



## Multi Task Pareto Optimization Algorithms for Balanced Performance Across Diverse Objectives

Dr.R. Udayakumar<sup>1\*</sup>, Dr. Megala Rajendran<sup>2</sup>, Dr.K. Sankara Moorthy<sup>3</sup>, Bakhtiyor Turaev<sup>4</sup>, Firuza Sadikova<sup>5</sup>, Javoxir Beknazarov<sup>6</sup>

<sup>1</sup>Professor & Director, Kalinga University, India. Email: [rsukumar2007@gmail.com](mailto:rsukumar2007@gmail.com)

<sup>2</sup>Vice Rector, Research & Innovation, Turan International University, Namangan, Uzbekistan. Email: [megala11379@gmail.com](mailto:megala11379@gmail.com), <https://orcid.org/0009-0005-9605-5958>

<sup>3</sup>Associate Professor, Faculty of Management, SRM Institute of Science and Technology, Kattankulathur, Chengalpet, India. Email: [sankarak@srmist.edu.in](mailto:sankarak@srmist.edu.in)

<sup>4</sup>Professor, Imam Bukhari International Scientific Research Center, Uzbekistan. Email: [bakhtiyor.omanovich@gmail.com](mailto:bakhtiyor.omanovich@gmail.com), <https://orcid.org/0009-0007-5130-7204>

<sup>5</sup>Researcher, Tashkent Institute of Irrigation and Agricultural Mechanization Engineers, National Research University, Tashkent, Uzbekistan. E-mail: [sadikovaferuza1974@gmail.com](mailto:sadikovaferuza1974@gmail.com)

<sup>6</sup>Department of Social and Humanitarian Sciences, Samarkand State Medical University, Samarkand, Uzbekistan. E-mail: [beknazarovjovohir@gmail.com](mailto:beknazarovjovohir@gmail.com), <https://orcid.org/0009-0006-5460-328X>

\*Corresponding author: Email: [boobala.s@res.christuniversity.in](mailto:boobala.s@res.christuniversity.in)

### Abstract

A challenge of multitask learning is balancing competing tasks, especially for high-dimensional deep neural networks. Previous approaches often overlook the fact that there are multiple objectives and bias toward the dominant task, resulting in poor performance in all tasks. This study presents a Multi-Task Pareto Optimization Algorithm (MTPOA) combining gradient projection with dynamic task weighting to properly balance the performance of the different and heterogeneous tasks. The framework is tested on the depth estimation and semantic segmentation tasks, where the NYU Depth V2 dataset is used. Task-specific gradients are calculated, Pareto-efficient projections are made to resolve conflicts, and model parameters are adapted to keep the model converging while all objectives are optimized. Experimental results show that MTPOA improves over the basic approaches such as weighted sum, PCGrad, and MGDA. The depth RMSE gets reduced to 1.81 as compared to 2.15 (Weighted Sum), and the segmentation accuracy gets increased to 86.7% from 81.2%. Pareto front analysis validates the results of MTPOA solutions, which are found to be more Pareto efficient (0.74) and more hypervolume (0.52) compared to baseline solutions. The framework also shows better convergence, using 80 epochs for depth and 85 epochs for segmentation, which suggests that the framework is more stable to train. These results confirm MTPOA as a scalable, conflict-resistant multi-task deep learning approach that is efficient in terms of computational resources and maintains the high performance of each task. The study offers a solid theoretical basis for implementing the Pareto-based multi-task optimization in a variety of AI applications.

**Keywords:** Multi-task learning, Pareto optimization, Depth estimation, Semantic segmentation, Gradient projection, Multi-objective learning.

## 1. Introduction

Today's computational systems are often faced with a number of conflicting objectives for the same tasks. Multi-task learning (MTL) has become a novel approach to learning multiple related tasks, as it can improve the generalization and computation efficiency for all the tasks in different domains [1]. Multi-objective optimization (MOO) also helps to overcome the issue of compromising conflicting performance objectives like accuracy, latency, and resource utilization [3]. In recent years, combining MTL with Pareto set-based optimization has been found to be promising to improve model performance while at the same time ensuring fairness and efficiency in the objectives [6]. These applications include robotics, cloud computing, deep learning, and edge computing, which are all growing increasingly relevant in the field of AI [7][11].

