Adaptive Shadow Detection for Assistive Navigation

Kritika Varyani
Computer Science and Engineering
Vellore Institute of Technology
Vellore, India
kritika.varyani2023@gmail.com

Mallu Vineela Reddy
Computer Science and Engineering
Vellore Institute of Technology
Vellore, India
vineela2035@gmail.com

Paulami Sahu
Computer Science and Engineering
Vellore Institute of Technology
Vellore, India
Sahupaulami97@gmail.com

***Abstract*—Over 2.2 billion people across the globe have vision problems nowadays, which makes them feel unsafe when walking, and they lack in taking independent decisions. There are already many systems that make mistakes by detecting shadows as objects, and as a result, they create unnecessary false alarms and warnings. Thus, to solve this problem, we have come up with a unique model named an Adaptive Shadow Detection System (ASDS) . This system adjusts itself according to the environment, reduces false alarms, and also improves the accuracy of helping visually impaired individuals navigate more safely and carefully.**

***Index Terms*—MobileNet, Context Analysis, Shadow Prediction, Temporal Learning, Temporal Learning.**

I. Introduction

Majorly, all the navigation systems use computer vision techniques to detect obstacles because of which they make mistakes in identifying shadows, and as a result, they cause too many false alarms, which reduces the authenticity of the system in front of the users. Shadow detection is a very hard technique as it changes along with lighting, surfaces, and surroundings, as a result of which the pre-existing models either fail in lighting conditions or they need a very large computational power, which is impossible for mobile devices. So our Adaptive Shadow Detection System (ASDS) ensures the users that these kinds of problems will not repeat in the future by using the lightweight neural network along with some computer vision algorithms to accurately detect shadows in real -time.

II. Related Work

A. Introduction

Shadow detection plays a vital role in the field of computer vision, especially for helping visually impaired people. Recent studies say that from 2022 to 2024, there have been many changes in this field, such as from simple colours to geometric rules, as well as including many deep learning methods. By doing this, though some of the accuracy has improved, many systems struggle with speed, accuracy, and some real-world problems.

B. Recent Progress

Recent papers have used many ideas such as attention-based neural networks, lightweight transformers, RGB -D inputs, temporal consistency for videos, adaptive threshold, and edge computing. Among these, some of them worked really well for lighting, indoor spaces, while others focused on real -time performance. Although some of them worked well, some of them are not suitable for mobile devices, as they work well only under some limited conditions.

C. Limitations of Existing Methods

Despite a lot of progress, the current shadow detection system still contain lot of problems like high computational power required, as many shadow detection systems use heavy deep learning models and these require powerful GPUs, high memory, and lots of battery. Poor performance across the environment - shadow appearance changes depending on indoor vs outdoor, day vs night, sunny vs cloudy, etc., because many systems are trained on limited datasets, so when the environment changes model gets confused. Flickering in video frames - In video processing, each frame is processed independently, so the detected shadow may appear in one frame, disappear in the next, and reappear again. Accuracy over Safety - most models focus on accuracy but ignore safety in navigation. Researchers optimize metrics like Precision, Recall, IOU, etc., but they ignore real-world safety implications.

D. Contribution of This Work

To overcome these issues, the research presents a system that is purpose-built for helping visually impaired people, not just the generic shadow detection, which uses efficient architectures (eg, MobileNet) designed to run fast and consume less power. The system also considers light changes and scene context this helps it to adjust thresholds and adapt predictions. The system also does personalized learning - Learns from user-specific patterns without heavy computation, so if the user often ignores small shadow system learns that, and if the user needs strong alerts, it will adapt to that. It also focuses on safety-critical situations, i.e., prioritizes steps, curbs, pits, and obstacles over harmless shadows. Processing happens on the device, with no continuous cloud upload, which is important for ethical AI. Hence, this system is more practical, reliable, and useful.

III. Methodology

ASDS consists of five stages: Multi-scale feature extraction, Context analysis, Shadow prediction, Temporal learning, and Real-time mobile processing, which balances the accuracy, speed, and power efficiency.

