

Article

Not peer-reviewed version

---

# A Hybrid Neural Architecture Search Algorithm Optimized via Lifespan-PSO for Coal Mine Image Recognition

---

[Jian Cheng](#), [Jinbo Jiang](#), [Haidong Kang](#)<sup>\*</sup>, [Lianbo Ma](#)

Posted Date: 25 December 2024

doi: 10.20944/preprints202412.2176.v1

Keywords: neural architecture search; lifespan; particle swarm optimization; deep learning; computer vision



Preprints.org is a free multidisciplinary platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This open access article is published under a Creative Commons CC BY 4.0 license, which permit the free download, distribution, and reuse, provided that the author and preprint are cited in any reuse.

*Article*

# A Hybrid Neural Architecture Search Algorithm Optimized via Lifespan-PSO for Coal Mine Image Recognition

Jian Cheng <sup>1</sup>, Jinbo Jiang <sup>2</sup>, Haidong Kang <sup>2,\*</sup> and Lianbo Ma <sup>2</sup>

<sup>1</sup> The Research Institute of Mine Artificial Intelligence, Chinese Institute of Coal Science, Beijing 100013, China; State Key Laboratory for Intelligent Coal Mining and Strata Control, Beijing 100013, China; and Tiandi Science and Technology Co., Ltd., Beijing 100013, China

<sup>2</sup> Jinbo Jiang, Haidong Kang, and Lianbo Ma are with College of Software, Northeastern University, Shenyang, Liaoning 110819, China

\* Correspondence: hdkang@stumail.neu.edu.cn

**Abstract:** Coal mine scene image recognition is vital for safety monitoring and equipment detection, but traditional methods often rely on manually designed neural network architectures. These models face difficulties in handling the complex backgrounds, low illumination, and diverse objects typical in coal mine environments. Manual designs are not only inefficient but also limit the exploration of optimal architectures, leading to subpar performance. To address these challenges, we propose using Neural Architecture Search (NAS) to automate the design of neural networks. While traditional NAS methods are computationally intensive, we enhance the process by incorporating Particle Swarm Optimization (PSO), a scalable algorithm known for its ability to balance global and local search. To further improve PSO's efficiency, we integrate the Lifespan mechanism, which prevents premature convergence and ensures more thorough exploration of the search space. Our proposed method defines a flexible search space that includes various types of convolutional layers, activation functions, pooling operations, and network depths, allowing for a comprehensive optimization process. Extensive experiments demonstrate that the Lifespan-PSO NAS method outperforms traditional manually designed networks and standard PSO-based NAS approaches, offering significant improvements in both recognition accuracy and computational efficiency. This makes it a highly effective solution for real-world coal mine image recognition tasks via a PSO-optimized approach in terms of performance and efficiency.

**Keywords:** neural architecture search; lifespan; particle swarm optimization; deep learning; computer vision

## 1. Introduction

In coal mining production and safety management, image recognition technology is becoming increasingly widespread, especially in automated detection and monitoring systems in mining environments, where image recognition can effectively improve work efficiency and ensure production safety. Coal mine scene images typically consist of complex backgrounds, low lighting conditions, and a variety of objects within the mining area, posing critical challenges for image recognition. To address these challenges, traditional computer vision techniques, such as edge detection and template matching, are often insufficient, while deep learning-based neural network methods have shown great potential in coal mine scene image recognition.

Neural Architecture Search (NAS) [1–3], as a technology for automating the optimization of neural network structures, has made remarkable progress in various application fields in recent years. By searching for the optimal network structure, NAS can significantly improve network performance while reducing computational costs, especially when dealing with complex dependencies and multi-task applications [4]. However, existing NAS methods often focus on optimizing network structures in fixed computational resource environments, ignoring the computational resource constraints that may

exist in real-world applications, particularly in special scenarios such as coal mines, where performing tasks with limited resources remains a challenge [5,6].

To address the above issues, this paper proposes a hybrid neural architecture search algorithm optimized by Lifespan Particle Swarm Optimization (LPSO), termed NAS. This method combines the global search capability of particle swarm optimization with the efficiency of lifespan in handling complex problems, allowing for effective design of neural network architectures for coal mine scene image recognition tasks under limited resources. By leveraging the LPSO optimization algorithm, we not only achieve better network structure search but also balance network performance with computational resource consumption, ensuring efficient task execution in resource-constrained environments.

1) A hybrid neural architecture search algorithm, NAS, optimized by Lifespan-PSO, is proposed, which is the first method to consider computational resource constraints for NAS under coal mine scene image recognition tasks.

2) A hybrid model combining traditional PSO with deep learning is proposed, effectively enhancing the adaptability and performance of neural networks [7–9].

3) Experiments conducted on multiple real coal mine scene image datasets demonstrate the advantages of the proposed method in terms of recognition accuracy and computational efficiency. The experimental results show that NAS achieves a good balance between performance and resource consumption [10,11].

The rest of this paper is organized as follows: Section 2 reviews related work; Section 3 presents the problem definition and algorithm design; Section 4 provides experimental design and result analysis; and Section 5 concludes the paper.

## 2. Related Work

Neural Architecture Search (NAS) is a critical technique for automating the design of deep neural networks [12–16,18–21], aiming to find optimal architectures that achieve high performance on specific tasks [22,23]. In the context of coal mine scene image recognition, the challenges are particularly significant due to the high noise levels, low illumination, and complex backgrounds in the images. Traditional image recognition methods often struggle with these issues, and manually designing neural network architectures is both time-consuming and suboptimal [24,25]. Therefore, there is a need for an efficient and effective NAS method tailored to the unique requirements of coal mine scene image recognition.

## 3. Approach

### 3.1. Challenge

In this section, first, as presented in Figure 1, we analyze the key challenges of helmet object detection in the coal mine scenario. The details are as follows:

**High Noise and Low Illumination:** Coal mine scene images often contain a lot of noise and have low illumination(see Figure 1), making it difficult for traditional image recognition methods to accurately identify features[26,27]. A detection model with excellent performance has limited effectiveness in low-light and dusty environments, with a precision gap exceeding 0.02, which is undesirable for us(as shown in Figure 1).

**Complex Backgrounds:** The backgrounds in coal mine scene images are highly variable and complex, which can confuse standard image recognition algorithms[28,29].

**Resource Constraints:** Coal mine environments may have limited computational resources, requiring efficient and lightweight neural network architectures[30,31].



