International Journal of Advanced Multidisciplinary Research and Educational Development
Volume 2 - Issue 6 | November - December 2026 | www.ijamred.com
ISSN: 3107-6513

A Confidence-Aware Fusion Layer for Hazard Assessment in Autonomous Lunar Landing

V. Nithyasri
Department of Computer Science and Engineering (AI & ML)
Arunai Engineering College
Tiruvannamalai, Tamil Nadu, India
Email: nithyasrivijayan17@gmail.com

***Abstract*—**

Autonomous lunar landing requires a spacecraft to identify safe landing areas using camera and sensor data. This is difficult because lighting, shadows, and sensor noise affect how the terrain is interpreted, and AI-based hazard detectors can produce uncertain or incorrect predictions. When hazard detection is reduced to a simple Safe or Hazard label, confidence scores and disagreement between detectors are not used in the final decision. This paper proposes a Smart Confidence Fusion Layer placed between the AI hazard detectors and the final landing decision. The layer takes four inputs: hazard prediction, confidence score, detector agreement, and terrain information. Separate detectors cover rocks, craters, and slopes or rough terrain. When the detectors agree, confidence in the result increases; when they disagree, the site is classified as Uncertain instead of being accepted directly. The final output has three levels, Safe, Uncertain, and Hazardous, which may support actions such as continuing the landing, re-checking the area, slowing down, or avoiding the location. The work is a conceptual framework that has not been implemented on a spacecraft, and no experimental results are claimed.

***Keywords*—Autonomous lunar landing, hazard detection, confidence estimation, detector agreement, decision fusion, uncertainty-aware decision making, terrain assessment.**

I. Introduction

Autonomous lunar landing requires the spacecraft to assess the terrain below it without direct human control. Large rocks, craters, and steep or uneven slopes can endanger the landing, so selecting a suitable landing area is an important part of the landing process. NASA's Autonomous Landing and Hazard Avoidance Technology (ALHAT) project is one example of research on improving landing safety and precision through onboard hazard detection [1], [2]. Earlier studies also report that the performance of onboard hazard detection depends on terrain characteristics, sensor quality, and the design of the detection algorithm [1].

Detecting hazards from images is difficult because lunar terrain looks very different under changing illumination. Low-angle sunlight can produce deep shadows that hide rocks or make some surface features look like craters, and sensor noise adds further uncertainty. Earlier work on passive imaging combined several visual cues for hazard detection, which suggests that a single visual feature may not always be sufficient [3].

AI-based detectors can help identify different types of hazards, but their predictions may still be uncertain. Separate detectors for rocks, craters, and slopes or rough terrain may disagree about the same region. If each detector returns only a Safe or Hazard label, the label does not show how confident the detector is, and disagreement between detectors may not reach the final landing decision. A single incorrect prediction could then influence the decision more than it should.

A confidence-aware decision stage can address these issues. Instead of asking only whether an area is safe, it can also consider how confident the system is in that assessment. Confidence estimation and calibration of neural network outputs are active research topics [6], and predictive uncertainty has been studied for deep learning models [7], [8]. Combining the outputs of several classifiers is an established approach [9], and multisensor fusion is widely used to integrate information from different sources [10].

This paper proposes a Smart Confidence Fusion Layer for autonomous lunar landing. The layer takes four inputs: the hazard prediction, the confidence score, the agreement between detectors, and terrain information. From these it produces one of three outputs: Safe, Uncertain, or Hazardous. The work is a conceptual framework and does not include an implementation or experimental results. The remainder of the paper presents the problem, the proposed architecture and decision logic, and the limitations and future directions.

II. Problem Statement

A lunar lander must decide whether a candidate landing site is suitable using only the camera and sensor data available onboard. Rocks, craters, and steep or rough terrain can each affect this decision, and each may be handled by a separate detector. The problem has four parts.

Multiple hazards. A landing site must be free from all relevant hazards. A site that the rock detector considers safe may still be unsafe because of a steep slope or rough terrain.

Imperfect detectors. Changes in lighting, shadows, and sensor noise make the lunar surface hard to interpret, so an AI-based detector may sometimes give an uncertain or incorrect prediction.

Missing uncertainty. A Safe or Hazard label does not show how reliable the prediction is. A Safe result with low confidence looks the same as a Safe result with high confidence.

Unused disagreement. When detectors give different predictions for the same site, the disagreement may indicate an uncertain observation or terrain that is difficult to interpret. This information is lost if each detector is considered separately.

These problems call for a mechanism that combines the detector predictions, confidence scores, detector agreement,

ISSN: 3107-6513

and terrain information before the final landing decision is made. The mechanism should also be able to label a site as Uncertain instead of forcing every prediction into Safe or Hazard.