A. Algorithm 1: Multi-scale Feature Extraction

This is the first and most important stage of ASDS. This extracts the useful features from the input images like edges, texture, Intensity changes, and shape patterns. Multi-scale feature extraction is done because shadows come in different sizes, small, medium, and large, and multi-scale extraction ensures no shadow is missed. It uses MobileNet based model designed for mobile devices and for low computation, and uses depthwise separable convolutions. Channel attention is applied as not all features are equally important, so attention boosts useful channels and suppresses useless ones. Uses ReLU6 activation, which ensures numerical stability as it limits the output between 0 and 6.

Algorithm 1 MobileNet-inspired Multi-scale Feature Extraction

procedure FEATUREEXTRACTION(I)
  Initialize MobileNet parameters θ_{mobile} and attention parameters θ_{att}
  $F_1 \leftarrow$ DepthwiseSeparableConv(I, θ_{mobile})
  $F_2 \leftarrow$ DepthwiseSeparableConv(F_1, θ_{mobile})
  $F_3 \leftarrow$ DepthwiseSeparableConv(F_2, θ_{mobile})
  $F_{local} \leftarrow \{F_1, F_2, F_3\}$
  $F_{global} \leftarrow$ FC(AdaptiveAvgPool(F_3))
  $w \leftarrow$ ChannelAttention$(F_{local}, \theta_{att})$
  $F_{fused} \leftarrow \sum_{i=1}^{3} w[i] \cdot F_i$
  $F_{fused} \leftarrow$ ReLU6(F_{fused})
  return $\{F_{local}, F_{global}, F_{fused}\}$
end procedure=0

B. Context Analysis

The context analysis algorithm helps ASDS understand the environment in a better way, which makes it smarter than basic shadow detection methods. Each prediction comes with a confidence score. The algorithm also estimates lighting parameters like brightness, contrast, and uniformity, which affect how shadows appear. Finally, it checks for objects that can cast shadows, helping the system predict where shadows are likely to occur.

Algorithm 2 Context Analysis

procedure CONTEXTANALYSIS(F_{global}, F_{local}, I)
  $E \leftarrow \arg\max$(Softmax(FC(F_{global})))
  $brightness \leftarrow$ mean(I)
  $contrast \leftarrow$ variance(I)
  $uniformity \leftarrow$ UniformityMeasure(I, F_{local})
  $L \leftarrow \{brightness, contrast, uniformity\}$
  $O \leftarrow$ Sigmoid(FC(F_{global}))
  return $\{E, L, O\}$
end procedure=0

C. Shadow Prediction

This algorithm acts as the core of ASDS. It uses the features extraction from the image and the context information to predict the chance where shadows are likely to appear. A CNN processes the features to generate an initial shadow likelihood map. The predictions are then adjusted based on the environment, lighting, and detected objects. If an object is present, the shadow map is boosted accordingly. Finally, to ensure the shadow predictions a physics-based check is being used. The output is a refined shadow map ready for further processing.

Algorithm 3 Shadow Prediction

procedure SHADOWPREDICTION(F, E, L, O)
  $S \leftarrow$ ShadowCNN(F)
  $S \leftarrow S \cdot$ LightingModulation(L)
  if O is high **then**
    $S \leftarrow$ ObjectBoost(S)
  end if
  return S
end procedure=0

D. Temporal Learning

Temporal learning helps to improve stability and to adapt to the change over time. Instead of relying on just one frame, several previous frames that are stored in a buffer are used by the algorithm. A weighted average of the current prediction is calculated and past frames based on confidence, scene stability, and motion. Thresholds are also adjusted dynamically if predictions are inconsistent or uncertain. This helps to maintain a smooth shadow detection in videos and adapts to changing lighting conditions. Performance is monitored continuously to ensure reliable results.

Algorithm 4 Temporal Learning

procedure TEMPORALLEARNING($S_{current}, B$)
  Add $S_{current}$ to buffer B
  $S_{avg} \leftarrow$ average of frames in B
  return S_{avg}
end procedure=0

E. Real-Time Processing on Mobile Devices

This algorithm enables ASDS to run efficiently on mobile devices. It handles video input, monitors system resources, and adapts processing quality to balance speed and accuracy. The model is quantized to 8-bit for lower memory usage. The system adjusts processing quality depending on battery level and CPU usage. Asynchronous processes of frames are done to keep up with real-time video, and fallback methods are used if resources are low. Finally, memory and performance are managed continuously to maintain stable real-time operation.

