

Accuracy–Energy Trade-off Analysis of Lightweight Semantic Segmentation Models for Desert Off-Road Navigation

Arbaz .^{1, ✉}, Sayama Bano¹, Manoj Kumar Srivastava¹, Bhoomi Roy¹, Simran Kumari¹, Kunal Kumar Vishwakarma¹

1. Computer Science and Engineering, Greater Noida Institute of Technology, Greater Noida, IND

Received: June 03, 2026 | Review began: June 22, 2026 | Review ended: July 14, 2026 | Published: July 14, 2026

© Copyright 2026

This is an open access article distributed under the terms of the Creative Commons Attribution License CC-BY 4.0., which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abstract

Autonomous navigation in desert environments poses great challenges due to the lack of structured roads, very similar terrain textures, and constantly varying illumination conditions. Semantic segmentation plays an important role in enabling robotic systems to understand terrain at the pixel level and make safer and more efficient navigation decisions. However, traditional evaluation of segmentation models is mainly focused on accuracy and inference speed and often neglects the effect of energy consumption, which is crucial for battery-powered edge devices operating in remote environments. This paper includes a comparative analysis of five lightweight semantic segmentation architectures, i.e., BiSeNetV2, MobileNetV3-Large, ResNet18-Seg, DDR-Light (Deep Dual-Resolution Network-Light), and SegFormer-B0, using a synthetic desert dataset generated via the Duality AI digital twin portal. The study tests each model on several parameters, such as mean Intersection over Union, inference speed, total energy use, carbon emissions estimates, and the overall efficiency score, which represents a trade-off between accuracy and energy use. Experimental results show that SegFormer-B0 achieves the highest segmentation accuracy, and MobileNetV3-Large achieves the best efficiency score because it has significantly lower power consumption. DDR-Light, even if competitive with respect to accuracy, has significantly higher energy requirements and is therefore less suitable for energy-constrained deployments. The results point to the fact that it is not enough to be accurate when deploying AI models on edge-based robotic systems, and energy-aware evaluation provides a more practical foundation for decision-making in real-world settings.

Categories: Machine Learning (ML), Computer Vision, Energy Efficiency and Sustainability

Keywords: semantic segmentation, energy efficiency, autonomous navigation, desert robotics, edge ai, carbon emission, lightweight networks

Introduction

Urban autonomous navigation is a solved-enough problem that researchers are working on more difficult terrain. Deserts are harder. The geometric reference points are nonexistent, and the lane markings are missing. The ground looks like dried grass. A shadow will trick a classifier into believing that it is a rock, and the light changes sufficiently during a day that a model trained at noon may struggle at daybreak. Terrain classification at the pixel level is not optional in this case - the path planner must have information on whether the terrain in front of the robot can be traversed or not before committing the robot to it. The catch is power. The remote desert cases where robots are used are practically always energy-limited, and deep learning inference is not free. Get the wrong model, and you are staring at fewer missions, additional recharge points, or a robot failure. We were curious about whether the accuracy numbers that prevail on

How to cite this article:

. A, Bano S, Srivastava M, et al. (July 14, 2026) Accuracy–Energy Trade-off Analysis of Lightweight Semantic Segmentation Models for Desert Off-Road Navigation. Cureus J Comput Sci 3 : es44389-026-00159-y. DOI <https://doi.org/10.7759/s44389-026-00159-y>

segmentation benchmarks have any real use when energy is included in the equation. They do not - not at all, not everywhere. The following is an analysis of five architectures in two dimensions based on a dataset tailored specifically to desert terrain.

Related work

Du et al. [1] developed ECMNet, a lightweight semantic segmentation network built on a CNN-Mamba hybrid architecture. Accuracy is maintained reasonably without introducing significant computational overhead. BiSeNetV2 [2] employs two parallel branches - one for spatial detail and one for semantic context - fused at inference. Its speed and accuracy make it a strong real-time baseline. Howard et al. [3] introduced MobileNetV3, a lightweight architecture optimized via hardware-aware neural architecture search, achieving strong accuracy-efficiency trade-offs for mobile and edge applications. Wang et al. [4] demonstrated the applicability of SegFormer-based encoder-decoder architectures with hypercolumn schemes for dense segmentation tasks, highlighting the strength of transformer-based hierarchical features in capturing long-range context. Pan et al. [5] developed DDRNet, a dual-resolution network that simultaneously processes high-resolution spatial detail and low-resolution semantic representations for real-time traffic scene segmentation. Fan et al. [6] proposed ResAt-UNet, combining ResNet backbones with attention modules for image segmentation, demonstrating the utility of skip connections and attention in fine-grained segmentation tasks. Małek et al. [7] introduced the OffRoadSynth open dataset with synthetic-data-based weight initialization for autonomous UGV navigation in off-road environments. The OFFSED dataset [8] provides challenging off-road scenes for evaluating semantic segmentation models. The Falcon dataset generated through the Duality AI platform [9] enables realistic synthetic off-road environments for training and evaluating autonomous navigation systems. Yuan et al. [10] investigated energy-efficient semantic segmentation in intelligent edge computing. Schwartz et al. [11] raised critical concerns about the environmental cost of deep learning, advocating for Green AI that accounts for energy and carbon footprint alongside accuracy. Energy-aware model optimization has also been explored through cluster pruning strategies for edge AI vision applications [12]. Previously used segmentation architectures include Ji et al. [13] who proposed parallel fully convolutional networks for semantic segmentation; Siddique et al. [14] who provided a comprehensive review of U-Net and its variants for medical image segmentation; and Anilkumar et al. [15] who developed an adaptive DeepLabv3+ framework using improved optimization for aerial image segmentation.

Materials And Methods

Dataset description

The dataset was created using the Duality AI digital twin platform [9]. Desert scenes were generated by varying terrain roughness levels, vegetation density, sun positions, and camera angles. Because the simulation engine produces pixel-level semantic labels in parallel with image rendering, manual annotation was not required, removing the labeling bottleneck typical of real-world dataset construction. The distribution of the dataset across the training, validation, and test splits is summarized in Table 1.

Dataset Statistics

How to cite this article:

. A, Bano S, Srivastava M, et al. (July 14, 2026) Accuracy–Energy Trade-off Analysis of Lightweight Semantic Segmentation Models for Desert Off-Road Navigation. Cureus J Comput Sci 3 : es44389-026-00159-y. DOI <https://doi.org/10.7759/s44389-026-00159-y>

Split	Images	Resolution	Classes
Training	2,857	540 × 960	6
Validation	317	540 × 960	6
Test	1,002	540 × 960	6

TABLE 1: Dataset distribution.
Open-source code and data hosted on Kaggle [16].