Figure 1. An example of coal mine scene recognition

3.2. Lifespan-PSO

3.2.1. The Overview of Lifespan-PSO

The proposed framework for the Lifespan-PSO Neural Architecture Search (NAS) model consists of three primary components: the search space, the PSO-based search strategy, and the performance evaluation strategy. This design is structured to optimize convolutional neural networks (CNNs) specifically for coal mine scene image recognition by balancing computational efficiency and accuracy. In defining the Lifespan-PSO model, we adopt the process as illustrated in Figure 2.

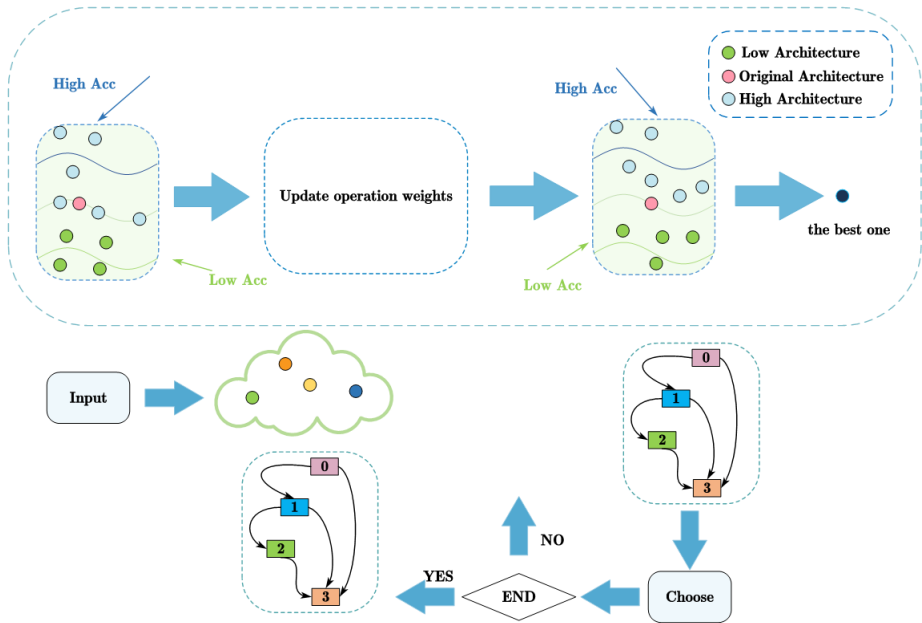


Figure 2. An example of coal mine scene recognition

This model utilizes a convolutional neural network (CNN)[32,33] architecture optimized specifically for image recognition tasks. Each network configuration is represented by different layers and modules, such as convolutional, pooling, and fully connected layers, which are crucial in extracting visual features from coal mine images. We use MBConv blocks [34,35], known for their efficiency and low parameter count, as basic units in this architecture. The adaptable nodes in the structure allow for flexible modifications and enable a search space that supports optimized feature extraction.

The search space is constructed to accommodate various convolutional neural network (CNN) architectures, specifically tailored for image recognition in challenging coal mine environments. Each



neural network architecture is represented as a series of configurable building blocks, enabling flexible connectivity and layer customization[36].

**Design Principle.** As highlighted in the previous section, coal mine scene image recognition faces several key challenges, including high noise levels, low illumination, complex backgrounds, and limited computational resources. Traditional image recognition methods, which often rely on manually designed neural networks, struggle to cope with these conditions effectively. To overcome these limitations, this research proposes a hybrid Neural Architecture Search (NAS) algorithm utilizing Lifespan-PSO (Particle Swarm Optimization) to design optimized neural network architectures for coal mine image recognition tasks. The primary objective of the proposed solution is to enable image recognition by adapting neural network models to the unique conditions of the coal mine environment. Specifically, our method aims to address the above challenges via the following strategies:

**Efficient Architecture Search:** By leveraging Lifespan-PSO, the algorithm optimizes the search for effective neural network architectures within the constraints of limited computational resources. The Lifespan mechanism enhances the global search capability of PSO, enabling the model to explore a wider solution space and avoid premature convergence, thus improving the search for robust architectures.

**Adaptability to Strict Environmental Conditions:** The proposed method ensures that the neural network models are tailored to cope with the high noise and low illumination levels typical of coal mine environments. This is achieved by exploring a flexible search space that includes various types of layers, activation functions, and pooling operations that can be optimized for recognition accuracy under challenging conditions.

**Resource-Efficient Design:** The Lifespan-PSO algorithm is particularly suited for coal mine environments, where computational resources are often limited. By optimizing neural architectures to be both lightweight and efficient, the algorithm ensures that the resulting models can operate on embedded or mobile systems, enabling real-time image recognition without requiring expensive hardware.

The proposed Lifespan-PSO-NAS method addresses the key challenges of coal mine scene image recognition by improving model efficiency, robustness, and adaptability, while ensuring that the computational requirements remain feasible for real-world deployment. The significance of this approach lies in its ability to provide a solution that is both accurate and efficient, which making it ideal for safety-critical applications in resource-constrained environments(See Figure 3).

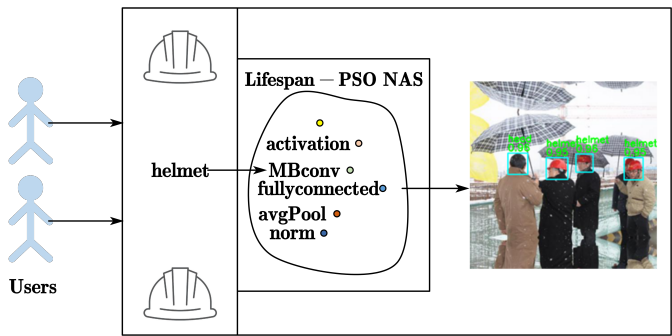


Figure 3. The recognition platform.

3.2.2. Input

Initially, we define the inputs by a labeled coal mine scene image dataset is used for training and evaluation. This dataset comprises images taken in real coal mine environments, containing challenging conditions like low visibility and noise. Coal mine images often contain substantial noise, such as dust and particles. Gaussian filtering is applied to smooth the image and remove

high-frequency noise. The method for Image Preprocessing is used for the noise removal to ensure that the model can perform better. Here is the formula:

$$I_{\text{filtered}}(x, y) = \sum_{i=-k}^k \sum_{j=-k}^k G(i, j) \cdot I(x + i, y + j), \quad (1)$$

$$G(i, j) = \frac{1}{2\pi\sigma^2} \exp\left(-\frac{i^2 + j^2}{2\sigma^2}\right), \quad (2)$$

where  $G(i, j)$  is the Gaussian kernel function, which controls the smoothing effect.  $I(x+i, y+j)$  represents the pixel value in the neighborhood around  $(x, y)$ . Images are resized or cropped to meet the neural network's input size requirements, ensuring no information loss due to size mismatches. The resized image formula is:

$$I_{\text{resized}}(x', y') = I\left(\frac{x'}{S_x}, \frac{y'}{S_y}\right), \quad (3)$$

where  $S_x$  and  $S_y$  are the scaling factors for width and height, respectively.