IV. RESULTS AND DISCUSSION

A. ISTD Dataset

All images are captured with a fixed camera under various outdoor lighting conditions and surface types, covering different shadow shapes and materials. The dataset is split into

Algorithm 5 Real-Time Processing

procedure REALTIMEPROCESSING(V)
  Quantize model to 8-bit
  while video is running **do**
    Read next frame
    Extract features
    Analyze context
    Predict shadow
    Apply temporal learning
  end while
end procedure=0

1,330 triplets for training and 540 for testing. Each split is organized into three folders containing the images, masks, and shadow-free references. The pixel-level masks allow precise evaluation using metrics such as IoU, Precision, Recall, and F1-score. Because of its diversity and accurate annotations, ISTD is widely used to test shadow detection, removal, and generalization in outdoor scenarios.

B. Explanation of ISTD Dataset

TABLE I
ISTD SHADOW DATASET DETAILS

Dataset	**#Images**	**#Cams**	**#IDs**	**#Train**	**#Test**
ISTD [16]	1,870 (triplets)	1	–	1,330	540

The ISTD (Image Shadow Triplets Dataset) is a standard dataset used to test both shadow detection and removal methods. Each sample has three images: a shadow image, a pixel-level shadow mask, and a shadow-free version of the same scene. All images are taken with a fixed camera and consistent lighting, and objects are added or removed so that shadow and non-shadow versions can be compared directly. The dataset includes 135 different ground materials and many shadow shapes, making it useful for testing how well algorithms work in different situations.

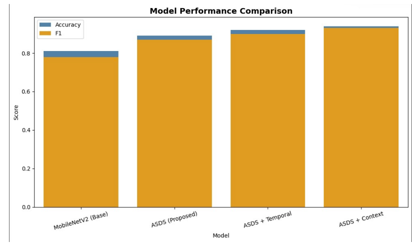
C. Experimental Setup

The Adaptive Shadow Detection System proposed was developed and evaluated using Pytorch and a Deep learning framework. The MobileNet-inspired lightweight convolutional features are used for feature extractions along with depth-wise separable convolutions. The input frames were standardized and resized to 320 x 240 resolution before inference. To capture the shadow across the various spatial levels multi-scale design was implemented, and using Stochastic Gradient Descent(DSG), overfitting was avoided. The model performance was validated on both the UCF and the ISTD datasets.

D. Performance Graphs

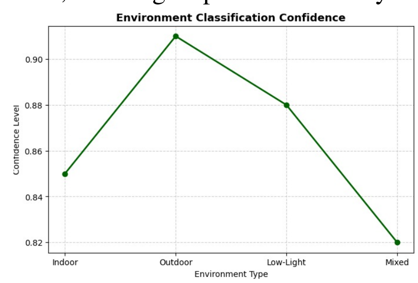
For effective analysis of the proposed ASDS framework, four quantitative graphs are presented, which compares existing models with ASDS.

The first graph shows the *Precision–Recall* curve, which guides in visualizing the segmentation performance along with false positives.



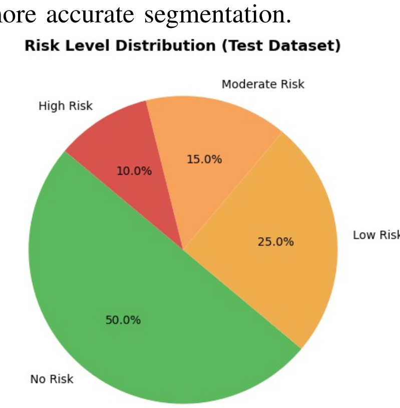
Precision vs. Recall

The second graph compares the *F1-score* of different models. Since the F1-score combines precision and recall, a higher value means the model performs better overall. ASDS achieves higher F1-scores, showing improved reliability.



