institute at the University of the West of Scotland (UWS), UK. Before joining

UWS, he was with Bradford and Strathclyde Universities in UK. He obtained his PhD and Master from Strathclyde. His research interests lie in the areas of applied AI, blockchain, intelligent systems, trust/security modelling in distributed systems, and scheduling/optimisation problems. He has published extensively with award winning papers, and has sat on organising/program committees of over 60 international conferences including as the General Chair and Programme Chair. He is a senior member of the IEEE.

1 Introduction

In deep learning (DL), image classification has achieved great strides in recent years. In particular, advancements in image analysis and key features identification allow to categorise a collection of images appropriately with a high level of classification accuracy (Hou et al., 2021). Spatial-exploitation convolutional neural networks (CNNs) (Bhatt et al., 2021), a subclass of CNNs (LeCun et al., 1989), are ideal for tasks related to image classification due to their simplified architecture.

Spatial-exploitation CNNs originate from LeCun's seminal work (LeCun et al., 1989). LeCun's simple architecture presents a small number of parameters, as such it could be considered as the first low complexity spatial-exploitation CNN.

However, modern spatial-exploitation CNNs present different architectures depending on the type of model that is selected, introducing varying levels of model complexity (Bhatt et al., 2021). Complex architectures pose the problem of an increase in time complexity (Simonyan and Zisserman, 2015; He et al., 2016; Xin et al., 2020; Tuba et al., 2021), accuracy (Joshi et al., 2021), and power consumption (Stamoulis et al., 2018). This problem is particularly true for computational systems with limited processing power (Şen and Özkurt, 2020), as classification tasks for large datasets on these devices are either prevented to be completed or require a considerable amount of time (Xin et al., 2020). To address this, we explore combining optimisation methods such as hyperparameter tuning (Liashchynskyi and Liashchynskyi, 2019) and human-in-the-loop (HITL) (Wu et al., 2022).

To show the versatility of combining the two methods, several heterogeneous datasets (Parvin and Hasan, 2020; Huang et al., 2017), such as CIFAR-10 (Krizhevsky, 2009), MNIST (LeCun et al., 1989), and Imagenette (Deng et al., 2009), are used in this paper as input for the optimised spatial-exploitation CNN model. Each dataset used within the study contains images that are either grey-scale or RGB (Makantasis et al., 2015). This variety allows to test the robustness of the optimised spatial-exploitation CNN model, as well as showing its potential applications. The CIFAR-10 dataset is used to train the initial CNN model, whilst the MNIST and Imagenette datasets demonstrate the applicability of the optimised model to different use cases.

There are various spatial-exploitation CNN models that have their roots in LeCun's work (LeCun et al., 1989) such as AlexNet (Krizhevsky et al., 2017), ResNet50 (Liu et al., 2020) and GoogLeNet (Szegedy et al., 2015) – otherwise known as Inception (Szegedy et al., 2015). We selected the AlexNet model to apply our optimisation technique to show how it can benefit different fields, due to the wide range of its applications, such as synthetic aperture radar (SAR) imagery classification (Wang et al., 2021), alcoholism detection (Wang et al., 2019), pathological brain detection (Luo et al., 2019), and lung cancer detection (Agarwal et al., 2021).

For the purpose of this paper hyperparameter tuning will be focused on attributes within each of the layers, such as the number of convolution filters, number of epochs, learning rate value, and the dropout rate in fully connected layers (Berg and Hjelmervik, 2021; Vrskova et al., 2021). The quantity of layers contained by the spatial-exploitation CNN model architecture will remain constant as well as the selected activation function, ReLU, due to its already known proficiency (Krizhevsky et al., 2017).

We found that hyperparameter tuning proves to be beneficial in terms of significant time complexity reduction of the spatial-exploitation CNN model, at the cost of classification accuracy. The application of HITL helps to maintain the time performance whilst striving to regain classification accuracy. To achieve this, the HITL process exclusively optimises the learning rate hyperparameter dynamically during run-time (Usman et al., 2021).

This paper contributes to spatial-exploitation CNNs by exploring the combination of deep learning techniques and a layer-wise filter-pruning approach using the AlexNet model. A novel technique is proposed for tuning the learning rate hyperparameter through human interaction in real-time during model testing.