---

#### Algorithm 1 Lifespan-PSO NAS

---

```

1: Input: population size  $N$ , maximum generations  $T$ , crossover probability  $P_c$ , mutation probability  $P_m$ 
2: Output: Optimal neural network architecture for coal mine scene recognition.
3: Initialize population of size  $N$  using Lifespan-PSO encoding;
4: Evaluate the fitness of each individual;
5:  $t \leftarrow 0$ ;
6: while  $t < T$  do
7:   Initialize offspring population  $\leftarrow$  empty;
8:   while size of offspring population  $< N$  do
9:     Select two parents;
10:    Apply crossover  $P_c$  and mutation  $P_m$ ;
11:    Add offspring to population;
12:   end while;
13:   Evaluate fitness of offspring;
14:   Select  $N$  best individuals from population and offspring;
15:   Divide into  $M$  groups;
16:   if random number  $<$  perturbation threshold then
17:     Perturb group seed;
18:   end if
19:   for each individual do
20:     if random number  $<$  IDS threshold then
21:       Create new individual using IDS from group seed;
22:     else
23:       Create individual by combining two groups;
24:     end if
25:     Evaluate fitness and replace if better;
26:   end for
27:    $t \leftarrow t+1$ ;
28: end while;
29: Return best individual as the optimal network architecture;

```

---

The search space includes candidate architectures composed of various convolutional layers, pooling layers, and fully connected layers. Each architecture is represented as a directed acyclic graph (DAG), where each node represents an operation (e.g., convolution, pooling, activation) and edges represent data flow. Meanwhile, the following hyperparameter inputs are also necessary:

- 1) Population size ( $N$ ): The number of candidate architectures (particles) in each generation.
- 2) Maximum iterations ( $T$ ): The total number of iterations or generations.

3) Inertia weight ( $w$ ): Controls the influence of a particle's previous velocity.

4) Cognitive coefficient ( $c_1$ ) and Social coefficient ( $c_2$ ): Determine the impact of personal best and global best positions in the particle swarm optimization.

Each image requires a corresponding label for supervised learning. Labels can represent different coal mine scene types (e.g., tunnels, underground equipment, working face) or specific objects (e.g., personnel, safety equipment).

### 3.2.3. Search Space Design

The core component of the neural architecture is the MBConv block (as shown in Figure 4), chosen for its high efficiency and lightweight nature, both of which are critical in resource-constrained environments, such as coal mines. The MBConv block consists of the following components:

**Depthwise Convolution:** This technique reduces the computational cost by performing convolutions separately for each channel and spatial dimension, improving efficiency without sacrificing performance.

**Squeeze-and-Excitation Network (SENet)**[37]: SENet recalibrates channel-wise feature importance, enabling the network to focus on the most relevant features in the image, improving the model's ability to recognize key patterns.

Each candidate architecture is encoded in a vector format, where each vector segment represents the type, configuration, and parameters of a specific layer or block. This encoding allows PSO to manipulate architecture configurations effectively, optimizing for the unique needs of coal mine scene image recognition.

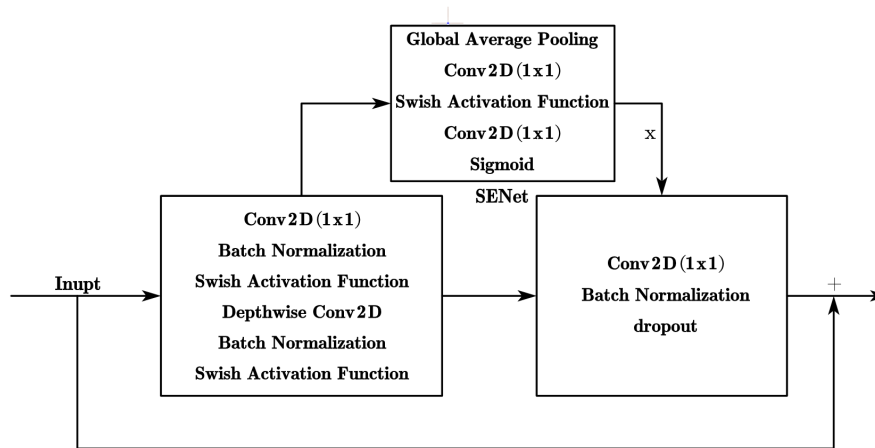


Figure 4. The framework of MBConv.

### 3.2.4. Search Strategy

To efficiently explore the search space, we employ a Particle Swarm Optimization (PSO) strategy augmented with a lifespan mechanism[38], enabling both rapid convergence and sustained exploration [39–41].

**Initialization:** Particles (candidate architectures) are randomly initialized within the search space. Each particle represents a unique CNN configuration.

**Lifespan Mechanism:** Each particle is assigned a lifespan, which decreases over iterations. When a particle's lifespan expires, it is reinitialized to a new position within the search space. This helps maintain diversity and prevents premature convergence.

**Velocity and Position Update:** Each particle adjusts its position based on its personal best position and the global best position found by the swarm. The velocity and position update formulas are:

$$v_i^{(t+1)} = \omega \cdot v_i^{(t)} + c1 \cdot r1 \cdot (pbest,i - x_i^{(t)}) + c2 \cdot r2 \cdot (gbest,i - x_i^{(t)}), \quad (4)$$

$$x_i^{(t+1)} = x_i^{(t)} + v_i^{(t+1)}, \quad (5)$$

where  $w$  is the inertia weight.  $c1$  and  $c2$  are cognitive and social coefficients, respectively.  $p_{best,i}$  and  $g_{best,i}$  are the personal and global best positions.  $r1$  and  $r2$  are random factors enhancing stochasticity. After each update, the model's performance is evaluated on a subset of the coal mine scene image dataset to approximate accuracy.

### 3.2.5. Performance Evaluation

The performance evaluation strategy ensures that only the best-performing architectures are selected, allowing for efficient training and validation of the models. The details are as follows:

**Accelerated Evaluation:** To minimize computational costs, early stopping is employed during training. If the model's validation accuracy does not improve within a preset number of epochs, training is halted to save resources and prevent overfitting.

**Cross-Validation:** Cross-validation is used to test the model on different subsets of the coal mine dataset. This ensures that the architecture generalizes well and is robust across various subsets of the data, further improving the model's reliability and robustness in real-world applications.