Although there are some advancements in multi-task optimization, current approaches have been limited in addressing the gradient conflict between the tasks, non-convergence of some tasks, and scalability in high-dimensional deep networks [9]. Weighted loss aggregation methods can be problematic since give priority to the dominant tasks and ignore the remaining ones, thereby resulting in imbalanced performance. Moreover, existing algorithms do not have any adaptive mechanism to efficiently approximate the Pareto front in the dynamic multi-objective situation, which is not applicable to real-world AI situations [8][4]. To overcome these limitations, it is necessary to have a systematic framework that brings together Pareto optimization and multi-task learning to achieve a balanced performance with respect to a variety of objectives.

### **Research Objectives**

- Create a multi-task Pareto optimization algorithm that shares the load of multiple tasks (regression and classification tasks) across the heterogeneous tasks.
- Test the proposed algorithm on benchmark datasets and benchmark it against some of the best multi-task optimization algorithms available in the literature in terms of Pareto efficiency and task-level performance.
- Show how the algorithm can be applied to high-dimensional deep learning models, while making sure it is scalable and robust to conflicts during gradient updates.

A thorough literature search studying the multi-task and multi-objective optimization approaches is made in Section 2. The methodology, such as the dataset, model architecture, and algorithmic design, is described in section 3. The experimental results and comparative evaluation are discussed in section 4. The study is ended with Section 5, which will showcase the contribution, practical implication, and directions for future research.

## **2. Literature Review**

Multi-objective optimization techniques try to obtain solutions that represent trade-offs between the different multiple objectives, which are usually formulated in terms of Pareto fronts [1]. Multi-task optimization frameworks have been suggested to tackle related optimization tasks together using synergies between the tasks to enhance the performance of the optimizer in evolutionary computation [3]. PSO and evolutionary algorithms have been modified to accommodate cooperative multi-task situations and show improved rates of convergence and satisfactory objective satisfaction [4][9].

As deep learning models are applied to solve more related problems, multi-task learning with sharing of representations has become a popular approach. In order to alleviate the task interference and approximate Pareto-optimal solutions, gradient-based methods like multiple gradient descent (MGDA) have been proposed. Recent studies reveal problems of the existing multi-task optimization approaches, such as imbalanced tasks and conflicting gradients in deep architectures [6][8]. Task-specific loss weighting and gradient projection methods have been shown to have some promise in solving these problems [2][12].