The next section explores related literature, followed by a description of the methodology used to optimise AlexNet as well as data handling procedures. The experimentation conducted on the new approach is then detailed, followed by the results obtained from it. We then conclude with remarks outlining the impact of the techniques used within this paper for the CNN model optimisation.

2 Related work

Wang et al. (2021) highlight that the amount of time required to perform classification tasks in CNN is an issue. To address time related issues in spatial exploitation CNNs, deep learning techniques such as hyperparameter tuning (Joshi et al., 2021) and human-in-the-loop (Wu et al., 2022) are used. These optimisation techniques can be used on spatial exploitation CNNs to enhance metrics such as time (Tuba et al., 2021), classification accuracy (Joshi et al., 2021), and power consumption (Stamoulis et al., 2018).

CNN models often contain multiple hyperparameters such as, number of layers, number of filters, learning rate, dropout, and stride values (Krizhevsky et al., 2017; Bhatt et al., 2021). Hyperparameter tuning allows to optimise a CNN model in a localised way by focusing on specific parameters of the CNN model.

Usman et al. (2021) and Budd et al. (2021) demonstrate that HITL can be applied to both the input and output stages of a CNN model for optimisation, whilst Kumar et al. (2016) theorises how human intervention within the training execution loop can also optimise the CNN model.

In this paper, HITL explores Kumar et al.'s (2016) theory with a focus on dynamically optimising the learning rate hyperparameter during model training with a human intervening during testing run-time of the CNN model. The purpose of this is to explore if human intervention at this point of the mode life-cycle can improve classification accuracy whilst maintaining the time efficiency.

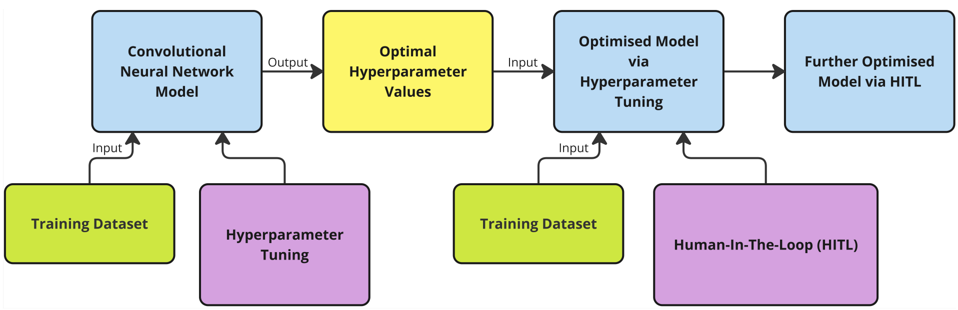
Combining the deep learning techniques of hyperparameter tuning and HITL is important because while hyperparameter tuning improves spatial exploitation CNN time efficiency, it comes to the cost of classification accuracy. In this paper we show that, by adding HITL techniques to the optimisation process, it is possible to bring back classification accuracy of the model while keeping the efficiency introduced by the hyperparameter tuning technique.

3 Research methodology

An iterative approach was taken in this paper combined with Şen and Özkurt (2020) comparative approach, comparing the standard AlexNet CNN model with an hyperparameter tuned version of itself (Berg and Hjelmervik, 2021; Vrskova et al., 2021).

A mapping of how the spatial exploitation CNN model AlexNet is gradually optimised can be seen in Figure 1, identifying the core components that are explained in detail throughout this paper.

Figure 1 An overview of the how deep learning techniques are used to optimise a CNN model (see online version for colours)



3.1 Data collection and results gathering

The initial draw of focus within this portion of the research project looks towards identifying sufficient datasets for use within the experimental part of the study. For appropriate datasets to be selected they required to be image-orientated primarily for application in the CNN models. Then datasets that were selected differentiated in terms of the number of channels pertaining to the selection of images which each possessed; as more explicitly detailed in the related work section (Huang et al., 2017; Akshay et al., 2020; Makantasis et al., 2015) – allowing for the model produced to be applied in more instances.

An integral portion of the studies data collection takes the form of the results produced in the experimentation conducted, as identified previously. These results will be collected from the standard un-tuned AlexNet model as well as the optimised model that has undergone hyperparameter-tuning. Throughout this optimisation process deep learning techniques are employed as the hyperparameter search space is detailed, allowing for a grid search to be applied (Fraccaroli et al., 2021) through the use of the Keras-tuner library (Abdelminaam et al., 2021).