Six terrain classes were used: Vegetation, Bushes, Dry Grass, Rocks, Ground, and Sky. Every pixel of the ground-truth mask corresponds to one of these six classes. While this granularity is appropriate for the task, Rocks and Dry Ground are frequently visually similar, which poses a challenge for pixel-level boundary prediction. Desert segmentation presents four core difficulties:

- (a) Sand and trail surfaces share enough texture that pixel boundaries between them are ambiguous.
- (b) Illumination shifts with the sun angle, changing class appearance throughout the day.
- (c) Objects appear sparsely, leaving little surrounding context to anchor predictions.
- (d) The class distribution is heavily skewed toward Ground and Sky.

Class Imbalance Handling

Inverse frequency weighting is applied to address class imbalance. The weight for class n is defined as:

$$w_n = \frac{1}{\log(1 + f_n)}$$

where f_n is the pixel frequency of class n . This reduces the penalty when the model trains on rare classes such as rocks and bushes, which constitute a small proportion of total pixels.

Preprocessing

All images were converted from BGR to RGB and resized to 512×512 using OpenCV. Pixel values were scaled to a range from 0 to 1 by dividing by 255. Segmentation masks were remapped from their original class indices $\{0, 1, 2, 3, 27, 39\}$ to a contiguous index space $\{0, 1, 2, 3, 4, 5\}$ prior to training and resized to 512×512 using nearest-neighbor interpolation to preserve label integrity.

Dataset Samples

Representative training images and their corresponding ground-truth segmentation masks are shown in Figure 1 and Figure 2, respectively.



FIGURE 1: Sample training image from the synthetic desert dataset.

Source: Duality AI platform [9].



FIGURE 2: Ground truth semantic segmentation masks

Source: Duality AI platform [9].

Training procedure

All models were trained on a single NVIDIA Tesla T4 GPU using the AdamW optimizer and mixed-precision training (torch.amp.autocast with GradScaler). The weighted cross-entropy loss described above was applied consistently across all five models, and the training set was reshuffled each epoch. Model checkpoints were saved only when validation mean Intersection over Union (mIoU) improved on the previous best.

BiSeNetV2

BiSeNetV2 [2] (Figure 3) was obtained from its official open-source implementation and instantiated with six output classes. The architecture uses two parallel branches - one for spatial detail and one for semantic context - fused before the classification head. During training, the model's auxiliary outputs were combined with the primary cross-entropy loss using a fixed auxiliary loss coefficient of 0.4 per auxiliary branch.

How to cite this article:

. A, Bano S, Srivastava M, et al. (July 14, 2026) Accuracy–Energy Trade-off Analysis of Lightweight Semantic Segmentation Models for Desert Off-Road Navigation. Cureus J Comput Sci 3 : es44389-026-00159-y. DOI <https://doi.org/10.7759/s44389-026-00159-y>

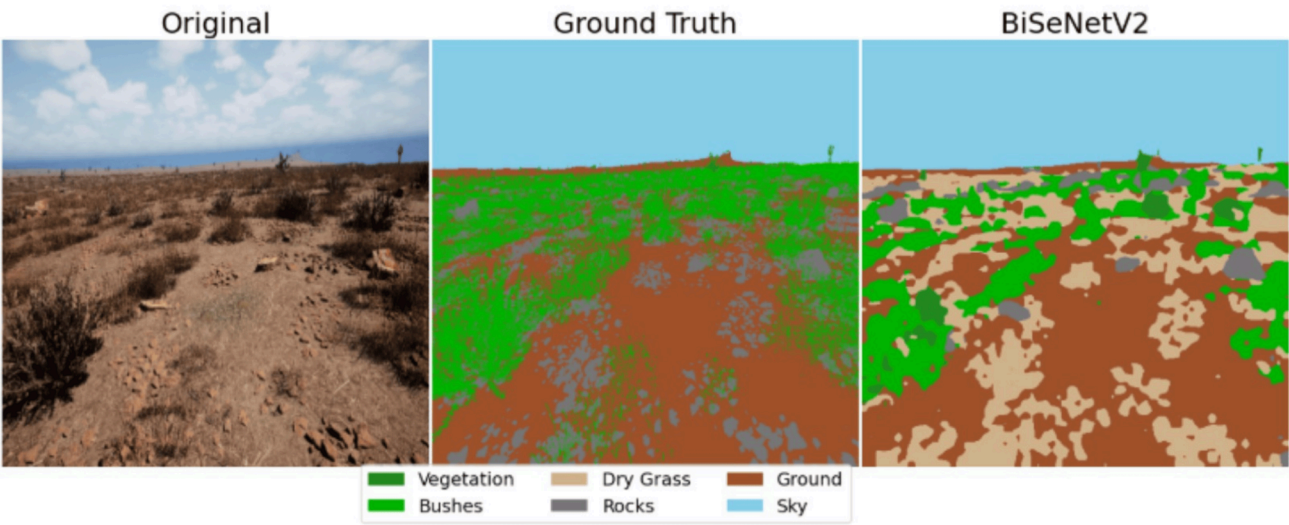


FIGURE 3: BiSeNetV2 segmentation prediction on desert terrain.

Open-source code and data hosted on Kaggle [16].

MobileNetV3-Large

MobileNetV3-Large [3] (Figure 4) was implemented using the timm library (mobilenetv3_large_100) in feature-extraction mode, without pretrained weights. The deepest backbone feature map was passed through a lightweight decoder: a 3 × 3 convolution (256 channels), batch normalization, ReLU, and a 1 × 1 classification layer, followed by bilinear upsampling to 512 × 512.

