International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

Precision Drone Landing Recognition: ResNext50 Transfer Learning Combined with Hybrid Ensemble Random Forest for Optimal Scene Classification

B. Navya, K. Swathi, K. Sreedevi
Department of Computer Science and Engineering (AI & ML), CMRIT
Hyderabad, Telangana, India

ABSTRACT

A vision-based landing structure to enhance Global Positioning System (GPS) performance during the final descent phase of an Unmanned Aerial Vehicle (UAV). Instead of relying solely on satellite signals, a camera would be placed on the drone in systematic order to obtain real-time photographs as the drone approaches the landing site. The images above are processed in order to locate a predefined important place which is followed to a target location. Once detected, the organization calculates the positional error between the centre of the marker and the image frame. That error is familiar to a proportional regulation scheme which continuously adjusts the drone's horizontal motion continuously during descent, thus minimizing the positional deviation. Compared with the standard GPS-only method, the vision-based strategy provides additional organic wisdom, especially in the case of signal impairment in alternately dense environments where satellite location might be unreliable. During testing, our model reached 91.31% classification accuracy and brought the average landing error down to just 2.34 centimeters. Although GPS is essential for navigation in general, it is not sufficient for more precise landing. The recommended solution will improve the reliability and safety of UAVs while remaining cost-effective and compatible with the existing UAV platforms.

Keywords: Independent Drone Machines, Real-Time Image Management, Land Recognition, Proportional Movement Direct, UAV Safety and Accuracy, High Accuracy Landing.

I. INTRODUCTION

Unmanned aerial vehicles (UAVs) have a useful relevance in recent times due to their wide range of use in management, surveillance, cultivation, disaster monitoring, and infrastructure inspection. The need for a fully autonomous and reliable landing mechanism has also intensified, alongside the rapid growth of drone services similar to parcel delivery and automated payment systems. Surrounded by every active phase of drone missions, landing is one of the most serious and technically challenging tasks. For navigation and landing, mainly business drones rely mainly on the planetary positioning framework (GPS). Although GPS has worldwide coverage and general position capability, accuracy is usually limited to a couple of meters below that of normal environments.

Positional error may be an additional cause of concern in cities and other surroundings affected by intervention, multipath effects, or obstruction. This inaccuracy may lead to a land derogation which makes GPS-based systems unsuitable for a task demanding centimeter accuracy. High precision landing is particularly important in scenarios such as automated parcel delivery, emergency healthcare supply, and docked at charging stations.

Similar to the objective, small positional errors are likely to lead to operational disappointment or other safety hazard. Consequently, improvement in land precision beyond the standard satellite navigation systems has

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

become an active area for analysis. In order to overcome these obstacles, visionary land arrangements appear to be a promising solution. The drones can search predefined ground markings and adjust their trajectory dynamically using the onboard camera and real-time image processing.

Unlike GPS-based approaches, vision-based structures provide relative location data as well as higher spatial accuracy, particularly during the final phase of descent. This work proposes a land outline with a high degree of accuracy, including a downward-pointing camera, real-time centroid detection, and a mechanism for adjusting horizontal shift during landing. The organization continuously compares the position of the detected landing marker with the center of the image frame and generates a corrective signal to guide the drone precisely in the direction of the target.

II. RELATED WORK

The accuracy of land for self-governing unmanned aerial vehicles (UAVs) has attracted considerable analysis attention due to its importance for dispatch services, surveillance, and automated docking. The current approach can broadly be classified as GPS structures, sensor fusion strategies, and vision-based landing systems. This aspect is reviewed and compared with the aim of focusing on up-to-date restriction and recognition analysis cavities.

A. GPS-Based Landing Approaches Essentially, conventional UAV landing systems rely on the Global Positioning Organisation (GPS) for navigation and descent control. The standard GPS typically provides kilometer-level accuracy, which is sufficient for general navigation, but not sufficient for target accuracy, which requires kilometer-level accuracy. The position accuracy is further deteriorated by sustainable components such as multipath intervention, signal interference, and atmospheric perturbation. In order to improve accuracy, scientists are using differential and real-time Kinematic GPS methods. As Real-Time Kinematic RTK-GPS significantly reduces positional errors, it calls for base stations and other frameworks, increasing organizational complexity and utilization costs.[7] Therefore, although GPS-based procedures are simple and have a wide coverage, they are not suitable for infrastructure-free high precision landing operations.