3.2 Data analysis and interpretation

Due to the chosen search algorithm for the optimisation process being grid search, this allows for each individual hyperparameter to be examined separately and in turn identify their specific impact on the model performance. This analysis aids in isolating statistics for the parameters in the search space that benefited more performance the more than others as well as noticing redundant parameters in the search space that could not be optimised any further.

Once data has been collated from both of the aforementioned models responsible for processing the previously mentioned datasets, this will then allow for improved interpretation of the results retrieved as comparisons will be able to be conducted on the resultant datasets. The key values which would be inspected to achieve this comparison would be the time complexity and the classification accuracy across the number of epochs that each model will process.

Presenting what the findings mean from the analysis portion of the study is the key objective when interpreting the data comparison. After both models have been looked at comparatively, there should be a clear insight as to how the hyperparameters have either aided or hindered performance, allowing for the optimised model to be refined further in a process that intends to be iterative in nature and thus relating to the methodology outlining the project as a whole.

3.3 Hyperparameter tuning

To conduct hyperparameter tuning a tuning framework is required, for this paper Keras-tuner was selected due to its common use when applying the hyperparameter tuning deep learning technique (Manaswi, 2018) and for its availability of search algorithms.

Search algorithms found within the Keras tuning framework are differing methods of hyperparameter tuning. The algorithms that are commonly utilised within CNN optimisation vary between grid search, random search, hyperband, and Bayesian optimisation (Putatunda and Rama, 2018; Liashchynskyi and Liashchynskyi, 2019; Joshi et al., 2021). Grid search is the conventional method of hyperparameter optimisation that conducts an exhaustive search over the entire hyperparameter search space; the search space is the hyperparameters selected for tuning. Due to the exhaustive nature of this approach, it can be slower than the alternatives. Despite the slower time to tune the hyperparameters, grid search is appropriate to use as it looks at all combinations within the hyperparameter search space. This ensures the most optimal combination within the search space is found, unlike random search that has a chance to overlook the optimal combination.

In terms of how the tuning was carried out once a framework and search algorithm were decided upon, consideration was given towards the work conducted by Tajbakhsh et al. (2016). This focused on full training vs. fine-tuning in a layer-wise approach, and ultimately is the reasoning behind showing this is a viable method of tuning. This contributed to the choice to utilise a layer-wise hyperparameter tuning in this paper.

3.4 Human-in-the-loop discussion

The framework that has been discussed thus far is met with the addition of the HITL machine learning technique. The purpose of this implementation was to understand whether it was possible to improve classification accuracy. This was whilst maintaining the previously enhanced time complexity of the CNN model that was achieved from hyperparameter tuning. As seen later in the paper, although the model time complexity improved, there was a slight decline in the final classification accuracy. HITL intends to address this recognised deficit.

The work of Usman et al. (2021) where implementation of a HITL learning process is involved in CNN model enhancement draws parallels to the work in this paper. However, inclusion of the human in this HITL process is focused during the run-time of the CNN model instead. The design of this process involved monitoring the learning rate hyperparameter whilst a dynamic learning rate was applied, altering after each epoch. The value of the learning rate could either be incremented or decremented after each epoch. Ultimately, a decaying learning rate (Goodfellow et al., 2016) proved to give the only recognisable benefit. The issue that had to be addressed at this stage was understanding what method of decaying the learning rate was best suited to the CNN model. For this to be achieved, learning rates were decremented in various implementations when applied to the CNN model. During run-time of the model the human in the process would identify whether values either became stagnant or began to have negative impact upon the classification accuracy or time complexity of the model. The human would know this by observing the resultant time and classification accuracy values given after each epoch.