III. Existing Approach and Limitations

A. Existing Decision Approach

The baseline considered in this paper is a decision scheme in which each hazard detector returns only a Safe or Hazard label. Confidence scores and disagreement between detectors are not used in the final decision. As a result, a single incorrect prediction can influence the landing decision, and an incorrect Safe prediction is especially serious because it can create a risk during landing. This paper does not claim that all existing lunar landing systems follow this scheme.

B. Related Background

The following points come from general background literature and are separate from the baseline above. Practical hazard detection and avoidance involves a complete process of sensing, algorithms, vehicle behaviour, and guidance, rather than a single classifier [1]. Landing site selection has also been treated as a separate process in which candidate locations are evaluated before a landing point is chosen [4]. Different sensing approaches have been explored for landing, including imaging lidar in the ALHAT project [2] and passive imaging with multiple visual cues [3], and hazard detection has been studied for lunar missions such as Chang'e-3 [5]. This paper does not compare the proposed method with these systems. They are cited only to describe the setting in which a confidence-based decision layer could be applied.

Machine learning research has also shown that the confidence values of modern neural networks may be poorly calibrated [6]. A fusion layer therefore should not assume that a reported confidence value is a reliable probability.

IV. Proposed System

The proposed Smart Confidence Fusion Layer is placed between the hazard detectors and the final landing decision. The overall flow is:

Lunar images / sensor data → Rock detector → Crater detector → Slope / terrain detector → Confidence and agreement information → Confidence Fusion Layer → Safe / Uncertain / Hazardous.

A. Input Data

The input consists of lunar surface images and other sensor data collected during descent. The exact sensor type and image resolution are not fixed in this conceptual framework and would be chosen during later development and validation.

B. Hazard Detection

Three AI-based detectors are considered: a rock detector, a crater detector, and a slope or rough-terrain detector. They are treated as separate conceptual components, and this paper reports no training or testing of them. Each detector evaluates a candidate landing site and gives a prediction of Safe or Hazard from its own point of view, as a standalone detector would. The change proposed here is in how these outputs are handled before the final decision.

C. Confidence Estimation

Along with its prediction, each detector gives a confidence score, such as 90%. The fusion layer uses this score to decide how strongly each prediction should influence the final decision. A value such as 90% is treated as an indicator of confidence, not as a guaranteed 90% probability, because reported confidence values may not be well calibrated [6]. A practical implementation would need to evaluate the confidence values of each detector and could use suitable uncertainty estimation methods [7], [8]. This paper does not select a particular method.

D. Detector Agreement

The fusion layer also checks whether the detectors give consistent results for the same site. Agreement provides supporting evidence and increases confidence in the result. Disagreement indicates uncertainty and means the site may need further checking, but it is not treated as proof of a hazard. The framework therefore uses Uncertain as an intermediate decision between Safe and Hazardous.

E. Terrain Information

Terrain information, such as the slope or roughness of the surface, is the fourth input. It gives additional context for the detector predictions. For example, terrain measurements may indicate a possible hazard even when the rock and crater detectors do not identify one. The exact terrain measurements are not specified and are not fixed in this conceptual framework.

F. Confidence Fusion and Safety Decision

The Confidence Fusion Layer receives the four inputs (hazard prediction, confidence score, detector agreement, and terrain information) and combines them into an overall assessment of the candidate site, as described in Section VI. The result is one of three levels: Safe, Uncertain, or Hazardous. The actions associated with these levels are described in Section VIII.

V. System Architecture

Fig. 1 shows the proposed architecture. It has four main blocks connected in sequence.

Input. Lunar surface images and sensor data collected during descent, which represent the raw information about the terrain.

AI Hazard Detection. The rock, crater, and slope or rough-terrain detectors. Each processes the available input and produces a prediction with a confidence score.

Confidence Fusion Layer. The main proposed component. It combines the hazard predictions, confidence scores, detector agreement, and terrain information to assess the reliability of the landing site.

Safety Decision. The final output, classified as Safe, Uncertain, or Hazardous. A Safe result can support continuing the landing, an Uncertain result can trigger re-checking or slowing down, and a Hazardous result can support avoiding the area or selecting another landing point.