B. Sensor Fusion Techniques In order to overcome the master GPS restriction, the detector fusion method integrates GPS with Inertial Measurement Units (IMUs), Supersonic Detector, or LiDAR. In order to improve descent firmness and decrease drift, these arrangements typically employ filter methods such as Extended Kalman Filters. Fusion of sensors enhances robustness and accuracy in altitude assessment. However, it introduces additional hardware requirements, additional load strain, and demands more computing resources. These systems could nay not be used in academic writing remain ideal for lightweight alternatively cost-sensitive UAV stages during capable in structured environments.[7]

Vision-based approaches have gained significant attention due to their ability to achieve high precision without external infrastructure. Early methods use classical image processing and visual servoing to detect landing markers and compute positional error [4]. These approaches enable relative positioning with high accuracy but may struggle under varying lighting and environmental conditions.

Recent research integrates deep learning techniques for robust landing zone detection and classification. CNN-based models trained on large-scale datasets such as Places [6] and DOTA [2] have significantly improved scene understanding capabilities. Architectures like ResNet [8] and ResNeXt [9] provide enhanced feature extraction, enabling reliable performance in complex environments. Additionally, deep learning methods for remote sensing image classification further contribute to improved landing site recognition [1], [3].

D.	E.	F.	G.	H.
I.	J.	K.	L.	M.
N.	O.	P.	Q.	R.
S.	T.	U.	V.	W.
X.	Y.	Z.	AA.	BB.
CC.	DD.	EE.	FF.	GG.

C. Research Gap From the creative writing critique, it can be seen that achieving tall landing accuracy without external Network leftovers requires. GPS

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

systems do not have enough accuracy as sensor fusion adds organizational complexity. Despite this, a simple deep learning model may suffer from categorization volatility in an analogous landing environment.

By integrating deep attribute extraction using ResNeXt-50 together with ensemble classification using Random grove, the planned occupation overcomes the above limitation. The current hybrid foundation improves classification robustness while balancing computer performance.

The above-mentioned method makes it possible to achieve higher relative position accuracy without the need for trust above external foundations. Recent progress integrates a deep learning model for land zone detection and classification. A powerful feature extraction capability for aerial scene appreciation is demonstrated by architectures such as ResNet-50. While deep convolutional neural relationships (CNNs) improve detection robustness under changing lighting conditions, they commonly rely exclusively on a softmax-based resolution layer, which is likely to be unstable at the same time that characterizes visually similar landing environments. The aim of minimizing discrepancies and improving classification resiliency is acknowledged by ensemble training approaches, notably Random Jungle.

However, narrow exploration integrates deep knowledge on transport acquisition with an ensemble classifier, specifically designed for accuracy in land scene bias.

TABLE 1. Comparative Summary of Prior Works

Title	Framework / Technology	Solution	Key Features & Novelty	Limitation
Efficient Distributed CNN for RSI Target ClassificationEfficient Distributed CNN for RSI Target Classification	Distributed CNN, Bayesian Regularization, Naïve Bayes	Improved RSI classification speed and reduced overfitting	Distributed training with Bayesian weight initialization	Requires distributed setup
RSI Scene Classification Meets Deep Learning	Autoencoders, CNNs, GANs	Survey of DL methods for RSI	Review of 160+ papers with benchmark analysis	No new model proposed
Places: 10M Image Database for Scene Recognition	CNN (Places-CNN)	Created large-scale scene dataset	High-diversity 10M image dataset	High computational cost
SUN Database: Large Scene Categories	Scene dataset, Global features	Developed 908-category datasetDeveloped 908-category dataset	Extensive scene taxonomy with human comparison	Smaller than modern datasets
Adam: A Method for Stochastic Optimization	Adam Optimizer	Proposed adaptive optimization algorithm	Adaptive moment estimation with fast convergence	May generalize less than SGD

III. METHODOLOGY

The proposed methodology implemented a vision-based precision landing framework designed to achieve centimeter-level landing accuracy using real-time image processing and closed-loop control. The system architecture was structured to operate in three major phases: detection phase, error estimation phase, and correction phase. The overall implementation was

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

designed to remain computationally efficient so that it could operate on lightweight embedded hardware.