4 Experimentation and results

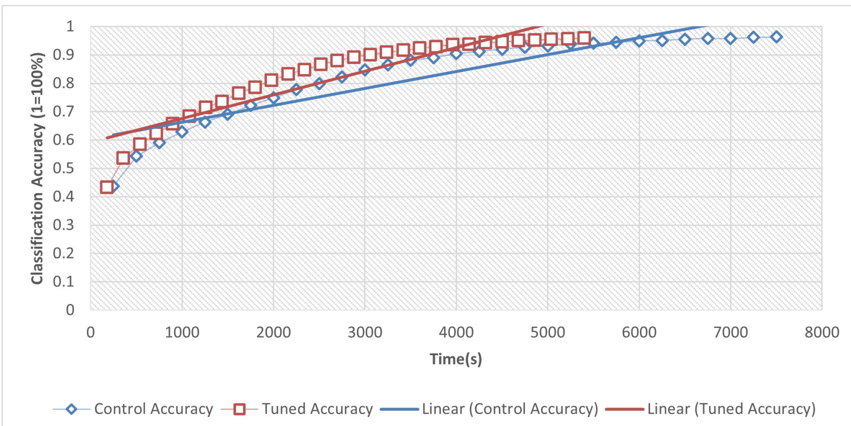
Preliminary experimentation was conducted to observe the effect of hyperparameter tuning upon time and accuracy for each individual hyperparameter applied. Table 1 displays the obtained results showing the values for hyperparameters of dropout rate, learning rate and number of filters. It should be noted that the initial dropout rate value of 0.4 was taken from an original AlexNet study (Guo et al., 2017). After tuning, this value remained consistent from this paper and therefore was not included in Table 1 containing each optimisation from the hyperparameter search space. The results are incrementally shown to illustrate the individual effect each had throughout the optimisation process.

After the preliminary results were explored and the optimal values for differing hyperparameters were obtained, this then allowed for a more detailed exploration of the changed model architecture. Figure 2 illustrates a comparison over 30 epochs between a hyperparameter tuned version of the AlexNet model and its equivalent AlexNet model control experiment which has no modification through hyperparameter tuning. The experimentation of each of the discussed models were carried out over the course of 30 epochs and the hyperparameter value of the number of filters for each model and their layers was of focus at this point. After already previously identifying an optimal hyperparameter value for dropout rate and learning rate. A linear trendline was applied to the results obtained from each model, where it is seen that the relation between accuracy and time is more favoured within the tuned AlexNet model.

Table 1 The values for hyperparameters during CNN model optimisation, whilst observing classification accuracy and time complexity

CNN model variations	*Epochs*	*Time*	*Acc (%)*	*LR*	*DR*
Standard AlexNet model as control	5	24 m 47 s	63.8%	0.001	0.4
Optimised learning rate	5	20 m 20 s	66.1%	0.0001	0.4
Optimised layer 1 filter (64)	5	19 m 40 s	66.3%	0.0001	0.4
Optimised layer 2 filter (224)	5	17 m 54 s	66.1%	0.0001	0.4
Optimised layer 3 filter (288)	5	17 m 30 s	65.6%	0.0001	0.4
Optimised layer 4 filter (288)	5	15 m 45 s	65.2%	0.0001	0.4
Optimised layer 5 filter (224)	5	15 m	65.3%	0.0001	0.4

Figure 2 Comparison of hyperparameter tuned and untuned version of the AlexNet model with CIFAR-10 dataset (see online version for colours)



The overall impact of the hyperparameter tuning allowed for a time complexity optimisation of 28% whilst only losing a marginal amount of classification accuracy; the resultant being less than a single percentile. Similarly, when applying this enhanced CNN model to other datasets it was found that the same trend of time reduction to run all epochs continued despite the alteration to the dataset and the respective quantity of channels it possessed. Figure 3 illustrates the optimised time for the MNIST dataset whilst Figure 4 does similar for the ImageNet subset, Imagenette. The optimised times were 28.38% and 27.21% respectively whilst maintaining less than a loss of less than 1% to classification accuracy.

This implies that through following the methodology applied within this paper less computationally demanding model architectures can be achieved with the application of deep learning techniques such as hyperparameter tuning. As a result, the learning rate optimisation aided in increased classification accuracy with a slight improvement to time taken to run the model, whilst the reduction to the number of filters in each layer significantly aided in the time complexity of each model decreasing. This is effectively applying a method known as pruning (Li et al., 2017), in which the redundant parameters have been removed from the model therefore allowing for compression of the CNN model consequently.