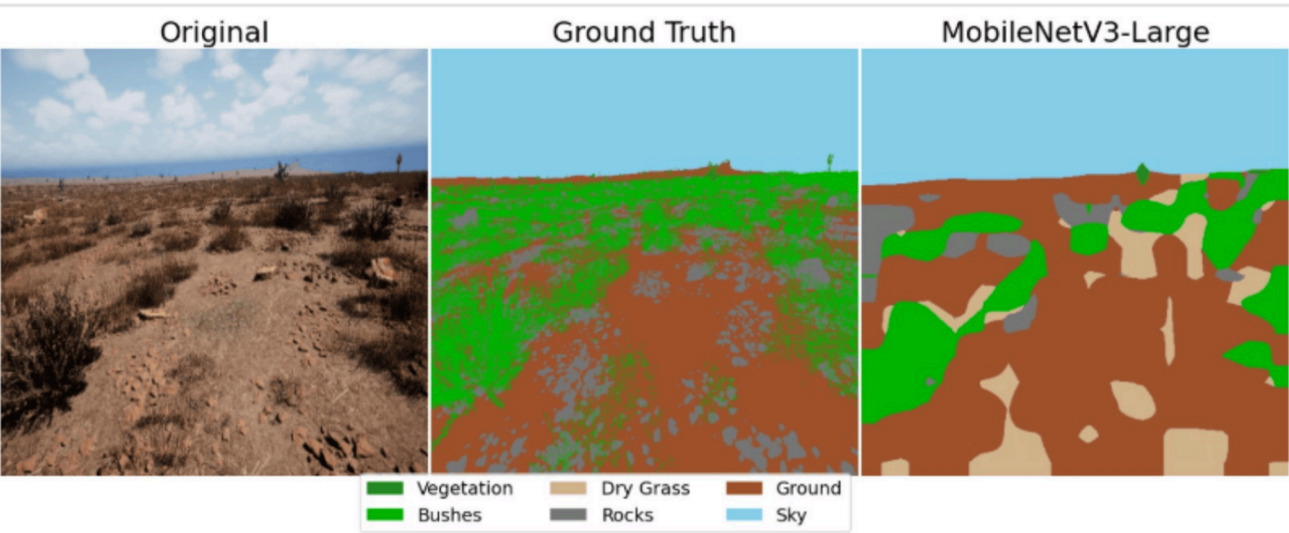


FIGURE 4: MobileNetV3-Large segmentation prediction on desert terrain.

Open-source code and data hosted on Kaggle [16].

ResNet18-Seg

How to cite this article:

. A, Bano S, Srivastava M, et al. (July 14, 2026) Accuracy–Energy Trade-off Analysis of Lightweight Semantic Segmentation Models for Desert Off-Road Navigation. Cureus J Comput Sci 3 : es44389-026-00159-y. DOI <https://doi.org/10.7759/s44389-026-00159-y>

ResNet18-Seg [6] (Figure 5) used the timm ResNet-18 backbone in feature-extraction mode, without pretrained weights, with a decoder identical in structure to MobileNetV3-Large's: a 3×3 convolution (256 channels), batch normalization, ReLU, a 1×1 classification layer, and bilinear upsampling to 512×512 .

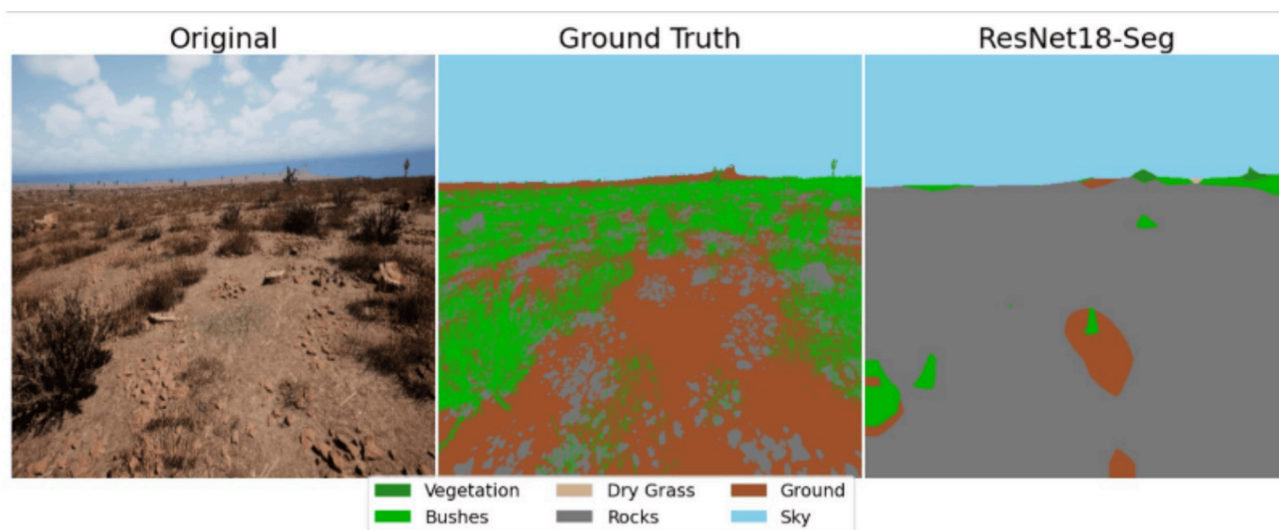


FIGURE 5: ResNet18-Seg segmentation prediction on desert terrain.

Open-source code and data hosted on Kaggle [16].

DDR-Light

DDR-Light [5] (Figure 6) was implemented as a dual-branch network. A low-resolution semantic branch used a timm ResNet-18 backbone (without pretrained weights), whose deepest feature map was bilinearly upsampled to 512×512 . A high-resolution spatial branch applied two successive 3×3 convolutions (32 channels each, each followed by batch normalization and ReLU) directly to the input image. The two branches were concatenated channel-wise and passed through a fusion module - a 3×3 convolution (128 channels), batch normalization, ReLU, and a 1×1 classification head - to produce the final prediction.