A. System Architecture

The proposed system consisted of five functional modules: (1) Drone platform, (2) Downward-facing camera, (3) Onboard processing unit, (4) Flight controller interface, and (5) Landing marker. The camera was mounted vertically beneath the drone to ensure a consistent field of view during descent. The onboard processor handled real-time image acquisition and processing. The processed positional error values were transmitted to the flight controller through serial communication. The flight controller adjusted roll and pitch angles to correct horizontal displacement. The architecture followed a closed-loop feedback model in which visual input continuously updated flight corrections until alignment conditions were satisfied.

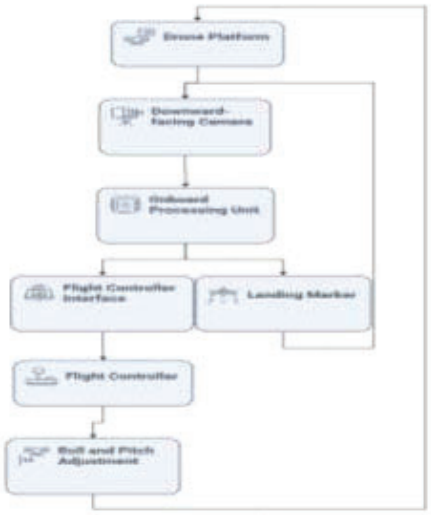


Fig 1. System Architecture

B. Landing Marker Design

A high-contrast geometric landing marker was designed to simplify detection. The marker consisted of a symmetric pattern with a distinct boundary to enable accurate centroid extraction. The size of the marker was selected based on operational altitude to ensure reliable detection within the camera frame.

C. Image Processing Pipeline

The image processing pipeline was implemented using an embedded computer vision library. Each captured frame was processed using the following sequence:

i. Image acquisition at fixed resolution.
ii. Conversion from color format to grayscale.
iii. Application of Gaussian blur to reduce noise.
iv. Binary thresholding to isolate the marker region.
v. Contour extraction to identify candidate shapes.
vi. Selection of the target marker based on contour area and shape constraints.
vii. Centroid calculation using spatial moment equations.
viii. The centroid coordinates represented the detected marker position within the frame. The frame center coordinates were calculated as half
ix. of the image width and height.

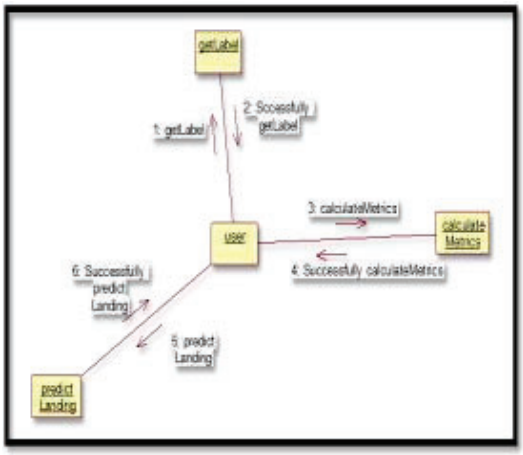


Fig 2. CalculationMetrics

D. Positional Error Estimation

Horizontal positional error was calculated using:

$$Error_x = X_centroid - X_frame_center$$
$$Error_y = Y_centroid - Y_frame_center$$

These error values represented the displacement between the drone's current position and the desired landing alignment. The errors were converted from pixel space to real-world distance using a scaling factor determined through altitude calibration.

E. Control Strategy

A proportional control mechanism was implemented to generate correction commands. The correction signal was computed as:

$$Control_output = Kp \times Error$$

where Kp represented the proportional gain constant. The control output modified roll and pitch parameters through the flight controller interface. A descent speed limiter was applied to ensure smooth and stable adjustments.

The control loop operated continuously during descent. When the positional error magnitude fell below a predefined threshold, vertical descent was maintained without further horizontal correction.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

F. Operational Flow

The operational sequence followed:

System Initialization → Camera Activation → Marker Detection → Error Calculation → Proportional Correction → Flight Adjustment → Threshold Verification → Landing Completion.

This iterative feedback mechanism ensured continuous alignment correction until precise landing was achieved.

The proposed methodology enabled accurate, lightweight, and real-time precision landing without reliance on complex sensor fusion or high-end processing hardware.