ISSN: 3107-6513

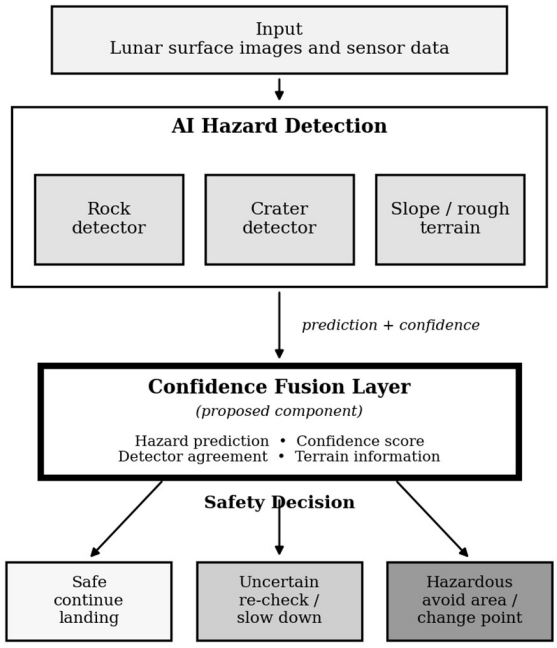

Fig. 1. Architecture of the proposed Smart Confidence Fusion Layer. The heavier border marks the proposed component.

VI. Confidence Fusion Mechanism

The fusion layer is described at a conceptual level. This work does not define a validated mathematical model or an experimental formula.

A. Core Ideas

Agreement increases confidence. If the rock, crater, and terrain detectors reach the same conclusion with sufficient confidence, the assessment of the landing site becomes more reliable.

Disagreement increases uncertainty. If one detector reports Hazard while the others report Safe, the site should not be accepted directly as Safe and requires further checking.

A confident Safe result is not automatically accepted. A strong Safe prediction from one detector cannot cancel a hazard identified by another. Each detector covers a different type of hazard, so a Safe result from the rock detector says nothing about slope-related hazards.

Confidence also matters for Hazard predictions. A highly confident Hazard prediction can lead to a Hazardous decision. A low-confidence Hazard prediction is treated as a warning and is not sufficient by itself to reject the site.

B. Notation and Decision Rules

Let i denote one of the three detectors: rock, crater, or slope/terrain. For detector i, let p_i be its prediction, which is either Safe or Hazard, and let c_i be its confidence score, with a value from 0 to 1 (written as percentages in the example of Section VII). Let τ_H be the hazard confidence threshold and τ_S the safety confidence threshold. Let F_T be true when the terrain information indicates a possible problem. The conceptual decision rules are:

R1. Hazardous if any detector i predicts p_i = Hazard with $c_i \geq \tau_H$.

R2. Safe if all detectors predict p_i = Safe with $c_i \geq \tau_S$, and F_T is false.

R3. Uncertain in all other cases, including detector disagreement, low-confidence predictions, and ambiguous information.

R1 and R2 cannot hold at the same time, so every site falls into exactly one of the three outputs. These rules describe the proposed framework and are not a validated experimental model. The values of τ_H and τ_S, and the way terrain information is included in the decision, are design choices that would need to be selected and evaluated with real data. Learned or weighted fusion, in which fixed thresholds are replaced by a learned mechanism, is a possible future direction (Section XII).

VII. Working Example (Illustrative)

The example below only explains how the proposed mechanism could work. The values are illustrative and are not results from any dataset, experiment, or flight.

TABLE I
Illustrative Fusion Example for One Landing Site (Values Are Illustrative, Not Experimental)

Detector	Prediction	Confidence (illustrative)
Rock detector	Safe	90%
Crater detector	Safe	85%
Terrain detector	Hazard	60%
Fusion layer output	**Uncertain**	–

Two of the three detectors report Safe with fairly high confidence. A decision based only on majority voting, or only on the detector with the highest confidence, might accept this site as safe. The proposed layer does not decide in this way.

Here the terrain detector reports a possible hazard, so the detectors disagree and rule R2 cannot be satisfied. The 60% confidence of the Hazard prediction is not treated as high enough to classify the site directly as Hazardous, but it is also not low enough to ignore. Under rule R1 the output would be Hazardous only if 60% reached τ_H. The example assumes that τ_H is above 60%, so rule R3 applies and the output is Uncertain. The result therefore depends on the conceptual threshold, which is not fixed in this paper.

An Uncertain result can trigger additional checking of the landing area, such as evaluating the site again with new sensor data, slowing the descent to collect more observations, or considering another candidate location. The purpose of the example is to show that a Safe prediction should not be accepted without considering confidence and disagreement between the detectors.

VIII. Decision Logic

TABLE II
Decision Levels and Possible Responses

Condition	Decision	Possible response
All detectors indicate Safe with adequate confidence, and terrain information shows no problem	Safe	Continue landing
Detectors disagree, or confidence is insufficient	Uncertain	Re-check / slow down
Hazard indicated with sufficient confidence	Hazardous	Avoid area / change landing point

Safe is the only output that allows the landing to continue as planned. A Hazardous result removes the site from further consideration, and Uncertain acts as an intermediate state when the evidence is not enough to choose between Safe and Hazardous. The listed responses are possible actions only and are not implemented control actions. The actual guidance response would depend on the vehicle and mission requirements and is outside the scope of this paper. Selecting an alternative landing point is also a separate problem [4].