How to cite this article:

. A, Bano S, Srivastava M, et al. (July 14, 2026) Accuracy–Energy Trade-off Analysis of Lightweight Semantic Segmentation Models for Desert Off-Road Navigation. Cureus J Comput Sci 3 : es44389-026-00159-y. DOI <https://doi.org/10.7759/s44389-026-00159-y>

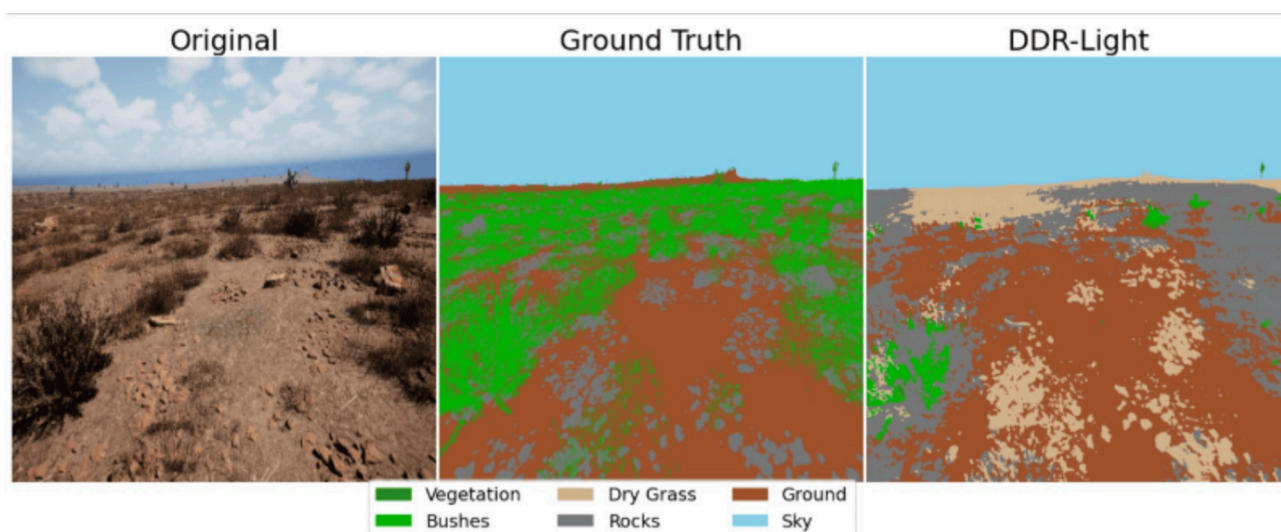


FIGURE 6: DDR-Light segmentation prediction on desert terrain.

Open-source code and data hosted on Kaggle [16].

SegFormer-B0

SegFormer-B0 [4] (Figure 7) was implemented using the HuggingFace Transformers library (nvidia/segformer-b0-finetuned-ade-512-512), fine-tuned for six output classes with `ignore_mismatched_sizes=True` to accommodate the change in output channels. Output logits were upsampled to 512×512 via bilinear interpolation.

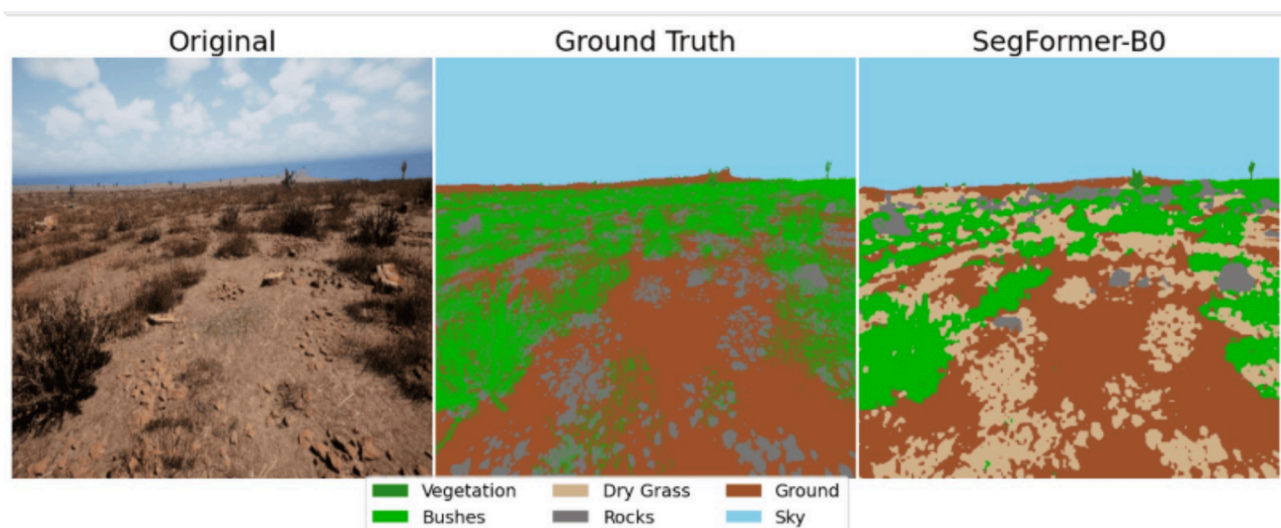


FIGURE 7: SegFormer-B0 segmentation prediction on desert terrain.

Open-source code and data hosted on Kaggle [16].

Evaluation metrics

mIoU: Computed per class as $\frac{TP}{TP+FP+FN}$ and averaged across all six terrain classes, accumulated over the full evaluation split.

How to cite this article:

. A, Bano S, Srivastava M, et al. (July 14, 2026) Accuracy–Energy Trade-off Analysis of Lightweight Semantic Segmentation Models for Desert Off-Road Navigation. Cureus J Comput Sci 3 : es44389-026-00159-y. DOI <https://doi.org/10.7759/s44389-026-00159-y>

Frames Per Second (FPS): Measured by recording the wall-clock time to run the entire test set (batch size 1) through the model with gradient computation disabled (`torch.no_grad()`), then dividing the number of test images by the elapsed time.

Energy Consumption (Joules): Measured using the NVIDIA Management Library (`pynvml`). For each test batch, GPU power draw was sampled immediately before and immediately after the forward pass via `nvmlDeviceGetPowerUsage()`; the average of the two readings was multiplied by the batch's inference time to obtain energy per batch, summed across all test batches to obtain total energy.

Carbon Emission (kg CO₂): Total energy was converted from joules to kilowatt-hours ($\div 3.6 \times 10^6$) and multiplied by a regional carbon intensity factor of 0.708 kg CO₂/kWh.

Efficiency Score (mIoU/Energy): The ratio of test mIoU to total inference energy, capturing the accuracy-per-joule trade-off relevant to battery-powered robotic platforms.

Results And Discussion

Quantitative performance

The best mIoU score of 0.3933 was achieved by SegFormer-B0 (Table 2). Transformer attention helps with the ambiguity of surface boundaries that trip up CNN models. MobileNetV3-Large had the best efficiency score at 0.00092 - roughly twice the accuracy per joule of BiSeNetV2 and roughly 13 times that of DDR-Light. DDR-Light consumed 4,633 J overall, approximately 5× higher than MobileNetV3-Large relative to the accuracy it provides.

