Optimizing Face Embedding Sizes and Accuracy in Facial Recognition Systems

Wui Chung Alton Chau, Hao Feng Chan, Dexter Sing Fong Leong, Shakir Hussain Naushad Mohamed
School of Engineering, Deakin University, Australia

Andi Prademon Yunus
Telkom University, Indonesia

Takao Ito
Hiroshima University, Japan

Zheng Cai, Xinjie Deng, Yit Hong Choo*
Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University, Australia
Email: y.choo@deakin.edu.au

Abstract

Face recognition technology is integral to security, access control and identity verification in finance, healthcare and transportation. It protects personal data, secures online transactions, controls access to areas, and helps prevent identity theft. This paper proposes a novel hybrid optimization algorithm, Moss Particle Swarm Optimization (MPSO) to perform hyperparameter tuning, aiming to identify the neural network, alongside Triple Loss metrics for efficient recognition. The proposed method is tested with the Labeled Faces in the Wild (LFW) dataset, demonstrating its effectiveness in improving facial recognition systems.

Keywords: Face Recognition, Optimization, Hyperparameter Tuning, Embedding Sizes, Deep Learning

1. Introduction

Face recognition technology has emerged as a cornerstone of the computer vision industry, with diverse applications spanning mobile phone security, identity verification, and access control [1]. The performance of these systems relies heavily on the optimal representation of facial features, which involves complex trade-offs between computational efficiency and recognition accuracy [2]. This balance is achieved through the careful tuning of model configurations, particularly embedding sizes and hyperparameters, which determine how well a model generalizes to unseen data. By refining these aspects, face recognition systems can achieve robust performance, safeguarding sensitive personal data and ensuring reliable operations across industries such as finance, healthcare, and transportation [3].

Achieving state-of-the-art performance in face recognition systems often involves leveraging deep learning architectures like FaceNet [4], ResNet [5], and VGGFace [6]. These models employ sophisticated loss functions such as Center Loss [7], Cross-Entropy Loss [8], and Triplet Loss [9] to learn discriminative facial feature embeddings. However, beyond architectural design, hyperparameter tuning plays a pivotal role in optimizing these systems. Hyperparameters define the configuration before training and influence both the training process and the final predictive performance [10]. Effective hyperparameter optimization ensures that models not only perform well on training data but also generalize to challenging real-world scenarios. In practical applications, these models must be capable of accurately recognizing faces under a wide variety of conditions such as low lighting, partial occlusions, and changes in facial expression without requiring significant retraining. This level of robustness can only be achieved through the optimization of both the model architecture and the hyperparameters. Hyperparameter Tuning has also increasingly been tackled using metaheuristic algorithms due to their speed, flexibility, and ability to explore complex and high dimensional search spaces efficiently. Metaheuristic algorithms are particularly suited for such tasks due to their flexibility and ability to navigate large, complex search spaces that traditional methods struggle with [11]. By applying techniques like Genetic Algorithms (GA) [12], Whale Optimization Algorithm (WOA) [13] and Harris' Hawks Optimization (HHO) [14], researchers can fine-tune deep learning models, ensuring that they maintain high performance across a range of real-world conditions. These metaheuristic algorithms are particularly useful in machine learning for neural architecture search [15], hyperparameter tuning, and feature selection. Thanathamathee et al. [16] used grid search and nested cross-validation to enhance facial and masked facial recognition, while Ozcan et al. [17] applied Particle Swarm Optimization (PSO) [18] in conjunction with transfer learning to improve expression recognition systems. The emergence of advanced algorithms like Moss Growth Optimization (MGO) [19] further highlights the importance of exploring innovative approaches for hyperparameter tuning, as these techniques have demonstrated success in solving complex optimization problems across various domains. However, choosing the right algorithm is challenging due to problem-specific characteristics and performance metrics. The complexity of modern machine learning models makes manual tuning impractical and increases the risk of suboptimal solutions. Thus, efficient and robust hyperparameter tuning methods are essential to improve model performance.

This study introduces a new approach, Hybrid Moss Particle Swarm Optimization (HMPSO). In Section 2, a review on metaheuristic algorithms, hyperparameter tuning on facial recognition systems are provided. In Section 3, the proposed hybridization of PSO and MGO, into HMPSO, is explained in detail. In Section 4, the effectiveness of HMPSO in enhancing the accuracy of the ResNet-18 by hyperparameter tuning is evaluated, compared, discussed. Concluding remarks for future work are presented in Section 5.