IX. Advantages

The following advantages follow from the design of the concept and have not been verified by experiments. The framework is intended to provide:

- **Reduced dependence on a single detector.** The final decision is not based on only one model.
- **Explicit handling of uncertainty.** An Uncertain output is available when the information is not sufficient for a clear decision.
- **Use of detector agreement and disagreement.** Differences between detector outputs are treated as information rather than ignored.
- **More cautious handling of conflicting information.** Sites with uncertain or conflicting observations are not automatically accepted as Safe.
- **Support for additional checking.** An Uncertain result can indicate that the area should be checked again before the final decision.

These are expected benefits only. Whether they improve landing safety can be determined only through further testing.

X. Potential Applications

The concept could potentially be used in the following areas. None of these applications has been tested in this work.

- Lunar landers, where the terrain must be assessed during the final stages of descent.
- Other planetary landers operating over unfamiliar or uncertain terrain.
- Autonomous rovers and navigation systems, where terrain conditions must be evaluated before movement.
- UAV landing systems operating where the landing surface is not known in advance.

XI. Limitations

The main limitations of this work are as follows.

- The work is a conceptual framework without an implementation. It has not been demonstrated on the Moon or tested or flown on a spacecraft, and it reports no experimental performance results.
- System performance would depend on the quality of the individual detectors, the available sensor data, and the design of the fusion mechanism.
- Detector confidence values may not be reliable without proper calibration [6].
- Real lunar imagery and the computing limits of onboard systems would need to be considered and evaluated.
- The confidence values in the example (90%, 85%, and 60%) are illustrative. The thresholds τ_H and τ_S, and the method for including terrain information, have not been finalized or tuned.
- The Uncertain category may add cost. If too many sites are classified as Uncertain, the landing process could become slower or more complex. This trade-off has not been studied.

XII. Future Scope

The framework could be extended in the following ways.

Learned and weighted fusion. Instead of fixed thresholds, weights learned from data could give more reliable detectors a greater influence on the final decision. Existing research on classifier combination provides a basis for this direction [9].

Testing with real lunar imagery. The framework should be evaluated with real lunar images under the lighting and shadow conditions that are relevant to landing.

Real-time onboard deployment. The fusion layer would need to operate within the processing and memory limits of an actual lunar lander.

Confidence calibration. The calibration of each detector's confidence values should be evaluated and improved before they are used in fusion [6]–[8].

XIII. Conclusion

This paper presented a Smart Confidence Fusion Layer for autonomous lunar landing. The layer combines the predictions of rock, crater, and slope/terrain detectors with their confidence scores, the agreement between detectors, and terrain information, and produces one of three outputs: Safe, Uncertain, or Hazardous. The main idea is that a landing decision should depend not only on what the detectors predict but also on how confident and consistent those predictions are. The framework is conceptual and has not been implemented or tested on a spacecraft, so its effect on landing safety remains to be evaluated with real lunar imagery and realistic onboard computing constraints.

References

[1] A. E. Johnson, A. Huertas, R. A. Werner, and J. F. Montgomery, "Analysis of on-board hazard detection and avoidance for safe lunar landing," in *Proc. IEEE Aerosp. Conf.*, Big Sky, MT, USA, 2008, pp. 1–9, doi: 10.1109/AERO.2008.4526301.

[2] C. D. Epp and T. B. Smith, "Autonomous precision landing and hazard detection and avoidance technology (ALHAT)," in *Proc. IEEE Aerosp. Conf.*, 2007, pp. 1–7.

[3] A. Huertas, Y. Cheng, and R. Madison, "Passive imaging based multi-cue hazard detection for spacecraft safe landing," NASA Tech. Rep. Server, 2006.

[4] B. E. Cohanim and B. K. Collins, "Landing point designation algorithm for lunar landing," *J. Spacecraft Rockets*, vol. 46, no. 4, pp. 858–864, 2009.

[5] X. Jiang, S. Li, and T. Tao, "Innovative hazard detection and avoidance guidance for safe lunar landing," *Proc. Inst. Mech. Eng. G, J. Aerosp. Eng.*, 2016, doi: 10.1177/0954410015625671.

[6] C. Guo, G. Pleiss, Y. Sun, and K. Q. Weinberger, "On calibration of modern neural networks," in *Proc. 34th Int. Conf. Mach. Learn. (ICML)*, 2017, pp. 1321–1330.