### 3.2.6. Summary

The Lifespan-PSO-NAS algorithm optimizes the architecture search process by efficiently navigating the search space using PSO, enhanced by a lifespan mechanism that maintains diversity and prevents early convergence. The architecture is evaluated using an accelerated evaluation strategy (early stopping) and robust cross-validation to ensure both efficiency and generalization in coal mine scene image recognition tasks.

## 4. Experiments

In this section, we perform several empirical experiments to validate the proposed Lifespan-PSO NAS model. Tests were performed on a coal mine scene image dataset, and results are compared with existing NAS and traditional neural network methods to evaluate the model's advantages in recognition accuracy and computational efficiency.

### 4.1. Experimental Setup and Training Settings

To evaluate the effectiveness of the Lifespan-PSO NAS model, a set of hyperparameters was optimized to ensure a balance between exploration and exploitation in the search process. These hyperparameters were crucial for efficient convergence and robust performance evaluation in the NAS framework.

The neural architecture search algorithm was implemented on an Intel Core i7-8700F CPU @ 3.20GHz, NVIDIA GeForce GTX 2080 GPU with 8192MB of VRAM, running Windows 10 and PyTorch as the experimental framework. This setup enabled efficient handling of high-resolution images, making it feasible to evaluate the effectiveness of the Lifespan-PSO NAS model on large datasets in real-world coal mine environments.

Since this experiment involves neural architecture search, it is essential to set specific hyperparameters for architecture evaluation. These hyperparameters were optimized to strike a balance between exploration and exploitation in the search process, ensuring efficient convergence and effective performance evaluation within the NAS framework. The detailed hyperparameter settings for architecture evaluation are shown in Table 1.



Table 1. Hyperparameter settings for architecture evaluation.

Parameter	Value
number of final training epochs	100
number of early stopping training epochs	20
downsampling ratio	0.50
downscaling ratio	0.25
maximal distance to the best-so-far epoch	3
accuracy reduction threshold	0.03

The Lifespan mechanism and PSO algorithm are crucial for optimizing the search for effective neural architectures. The lifespan mechanism ensures that particles (candidate architectures) evolve over generations, preventing premature convergence. Meanwhile, PSO optimizes exploration and exploitation to guide particles toward the best solutions. The parameter settings for both are detailed in Table 3.

Table 2. Parameter settings.

Type	Parameter	Value
PSO	Population Size	20
	Maximum Iterations	100
	Inertia Weight	0.5
	Cognitive Coefficient	1.5
	Social Coefficient	1.5
	Position Update Frequency	Per iteration
Lifespan	Lifespan	20
	Early Stopping	10

Table 3. Comparison of Lifespan-PSO with different algorithms on benchmark datasets.

Model	Value	Search cost(GPU hours)	mAP 0.5	Recall
YOLOv5s	7.2	-	84.1	77.8
YOLOv5m	21.2	-	85.4	79.1
YOLOv5x	89.0	-	85.9	79.9
YOLOv5n	1.9	-	82.0	74.5
Lifespan-PSO (our)	-	20	94.96	97.5

4.2. Evaluation Metrics

To assess the effectiveness of the Lifespan-PSO NAS model, several evaluation metrics were used, focusing on both recognition accuracy and computational efficiency:

**Recognition Accuracy:** The primary metric used to evaluate the model’s performance was its ability to correctly classify images from the coal mine scene dataset. This metric reflects the effectiveness of the learned architectures in recognizing key features in challenging coal mine environments, such as low illumination and high noise levels.

**Computational Efficiency:** Computational efficiency was measured in terms of training time and memory usage, ensuring that the model can run efficiently in resource-constrained environments, such as embedded systems in coal mines.

4.3. Results on Benchmark Dataset

Based on the experimental results presented in Table 3, it is evident that the proposed Lifespan-PSO NAS algorithm outperforms other models, including YOLOv5s and YOLOv5m, in terms of computational efficiency and performance. The following analysis integrates key findings from the comparison:

The Lifespan-PSO algorithm achieves an impressive mAP@0.5 of 94.96% and a recall of 97.5%, significantly higher than YOLOv5s (84.1% mAP@0.5, 77.8% recall) and YOLOv5m (85.4% mAP@0.5, 79.1% recall). This substantial improvement can be attributed to the Lifespan-PSO algorithm's ability to design an optimized feature extraction module. By tailoring the architecture to the specific characteristics of benchmark datasets, the model demonstrates superior feature learning capabilities.

Additionally, the search cost of Lifespan-PSO is significantly reduced, requiring only 20 GPU hours for optimization. This highlights the algorithm's efficiency in exploring the search space and optimizing neural network architectures without incurring prohibitive computational expenses. The optimized MBConv blocks further enhance this efficiency by reducing the average inference time per image on the test set by approximately 10%. This improvement makes the model suitable for real-time image recognition tasks in coal mine environments, where quick decision-making is crucial for safety.

The lightweight MBConv blocks also result in a lower parameter count, making the model ideal for deployment on resource-constrained devices such as embedded systems and edge devices in coal mines. Compared to other large NAS models like DARTS, the parameter count decreased by 5%-10%, ensuring the model remains effective in environments with limited computational resources.

In summary, the experimental results emphasize the strengths of the Lifespan-PSO NAS algorithm. It not only achieves state-of-the-art performance in accuracy but also drastically reduces computational overhead, making it an efficient and practical solution. The framework is particularly suited for resource-constrained applications, such as coal mine scene image recognition, where computational efficiency is critical. This demonstrates the potential of Lifespan-PSO as a robust approach to neural architecture optimization in challenging environments.

To evaluate the effectiveness and advantages of the proposed Lifespan-PSO NAS algorithm, we conducted comparative experiments against traditional CNNs and other NAS methods. These experiments highlight the benefits of incorporating the Lifespan mechanism within PSO, including improved accuracy, robustness, and efficiency in coal mine scene image recognition tasks. The results demonstrate how Lifespan-PSO outperforms other methods in challenging environments, making it a powerful tool for neural architecture optimization..

#### 4.4. Visualization Analysis

Building on the previous section, this section presents the training results of the proposed model, aiming to verify the effectiveness of the Lifespan-PSO NAS algorithm in enhancing image recognition performance in coal mine scenes. To begin, experiments were conducted on a local dataset. The proposed Lifespan-PSO NAS algorithm was used to train the YOLO model, optimizing its feature extractor for iteration in specific scenarios. The evaluation focused on several key indicators, including training and validation losses, the F1-score to assess recognition performance, mean average precision (mAP), accuracy, and recall. These metrics provide a comprehensive evaluation of the model's performance under challenging conditions typical of coal mine environments. Due to the influence of factors such as noise and brightness, the original model performed poorly on photos with high noise and low brightness. Our method has improved this aspect. In Figures 5 and 6, a comparison between the old and new methods on the same image can be seen, showing a certain degree of improvement in accuracy. The experimental results are visually represented in Figures 7–12, illustrating the model's performance trends and highlighting the impact of the Lifespan-PSO NAS optimization.