2. Deep Learning CNN model

Deep Convolutional Neural Networks (DCNNs) are widely used in computer vision tasks to identify patterns in images and videos. In facial recognition and image processing, DCNNs like ResNet-18 [5] shown in Figure 1 can effectively identify and extract features from facial images, making them a robust solution for tasks such as facial recognition.

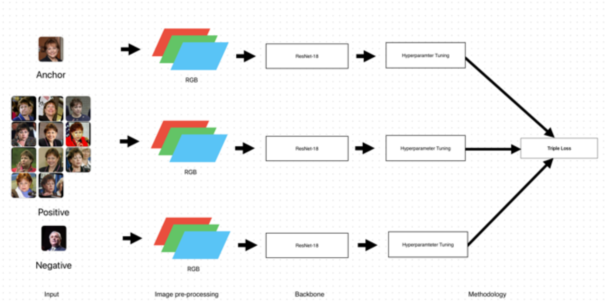


Figure 1 The Structure of the ResNet-18 with Hyperparameter Tuning and Triple Loss Metrics

The Labelled Faces in the Wild (LFW) dataset is divided into Anchor, Positive, and Negative triplets. Each image is resized to a new resolution in RGB colors. The ResNet-18 model generates embedding vectors for each image in the triplet, and hyperparameter tuning is used to optimize the embedding size and other parameters. These embeddings are then passed to the Triplet Loss function for results.

2.1 Facial Image Pre-processing

Image pre-processing involves transforming raw image data into a usable and meaningful format, eliminating unwanted distortions and enhancing specific features critical for computer vision applications. It serves as an essential step in preparing image data before inputting it into machine learning models. Instead of using grayscale images, RGB (Red, Green, Blue) color images are typically employed, as they provide a more detailed and comprehensive representation of facial features. In RGB images, each pixel is represented by three intensity values corresponding to the red, green, and blue color channels, enabling the capture of a broader range of colors and subtle variations in facial features.

2.2 ResNet-18

The ResNet-18 architecture is a lightweight deep neural network ideal for face recognition tasks. It balances high accuracy with computational efficiency, making it suitable for resource-constrained environments. ResNet-18 extracts face embeddings through optimized convolutional layers, which are then used in a triplet loss function to ensure embeddings of the same identity are closer in feature space. Its compact design and strong performance make it well-suited for real-world face recognition applications.

2.3 Embeddings and Triple Loss Function

Face embeddings represent facial features as vector arrays for comparison via distance, similarity, or search. This study examines embedding sizes between 64 and 256. Smaller embeddings reduce latency, while larger ones improve accuracy but require more resources. Using ResNet-18 and the triplet loss function, the model optimizes the embedding space for facial recognition. Triplet loss ensures embeddings of the same identity are closed while keeping different identities apart, is expressed as:

$$L = \max\left(d(a,p) - d(a,n) + m, 0\right) \qquad (1)$$

where a, p, and n are the anchor, positive, and negative samples respectively; d(a,p) and d(a,n) are the distance functions; and m is the margin ensuring sufficient separation. This setup enhances identity discrimination for verification and identification tasks. The lightweight ResNet-18 architecture ensures efficient training and embedding optimization. Experiments focus on minimizing triplet loss while maintaining high accuracy, optimizing performance across embedding sizes for scalable and effective facial recognition systems.

2.4 LFW Dataset

The Labelled Faces in the Wild dataset (LFW) [20] is designed for studies in unconstrained face recognition. It comprises over 13,000 labelled face images, including 1,680 individuals with multiple distinct photos, all collected from the web. For this study, 2,950 selected images were selected to evaluate the face recognition performance of the ResNet-18. The original image resolution of these images 250x250 are resized into a smaller resolution for the computational efficiency of the embedding size chosen by the ResNet-18 model.