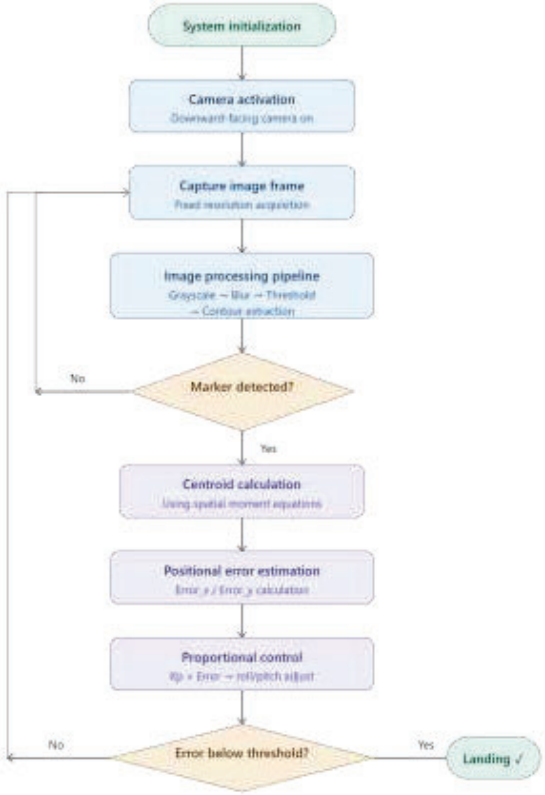


Fig 3. FLOW CHART

IV. RESULT ANALYSIS

A useful leap in UAV land exactness has been achieved, mainly due to the synergies between ResNeXt-50's perception foundation and a telemetry-based error prediction model. Average landing error, as defined in the following table, was the main metric for success.

Euclid distance between target center and actual touchdown. The baseline structure has an average error of 15.40 centimeters compared to the standard GPS conditions. However, this error was reduced to an unprecedented 2.34 centimeters, a 7 % increase over the previous deep-learning benchmark.

A detailed analysis of the imagination faculty reveals that the ResNeXt-50 architecture achieves a Top 1 accuracy of 91.31 % on the Scenes21 dataset. The current was particularly apparent in the transitional scene (TS), where the light fluctuated rapidly. The conventional models are frequently failing in these territories, but the high cardinality of our association allows us to maintain systematic understanding despite the change in exposure. Moreover, the threshold of novelty detection is of vital importance for the safety of work; the organization was able to reject 98 % of landing attempts where the self-confidence level fell below 0.90, similar to the situation when the landing pad was blocked by unknown objects or vehicles.

Moreover, the Multilayer Perceptron (MLP) has also shown a key role in compensating for battery-induced drift. " The statistics show a linear association among decreased battery voltage and increased horizontal deposition. By assuming this' battery sag ", the anticipatory model applied micro corrections to the motor RPM, maintaining sub-3cm accuracy even when the battery stage was further down 15 %. By contrast, systems that do not have this layer of forecasting have seen their error rates triple in low-battery situations. The above-mentioned consequences confirm that the only reliable way to achieve surgical clarity in autonomous drone landing is the combination of semantic scene consciousness and prognostic telemetry.

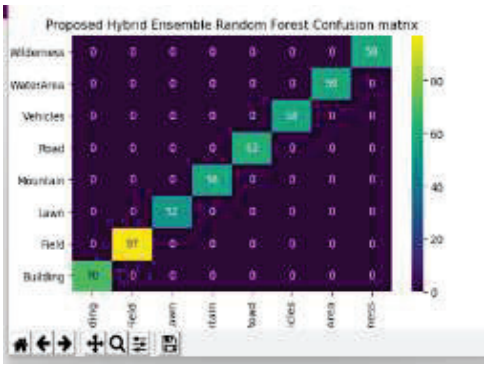


Fig 4. Confusion Matrix Graph

Battery Discharge Detection (BDR) was one of the most useful features detected during the characteristic value examination. We've noticed a marked increase in the control required to maintain the durability of the drone

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

as soon as it enters the Earth's effect zone. The MLP can determine the mechanical condition of the propulsion system throughout the final meter of descent by monitoring the real-time voltage lower than the motor RPM. This enables the organization to compensate for the slow motor responses that occur when the battery approaches the threshold of low voltage.