Figure 3 The classification accuracy for applying the tuned and preliminary CNN model on MNIST dataset (see online version for colours)

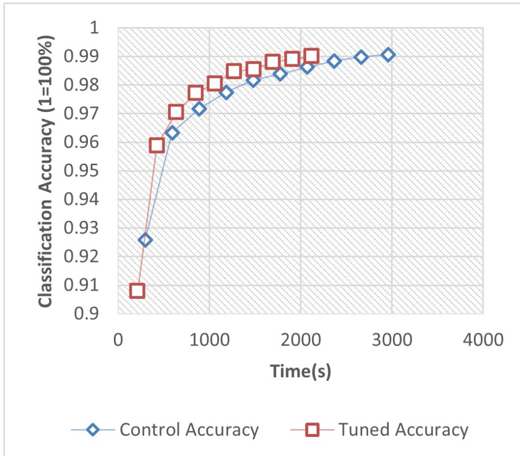
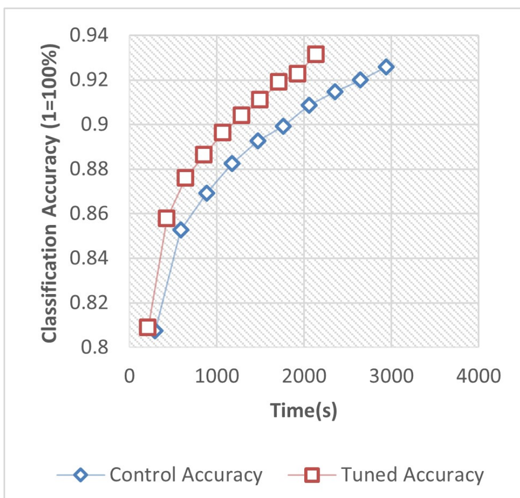


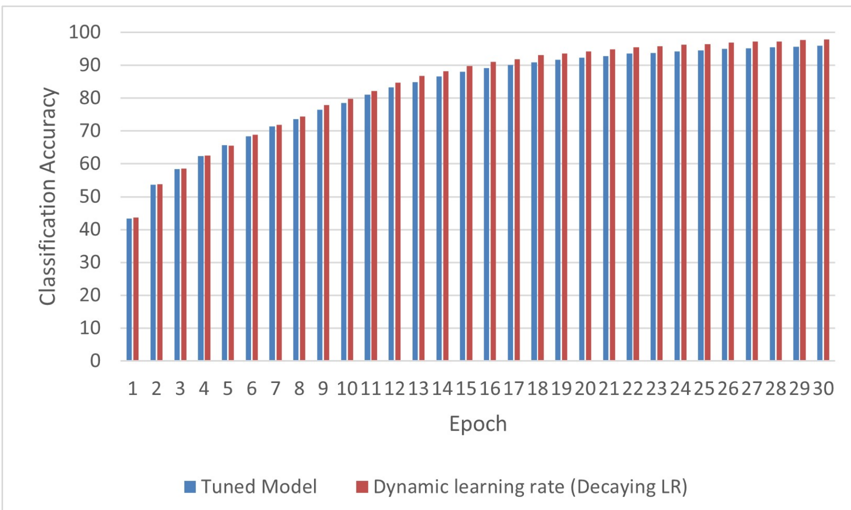
Figure 4 The classification accuracy for applying the tuned and preliminary CNN model on Imagenette dataset (see online version for colours)



Human-in-the-loop (HITL) learning was implemented in an attempt to further the previously mentioned gained efficiencies, specifically this was focused on the optimisation of the learning rate hyperparameter. The approach taken to achieve this was by applying learning rate schedules to the model, allowing for the learning rate to dynamically alter during the training process (Goodfellow et al., 2016). This is where the HITL technique becomes more apparent as this was required to identify the drawbacks or benefits to using certain learning rate schedules.

It was discovered that incrementing the learning rate over time negatively impacted the classification accuracy regardless of the delta between each epoch. In contrast, when decrementing the learning rate, otherwise known as a decaying learning rate, this seen negative effect as well. This occurred when the delta was too large or when the decay transitioned back to a constant. However, there was one case where benefit was seen to implementing a decaying learning rate. Specifically, the decaying value was derived from the beginning learning rate divided by the quantity of epochs that the model would execute; this caused for a small gradual decay between each epoch. As seen in Figure 5, the model that had the dynamic learning rate attained a higher classification accuracy compared to the previous tuned model without a dynamic learning rate. Almost every epoch experienced an increase to classification accuracy between 1.47–2.34% whilst not experiencing any alteration to the total time taken to run the model.