3. Methodology

The proposed Hybrid Moss Growth Particle Swarm Optimization (HMPSO) algorithm enhances the optimization performance of the ResNet-18 by combining key elements of Particle Swarm Optimization (PSO) and Moss Growth Optimization (MGO). In Figure 2, it demonstrates the structure of the HMPSO. The PSO component provides global search capabilities by guiding particles using personal and global best positions, while MGO introduces adaptive growth dynamics to explore diverse regions and prevent premature convergence. By combining PSO's convergence ability with MGO's adaptability, HMPSO balances exploration and exploitation, improving search efficiency and solution quality. This hybrid approach enhances robustness and flexibility in navigating complex optimization landscapes, making it effective for finding optimal configurations for the ResNet-18 model.

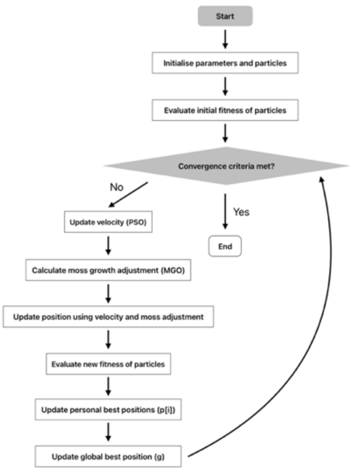


Figure 2 Flowchart of the Hybrid Moss Particle Swarm Optimization (HMPSO)

3.1 Velocity Update

Particle Swarm Optimization (PSO) is used to control the particle velocities and ensure effective exploration of the search space. PSO employs the concept of inertia weights, cognitive factors, and social factors to guide each particle based on its personal best position and the global best position found by the swarm. The velocity update equation in PSO helps particles move through the search space, allowing them to converge toward an optimal solution over time. This is given by the following equation:

$$v_i(t+1) = w \cdot v_i(t) + c_1 \cdot r_1 \cdot \left(p_i(t) - x_i(t)\right) + c_2 \cdot r_2 \cdot \left(g(t) - x_i(t)\right) \quad (2)$$

Where $v_i(t)$ is the velocity of particle i at time t, $x_i(t)$ is the position of particle i at time t, w is the inertia weight, c_1, c_2 are the cognitive and social scaling factors, r_1, r_2 are the random numbers between 0 and 1, $p_i(t)$ is the personal best position of particle I, g(t) is the global best position across all particles at time t. This equation drives the particles to explore the search space effectively and converge toward the best-found solutions, maintaining a balance between exploration and exploitation.

3.2 Moss Growth Update

The addition of moss growth dynamics introduces an additional element of growth adjustment in the position update. This adjustment is based on a dynamic model that simulates the growth of moss over time. The moss growth model enables particles to experience an additional growth force that alters their position, enhancing both the exploration and convergence processes. The moss growth adjustment is given by the following formula:

$$\Delta x_i = \frac{K}{1 + (\frac{K - x_i}{x_i})e^{rt}} - x_i \quad (3)$$

Where Δx_i is the moss growth adjustment for the particle I, K is the constant related to the moss growth model, r is the rate constant associated with moss growth dynamics, t is the time or iteration index, e^{rt} is the exponential growth term, simulating moss growth over time, x_i is the current position of particle i. This equation introduces an exponential growth effect that alters the particle's position, enabling enhanced exploration of the search space by simulating the natural expansion of moss.

3.3 Position Update Rule

The overall position update in Hybrid MPSO combines the classical PSO position update with the moss growth adjustment. This results in the following combined position update equation:

$$x_i(t+1) = x_i(t) + v_i(t+1) + \Delta x_i \quad (4)$$

Where $x_i(t+1)$ is the updated position of particle i at time t+1, $v_i(t+1)$ is the updated velocity based on PSO equation, Δx_i is the moss growth adjustment based on the moss growth model. This combined position update rule ensures that particles are guided both by their velocities, based on their personal and global best positions, and by the moss growth dynamics, which provide an additional layer of exploration.

4. Experimental Results and Discussion

The result demonstrates the effectiveness of the proposed novel approach, Hybrid Moss Particle Swarm Optimization (HMPSO) by comparing with the Particle Swarm Optimization (PSO) and the Moss Growth Optimization (MGO). As shown in Table 1, 2 and 3, each optimization runs 3 times to get 3 different sets of best hyperparameters. And Figure 3, 4 and 5 shows the training history of loss and accuracy of PSO, MGO and HMPSO respectively. In addition, comparing the average accuracies between HMPSO, PSO and MGO shown in Table 4, we conducted that HMPSO has better sets of hyperparameters chosen and achieve the highest average accuracy with a 97.23%, PSO the second with 94.82 % and MGO the last with 85.71%. According to Figure 5, the first run of HMPSO has the reached 100% accuracy during training in the last epoch a few times, implies that the model is performing well.