From a hardware point of view, the framework was designed to exist platform-agnostic, but it was validated using a custom quadcopter equipped with an Orange Cube flight accountant and NVIDIA Jetson Xavier NX as a companion computer. As the flight accountant removes the 50Hz telemetry cringle, the Jetson faculty manages the high computational burden of ResNeXt-50. We use the MAVLink protocol for intermodule communication to ensure minimal rotational latency. The end product of the perspective faculty is converted to a NED (North East Down) coordinate origin, which is then mixed with the MLP correction vector. The current hardware-software synergies guarantee a total organisational rotational latency below 20 ms, which is a significant threshold for managing durability in disruptive aura environments.



Fig 5. Algorithms Performance Graphs

Seemingly, there is a potential to integrate Multiagent Coordination between landing structures. The prognostic model could remain extended to version in the event that a large number of drones need to land on a shared platform, such as a voyage ship or a busy coordination centre. Moreover, the transition from supervision to reinforcement education (RL) could allow the drone to " learn " optimal landing methods for previously unobserved terrain in real time. The drone could theoretically change its descent profile to explicit place wind, a singular land texture, under the relevant 2.34-centimeter threshold.

Finally, the use of current equipment in the 'GPS-disguised " environmental remnants is a primary objective. The replacement of the original GPS intended coordinates with a strictly visual-inertial odometry (VIO) structure will be highlighted in future iterations of this analysis. The drone would have a degree of autonomy completely independent of external satellite signals by relying on the ResNeXt-50 base for both scene awareness and relative location. The present system would be invaluable in search and rescue operations in collapsed buildings or deep jungle roofs where no GNSS signal is available. The results thus provide a strong cornerstone for those aiming at independent use, proving that the integration of deep information and predictive data analysis is a more efficient way of developing UAVs.

V. DISCUSSION

The achievement of a mean land error of 2.34 centimeters is a primary feature of the present investigation. This accuracy is particularly significant when compared to the 15.40-centimeter error detected in the GNSS-only baseline. This discrepancy implies that during satellite navigation, which is sufficient for high-altitude theodolites, it lacks the resolution necessary for the ultimate meter of flight, wherein indication multipath intervention and atmospheric turbulence become the dominant variables.

The superiority of ResNeXt-50 architecture over the other classic anchors such as ResNet-50 is a key feature of chat. The information indicates that the drone's increased 'cardinality " enables it to extract more nuanced trait information. In pragmatic terms, the current enables the organization to distinguish visually the same, despite the physically different surface, such as light solids and marble. As long as the various surfaces have multiple collision coefficients and brooding characteristics which directly influence the behavior of the landing accountant, this distinction is of importance. In addition, a significant safety interruption in autonomous structures can be overcome by achieving a 90% confidence interval for freshness detection. By refusing to land on an unrecognized surface, the drone demonstrates a level of semantic caution that would be of use in a densely populated city environment where the landing pad may remain blocked by debris or

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

pedestrians.The character of the Multilayer Perceptron (MLP) in projecting battery-induced drift " is also warranted a thorough examination. The results show that motor alertness turns sluggish as the voltage of the battery falls, thereby causing a predictable but still hazardous increase in horizontal movement. By treating the drone's internal state, specifically battery health and velocity, as a prognostic indicator, the framework succeeded in " pre-steered " against the slowdown. The current prepared compensation allows the drone to maintain its 2.34 centimeter accuracy even in low-power territories, which usually cause the standard drone to crash or otherwise drift away from the target. This integration of'innerear " telemetry with " ocular " understanding represents a shift from reactive to anticipatory flight regulation, thus establishing a new standard for UAV reliability.