Figure 5 The classification accuracy comparison between the previously tuned model alongside a version of the tuned model with a dynamic learning rate (see online version for colours)



5 Conclusions

It can be noticed that a brute force approach was taken in the hyperparameter tuning segment of this paper, due to the usage of a grid search algorithm. This allows for the ideal of a less computationally intensive convolutional neural network to be discovered whilst also positively impacting upon the execution time of the model by a significant degree. Importantly, this was achieved as a level of classification accuracy was maintained. This is a success within itself as there is normally a trade-off between sacrificing classification accuracy at the cost of improving time. The latter half of the study that involved HITL learning took the previous success and allowed for classification accuracy to be increased, whilst keeping the previously reduced time complexity at a constant. This has an overwhelmingly positive impact on the outcome of the CNN model architecture as both classification accuracy and the time complexity benefited from the techniques implemented. Considering the applications of the optimised model in conjunction with the different datasets that are applicable for use, it can be concluded that there is a strong potential for the findings to be discovered that pertain to a more diverse fashion for moving ahead with future work.

Acknowledgements

Thank you to the generosity of the Scottish International Education Trust and Glasgow Educational and Marshall Trust who partially funded my research. A special thanks to Dr Marco Gilardi for his advice and feedback throughout this project.

References

Abdelminaam, D.S., Ismail, F.H., Taha, M., Taha, A., Houssein, E.H. and Nabil, A. (2021) 'CoAID-DEEP: an optimized intelligent framework for automated detecting COVID-19 misleading information on Twitter', *IEEE Access*, Vol. 9, pp.27840–27867, DOI: 10.1109/ACCESS.2021.3058066.

Agarwal, A., Patni, K. and Rajeswari, D. (2021) 'Lung cancer detection and classification based on AlexNet CNN', *Proceedings of the 6th International Conference on Communication and Electronics Systems (ICCES)*, Coimbatore, India, pp.1390–1397, DOI: 10.1109/ICCES51350.2021.9489033.

Akshay, S., Mytravarun, T.K., Manohar, N. and Pranav, M.A. (2020) 'Satellite image classification for detecting unused landscape using CNN', *Proceedings of the International Conference on Electronics and Sustainable Communication Systems (ICESC)*, Coimbatore, India, pp.215–222, DOI: 10.1109/ICESC48915.2020.9155859.

Berg, H. and Hjelmervik, K.T. (2021) 'Deep learning on active sonar data using Bayesian optimization for hyperparameter tuning', *Proceedings of the 25th International Conference on Pattern Recognition (ICPR)*, Milan, Italy, pp.6546–6553, DOI: 10.1109/ICPR48806.2021.9412347.

Bhatt, D., Patel, C., Talsania, H., Patel, J., Vaghela, R., Pandya, S., Modi, K. and Ghayvat, H. (2021) 'CNN variants for computer vision: history, architecture, application, challenges and future scope', *Electronics*, Vol. 10, No. 20, p.2470, DOI: 10.3390/electronics10202470.

Budd, S., Robinson, E.C. and Kainz, B. (2021) 'A survey on active learning and human-in-the-loop deep learning for medical image analysis', *Medical Image Analysis*, Vol. 71, p.102062, DOI: 10.1016/j.media.2021.102062.

Deng, J., Dong, W., Socher, R., Li, L.J., Li, K. and Fei-Fei, L. (2009) 'ImageNet: a large-scale hierarchical image database', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Miami, FL, USA, pp.248–255, DOI: 10.1109/CVPR.2009.5206848.

Fraccaroli, M., Lamma, E. and Riguzzi, F. (2021) 'Automatic Setting of DNN hyper-parameters by mixing Bayesian optimization and tuning rules', *Proceedings of the International Conference on Machine Learning, Optimization, and Data Science*, Siena, Italy, pp.477–488, DOI: 10.1007/978-3-030-64583-0_43.

Goodfellow, I., Bengio, Y. and Courville, A. (2016) *Deep Learning*, MIT Press, Cambridge, MA.

Guo, T., Dong, J., Li, H. and Gao, Y. (2017) 'Simple convolutional neural network on image classification', *Proceedings of the IEEE 2nd International Conference on Big Data Analysis (ICBDA)*, Beijing, China, pp.721–724, DOI: 10.1109/ICBDA.2017.8078730.

He, K., Zhang, X., Ren, S. and Sun, J. (2016) 'Deep residual learning for image recognition', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, pp.770–778, DOI: 10.1109/CVPR.2016.90.