According to Figure 5, the training and validate accuracy vs 20 epochs graph, shows that the HMPSO has reached above 90% accuracy a few times during the last epochs when training the validating the model, as well as PSO shown in Figure 3 also achieved above 90% accuracies during training and validating but HMPSO is slightly a bit more consistent than PSO, as the second run of PSO drops to an accuracy of 86.59%. According to Figure 5, the training and validation loss versus 20 epochs graph shows a clear trend of the ResNet-18 has been learning effectively. The training loss starts at a relatively high value of 0.3920, steadily decreasing to 0.0316 by the 20th epoch, demonstrating the consistent improvement in fitting the training data. Similarly, the validation loss follows a complementary trend, indicating that the model is not only reducing errors on the training set but is also generalizing well to unseen data. This alignment between training and validation loss trends highlights that the optimization process is functioning effectively, with no apparent overfitting or underfitting during the training period.

As a result, the Hybrid Moss Particle Swarm Optimization (HMPSO) outperforms the Particle Swarm Optimization (PSO) and the Moss Growth Optimization (MGO), achieving the highest average accuracy of **97.23%** and demonstrating greater consistency across runs. The training and validation loss trends confirm effective learning and generalization without overfitting, highlighting HMPSO's potential as a robust hyperparameter optimization method for deep learning tasks.

Table 1 Hyperparameter Tuning and Accuracy Optimization performed by PSO.

Particle Swarm Optimizer (PSO)	Run 1	Run 2	Run 3
Embedding Size	146	224	100
Margin	0.7437	1.0	0.4006
Batch Size	46	98	95
Learning Rate	0.0001	0.01	0.0003
Loss	**0.0495**	**0.4187**	**0.0161**
Accuracy (%)	**98.63**	**86.59**	**99.25**
Avg Accuracy (%)		**94.82**	

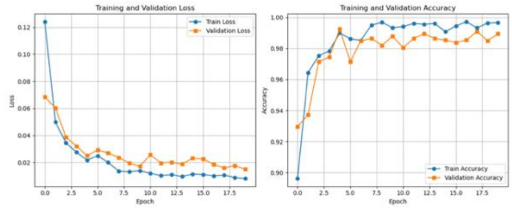


Figure 3 Train and Validate Loss vs Epochs graph (Left), Train and Validate vs Epochs graph (Right) of PSO.

Table 2 Hyperparameter Tuning and Accuracy Optimization performed by MGO

Moss Growth Optimizer (MGO)	Run 1	Run 2	Run 3
Embedding Size	22	72	197
Margin	0.2908	0.9989	0.4819
Batch Size	32	16	83
Learning Rate	0.0069	0.0057	0.0069
Loss	**0.0967**	**0.0734**	**0.1863**
Accuracy (%)	**86.77**	**87.03**	**83.33**
Avg Accuracy (%)		**85.71**	

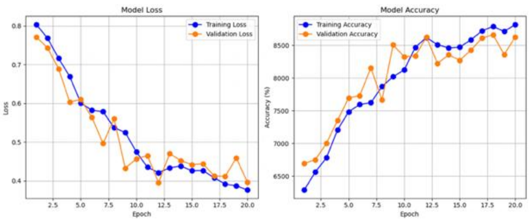


Figure 4 Train and Validate Loss vs Epochs graph (Left), Train and Validate vs Epochs graph (Right) of MGO

Table 3 Hyperparameter Tuning and Accuracy Optimization performed by HMPSO

Hybrid Moss Particle Swarm Optimizer (HMPSO)	Run 1	Run 2	Run 3
Embedding Size	141	64	224
Margin	0.8030	0.9753	0.3226
Batch Size	49	16	102.3274
Learning Rate	0.0001	0.0001	0.0022
Loss	**0.0592**	**0.1134**	**0.0454**
Accuracy (%)	**99.22**	**98.12**	**94.35**
Avg Accuracy (%)		**97.23**	