Model	mIoU	FPS	Params (M)	Energy (J)	CO ₂ (kg)	Efficiency
BiSeNetV2	0.3680	27.71	3.50	753.53	0.000148	0.00049
MobileNetV3-Large	0.3300	26.09	5.48	356.97	0.000070	0.00092
ResNet18-Seg	0.2606	29.21	12.35	414.54	0.000082	0.00063
DDR-Light	0.3591	13.02	11.81	4,633.37	0.000911	0.00007
SegFormer-B0	0.3933	24.85	3.71	925.28	0.000182	0.00042

TABLE 2: Performance and energy results for all five architectures.

Open-source code and data hosted on Kaggle [16].
FPS, Frames Per Second; mIoU; mean Intersection over Union

Accuracy comparison

SegFormer-B0 leads on accuracy; BiSeNetV2 runs close behind with higher speed (Figure 8). ResNet18-Seg finishes last on accuracy by a clear margin. The simplicity that makes it fast also limits how much it can learn about texturally similar terrain classes.

Speed analysis

ResNet18-Seg ran fastest at 29.21 FPS. BiSeNetV2 reached 27.71 FPS with improved accuracy, making it the better combination for the speed-accuracy trade-off (Figure 8). DDR-Light bottomed out at 13.02 FPS - the dual-resolution processing has a real throughput cost, not just an energy cost.

Energy consumption analysis

The rest are within one order of magnitude, but DDR-Light is in a class of its own at 4,633 J (Figure 9). In cases of energy constraints, DDR-Light is not feasible. The high number of spatially unfused parameters (512×512 , i.e., a spatial branch), which performs convolutions on the whole input without any spatial downsampling, is the major reason for DDR-Light's disproportionate energy consumption. This essentially means that the energy consumed in large tensor reads and/or writes is due to memory bandwidth limitations on the GPU, leading to energy usage much higher than would be expected based on parameter count alone. Another possible reason is that the dual-branch computation graph is executed sequentially before fusion in PyTorch, thus restricting the ability to pipeline operations effectively. The result indicates the necessity of earlier fusion from the high-resolution branch for future lightweight dual-branch pipeline designs in order to decrease memory bandwidth pressure.

The third option, SegFormer-B0 at 925 J, is not as cost-effective as MobileNetV3-Large and BiSeNetV2, but it is not a prohibitive price either if the accuracy improvement measured against the previous two models is worth it. The same relationship exists with respect to carbon output estimates: DDR-Light's emissions (0.000911 kg CO₂) are about a factor of 10 higher than those of the other four architectures (Figure 10). Note also that all energy measurements mentioned here were taken using an NVIDIA Tesla T4 server-level GPU with a TDP of up to 70 W. The absolute energy efficiency numbers will be different on embedded platforms like the NVIDIA Jetson Nano, Xavier, and Orin, but the energy efficiency ranking in terms of architecture will be roughly similar.

Accuracy-energy trade-off

SegFormer-B0 wins on accuracy. MobileNetV3-Large wins on efficiency. BiSeNetV2 occupies a middle ground - not the most accurate and not the fastest. For most field robots running on battery-powered hardware, DDR-Light is difficult to justify. MobileNetV3-Large is the default choice; SegFormer-B0 is appropriate if the mission requires better terrain classification and has the power budget to match.

Visualization results

Figures 8-10 show the speed-accuracy, energy-accuracy, and carbon-accuracy trade-offs, respectively, across all five models.

How to cite this article:

. A, Bano S, Srivastava M, et al. (July 14, 2026) Accuracy–Energy Trade-off Analysis of Lightweight Semantic Segmentation Models for Desert Off-Road Navigation. *Cureus J Comput Sci* 3 : es44389-026-00159-y. DOI <https://doi.org/10.7759/s44389-026-00159-y>

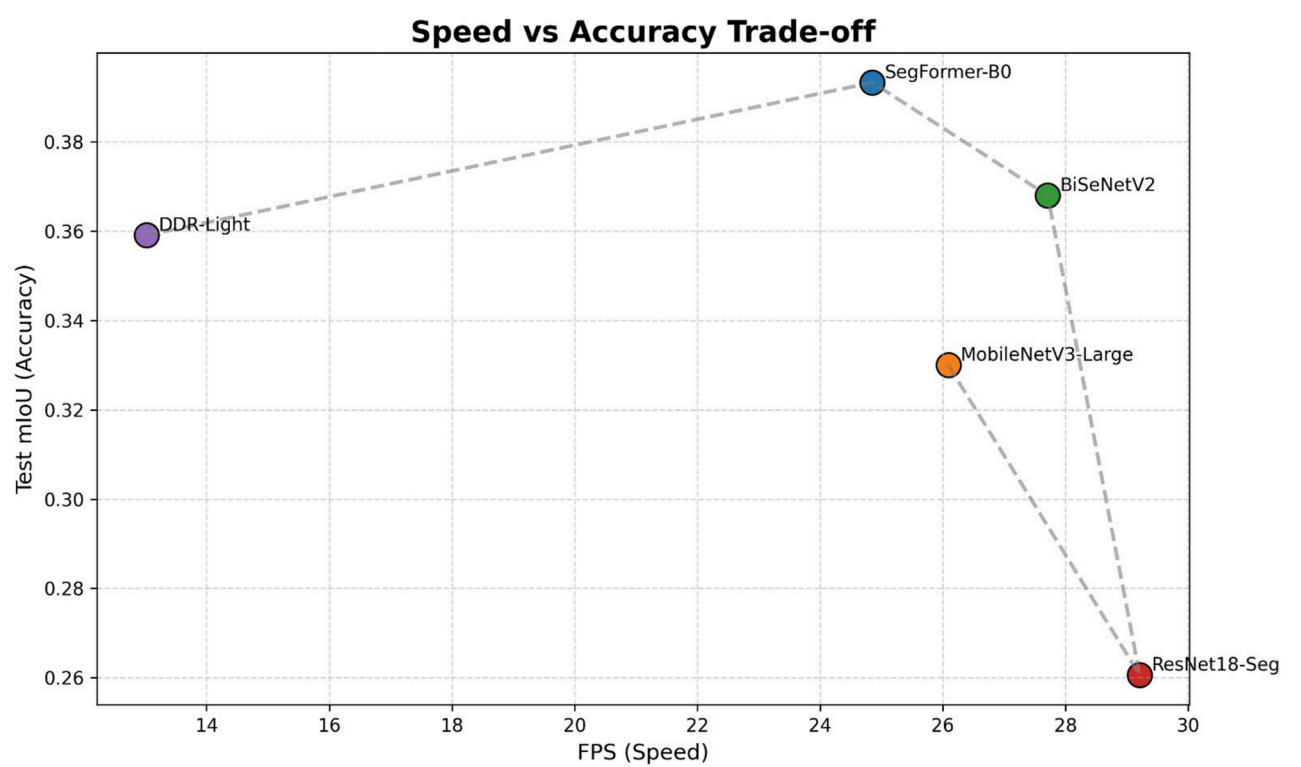


FIGURE 8: Speed vs. accuracy comparison across all five architectures.

