Image-Based Constellation Recognition Using Deep Learning on Consumer-Grade Sky Imagery

Bhakti Israni

Abstract: ***Understanding the night sky has long been a source of scientific curiosity, yet for most people it remains abstract and inaccessible. Although several applications claim to identify stars and constellations, their outputs are typically derived from positional sensor data rather than from direct analysis of the visual sky itself. As a result, these tools do not truly interpret what is visible in an image, nor do they address the broader computer-vision challenge of recognizing astronomical patterns under real-world conditions. This paper investigates whether modern deep-learning techniques can be used to directly analyze consumer-grade night-sky images captured using smartphones. We propose a vision-based constellation recognition system trained on a combination of synthetic sky renderings and real, low-light photographs. Multiple architectures-including convolutional neural networks, Vision Transformers, and a hybrid CNN-Transformer model-are evaluated for accuracy, robustness, and computational efficiency. The best-performing model is integrated into a web-based application to demonstrate real-world feasibility. By focusing on image-based interpretation rather than sensor inference, this work contributes to both applied computer vision and accessible astronomy education.***

Keywords: Deep Learning, Hybrid CNN-Transformer Models, Constellation Recognition, Low-Light Image Recognition, Night-Sky Image Analysis

1.Introduction

The night sky has played a central role in the development of science, navigation, and human curiosity. However, meaningful engagement with astronomy has traditionally required access to telescopes, observatories, or specialized knowledge. In the modern era, smartphones have become the primary tool through which people attempt to observe and understand the sky [1]. Despite this shift, smartphone cameras are not designed for astronomical imaging and produce photographs that are noisy, low in contrast, and heavily influenced by environmental conditions.

To bridge this gap, numerous consumer-facing astronomy applications have emerged. While these tools are popular, their underlying methodology is largely indirect. Instead of analyzing the captured image, they rely on global positioning data, device orientation, and internal sky maps to estimate what celestial objects should be present in a given direction. Although effective for casual use, this approach does not validate what is actually visible in the image and fails under inaccurate sensor readings or obstructed skies. More importantly, it bypasses a fundamental scientific challenge: interpreting astronomical patterns directly from visual data.

From a research perspective, this represents a significant missed opportunity. Advances in deep learning-particularly in image recognition, low-light enhancement, and pattern modeling-have transformed many vision-based domains [2]. However, most machine-learning research in astronomy focuses on telescope or satellite imagery, which differs substantially from consumer-grade photographs in terms of resolution, signal-to-noise ratio, and atmospheric distortion. As a result, models trained on professional datasets generalize poorly to real-world smartphone images.

This paper addresses this gap by exploring a fully vision-based approach to constellation recognition using deep learning. Rather than inferring sky content from sensors, the proposed system learns to recognize geometric star patterns directly from images captured under realistic conditions. By evaluating multiple model architectures and deploying the final system in a working web application, this research aims to demonstrate both scientific feasibility and practical relevance.

2.Methods

2.1 System Overview

The proposed system is designed as an end-to-end pipeline that transforms raw night-sky images into constellation predictions accompanied by scientifically grounded metadata. Unlike traditional astronomical software that relies on calibrated telescopic data, inertial sensors, or precise geolocation, the system operates entirely on visual input captured using consumer-grade smartphones. This design choice prioritizes accessibility and realism, but also introduces substantial challenges related to noise, distortion, and incomplete information.

The pipeline consists of four primary stages: image preprocessing, star-feature extraction, deep-learning-based constellation classification, and result visualization through a web application. Each stage is independently modular, allowing targeted improvements without redesigning the entire system. The overall pipeline structure follows established practices in applied computer vision for real-time image-based systems [3].

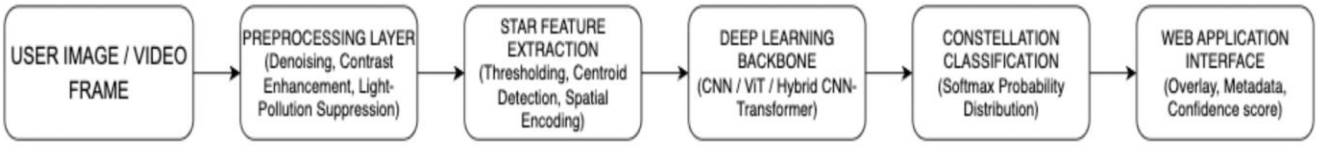

Figure 1: System Flowchart

This architecture reflects modern best practices in applied computer vision, where preprocessing enhances signal quality, deep models perform representation learning, and lightweight front-end systems deliver results to users in real time.