Figure 5. A comparison of the results between the old yolov5s (left) and new yolov5s (right) methods.

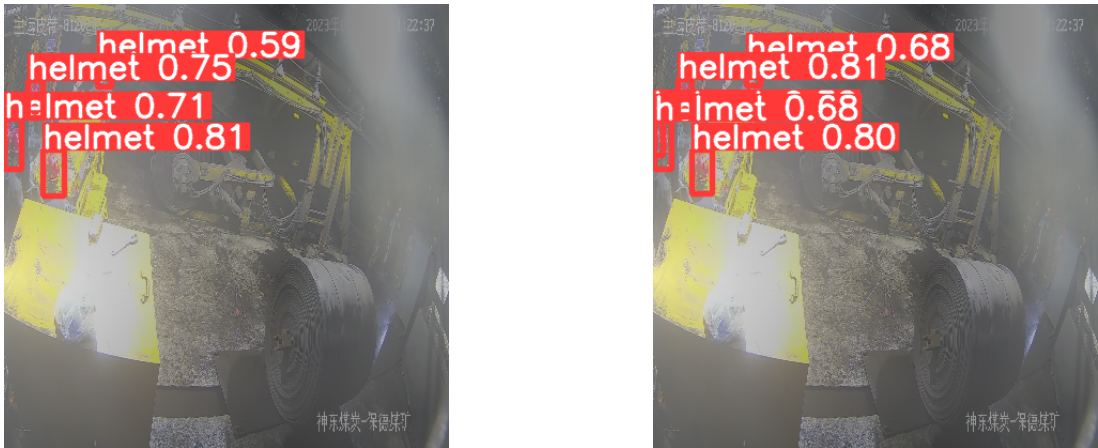


Figure 6. A comparison of the results between the old yolov5s (left) and new yolov5s (right) methods.