Open-source code and data hosted on Kaggle [16].

How to cite this article:

. A, Bano S, Srivastava M, et al. (July 14, 2026) Accuracy–Energy Trade-off Analysis of Lightweight Semantic Segmentation Models for Desert Off-Road Navigation. Cureus J Comput Sci 3 : es44389-026-00159-y. DOI <https://doi.org/10.7759/s44389-026-00159-y>

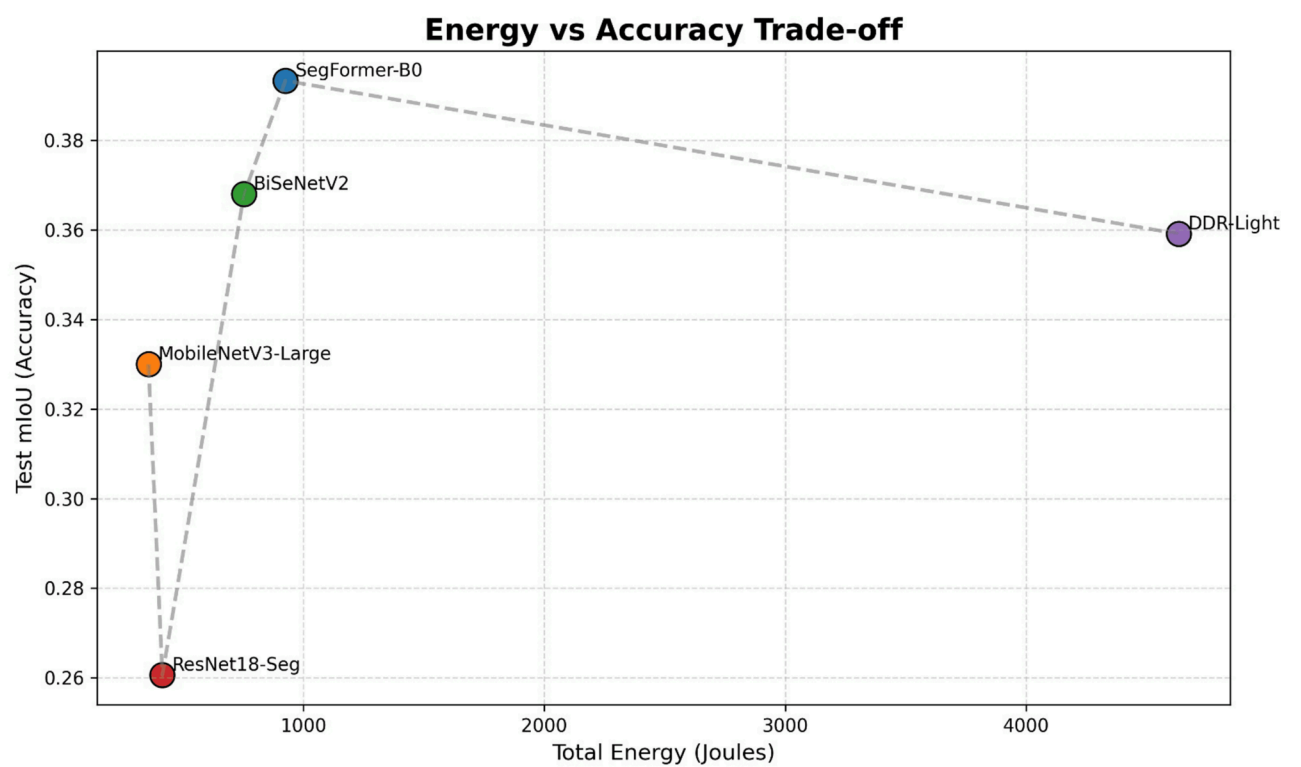


FIGURE 9: Energy consumption vs. accuracy for all five architectures.

Open-source code and data hosted on Kaggle [16].

How to cite this article:

. A, Bano S, Srivastava M, et al. (July 14, 2026) Accuracy–Energy Trade-off Analysis of Lightweight Semantic Segmentation Models for Desert Off-Road Navigation. Cureus J Comput Sci 3 : es44389-026-00159-y. DOI <https://doi.org/10.7759/s44389-026-00159-y>