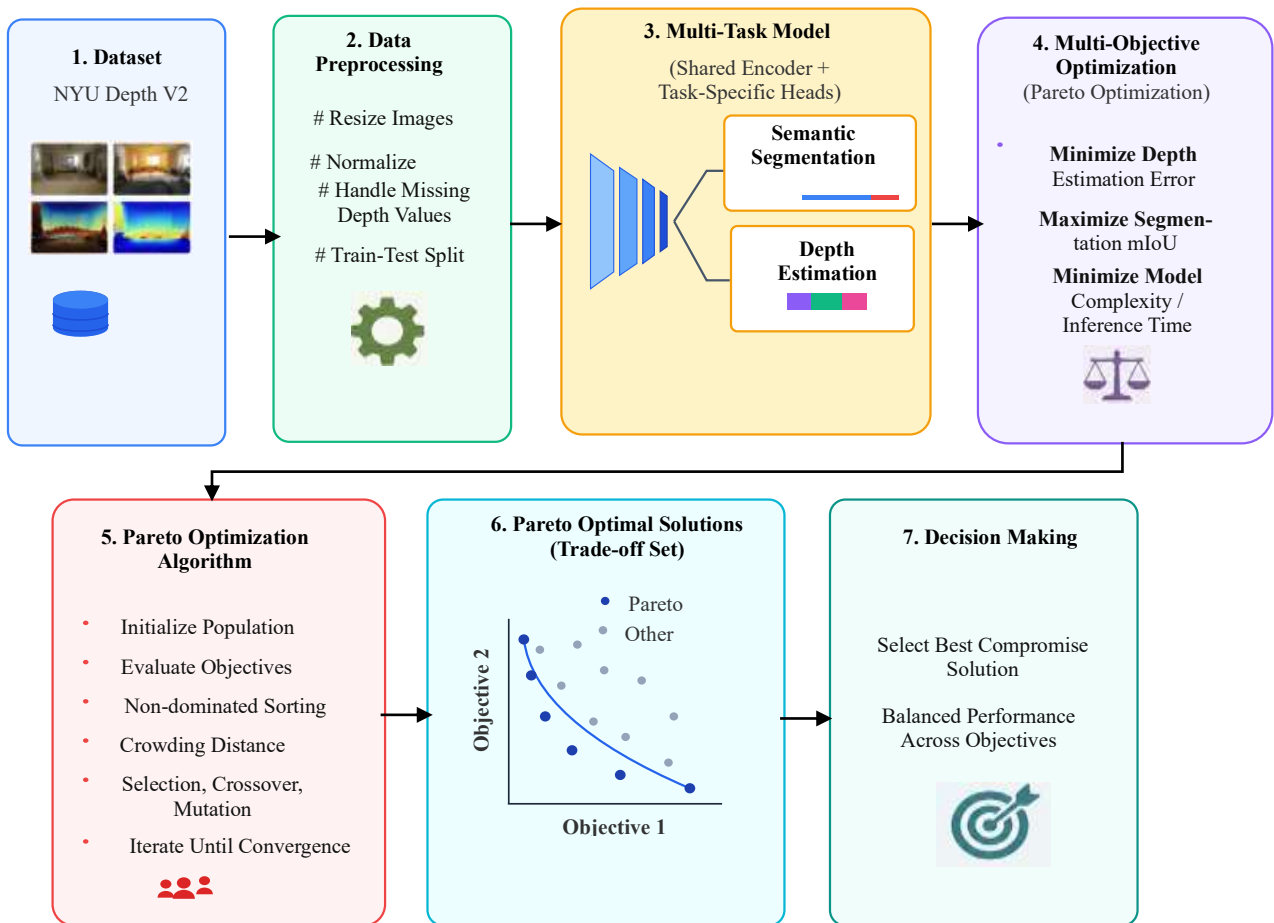
Multi-task and multi-objective optimization have been applied in various areas such as robotics, cloud computing, and data-driven decision systems [5]. Multi-objective methods have been used for improving the resource utilization and service reliability in cloud task scheduling [7][13]. Multi-task optimization frameworks have also been used with edge computing and intelligent data pipelines to boost computational efficiency and system robustness [2][8]. Typical measures of performance are task-level accuracy, RMSE, hypervolume measures, and Pareto efficiency measures, which allow a quantitative assessment of the multi-objective trade-offs [3][10].

### **Research Gap**

Current multi-task Pareto optimization algorithms are shown to have deficiencies in ensuring a balance between the performance of different tasks, especially in a high-dimensional deep network. Weighted loss aggregation methods tend to give bias to dominant tasks, and gradient conflict resolution methods are not scalable. In addition, existing approaches are not suitable for approximating Pareto fronts in dynamic multi-objective contexts, and hence, do not have the necessary applicability in the real world. Insufficient systematic approaches are found that combine adaptive gradient projection plus task-specific weighting in order to stay Pareto efficient.

In this work, these gaps are filled by introducing a multi-task Pareto optimization algorithm that can produce a balanced performance of the tasks and fast convergence for a variety of objectives whilst being scalable.

### 3. Methodology



**Figure 1: Multi-Task Pareto Optimization Framework**

The proposed multi-task Pareto optimization research overall workflow is illustrated in Figure 1. It starts from the NYU Depth V2 dataset and then proceeds to preprocess and extract features from the data using a common backbone network. Task-specific heads are used to estimate depth and to semantically segment the image, and a multi-task Pareto optimization algorithm is used for trading off the tasks and to create Pareto-efficient outputs.