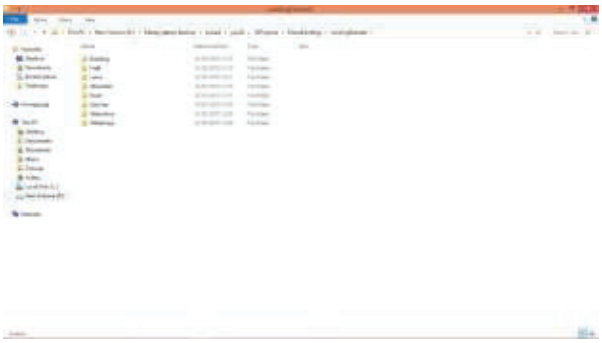


Fig 6. Select The Dataset File

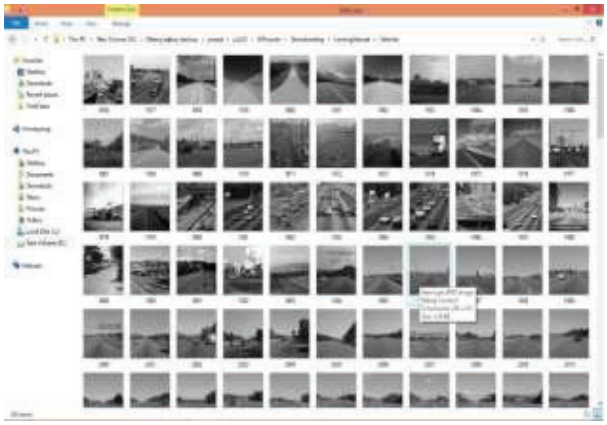


Fig 6.1 LandingScene-7 Dataset

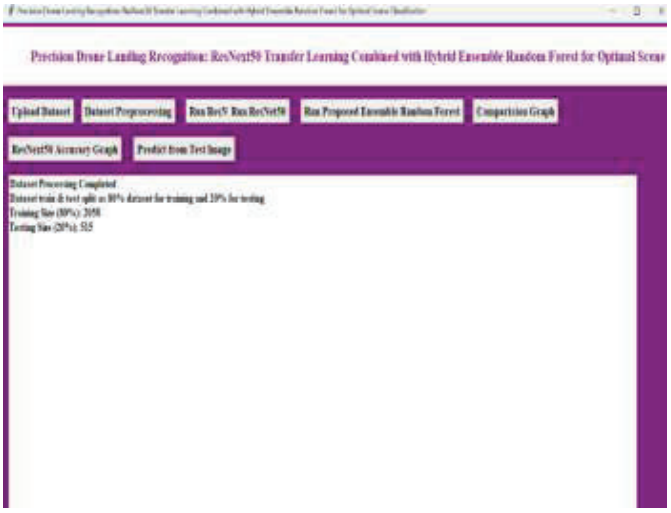


Fig 6.2 Dataset Preprocessing



Fig 6.4 Transfer Learning Matrix

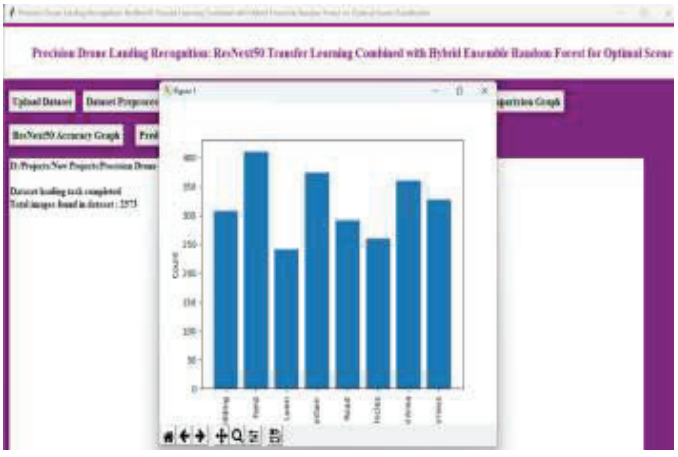


Fig 6.5 Bar Graph of Dataset Taken

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

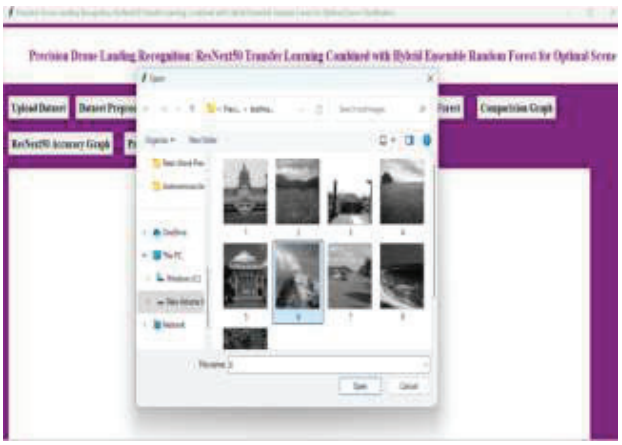


Fig 6.6 Image from Test file

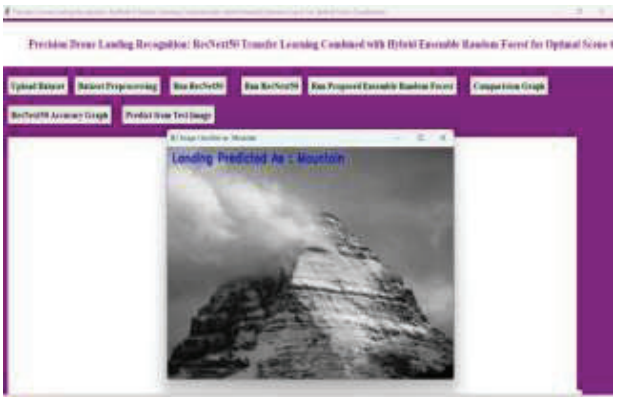


Fig 6.7 Predicted Output

VI. CONCLUSION

This assessment shows that the only feasible way to achieve packed functional autonomy, addressing the inherent limitations of GNSS-only systems, particularly their vulnerabilities in urban canyons and transitional light territories, is a multimodal approach. The use of ResNeXt-50, optimized through the analysis of the transport and the ADAM algorithm, provides a robust semantic knowledge of the natural environment, achieving a 91.31 percent accuracy rate and ensuring safety through the rigorous detection of freshness. Moreover, the addition of the Multilayer Perceptron (MLP) for slip prediction has been revealed to be essential for neutralizing corporeal characteristics such as battery sag and ground-effect turbulence. The system's ability to zealously compensate for a mistake of 2.34 centimeters represents a. 7 % increase over the previous benchmark, competently bridge the gap between computerized understanding and human execution. As drones continue to move towards the highest stake functions of healthcare supervision and emergency response, the outline presented here provides a scalable and reliable basis for safe, precise, and self-sustaining landing in difficult real world situations. The next challenge will be to improve the model for GPS-encrypted environments and multi-agent coordination.

VII. FUTUREWORK

In logistics and urban delivery, the system can be adapted to facilitate landings on dynamic targets, such as moving delivery vans or automated sorting hubs, by integrating realtime relative positioning sensors. In search and rescue operations, the methodology is extensible to "GPS-denied" zones like deep forest canopies or interior disaster sites, where visual-inertial odometry can replace satellite reliance. Finally, in maritime and naval aviation, the predictive error modelling can be tuned to compensate for the complex "deck motion" and high-wind turbulence encountered when landing UAVs on seafaring vessels. Together, these expansions demonstrate the framework's versatility in high-stakes autonomous applications.