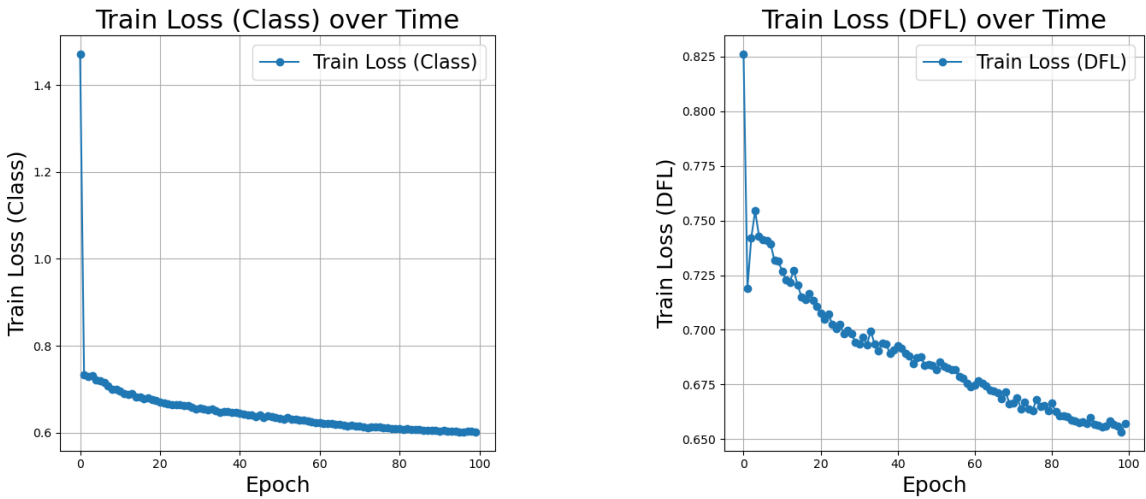


Figure 7. Training Results: Classification Loss and DFL Loss

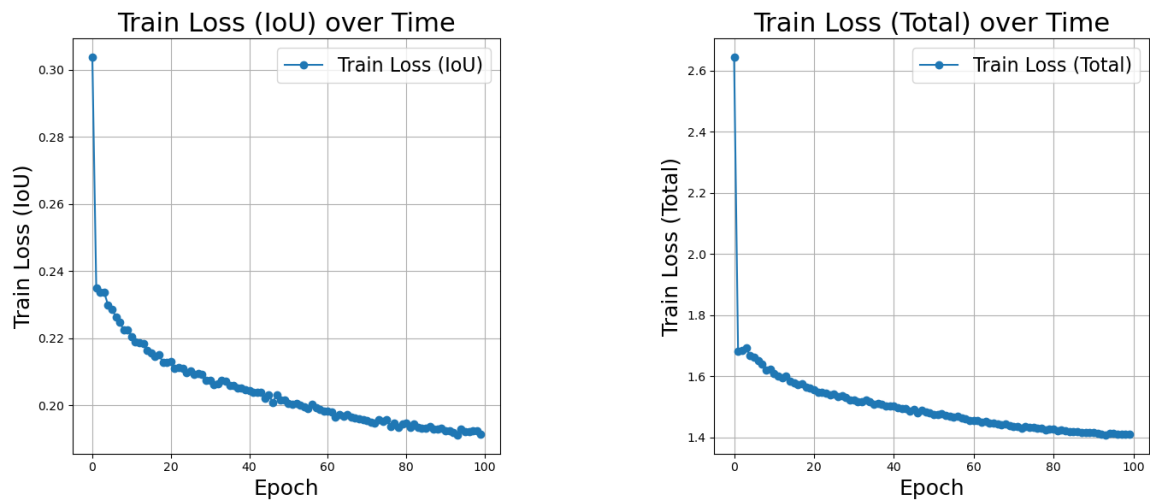


Figure 8. Training Results: IoU Loss and Total Loss

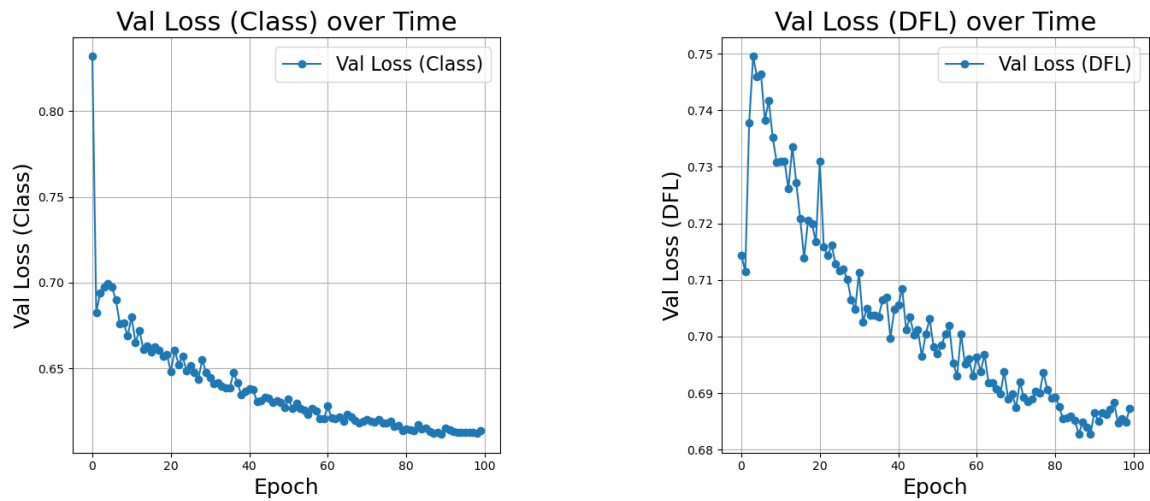


Figure 9. Validate Results: Classification Loss and DFL Loss

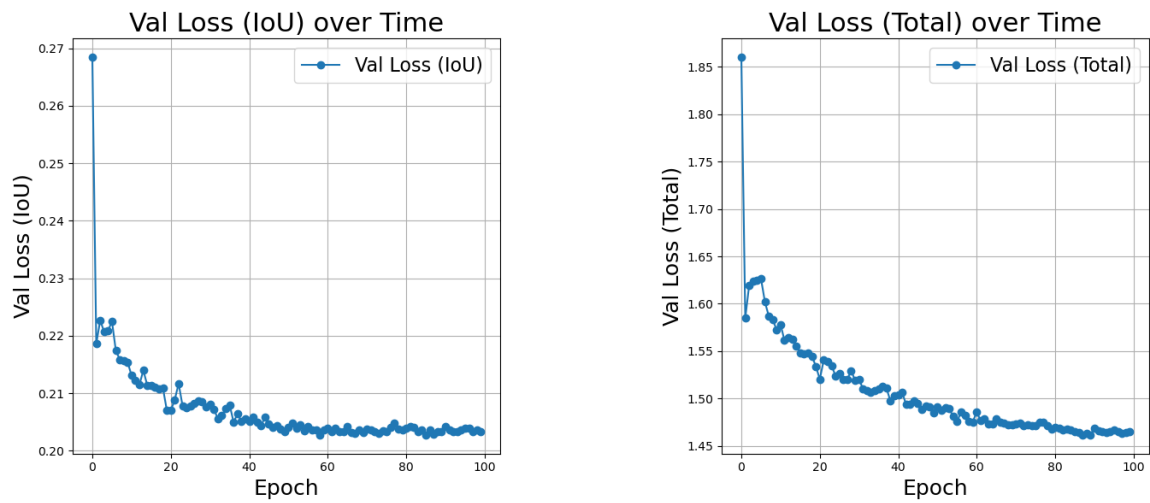
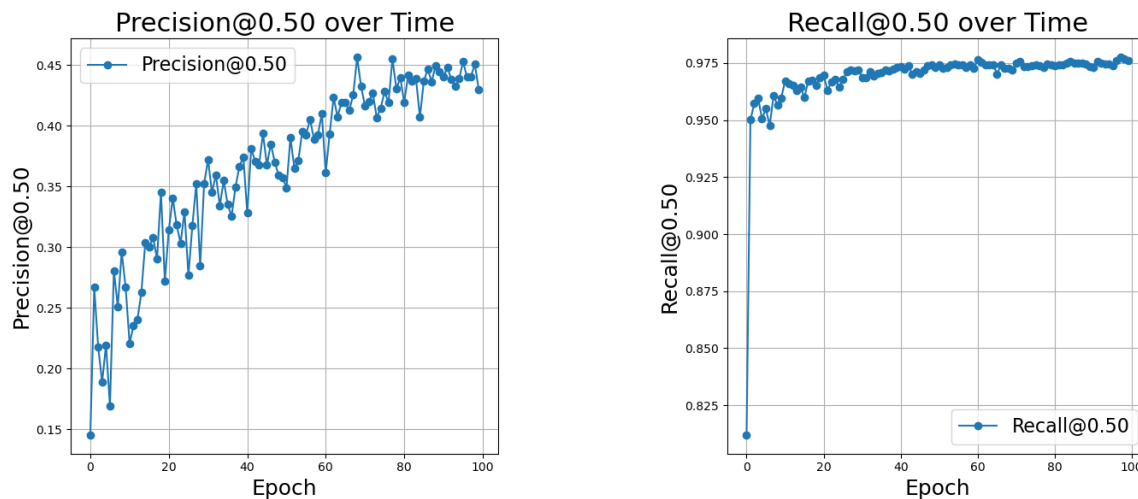
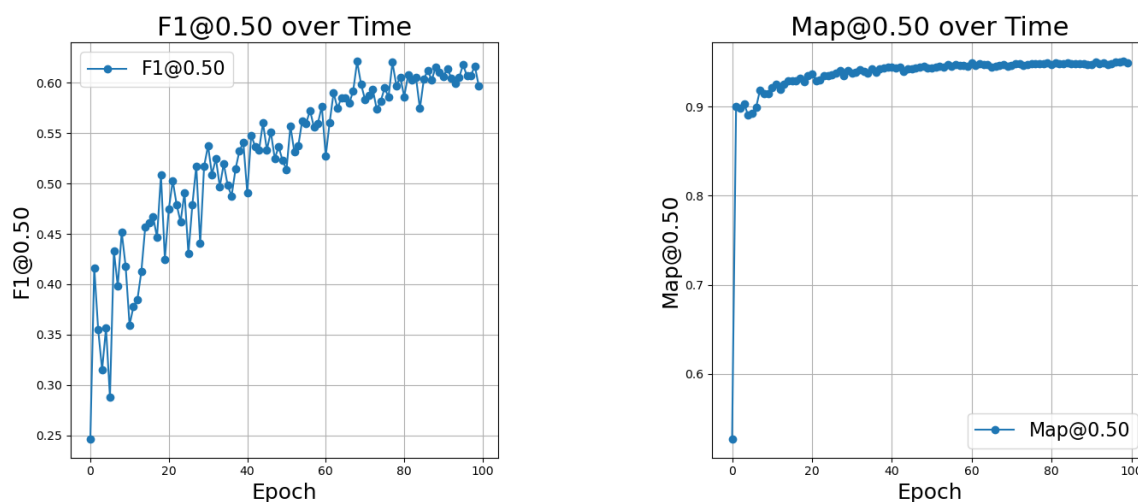


Figure 10. Validate Results: IoU Loss and Total Loss



**Figure 11.** Results: Precision@0.50 and Recall@0.50



**Figure 12.** Results: F1@0.50 and Map@0.50

## 5. Discussion

### 5.1. Advantages

The proposed Lifespan-PSO NAS algorithm demonstrates significant advantages and innovations. By integrating Particle Swarm Optimization (PSO) with a lifespan mechanism, it effectively prevents premature convergence and maintains population diversity, ensuring thorough exploration of the search space. This is particularly beneficial for complex tasks like Helmet detect in coal mine scene, where finding optimal neural architectures is crucial.