Hou, Q., Zhou, D. and Feng, J. (2021) 'Coordinate attention for efficient mobile network design', *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, Nashville, TN, USA, pp.13713–13722, DOI: 10.1109/CVPR46437.2021.01350.

Huang, C., Liang, B., Li, W. and Han, S. (2017) 'A convolutional neural network architecture for vehicle logo recognition', *Proceedings of the IEEE International Conference on Unmanned Systems (ICUS)*, Beijing, China, pp.282–287, DOI: 10.1109/ICUS.2017.8278355.

Joshi, S., Owens, J. A., Shah, S. and Munasinghe, T. (2021) 'Analysis of preprocessing techniques, Keras Tuner, and transfer learning on Cloud Street image data', *Proceedings of the IEEE International Conference on Big Data (Big Data)*, Orlando, FL, USA, pp.4165–4168, DOI: 10.1109/BigData52589.2021.9671878.

Krizhevsky, A. (2009) *Learning Multiple Layers of Features from Tiny Images*, Unpublished Master's thesis, University of Toronto, Toronto, Canada [online] https://www.cs.toronto.edu/~kriz/ (accessed 9 April 2023).

Krizhevsky, A., Sutskever, I. and Hinton, G.E. (2017) 'ImageNet classification with deep convolutional neural networks', *Communications of the ACM (CACM)*, Vol. 60, No. 6, pp.84–90, DOI: 10.1145/3065386.

Kumar, A., McCann, R., Naughton, J. and Patel, J.M. (2016) 'Model selection management systems: the next frontier of advanced analytics', *ACM SIGMOD Record*, Vol. 44, No. 4, pp.17–22, DOI: 10.1145/2935694.2935698.

LeCun, Y., Jackel L.D., Boser, B., Denker, J.S., Graf H.P., Guyon, I., Henderson, D., Howard, R.E. and Hubbard, W. (1989) 'Handwritten digit recognition: applications of neural network chips and automatic learning', *IEEE Communications Magazine*, Vol. 27, No. 11, pp.41–46, DOI: 10.1109/35.41400.

LeCun, Y., Boser, B., Denker, J.S., Henderson, D., Howard, R.E., Hubbard, W. and Jackel, L.D. (1989) 'Backpropagation applied to handwritten zip code recognition', *Neural Computation*, Vol. 1, No. 4, pp.541–551, DOI: 10.1162/neco.1989.1.4.541.

Li, H., Kadav, A., Durdanovic, I., Samet, H. and Graf, H.P. (2017) 'Pruning filters for efficient convnets', *Proceedings of the 5th International Conference on Learning Representations (ICLR)*, Toulon, France.

Liashchynskyi, P. and Liashchynskyi, P. (2019) *Grid Search, Random Search, Genetic Algorithm: A Big Comparison for NAS* [online] http://arxiv.org/abs/1912.06059 (accessed 9 April 2023).

Liu, W., Li, W. and Gong, W. (2020) 'Ensemble of fine-tuned convolutional neural networks for urine sediment microscopic image classification', *IET Computer Vision*, Vol. 14, No. 1, pp.18–25, DOI: 10.1049/iet-cvi.2018.5829.

Lu, S., Lu, Z. and Zhang, Y.D. (2019) 'Pathological brain detection based on AlexNet and transfer learning', *Journal of Computational Science*, Vol. 30, pp.41–47, DOI: 10.1016/j.jocs.2018.11.008.

Makantasis, K., Karantzalos, K., Doulamis, A. and Doulamis, N. (2015) 'Deep supervised learning for hyperspectral data classification through convolutional neural networks', *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS)*, Milan, Italy, pp.4959–4962, DOI: 10.1109/IGARSS.2015.7326945.

Manaswi, N.K. (2018) 'Understanding and working with Keras', *Deep Learning with Applications Using Python*, pp.31–43, Apress, Berkeley, CA, DOI: 10.1007/978-1-4842-3516-4_2.

Mehta, S., Hajishirzi, H. and Rastegari, M. (2022) 'DiCENet: dimension-wise convolutions for efficient networks', *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol. 44, No. 5, pp.2416–2425, DOI: 10.1109/TPAMI.2020.3041871.