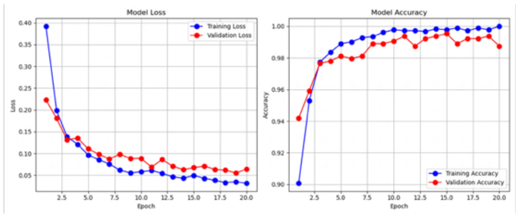


Figure 5 Train and Validate Loss vs Epochs graph (Left), Train and Validate vs Epochs graph (Right) of HMPSO

Table 4 Comparison of Average Accuracy in PSO, MGO and HMPSO with ResNet-18 model

Optimizers	PSO	MGO	HMPSO
Avg Accuracy (%)	94.82	85.71	**97.23**

5. Conclusion

In conclusion, the HMPSO algorithm introduced in this study represents a significant advancement in optimizing hyperparameters for deep learning models, specifically applied to the ResNet-18 architecture for face recognition. By combining PSO and MGO, HMPSO creates a powerful hybrid approach that enhances the accuracy and efficiency of hyperparameter tuning. It focuses on key hyperparameters such as face embedding sizes, the margin from the triple loss function, batch size, and learning rate, which are crucial for model performance. The experimental results demonstrated that HMPSO significantly outperforms traditional optimization methods, showing better accuracy and computational efficiency on the LFW dataset. This combination of PSO's global search and MGO's local search offers a robust solution, improving security, identity verification, and biometric authentication in real-world applications. Additionally, metaheuristic algorithms like PSO and MGO play a crucial role in solving complex optimization problems by exploring large solution spaces more effectively. This hybrid method shows potential for expanding into other domains such as natural language processing, computer vision, and reinforcement learning, where fine-tuning hyperparameters is essential. Overall, this research emphasizes the value of innovative optimization strategies for improving deep learning models, particularly in security and identity verification systems, paving the way for further advancements in various fields requiring intelligent decision-making and pattern recognition.

References

1. D. Kothadiya, A. Chaudhari, R. Macwan, K. Patel, and C. Bhatt, "The convergence of deep learning and computer vision: smart city applications and research challenges," in 3rd International Conference on Integrated Intelligent Computing Communication & Sec
2. R. Archana and P. E. Jeevaraj, "Deep learning models for digital image processing: a review," *Artificial Intelligence Review,* vol. 57, no. 1, p. 11, 2024.
3. I. Dijmărescu, M. Iatagan, I. Hurloiu, M. Geamănu, C. Rusescu, and A. Dijmărescu, "Neuromanagement decision making in facial recognition biometric authentication as a mobile payment technology in retail, restaurant, and hotel business models," Oeconomia C
4. F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR),* pp. 815-823, 2015.
5. K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 770-778.
6. O. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *BMVC 2015-Proceedings of the British Machine Vision Conference 2015*, 2015: British Machine Vision Association.
7. Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," in Computer vision–ECCV 2016: 14th European conference, amsterdam, the netherlands, October 11–14, 2016, proceedings, part VII 14, 2016: Springer
8. Z. Zhang and M. R. Sabuncu, "Generalized cross entropy loss for training deep neural networks with noisy labels," presented at the Proceedings of the 32nd International Conference on Neural Information Processing Systems, Montréal, Canada, 2018.
9. I. Haider, H.-J. Yang, G.-S. Lee, and S.-H. Kim, "Robust human face emotion classification using triplet-loss-based deep CNN features and SVM," *Sensors,* vol. 23, no. 10, p. 4770, 2023.
10. T. Yu and H. Zhu, "Hyper-parameter optimization: A review of algorithms and applications," *arXiv preprint arXiv:2003.05689,* 2020.
11. F. Peres and M. Castelli, "Combinatorial optimization problems and metaheuristics: Review, challenges, design, and development," *Applied Sciences,* vol. 11, no. 14, p. 6449, 2021.
12. A. Sohail, "Genetic algorithms in the fields of artificial intelligence and data sciences," *Annals of Data Science,* vol. 10, no. 4, pp. 1007-1018, 2023.
13. Z. Cai, Y. H. Choo, V. Le, C. P. Lim, and M. Liao, "Enhancing the Whale Optimisation Algorithm with sub-population and hybrid techniques for single-and multi-objective optimisation," *Soft Computing,* vol. 28, no. 5, pp. 3941-3971, 2024.
14. Y. H. Choo, Z. Cai, V. Le, M. Johnstone, D. Creighton, and C. P. Lim, "Enhancing the Harris' Hawk optimiser for single- and multi-objective optimisation," *Soft Computing,* vol. 27, no. 22, pp. 16675-16715, 2023/11/01 2023, doi: 10.1007/s00500-023-08952-w.
15. T. Elsken, J. H. Metzen, and F. Hutter, "Neural architecture search: a survey," *J. Mach. Learn. Res.,* vol. 20, no. 1, pp. 1997–2017, 2019.
16. P. Thanathamathee, S. Sawangarreerak, P. Kongkla, and D. N. M. Nizam, "An optimized machine learning and deep learning framework for facial and masked facial recognition," *Emerging Science Journal,* vol. 7, no. 4, pp. 1173-1187, 2023.
17. T. Ozcan and A. Basturk, "Static facial expression recognition using convolutional neural networks based on transfer learning and hyperparameter optimization," *Multimedia Tools and Applications,* vol. 79, no. 35, pp. 26587-26604, 2020.
18. J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of ICNN'95-international conference on neural networks*, 1995, vol. 4: ieee, pp. 1942-1948.
19. B. Zheng, Y. Chen, C. Wang, A. A. Heidari, L. Liu, and H. Chen, "The moss growth optimization (MGO): concepts and performance," *Journal of Computational Design and Engineering,* vol. 11, no. 5, pp. 184-221, 2024, doi: 10.1093/jcde/qwae080.
20. G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled Faces in the Wild: A Database forStudying Face Recognition in Unconstrained Environments," presented at the Workshop on Faces in 'Real-Life' Images: Detection, Alignment, and Recognition, Ma