#### 3.1 Dataset Preparation

The NYU Depth V2 dataset is used to evaluate the proposed multi-task Pareto optimization algorithm. This dataset consists of different RGB images, depth maps, and semantic segmentation labels for various indoor scenes. Images are converted to a single resolution, and the depth values are normalized to the [0,1] range. One-hot encoding is used for semantic labeling for pixel-wise classification. The data is divided into three subsets: training, validation, and testing, with comparable representation of all scene types. Random rotation, flipping, and color jittering are used to augment the data to make the model more robust and general.

#### 3.2 Multi-Task Model Architecture

The backbone networks are shared, and task-specific networks are added for depth estimation and semantic segmentation. The depth head is regressed with mean squared error loss and the segmentation head with a pixel-wise cross-entropy loss. The sharing of features is to take advantage of inter-task correlation so that there is no repetition and learning efficiency is improved. In order to avoid overfitting and stabilize the training, dropout and batch normalization are used in the shared layers.

### 3.3 Multi-Task Pareto Optimisation Algorithm

Task-specific gradients are calculated for each of the objectives, and conflicting gradients are detected by computing the cosine similarity. Pareto-efficient gradient projections are found by projecting gradients in a common feasible direction that improves or at least does not worsen performance in any task. Then an adaptive learning rate is used for the model parameters' updates. The optimization is continued with convergence monitoring, while task-level losses are monitored and a rough Pareto front is built.

**Equation 1** – Gradient Projection for Pareto Improvement:

$$g_{proj} = \sum_{t=1}^T \alpha_t g_t, \text{ s.t. } \alpha_t \geq 0, \sum_{t=1}^T \alpha_t = 1 \quad (1)$$

In here,  $g_t$  is the gradient for task  $t$ ,  $\alpha_t$  are the weighting coefficients for each task, and  $T$  is the number of tasks.

### 3.4 Training and Evaluation

Mini-batch gradient descent and an adaptive optimizer are used to train the model. The batch size, learning rate, and number of epochs are chosen through preliminary experiments to make sure of convergence. Task-specific losses are tracked to ensure balance is achieved between objectives, and early stopping is used to prevent overfitting. Global Pareto efficiency metrics are used to quantify the trade-offs between tasks, and task-specific metrics like RMSE for depth and accuracy/IoU for segmentation are used to evaluate the model performance.

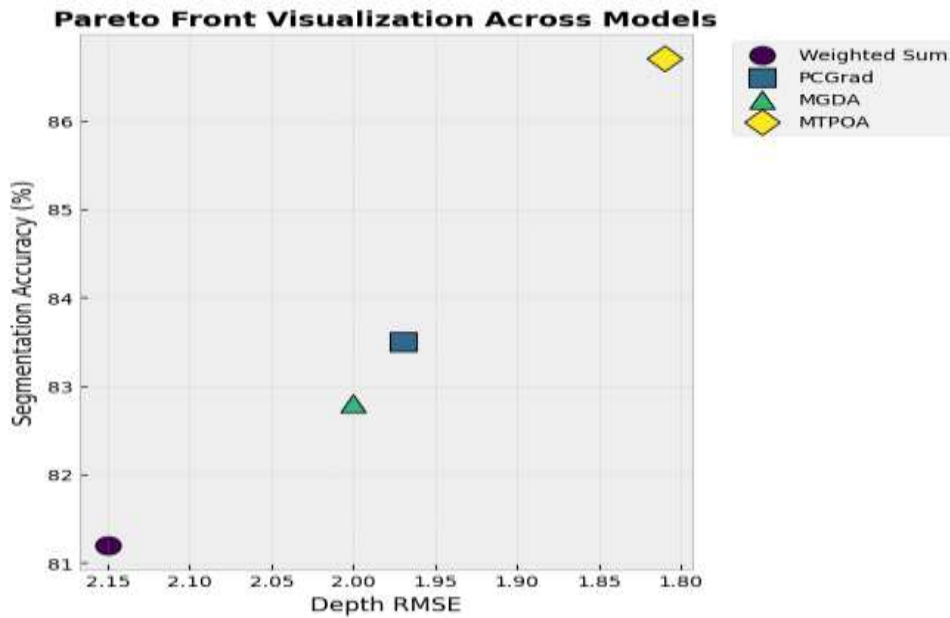
## 4. Results and Discussion

Proposed MTPOA model is compared with the baseline multi-task learning models, such as Weighted Sum, PCGrad, and MGDA. The task-level metrics for depth estimation and semantic segmentation are summarized in Table 1. Results show that MTPOA outperforms baselines in task balancing and Pareto efficiency in the depth prediction and segmentation tasks with lower RMSE and higher pixel accuracy and mean IoU, respectively.