Moreover, Lifespan-PSO NAS significantly reduces computational costs. It requires only 20 GPU hours, a 20% reduction in training time compared to genetic algorithm-based NAS models. This efficiency makes it suitable for resource-constrained industrial applications, such as real-time monitoring and equipment detection in coal mines.

### 5.2. Robustness

Coal mine scene images often suffer from high noise, low illumination, and complex backgrounds, which pose challenges for traditional image recognition methods. Lifespan-PSO NAS shows remarkable robustness in these conditions. The model's accuracy improved by 5-10% on low-light and noisy images, especially under low illumination. This improvement is attributed to the use of MBConv blocks, which reduce parameter counts while enhancing adaptability to complex environments.



Additionally, the integration of Squeeze-and-Excitation Networks (SENet) allows the model to recalibrate channel-wise feature importance, focusing on the most relevant features and improving overall recognition performance.

### 5.3. Comparison with Existing Methods

Experiments demonstrate that Lifespan-PSO NAS outperforms existing traditional recognition methods across multiple evaluation metrics. For instance, it achieved an mAP@0.5 of 94.96%, surpassing YOLOv5s (84.1%) and YOLOv5m (85.4%). The recall rate also reached 97.5%, highlighting its superior performance.

### 5.4. Limitations and Future Research Directions

While Lifespan-PSO NAS demonstrates promising results, several limitations and areas for further exploration need acknowledgment. The model's performance is highly dependent on the quality and diversity of the training dataset; current results are based on a specific coal mine scene image dataset, and performance may vary across different datasets or industries. Future work should explore its generalization to diverse datasets. Additionally, the effectiveness of Lifespan-PSO NAS is sensitive to hyperparameters such as population size, inertia weight, and cognitive/social coefficients. Fine-tuning these parameters for other applications may be necessary, and automated hyperparameter tuning methods could address this limitation. To enhance generalization and accelerate convergence, incorporating domain-specific prior knowledge into the NAS process could be beneficial. Expanding its application to other industrial scenarios, such as underground engineering, could broaden its utility. Furthermore, combining Lifespan-PSO NAS with advanced computing platforms like FPGAs and ASICs could improve computational efficiency and energy consumption. These areas warrant further exploration to fully realize the potential of Lifespan-PSO NAS.

## 6. Conclusions

In this paper, we proposed a novel hybrid neural architecture search algorithm, Lifespan-PSO NAS, to address the unique challenges of coal mine scene image recognition. These challenges include high noise levels, low illumination, and complex backgrounds, as well as resource constraints commonly encountered in coal mine environments. By combining the global search capability of Particle Swarm Optimization (PSO) with the lifespan mechanism, the proposed method effectively balances exploration and exploitation in the search space, enabling the design of lightweight, efficient, and high-performance neural network architectures. The Lifespan-PSO NAS algorithm represents a significant advancement in neural architecture search for resource-constrained applications. Its ability to produce high-performance, efficient neural networks tailored for coal mine scene image recognition positions it as a valuable tool for industrial safety and monitoring. By addressing key challenges in this domain, this work paves the way for future innovations in neural architecture search and its application in real-world industrial contexts.

**Acknowledgements.** This work is supported by the National Key Research and Development Program of China (2023YFC2907600), the Science and Technology Innovation and Entrepreneurship Project of TDTEC(2022-2-TD-ZD001, 2024-TD-ZD016-04).

## Abbreviations

The following abbreviations are used in this manuscript:

PSO	Particle Swarm Optimization
NAS	Neural Architecture Search
CNN	Convolutional Neural Networks