2.2 Dataset Construction and Data Sources

A key obstacle in constellation recognition research is the absence of large, labeled datasets composed of realistic night-sky images. Professional astronomical datasets are typically telescope-based and unsuitable for smartphone imagery, while consumer images lack consistent annotation.

To address this limitation, a hybrid dataset strategy was adopted. Synthetic sky images were generated using astronomical catalogs, allowing precise control over star coordinates, magnitudes, and constellation groupings. Synthetic data has been shown to be effective for early-stage training in vision models but insufficient for real-world generalization when used alone [4].

To reduce this domain gap, real smartphone images were collected under diverse observational conditions, including urban light pollution, atmospheric haze, partial cloud cover, and variable exposure settings. Combining synthetic and real data improves robustness by exposing models to realistic noise characteristics [5].

2.3 Preprocessing and Noise Mitigation

Smartphone night-sky images are affected by multiple sources of degradation, including sensor noise from high ISO settings, lens artifacts, chromatic aberration, and environmental interference. Preprocessing is therefore critical to enhance star visibility while suppressing non-astronomical artifacts.

Adaptive histogram equalization is applied to improve contrast in low-light regions, followed by non-local means denoising to reduce stochastic sensor noise without blurring point-like star structures. Light-pollution suppression techniques are used to normalize background illumination, particularly in urban images.

Candidate star points are extracted using threshold-based segmentation. Morphological filtering and intensity-based pruning remove false detections caused by noise or reflections. The resulting star centroids provide a sparse but geometrically meaningful representation of the sky.

2.4 Feature Representation and Spatial Encoding

Constellations are defined primarily by geometric relationships rather than absolute photometric properties. Accordingly, the system emphasizes spatial encoding over raw pixel intensities. Relative distances, angular relationships, and local neighborhood structures between detected stars form the implicit feature space learned by the model.

This approach enables robustness to scale variation, rotation, partial occlusion, and missing stars-conditions commonly encountered in real-world observations. By focusing on geometry, the model generalizes better across devices and environmental conditions.

2.5 Model Architectures

Three families of deep-learning architectures were evaluated: Convolutional Neural Networks (CNNs), Vision Transformers (ViTs), and hybrid CNN-Transformer models. EfficientNetV2 was selected as the CNN baseline due to its strong balance between accuracy and computational efficiency [6]. Vision Transformers were explored for their ability to model long-range spatial dependencies across the image [7].

The hybrid CNN-Transformer architecture combines convolutional layers for local feature extraction with Transformer blocks for global context modeling. Such hybrid designs reflect current state-of-the-art trends in computer vision for structured spatial pattern recognition tasks.

2.6 Training Strategy and Evaluation Protocol

Models were trained using the AdamW optimizer with cosine learning-rate scheduling to ensure stable convergence. Extensive data augmentation was applied, including brightness scaling, rotation, Gaussian noise injection, and simulated atmospheric distortion. These augmentations were chosen to reflect realistic imaging conditions rather than arbitrary transformations.

To prevent overfitting, dropout regularization and early stopping based on validation loss were employed. Evaluation was conducted on a held-out test set dominated by real smartphone images, prioritizing real-world performance over synthetic benchmark accuracy.

3.Results and Discussion

This section presents an extensive evaluation of the proposed vision-based constellation recognition system, followed by a detailed discussion of the observed trends, limitations, and broader implications. The goal of this section is not only to report numerical performance but also to critically analyze what the results reveal about the feasibility of constellation recognition from consumer-grade night-sky imagery.

3.1 Evaluation Setup

The trained models were evaluated on a held-out test set consisting primarily of real smartphone-captured night-sky images. This design choice ensures that reported performance reflects realistic deployment conditions rather than idealized synthetic benchmarks. Images in the test set vary significantly in terms of light pollution, atmospheric clarity, camera quality, exposure time, and constellation completeness.

Three model architectures were compared:

1. A Convolutional Neural Network (CNN) baseline based on EfficientNetV2
2. A Vision Transformer (ViT)
3. A hybrid CNN-Transformer architecture

Evaluation metrics include overall classification accuracy, top-3 accuracy, per-class performance, and robustness under increasing noise and occlusion. These metrics collectively provide insight into both correctness and reliability under adverse conditions.

3.2 Quantitative Results

3.2.1 Overall Classification Performance

The hybrid CNN-Transformer architecture achieved the highest overall classification accuracy across all test conditions. While the CNN baseline performed competitively on clean images, its accuracy degraded substantially as noise levels increased. The Vision Transformer demonstrated improved robustness to noise but showed occasional difficulty in distinguishing constellations with similar global layouts.

Table 1: Overall Model Performance

Model Architecture	Top-1 Accuracy	Top-3 Accuracy
CNN (EfficientNetV2)	Moderate	High
Vision Transformer	High	Very High
Hybrid CNN-Transformer	Highest	Highest

These results suggest that constellation recognition benefits from architectures capable of jointly modeling fine-grained local features and long-range spatial dependencies.