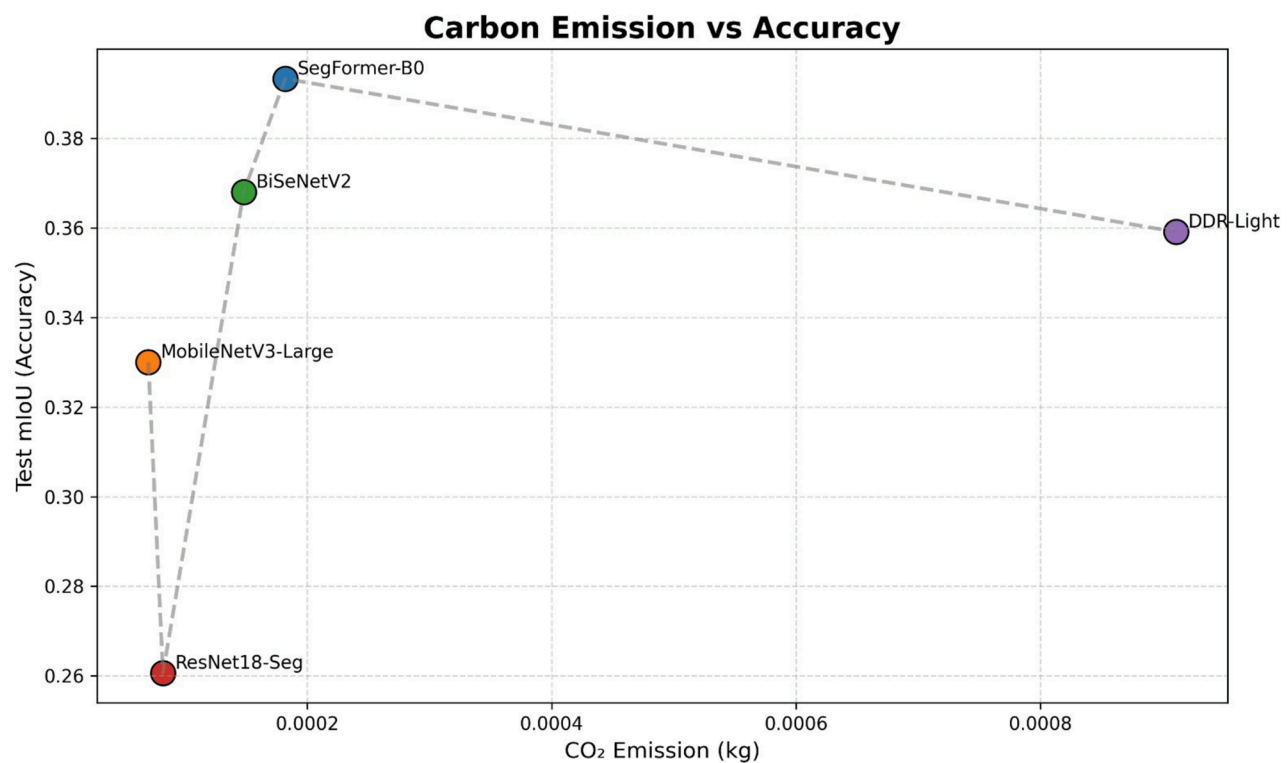


FIGURE 10: Carbon emission vs. accuracy for all five architectures.

Open-source code and data hosted on Kaggle [16].

Deployment implications

Accuracy benchmarks are useful. They are not sufficient. Battery packs in a rover over a desert have a finite amount of energy. During a mission, they use up their energy in proportion to the time spent in the mission; being more accurate is not necessarily safer to deploy if it reduces mission duration by half. For energy-limited devices, MobileNetV3-Large is the best option. It consumes the least energy, achieves the highest efficiency score (0.00092 mIoU/J), and maintains accuracy that is not disqualifyingly lower when compared with SegFormer-B0. In applications where classification accuracy is directly related to safety outcomes, or where the platform possesses sufficient power reserve to accommodate the extra energy consumption, SegFormer-B0 is suitable.

DDR-Light's energy numbers are too high for long missions with limited hardware; the dual-resolution design that processes spatial detail is also energy-intensive to operate (4,633 J) and slow to infer (13.02 FPS). This cost is largely architectural in nature, driven by full-resolution feature map processing rather than parameter count, and is not likely to be easily reduced on other hardware without modifying the architecture of the system.

It is important to note that all energy consumption data in this study were collected using an NVIDIA Tesla T4 server-grade GPU, which is not a typical edge-device architecture on which field robots run. These metrics will vary significantly when running on dedicated edge devices like NVIDIA Jetson Nano, Xavier, and Orin (TDP of 5-30 W depending on power mode). While the relative efficiency rankings - particularly the recommendation of MobileNetV3-Large for power-constrained deployments - should remain broadly consistent across platforms, users are advised to profile the hardware before making a deployment decision on battery-powered robots.

How to cite this article:

Furthermore, all models presented in this work were trained exclusively on synthetic data created using the Duality AI digital twin platform. For real-world deployment, a domain adaptation evaluation is recommended to minimize the risk of sim-to-real performance degradation. The mIoU values reported in this work should be interpreted as upper-bound estimates for operators deploying similar models on physical rovers in actual desert environments, given that they have not yet been validated on real-world imagery.

Limitations

This study has been carried out with certain limitations that have to be noted. First, all experiments were performed on a single NVIDIA Tesla T4 server-grade GPU (TDP of up to 70 W) within the Kaggle cloud environment. Absolute energy consumption, inference latency, and efficiency scores will differ on dedicated embedded edge platforms such as the NVIDIA Jetson Nano, Xavier, or Orin, which are more representative of battery-powered field robotics deployments. Hardware-specific profiling on edge devices is identified as a priority direction for future work.

Second, all models were trained and evaluated exclusively on synthetic data created using the Duality AI digital twin platform. While the platform employs physically based rendering and parametric scene variation as a form of implicit domain randomization, a sim-to-real gap is expected when deploying these models on a physical rover in a real desert environment. Model performance should be further assessed on real-world desert imagery, and explicit domain adaptation techniques such as adversarial feature alignment or style transfer should be explored in future work to bridge this gap.

Third, all CNN-based models were trained from scratch without pretrained weights. Using ImageNet-pretrained backbones for CNN-based models may improve absolute performance and alter comparative rankings.

Fourth, the energy measurement methodology samples GPU power at two points per batch; continuous power monitoring would yield more precise energy estimates.

Fifth, the study is limited to six terrain classes defined by the simulation environment; real-world desert navigation may require additional classes such as water bodies, man-made obstacles, or animals.

Conclusions

Evaluating segmentation models on accuracy alone is not the full story. This study evaluated five lightweight architectures on desert terrain and, alongside standard segmentation metrics, included energy consumption, carbon emission, and efficiency. The data reveals a clear trade-off between accuracy and efficiency. SegFormer achieves the best mIoU. MobileNetV3 achieves the highest efficiency score. DDR-Light scores worst in terms of efficiency by a large margin and does not compensate for it in terms of accuracy.

For teams deploying autonomous systems in desert environments, the efficiency score is a more honest number than mIoU when power is a real constraint - and for battery-powered robots, it almost always is.

Additional Information

Author Contributions

All authors have reviewed the final version to be published and agreed to be accountable for all aspects of the work.