Authors Introduction

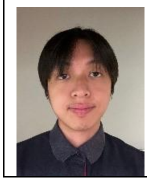
Mr. Hao Feng Chan

He is currently an undergraduate student majoring in Bachelor of Mechatronics Engineering (Honours) at Deakin University, Australia.

Mr. Shakir Hussain Naushad Mohamed

He is currently an undergraduate student majoring in Bachelor of Electrical and Electronics Engineering (Honours) at Deakin University, Australia.

Mr. Dexter Sing Fong Leong



He is currectly an undergrade pursuing Bachelors Of Mechatronic Engineering (Honours) in Deakin University, Australia.

Mr. Chau Wui Chung Alton

He is currently an undergraduate student majoring in Bachelor of Mechanical Engineering (Honours) at Deakin University, Australia.

Mr. Zheng Cai

He is currently a PhD candidate at Deakin University's Institute for Intelligent Systems Research and Innovation (IISRI). His research interests include multiobjective optimisation algorithms such as metaheuristic algorithms and evolutionary algorithms for scheduling problems. He is also exploring the integration of machine learning with optimisation algorithms.

Ms Deng Xinjie

She is pursuing a PhD in Information Technology at the Institute for Intelligent Systems Research and Innovation (IISRI), Deakin University, with her research ed on creating lightweight deep learning algorithms for computer vision applications.

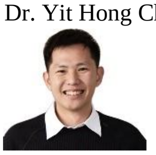
Dr. Andi Prademon，Yunus, Ph.D

He is an Assistant Professor at Telkom University and he received his PhD in Engineering from Mie Unviersity, Japan. His research focuses on applied and fundamental machine learning for motion and behavior computing. He also collaborates with industry partners to develop AI-based tools for language modeling and image analytics.

Dr. Takao Ito

He is Professor of Management of Technology (MOT) in Graduate School of Advanced Science and Engineering at Hiroshima University. His current research interests include automata theory, artificial intelligence, systems control, quantitative analysis of interfirm relationships using graph theory, and engineering approach of organizational structures using complex systems theory.

Dr. Yit Hong Choo



He has completed his PhD and is now a Research Fellow in Operations Analytics at Deakin University's Institute for Intelligent Systems Research and Innovation (IISRI), supported by the Rail Manufacturing Cooperative Research Centre (RMCRC). His research focuses on advanced multi-objective optimisation algorithms for complex maintenance scheduling in rolling stock. He collaborates with transportation industry partners to develop AI-based tools for video and image analytics.