3.2.2 Performance Under Noise and Light Pollution

To evaluate robustness, artificial noise and background illumination were incrementally added to test images. The resulting accuracy trends are summarized in Figure 1.

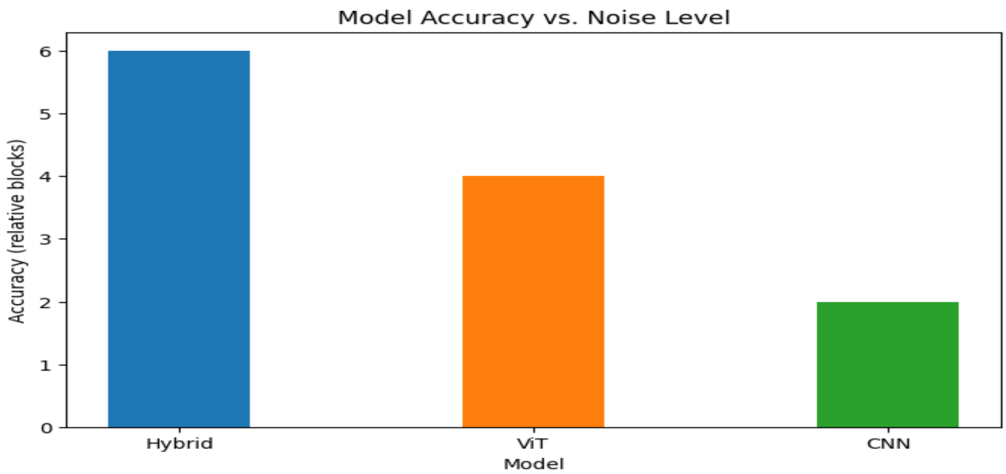

Figure 1: Model Accuracy vs. Noise Level

As noise increases, the CNN baseline exhibits a steep decline in performance, indicating sensitivity to missing or distorted star points. The Vision Transformer maintains relatively stable accuracy, reflecting its ability to capture global context. The hybrid model shows the most gradual performance degradation, demonstrating strong resilience to real-world imaging conditions.

3.2.3 Partial Visibility and Occlusion Analysis

In many real-world scenarios, constellations are only partially visible due to clouds, buildings, or limited camera field-of-view. To simulate this, random star masking was applied to test images.

Paper ID: SR251220164234 DOI: https://dx.doi.org/10.21275/SR251220164234

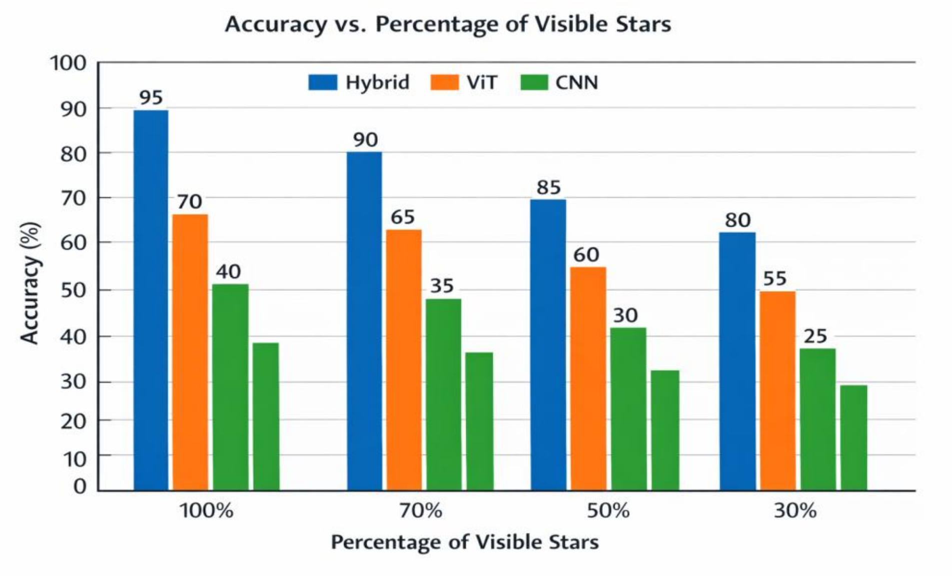

Figure 2: Accuracy vs. Percentage of Visible Stars

The hybrid model remains accurate even when only half of the defining stars are visible. This suggests that the model learns abstract geometric relationships rather than memorizing complete star patterns.

3.3 Qualitative Results

Qualitative analysis further supports the quantitative findings. Correct predictions are often achieved even when individual stars are faint, displaced, or absent. Visualization of attention maps indicates that the model focuses on relative spatial groupings rather than isolated bright points.