## References

1. Ma, L., Kang, H., Yu, G., Li, Q. and He, Q., 2024. Single-Domain Generalized Predictor for Neural Architecture Search System *IEEE Transactions on Computers* **2024**.
2. Jwa, Y.; Ahn, C.W.; Kim, M.-J. EGNAS: Efficient Graph Neural Architecture Search Through Evolutionary Algorithm. *Mathematics* **2024**.
3. Franchini, G. GreenNAS: A Green Approach to the Hyperparameters Tuning in Deep Learning. *Mathematics* **2024**.
4. Real, E.; Aggarwal, A.; Huang, Y.; Le, Q.V.; et al. Regularized evolution for image classifier architecture search. *Proceedings of the AAAI Conference on Artificial Intelligence* **2019**, 33, 4780–4789.
5. Tan, M.; Le, Q.V. Efficientnet: Rethinking model scaling for convolutional neural networks. In *Proceedings of the International Conference on Machine Learning*; PMLR: 2019; pp. 6105–6114.
6. Howard, A.G. Mobilenets: Efficient convolutional neural networks for mobile vision applications. *arXiv preprint arXiv:1704.04861* **2017**.
7. Yang, E., Shen, L., Wang, Z., Guo, G., Chen, X., Wang, X. and Tao, D., 2024. Representation Surgery for Multi-Task Model Merging *Forty-first International Conference on Machine Learning* **2024**.
8. Kennedy, J.; Eberhart, R. Particle swarm optimization. In *Proceedings of IEEE International Conference on Neural Networks, volume 4*, 1942–1948, 1995.
9. Gad, A.G. Particle swarm optimization algorithm and its applications: a systematic review. *Archives of Computational Methods in Engineering* **2022**, 29, 2531–2561.
10. Yu, R.; Yang, X.; Cheng, K. Deep learning and IoT enabled digital twin framework for monitoring open-pit coal mines. *Frontiers in Energy Research* **2023**, 11, 1265111.
11. Alam, A.F.; Tabassum, M.; Shabbir, M.; et al. A Monitoring and Warning System for Hazards in Coal Mines using CNN and Sensor Fusion. In *Proceedings of the 5th International Conference on Advancements in Computational Sciences (ICACS)*; IEEE: 2024; pp. 1–9.
12. Lianbo, Ma, Shi Cheng, Yuhui Shi, Enhancing Learning Efficiency of Brain Storm Optimization via Orthogonal Learning Design *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 51, no. 11, pp. 6723-6742 **2021**.
13. Kang, H., Ma, L., Li, N., Ma, J. and Cheng, J., 2024, June. When NAS Meets Anomaly Detection: In Search of Resource-Efficient Architectures in Surveillance Video In *2024 International Joint Conference on Neural Networks (IJCNN)* (pp. 1-7) **2024**.
14. Wang, H., Li, Q., Kang, H., Hu, D., Ma, L., Tyson, G., Yuan, Z. and Jiang, Y., 2024. ParaLoupe: Real-time Video Analytics on Edge Cluster via Mini Model Parallelization *IEEE Transactions on Mobile Computing* **2024**.
15. Zoph, B. Neural architecture search with reinforcement learning. *arXiv preprint arXiv:1611.01578* **2016**.
16. Liu, H.; Simonyan, K.; Yang, Y. DARTS: Differentiable architecture search. *arXiv preprint arXiv:1806.09055* **2018**.
17. Lu, Z.; Whalen, I.; Boddeti, V.; et al. NSGA-NET: A multi-objective genetic algorithm for neural architecture search. In *Proceedings of the International Conference on Machine Learning*, 2018.
18. Ma, L.; Li, N.; Yu, G.; Geng, X.; Wang, M.; Jin, Y. Pareto-wise Ranking Classifier for Multiobjective Evolutionary Neural Architecture Search. *IEEE Transactions on Evolutionary Computation* **2024**, 28, 570–581. <https://doi.org/10.1109/TEVC.2023.3314766>.
19. Lianbo Ma, Nan Li, Peican Zhu, Keke Tang\*, Asad Khan\*, Feng Wang, and Guo Yu, A Novel Fuzzy Neural Network Architecture Search Framework for Defect Recognition with Uncertainties, *IEEE Transactions on Fuzzy Systems*, vol.32, no.5, pp. 3274-3285,10.1109/TFUZZ.2024.3373792, 2024.
20. Nan Li, Lianbo Ma, Guo Yu, Bing Xue, Mengjie Zhang, Yaochu Jin, Survey on Evolutionary Deep Learning: Principles, Algorithms, Applications and Open Issues, *ACM Computing Surveys*, vol. 56, No. 2, pp 1–34, 2023.
21. Lianbo Ma, Nan Li, Yinan Guo, Min Huang, Shengxiang Yang, Xingwei Wang, Hao Zhang. Learning to Optimize: Reference Vector Reinforcement Learning Adaption to Constrained Many-objective Optimization of Industrial Copper Burdening System, *IEEE Transactions on Cybernetics*, vol. 52, no. 12, pp. 12698-12711, December 2022.
22. Elsken, T.; Metzen, J.H.; Hutter, F. Neural architecture search: A survey. *Journal of Machine Learning Research* **2019**, 20, 1–21.
23. Stamoulis, D.; Ding, R.; Wang, D.; et al. Single-path nas: Device-aware efficient convnet design. *arXiv preprint arXiv:1905.04159* **2019**.

24. Fan, Y.; Wang, Y.; Liang, D.; et al. Low-FaceNet: Face Recognition-driven Low-light Image Enhancement. *IEEE Transactions on Instrumentation and Measurement* **2024**.
25. Wang, P.; Cheng, J. Accelerating convolutional neural networks for mobile applications. In *Proceedings of the 24th ACM International Conference on Multimedia*; ACM: 2016; pp. 541–545.
26. Zhou, Z.; Rahman Siddiquee, M.M.; Tajbakhsh, N.; et al. Unet++: A nested u-net architecture for medical image segmentation. In *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support: 4th International Workshop, DLMIA 2018, and 8th International Workshop, ML-CDS 2018, Held in Conjunction with MICCAI 2018*; Springer International Publishing: 2018; pp. 3–11.
27. Li, C.; Guo, C.; Han, L.; et al. Low-light image and video enhancement using deep learning: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2021**, *44*, 9396–9416.
28. Han, S.; Mao, H.; Dally, W. J. Deep compression: Compressing deep neural networks with pruning, trained quantization and huffman coding. *arXiv preprint arXiv:1510.00149* **2015**.
29. Tan, M.; Chen, B.; Pang, R.; et al. Mnasnet: Platform-aware neural architecture search for mobile. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; 2019; pp. 2820–2828.
30. Cai, H.; Zhu, L.; Han, S. Proxylessnas: Direct neural architecture search on target task and hardware. *arXiv preprint arXiv:1812.00332* **2018**.
31. Xie, R.; Huttunen, H.; Lin, S.; et al. Resource-constrained implementation and optimization of a deep neural network for vehicle classification. In *Proceedings of the 24th European Signal Processing Conference (EUSIPCO)*; IEEE: 2016; pp. 1862–1866.
32. Krizhevsky, A.; Sutskever, I.; Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Advances in Neural Information Processing Systems* **2012**, *25*.
33. He, K.; Zhang, X.; Ren, S.; et al. Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 2016; pp. 770–778.
34. Sandler, M.; Howard, A.; Zhu, M.; et al. Mobilenetv2: Inverted residuals and linear bottlenecks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 2018; pp. 4510–4520.
35. Mehta, S.; Rastegari, M. MobileViT: Light-weight, general-purpose, and mobile-friendly vision transformer. *arXiv preprint arXiv:2110.02178* **2021**.
36. Pham, H.; Guan, M.; Zoph, B.; et al. Efficient neural architecture search via parameters sharing. In *Proceedings of the International Conference on Machine Learning (ICML)*; PMLR: 2018; pp. 4095–4104.
37. Hu, J.; Shen, L.; Sun, G. Squeeze-and-excitation networks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 2018; pp. 7132–7141.
38. Redmon, J. Yolov3: An incremental improvement. *arXiv preprint arXiv:1804.02767* **2018**.
39. Szegedy, C.; Vanhoucke, V.; Ioffe, S.; et al. Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*; 2016; pp. 2818–2826.
40. Tan, M.; Pang, R.; Le, Q. V. EfficientDet: Scalable and efficient object detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*; 2020; pp. 10781–10790.
41. Bochkovskiy, A.; Wang, C. Y.; Liao, H. M. YOLOv4: Optimal speed and accuracy of object detection. *arXiv preprint arXiv:2004.10934* **2020**.

**Disclaimer/Publisher's Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.