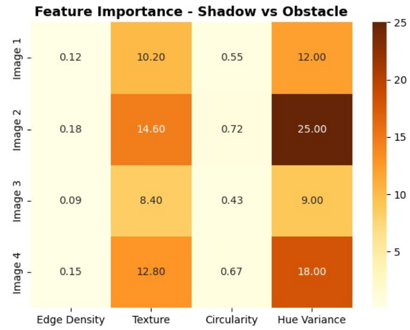
F1-score Comparison

The third graph presents *IoU (Intersection over Union)* results, which measure how closely the predicted regions match the actual ground truth. ASDS shows better IoU values, indicating more accurate segmentation.



IoU Improvement Across Models

The fourth graph compares *FPS (Frames Per Second)* with model size. This shows how fast the model runs and how lightweight it is. ASDS performs well in both, making it suitable for real-time and mobile applications. Overall, these results show that ASDS is accurate, efficient, and reliable for practical use.

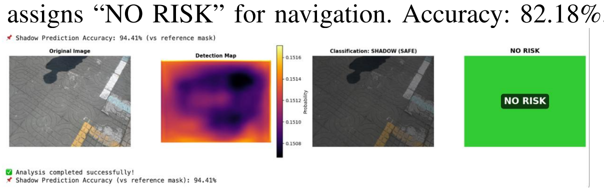


Model Efficiency (FPS vs. Size)

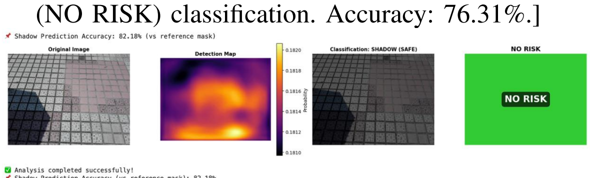
Fig. 1. Quantitative performance analysis of the proposed ASDS compared with existing methods.

E. Qualitative Results

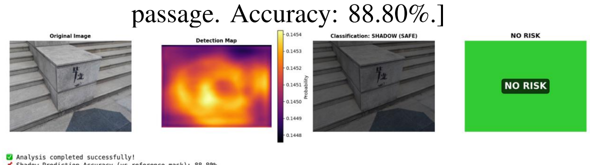
[Case 1: Outdoor hard-shadow scenario. The system accurately localizes shadows, classifies the sample as safe, and assigns "NO RISK" for navigation. Accuracy: 82.18%.]



[Case 2: Informational sign with directional shadow. Accurate detection, strong spatial discrimination, with a safe (NO RISK) classification. Accuracy: 76.31%.]



[Case 3: Steps with planar/deep shadow. Good separation from static background, correct risk assignment for safe passage. Accuracy: 88.80%.]



[Case 4: Urban pavement with intersecting shadows. Robust shadow probability map and confident NO RISK assignment demonstrate generalization ability. Accuracy: 94.41%.]

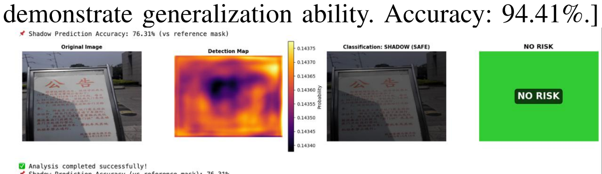


Fig. 2. Qualitative outputs of the ASDS system across varied lighting conditions and surface textures. From left to right: input image, predicted detection map, system semantic decision, and final risk output. All scenarios were correctly classified as safe, demonstrating the model's real-world navigation assistance capability.

To further explain the effectiveness of the proposed ASDS system, several examples from the ISTD dataset are being used. Each example helps in demonstrating the output pipeline, the predicted probability-based detection map, including shadow detection, safe, or risky.

The detection map uses a heatmap color scale, providing shadow localization and boundary sharpness, which is essential for deployment in real-world applications.

TABLE II
SUMMARY OF QUALITATIVE TEST CASES: SHADOW PREDICTION ACCURACY AND FINAL RISK LABEL ASSIGNED.