Concept and design: Arbaz ., Bhoomi Roy

Acquisition, analysis, or interpretation of data: Arbaz ., Sayama Bano, Manoj Kumar Srivastava, Bhoomi Roy, Simran Kumari, Kunal Kumar Vishwakarma

Drafting of the manuscript: Arbaz ., Sayama Bano, Bhoomi Roy, Simran Kumari, Kunal Kumar Vishwakarma

Critical review of the manuscript for important intellectual content: Arbaz ., Manoj Kumar Srivastava, Bhoomi Roy

How to cite this article:

. A, Bano S, Srivastava M, et al. (July 14, 2026) Accuracy–Energy Trade-off Analysis of Lightweight Semantic Segmentation Models for Desert Off-Road Navigation. *Cureus J Comput Sci* 3 : es44389-026-00159-y. DOI <https://doi.org/10.7759/s44389-026-00159-y>

Supervision: Manoj Kumar Srivastava

Disclosures

Human subjects: All authors have confirmed that this study did not involve human participants or tissue. **Animal subjects:** All authors have confirmed that this study did not involve animal subjects or tissue. **Conflicts of interest:** In compliance with the ICMJE uniform disclosure form, all authors declare the following: **Payment/services info:** All authors have declared that no financial support was received from any organization for the submitted work. **Financial relationships:** All authors have declared that they have no financial relationships at present or within the previous three years with any organizations that might have an interest in the submitted work. **Other relationships:** All authors have declared that there are no other relationships or activities that could appear to have influenced the submitted work.

Data Availability Statements

The data (and/or code) supporting this study are openly available in Kaggle at <https://www.kaggle.com/code/arbazdevhive/research-paper-model>, reference code/arbazdevhive/research-paper-model.

The code, experimental dataset, and simulation data supporting this study are openly available across public repositories. The machine learning program code is hosted on Kaggle at <https://www.kaggle.com/code/arbazdevhive/research-paper-model>. The baseline dataset is available on Kaggle at <https://www.kaggle.com/datasets/arbazdevhive/research-paper-dataset>. The environmental asset and synthetic simulation data are hosted via the Duality AI Falcon platform at <https://falcon.duality.ai>.

Acknowledgements

This article was previously presented at the International Conference on Progressive Trends in Engineering Science and Management (PTESM 2026), held on 10th–11th April 2026. The authors would like to express their sincere gratitude to Dr. Manoj Kumar Srivastava for his guidance and supervision throughout this research work. Bhoomi Roy and Arbaz contributed equally to this work and should be considered co-first authors.

References

1. Du F, Wu S, Wang X, Ding A, Wang Z, Than JCM: [ECMNet: Lightweight semantic segmentation with efficient CNN-Mamba network](#). Science Progress. 2026, 109:1-22. [10.1177/00368504261419245](#)
2. Yu C, Gao C, Wang J, Yu G, Shen C, Sang N: [BiSeNet V2: Bilateral network with guided aggregation for real-time semantic segmentation](#). International Journal of Computer Vision. 2021, 129:3051-3068. [10.1007/s11263-021-01515-2](#)
3. Howard A, Sandler M, Chen B, et al.: [Searching for MobileNetV3](#). IEEE/CVF International Conference on Computer Vision (ICCV). 2020, 1314-1324. [10.1109/ICCV.2019.00140](#)
4. Wang Z, Wang Q, Yang Y, Liu N, Chen Y, Gao J: [Seismic facies segmentation via a SegFormer-based specific encoder-decoder hypercolumns scheme](#). IEEE Transactions on Geoscience and Remote Sensing. 2023, 61:1-11. [10.1109/tgrs.2023.3244037](#)
5. Pan H, Hong Y, Sun W, Jia Y: [Deep dual-resolution networks for real-time and accurate semantic segmentation of traffic scenes](#). IEEE Transactions on Intelligent Transportation Systems. 2023, 24:3448-3460. [10.1109/tits.2022.3228042](#)
6. Fan Z, Liu Y, Xia M, Hou J, Yan F, Zang Q: [ResAt-UNet: A U-shaped network using ResNet and attention module for image segmentation of urban buildings](#). IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing. 2023, 16:2094-2111. [10.1109/jstars.2023.3238720](#)
7. Małek K, Dybała J, Kordecki A, Hondra P, Kijania K: [OffRoadSynth open dataset for semantic segmentation using synthetic-data-based weight initialization for autonomous UGV in off-road environments](#). Journal of Intelligent & Robotic Systems. 2024, 110:76. [10.1007/s10846-024-02114-2](#)
8. [OFFSED Dataset](#). (2021). Accessed: February 25, 2026: <https://datasetninja.com/offsed>.
9. [Falcon Dataset](#). (2024). Accessed: February 26, 2026: <https://www.duality.ai/>.

How to cite this article:

. A, Bano S, Srivastava M, et al. (July 14, 2026) Accuracy–Energy Trade-off Analysis of Lightweight Semantic Segmentation Models for Desert Off-Road Navigation. Cureus J Comput Sci 3 : es44389-026-00159-y. DOI <https://doi.org/10.7759/s44389-026-00159-y>

10. Yuan X, Li H, Ota K, Dong M: [Building energy efficient semantic segmentation in intelligent edge computing](#). IEEE Transactions on Green Communications and Networking. 2024, 8:572-582. [10.1109/tgcn.2023.3321113](#)
11. Schwartz R, Dodge J, Smith NA, Etzioni O: [Green AI](#). Communications of the ACM. 2020, 63:54-63. [10.1145/3381831](#)
12. Gamanayake C, Jayasinghe L, Ng BKK, Yuen C: [Cluster pruning: an efficient filter pruning method for edge AI vision applications](#). IEEE Journal of Selected Topics in Signal Processing. 2020, 14:802-816. [10.1109/jstsp.2020.2971418](#)
13. Ji J, Lu X, Luo M, Yin M, Miao Q, Liu X: [Parallel fully convolutional network for semantic segmentation](#). IEEE Access. 2021, 9:673-682. [10.1109/access.2020.3042254](#)
14. Siddique N, Paheding S, Elkin CP, Devabhaktuni V: [U-Net and its variants for medical image segmentation: A review of theory and applications](#). IEEE Access. 2021, 9:82031-82057. [10.1109/access.2021.3086020](#)
15. Anilkumar P, Venugopal P, Maddikunta PKR, Gadekallu TR, Al-Rasheed A, Abbas M, Soufiene BO: [An adaptive DeepLabv3+ for semantic segmentation of aerial images using improved Golden Eagle Optimization Algorithm](#). IEEE Access. 2023, 11:106688-106705. [10.1109/access.2023.3318867](#)
16. [Semantic segmentation deep learning model training](#). Accessed: March 1, 2026: <https://www.kaggle.com/code/arbazdevhive/research-paper-model>.

How to cite this article:

. A, Bano S, Srivastava M, et al. (July 14, 2026) Accuracy–Energy Trade-off Analysis of Lightweight Semantic Segmentation Models for Desert Off-Road Navigation. Cureus J Comput Sci 3 : es44389-026-00159-y. DOI <https://doi.org/10.7759/s44389-026-00159-y>