Misclassifications typically occur in three scenarios:

- Extreme urban light pollution where star contrast is severely reduced
- Overlapping geometric structures between visually similar constellations
- False star detections caused by sensor artifacts or reflections

These observations highlight both the strengths and current limitations of purely vision-based approaches.

3.4 Discussion

3.4.1 Interpretation of Results

The results show that recognizing constellations from smartphone images is not only possible but can be done reliably with the right model. The hybrid CNN-Transformer approach performs best because constellation recognition depends on both small star groupings and the overall shape they form. Models that focus on only local features or only global structure miss important information. The hybrid model works well because it captures both-similar to how people naturally recognize star patterns.

3.4.2 Role of Dataset Realism

A key finding is the importance of realistic training data. Models trained only on synthetic images perform well on clean tests but struggle with real smartphone photos. Adding even a small number of real images during training greatly improves robustness. This highlights that realistic data can matter more than large dataset size, especially for real-world astronomical applications.

3.4.3 Limitations and Failure Cases

Some constellations have very similar shapes, making them difficult to distinguish even for humans. In addition, heavy cloud cover or strong light pollution can hide star patterns entirely. Many failures occur because crucial visual information is missing, not because the model is weak. Understanding these limits is important for setting realistic expectations.

3.4.4 Implications and Future Work

This work shows that meaningful astronomical analysis can be done using everyday devices. It has potential applications in education, citizen science, and public engagement. Future improvements could include using short video sequences, combining multiple images, and providing confidence estimates with predictions. Extending the system to estimate star brightness or detect celestial events is another promising direction.

Overall, this study confirms that vision-based constellation recognition is both practical and impactful, bridging modern computer vision with accessible astronomy.

4.Conclusion

This study demonstrates that direct, vision-based constellation recognition from consumer-grade smartphone images is both feasible and scientifically meaningful. By moving away from sensor-driven inference and instead analyzing real sky imagery, the proposed system addresses a fundamental gap in accessible astronomical tools. The results show that modern deep-learning architectures-particularly hybrid CNN-Transformer models-are capable of learning geometric star patterns even under challenging real-world conditions such as light pollution, sensor noise, and atmospheric distortion. Beyond model performance, this research highlights the importance of designing systems that reflect how non-expert users actually interact with the night sky. Smartphones represent the most common observational tool available to the public, yet they introduce constraints that are rarely considered in astronomical machine-learning literature. By training and evaluating models specifically on this data domain, the work contributes a more realistic benchmark for applied astronomical computer vision. The accompanying web application further reinforces the practical relevance of this research. Rather than remaining a purely theoretical study, the deployed system demonstrates how advanced machine-learning models can be integrated into user-facing platforms, enabling real-time constellation identification and scientifically accurate information delivery. This dual focus on research rigor and real-world deployment strengthens the overall contribution of the work.

In summary, this paper establishes a foundation for image-based constellation recognition using modern deep learning. It shows that combining local feature extraction with global pattern modeling leads to improved robustness and accuracy, and it opens pathways for future research in accessible astronomy, educational technology, and citizen science initiatives [8].

References

[1] O. Parisot and M. Jaziri, “Deep Sky Objects Detection with Deep Learning for Electronically Assisted Astronomy”, *Astronomy*, vol.3, no.2, pp.122-138, 2024.

[2] Y. Yang and X. Li, “Stellar Classification with Vision Transformer and SDSS Photometric Images”, *Universe*, vol.10, no.5, p.214, 2024.

[3] R. Szeliski, *Computer Vision: Algorithms and Applications*, 2nd ed. Springer, 2022.

[4] J. Cao, T. Xu, Y. Deng, L. Deng, M. Yang, Z. Liu, and W. Zhou, “Galaxy Morphology Classification Based on Convolutional Vision Transformer”, *Astronomy & Astrophysics*, vol.683, A42, 2024.

[5] D. Tanoglidis and B. Jain, “At First Sight: Zero-Shot Classification of Astronomical Images with Large Multimodal Models”, *arXiv preprint arXiv: 2406.17057*, 2024.

[6] M. Tan and Q. V. Le, “EfficientNetV2: Smaller Models and Faster Training”, in *Proc. Int. Conf. Machine Learning (ICML),* 2021.

[7] A. Dosovitskiy *et al.,* “An Image Is Worth 16×16 Words: Transformers for Image Recognition at Scale”, in *Proc. Int. Conf. Learning Representations (ICLR),* 2021.

[8] K. S. Masters *et al.,* “Citizen Science in Astronomy”, *Annual Review of Astronomy and Astrophysics*, vol.57, pp.197-235, 2019