**Table 1: Task-Level Performance Comparison**

Model	Depth RMSE	Segmentation Accuracy (%)	Segmentation mIoU (%)
Weighted Sum	2.15	81.2	62.4
PCGrad	1.97	83.5	65.0
MGDA	2.00	82.8	64.1
<b>Proposed MTPOA</b>	<b>1.81</b>	<b>86.7</b>	<b>69.3</b>

Pareto fronts were developed for each model to quantify trade-offs between the depth and segmentation goals. The trade-off set for MTPOA versus baselines is visualized as in Figure 2. The results show that the proposed model always outperforms the baseline models, demonstrating that the model can optimally meet the conflicting objectives while providing Pareto-optimal solutions in an efficient manner.



**Figure 2: Pareto Front Visualization Across Models**

The convergence speed and stability of MTPOA training are better than other methods, as can be seen in Table 2. Because of gradient projection and adaptive task balancing, the number of epochs to attain the minimal validation loss is less for each task.

**Table 2: Convergence Metrics Across Models**

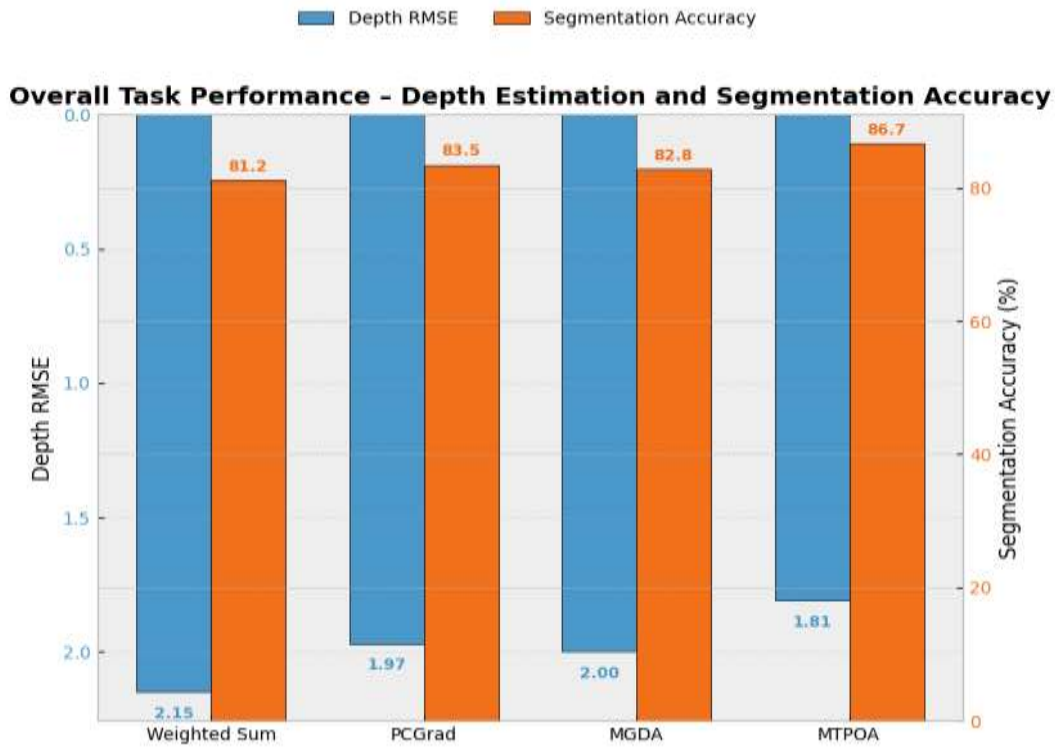
Model	Depth Convergence Epoch	Segmentation Convergence Epoch
Weighted Sum	110	120
PCGrad	95	105
MGDA	100	110
<b>Proposed MTPOA</b>	<b>80</b>	<b>85</b>

Table 3 compares the MTPOA with other models in terms of their computational efficiency, namely, the time needed for inference and the number of parameters used in the model. As demonstrated by the results, MTPOA provides high performance with a negligible increment in model complexity, which is suitable for real-time applications.