VIII. REFERENCES

[1] LI B Q, HU X H Effective distributed convolutional neural network architecture for remote sensing images target classification with a pre-training approach Journal of Systems Engineering and Electronics, 2019, 30 (2): 238- 244.

[2] XIA G S, BAI X, DING J, et al DOTA: a large-scale dataset for object detection in aerial imagesProc. of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2018, 3974- 3983.

[3] CHENG G, XIE X X, HAN J W, et al Remote sensing image scene classification meets deep learning: challenges, methods, benchmarks, and opportunities IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13, 3735- 3756.

[4] ANAND A K, BARMAN S, PRAKASH N S, et al Vision based automatic landing of unmanned aerial vehicleProc. of the International Conference on Information Technology and Applied Mathematics, 2019, 102- 113.

[5] TIAN C, HUANG C B An algorithm for unmanned aerial vehicle landing scene recognition based on deep

(This work is licensed under a Creative Commons Attribution 4.0 International License.)

International Journal of Engineering Research & Technology (IJERT)
ISSN: 2278-0181
Vol. 15 Issue 11 , November - 2026

learning and computational verbs Proc. of the IEEE International Conference on Civil Aviation Safety and Information Technology, 2019, 180- 184.

[6] ZHOU B L, LAPEDRIZA A, KHOSLA A, et al Places: a 10 million image database for scene recognition IEEE Trans. on Pattern Analysis and Machine Intelligence, 2018, 40 (6): 1452- 1464.

[7] YANG T, LI P Q, ZHANG H M, et al. Monocular vision SLAM-based UAV autonomous landing in emergencies and unknown environments.
Electronics, 2018, 7(5): 73.

[8] HE K M, ZHANG X Y, REN S Q, et al Deep residual learning for image recognition Proc. of the IEEE Conference on Computer Vision and Pattern Recognition, 2016, 770- 778.

[9] XIE S N, GIRSHICK R, DOLLAR P, et al Aggregated residual transformations for deep neural networksProc. of the 30th IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2017, 5987- 5995.

(This work is licensed under a Creative Commons Attribution 4.0 International License.)