Parvin, F. and Hasan, M.A.M. (2020) 'A comparative study of different types of convolutional neural networks for breast cancer histopathological image classification', *Proceedings of the IEEE Region 10 Symposium (TENSYMP)*, Dhaka, Bangladesh, pp.945–948, DOI: 10.1109/TENSYMP50017.2020.9230787.

Putatunda, S. and Rama, K. (2018) 'A comparative analysis of hyperopt as against other approaches for hyper-parameter optimization of XGBoost', *Proceedings of the 2018 International Conference on Signal Processing and Machine Learning*, New York, NY, USA, pp.6–10, DOI: 10.1145/3297067.3297080.

Şen, S.Y and Özkurt, N. (2020) 'Convolutional neural network hyperparameter tuning with Adam optimizer for ECG classification', *Proceedings of the IEEE Conference on Innovations in Intelligent Systems and Applications (ASYU)*, Istanbul, Turkey, pp.1–6, DOI: 10.1109/ASYU50717.2020.9259896.

Simonyan, K. and Zisserman, A. (2015) *Very Deep Convolutional Networks for Large-Scale Image Recognition*, https://doi.org/10.48550/arXiv.1409.1556.

Stamoulis, D., Cai, E., Juan, D.C. and Marculescu, D. (2018) 'Hyperpower: Power-and memory-constrained hyper-parameter optimization for neural networks', *Proceedings of the Design, Automation & Test in Europe Conference & Exhibition (DATE)*, Dresden, Germany, pp.19–24, DOI: 10.23919/DATE.2018.8341973.

Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D., Erhan, D., Vanhoucke, V. and Rabinovich, A. (2015) 'Going deeper with convolutions', *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, MA, USA, pp.1–9, DOI: 10.1109/CVPR.2015.7298594.

Tajbakhsh, N., Shin, J.Y., Gurudu, S.R., Hurst, R.T., Kendall, C.B., Gotway, M.B. and Liang, J. (2016) 'Convolutional neural networks for medical image analysis: full training or fine tuning?', *IEEE Transactions on Medical Imaging*, Vol. 35, No. 5, pp.1299–1312, DOI: 10.1109/TMI.2016.2535302.

Tuba, E., Bačanin, N., Strumberger, I. and Tuba, M. (2021) 'Convolutional neural networks hyperparameters tuning', in Pap, E. (Ed.): *Artificial Intelligence: Theory and Applications*, Springer, Cham, pp.65–84, DOI: 10.1007/978-3-030-72711-6_4.

Usman, M., Carie, A., Marapelli, B., Bedru, H.D. and Biswas, K. (2021) 'A human-in-the-loop probabilistic CNN-fuzzy logic framework for accident prediction in vehicular networks', *IEEE Sensors Journal*, Vol. 21, No. 14, pp.15496–15503, DOI: 10.1109/JSEN.2020.3023661.

Vrskova, R., Sykora, P., Kamencay, P., Hudec, R. and Radil, R. (2021) 'Hyperparameter tuning of ConvLSTM network models', *Proceedings of the 44th International Conference on Telecommunications and Signal Processing (TSP)*, Brno, Czech Republic, pp.15–18, DOI: 10.1109/TSP52935.2021.9522683.

Wang, S.H., Xie, S., Chen, X., Guttery, D.S., Tang, C., Sun, J. and Zhang, Y.D. (2019) 'Alcoholism identification based on an AlexNet transfer learning model', *Frontiers in Psychiatry*, Vol. 10, p.205, DOI: 10.3389/fpsyt.2019.00205.

Wang, X., Liu, J., Zhang, S., Deng, Q., Wang, Z., Li, Y. and Fan, J. (2021) 'Detection of oil spill using SAR imagery based on AlexNet model', *Computational Intelligence and Neuroscience 2021*, Vol. 2021, pp.1–14, DOI: 10.1155/2021/4812979.

Wu, X., Xiao, L., Sun, Y., Zhang, J., Ma, T. and He, L. (2022) 'A survey of human-in-the-loop for machine learning', *Future Generation Computer Systems*, Vol. 135, pp.364–381, DOI: 10.1016/j.future.2022.05.014.

Xin, R., Zhang, J. and Shao, Y. (2020) 'Complex network classification with convolutional neural network', *Tsinghua Science and Technology*, Vol. 25, No. 4, pp.447–457, DOI: 10.26599/TST.2019.9010055.