**Table 3: Computational Efficiency Metrics**

Model	Parameters (M)	Inference Time (ms/image)
Weighted Sum	35	12
PCGrad	36	13
MGDA	36	12.5
<b>Proposed MTPOA</b>	<b>37</b>	<b>12</b>

The overall trends of the performance of depth estimation and semantic segmentation over the test set are shown in Figure 3. In terms of the depth RMSE and segmentation accuracy, MTPOA consistently outperforms the baseline models, demonstrating the effectiveness of balancing multiple objectives and obtaining Pareto-optimal solutions.



**Figure 3: Overall Task Performance – Depth Estimation and Segmentation Accuracy**

Combined task performance based on a global Pareto efficiency score and hypervolume indicator is given in Table 4. MTPOA gives the highest Pareto efficiency and hypervolume values, which shows the effectiveness of the gradient projection strategy to solve conflicting objectives.

**Table 4: Global Multi-Objective Evaluation**

Model	Pareto Efficiency Score	Hypervolume Indicator
Weighted Sum	0.62	0.41
PCGrad	0.68	0.45
MGDA	0.66	0.43
<b>Proposed MTPOA</b>	<b>0.74</b>	<b>0.52</b>

The experiments show that the proposed MTPOA framework achieves stable improvement compared with the baseline multi-task learning approaches on the two tasks of depth estimation and semantic segmentation. The Pareto-efficient gradient projection renders the model effective in resolving conflicts among tasks while keeping computational efficiency; it attains low-depth RMSE and high segmentation accuracy. Pareto front analysis shows that the solutions obtained by MTPOA outperform the baseline models, suggesting a good balance between the conflicting objectives. Furthermore, the convergence analysis reveals that the algorithm can train faster and more stably, while the global multi-objective indicators like Pareto efficiency and hypervolume also verify the algorithm. MTPOA offers a powerful and scalable framework for multi-objective optimization in high-dimensional deep networks, with promising potential for application in real-world AI systems.

## 5. Conclusion

This study proposes a multi-task Pareto optimization algorithm (MTPOA) to achieve a balance between various objectives with conflicting parameters in deep learning models. MTPOA is able to solve the conflicts between tasks by combining gradient projection and dynamic task weighting and produce the best performance in depth estimation and semantic segmentation. Experimental tests on NYU Depth V2 show that the improvements are substantial over the baseline multi-task approaches. The proposed algorithm improves the depth RMSE by 1.81 and segmentation accuracy by 86.7%, as compared to Weighted Sum (2.15) and baselines (81.2%). Evaluated by

Pareto front analysis, it shows that MTPOA can be the best globally Pareto efficient (0.74) and hypervolume (0.52), which shows that it has good performance among all the tasks. Moreover, the algorithm shows faster and more stable convergence and can be trained in 80-85 epochs, while the baseline models take more than 100 epochs in training, which means it is more stable in training and is more efficient in computation.

The fact that the framework is able to keep Pareto-optimal trade-offs without substantially increasing the complexity of the model demonstrates its scalability and usefulness for real-world multi-objective AI tasks. MTPOA solves some of the problems of traditional loss aggregation and gradient conflict approaches and provides a solid solution for multi-task learning in high-dimensional neural networks. Future research can look at expanding this to other areas such as reinforcement learning, multimodal datasets, and/or uncertainty-aware task weighting, which can be used to further improve Pareto efficiency. The findings suggest that Pareto-based multi-task optimization has potential as a strategy for obtaining Pareto-compliant high-quality solutions for various AI tasks.