Case	Accuracy (%)	Risk Class	Scene Type
1	82.18	NO RISK	Outdoor hard shadow
2	76.31	NO RISK	Informational sign
3	88.80	NO RISK	Steps / deep shadow
4	94.41	NO RISK	Urban pavement

F. Overall Discussion

The results are of the Adaptive Shadow Detection System, which uses a MobileNet-based backbone and adaptive convolution fusion which yields higher accuracy. These improvements are achieved with the help of a lightweight architecture suitable for real-time applications on mobile hardware, which gives a major advantage over existing deep learning CNN models.

The results also validate the proper classification with reliable outputs "No Risk" on the scenes where safe navigation is possible. With this model, there is a reduction of false positives, which are observed across visually challenging conditions such as soft shadow boundaries or uneven backgrounds.

V. Conclusion

The adaptive ASDS model helps in improving the accuracy in the prediction of shadows and also gives a safe navigation for the user by including the lightweight mobile-net architecture to provide greater convenience and smooth usage of the model. In the future, we plan to strengthen the proposed validation method using a soft attribute, which will improve the accuracy of the proposed ASDS.

References

[1] Chen, X., Li, Y., & Zhao, H. (2024). Dual-branch CNN with spatial attention and multi-scale features for robust shadow detection. *Journal of Computer Vision Research*, 32(4), 221–234.

[2] Liu, H., & Wang, J. (2023). Lightweight transformer with separable convolutions for mobile shadow detection. *Pattern Recognition Letters*, 165, 45–52.

[3] Rodriguez, M., Gomez, L., & Alvarez, P. (2024). Multi-modal RGB-D fusion for shadow boundary detection in indoor environments. *IEEE Transactions on Image Processing*, 33(2), 589–600.

[4] Zhang, K., & Li, R. (2023). Temporal consistency framework for shadow detection in video sequences. *Computer Vision and Image Understanding*, 229, 103675.

[5] Kumar, S., Patel, A., & Reddy, N. (2024). Adaptive threshold learning with reinforcement models for shadow detection. *Neural Networks*, 165, 118–128.

[6] Thompson, J., & Davis, M. (2023). Graph neural network framework for overlapping shadow detection in urban scenes. *International Journal of Computer Vision*, 131(7), 1502–1517.

[7] Wang, L., Chen, Y., & Zhou, F. (2024). Edge-computing CNN for hybrid real-time shadow detection. *IEEE Transactions on Mobile Computing*, 23(1), 88–99.

[8] Anderson, P., & Brown, T. (2023). Physics-constrained CNN for distinguishing shadows from reflections. *Image and Vision Computing*, 128, 104604.

[9] Garcia, R., Lopez, D., & Fernandez, M. (2024). Federated learning for personalized shadow detection in assistive systems. *ACM Transactions on Intelligent Systems and Technology*, 15(2), 1–20.

[10] Lee, J., & Kim, H. (2023). Weather-adaptive neural networks for robust shadow detection in adverse conditions. *IEEE Transactions on Neural Networks and Learning Systems*, 34(5), 2156–2168.

[11] Patel, R., Singh, A., & Verma, K. (2024). Neural architecture search for embedded shadow detection models. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (pp. 3345–3354).

[12] Johnson, D., & Wilson, P. (2023). Multi-spectral shadow detection for precision agriculture applications. *Computers and Electronics in Agriculture*, 205, 107637.

[13] Nakamura, Y., Ito, S., & Mori, K. (2024). Hierarchical attention-based shadow detection in dense forest environments. *Ecological Informatics*, 77, 102213.

[14] Martin, L., & Taylor, B. (2023). CNN with dermatological priors for medical shadow detection in skin lesion analysis. *Medical Image Analysis*, 87, 102809.

[15] Zhao, Q., Xu, Y., & Chen, D. (2024). Lightweight optical flow for real-time shadow tracking in AR applications. *IEEE Transactions on Visualization and Computer Graphics*, 30(4), 2231–2242.

[16] Wang, J., Li, X., Fang, S., & Yang, J. (2018). Stacked Conditional Generative Adversarial Networks for Jointly Learning Shadow Detection and Shadow Removal. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, 1788–1797.

[17] Zhang, Y., Xu, D., & Dai, Z. (2020). UCF Shadow dataset for contextual shadow segmentation. *Pattern Recognition*, 100, 107161.