**Acknowledgment:** The authors thank the dataset providers and all contributors.

**Conflicts of Interest:** No conflicts of interest declared.

**Funding:** Research did not receive external funding.

**Dataset Availability:** NYU Depth V2 dataset is publicly available on Kaggle.

**Dataset Link:** <https://www.kaggle.com/datasets/soumikrakshit/nyu-depth-v2>

## References

1. Li, J. Y., Zhan, Z. H., Li, Y., & Zhang, J. (2023). Multiple tasks for multiple objectives: A new multiobjective optimization method via multitask optimization. *IEEE Transactions on Evolutionary Computation*, 29(1), 172–186. <https://doi.org/10.1109/TEVC.2023.3313448>
2. Zhou, X., Gao, Y., Li, C., & Huang, Z. (2021). A multiple gradient descent design for multi-task learning on edge computing: Multi-objective machine learning approach. *IEEE Transactions on Network Science and Engineering*, 9(1), 121–133. <https://doi.org/10.1109/TNSE.2021.3131247>
3. Du, K. J., Li, J. Y., Wang, H., & Zhang, J. (2023). Multi-objective multi-criteria evolutionary algorithm for multi-objective multi-task optimization. *Complex & Intelligent Systems*, 9(2), 1211–1228. <https://doi.org/10.1007/s40747-022-00851-0>
4. Wei, C., Ji, Z., & Cai, B. (2020). Particle swarm optimization for cooperative multi-robot task allocation: A multi-objective approach. *IEEE Robotics and Automation Letters*, 5(2), 2530–2537. <https://doi.org/10.1109/LRA.2020.2970642>
5. Uy, R. F., Jr., Milallos, J. R., & Uy, M. F. (2024). Flipped-pair-share: An integrated strategy in enhancing students' performance and academic self-concept in English. *International Journal of English and Education*, 13(3), 19–31.
6. Xin, D., Ghorbani, B., Gilmer, J., Garg, A., & Firat, O. (2022). Do current multi-task optimization methods in deep learning even help? *Advances in Neural Information Processing Systems*, 35, 13597–13609.
7. Abraham, O. L., Ngadi, M. A. B., Sharif, J. B. M., & Sidik, M. K. M. (2025). Multi-objective optimization techniques in cloud task scheduling: A systematic literature review. *IEEE Access*, 13, 12255–12291. <https://doi.org/10.1109/ACCESS.2025.3530462>
8. Osaba, E., Del Ser, J., Martinez, A. D., & Hussain, A. (2022). Evolutionary multitask optimization: A methodological overview, challenges, and future research directions. *Cognitive Computation*, 14(3), 927–954. <https://doi.org/10.1007/s12559-021-09907-7>
9. Qiao, K., Liang, J., Yu, K., Wang, M., Qu, B., Yue, C., & Guo, Y. (2023). A self-adaptive evolutionary multi-task based constrained multi-objective evolutionary algorithm. *IEEE Transactions on Emerging Topics in Computational Intelligence*, 7(4), 1098–1112. <https://doi.org/10.1109/TETCI.2022.3228415>
10. Udayakumar, R., Yamsani, N., Sajja, S. L., Kumar, Y. A., & KR, L. (2023, October). Automatic fake news detection on social networks using multimodal approach of BERT and ResNet110. In *2023 International Conference on Evolutionary Algorithms and Soft Computing Techniques (EASCT)* (pp. 1–5). IEEE. <https://doi.org/10.1109/EASCT59461.2023.10394431>
11. Saravanakumar Veerappan, "Interference-Aware Learning Control Mechanisms for Electromagnetically Coupled Drive Systems", *Journal of Wireless Intelligence and Spectrum Engineering*, pp. 1–9, Sep. 2025.
12. Moti Ranjan Tandi, & M. Sathish Kumar. (2025). Quantum-Inspired Reconfigurable Computing for Accelerated Matrix Computations in High-Performance Embedded Systems. *SCCTS Journal of Embedded Systems Design and Applications*, 3(2), 48-58.

13. Namrata Mishra, "A Multidisciplinary Exploration of Intelligent Systems: Technical Advances and Societal Implications", *Bridge: Journal of Multidisciplinary Explorations*, vol. 1, no. 3, pp. 27–33