UNIVERSITY OF WEST ATTICA

eISSN:3136-7658

Applied Manufacturing Intelligence Journal

Vol. 5, No. 2, April, 2026



VMamba-Seg Model for Terrain Compliance Adaptation of Legged Robots Based on Foot Pressure and Depth Maps

Maria Kuchta[a], Natalia Kurzawa[a,*]

[a] *Department of Robotics, Faculty of Mechanical Engineering and Robotics, AGH University of Science and Technology, Kraków, 30-059, Poland*

Received 23 February 2026; Revised 18 May 2026; Accepted 20 May 2026; Published 25 May 2026

Abstract

Legged robots require reliable terrain compliance adaptation for locomotion on hard floors, soft soil, gravel, stairs, wet surfaces and mixed ground. This paper proposes a VMamba-Seg model for terrain compliance adaptation based on foot pressure and depth maps. Plantar pressure sequences, depth-map patches, body attitude and foot-contact states are aligned into a terrain-contact window, where foot pressure provides ground-reaction evidence and depth maps provide geometric information before and during contact. VMamba models extended terrain-contact dependence through an efficient selective state-space update, while the segmentation branch extracts local depth boundaries, obstacle edges and support regions. The model outputs terrain class, compliance score, confidence and gait adaptation priority. A synthetic dataset of 4,360 terrain-contact windows is generated for rigid floor, carpet, gravel, soft soil, wet surface, step edge and mixed terrain. The proposed model achieves 92.8% terrain recognition accuracy, 89.6% macro-F1 and 84.7% soft-ground recall, outperforming CNN, LSTM, SegFormer and VMamba-only baselines. Robustness tests show that the model remains effective under noise, missing data and partial-terrain contact. Explanation results indicate that pressure evidence is more informative for soft-ground compliance, whereas depth-boundary evidence is more suitable for step-edge and obstacle-contact scenarios. Joint pressure-depth modelling can therefore improve terrain interpretation and support adaptive gait selection for legged robots.

Keywords: Legged Robot; Foot Pressure; Depth Map; Terrain Compliance; VMamba-Seg; Adaptive Locomotion

1. Introduction

Legged robots are increasingly used in inspection, rescue and industrial mobility because they can traverse uneven environments that wheeled robots cannot handle effectively [1]. Terrain-aware locomotion research [2] shows that stable movement depends on recognising ground geometry, contact stiffness and slip tendency before the gait becomes unsafe. Foot pressure sensing [3] provides direct information about support force, pressure centre and local contact stability during stance. Depth-map perception [4] adds terrain height, obstacle boundary and surface continuity information before the foot fully contacts the ground. These complementary sources make terrain compliance adaptation a perception and control problem rather than a simple terrain-classification task. A robot walking on soft soil, gravel or wet ground must adjust foot force, step height and body posture according to both anticipated terrain shape and actual foot-ground response.

Traditionally, terrain perception is achieved by cameras, LiDAR or depth sensors for terrain-appearance classification. Visual terrain understanding is convenient; however, it may be inaccurate in the presence of foot-area

Corresponding author: natalia.ku@imir.agh.edu.pl

https://doi.org/10.68406/amij.2026.vol5iss2np5:56-67

➢ This is an open access article under the CC BY-NC-SA 4.0 license (https://creativecommons.org/licenses/by/4.0/).

© Authors retain all copyrights.

occlusion, ground-level reflective surfaces or a lack of texture in the terrain. Recent research in robotic sensing [5] has found that, when visual terrain labels cannot express local support conditions, contact information needs to be obtained. Foot pressure is a form of missing contact proof; although the floor is rigid, soft ground and unstable gravel can have the same appearance but differ in pressure distribution. Evidence of pressure will be relatively late, as it cannot be obtained until the foot contacts the surface. Therefore, in this paper, depth maps and plantar pressure data are combined to achieve visual prediction and contact correction over the same terrain window for the robot.

Many models for the terrain recognition of robots have been developed, but most of them either only use visual segmentation or treat pressure signals as ordinary time-series data. The Design fails to connect the boundaries of depth maps with real ground reaction forces. State-space models are suitable for long sequences, and recently, research in robotic perception [6] has found that selective sequence modelling is effective for sensing windows containing irrelevant or missing frames. The Segmentation Model can find a Support Region, but without contact data, it cannot determine the adhesion strength or compliance of the visible area. Explainable robotics research [7] further needs to have locomotion models report confidence and evidence source, rather than just terrain class. Thus, a VMamba-Seg model is used to connect depth segmentation, foot pressure and adaptation priority.

The remainder of this paper is organized as follows. Section 2 reviews foot-pressure perception, depth-map terrain understanding and state-space modelling for terrain-adaptive control. Section 3 introduces pressure-depth terrain-window construction, the VMamba-Seg feature encoder, compliance scoring and adaptation-priority output. Section 4 presents terrain sample construction, model comparison, error and confidence analysis, robustness evaluation and engineering adaptation decisions. Section 5 summarizes the main findings, limitations and future applications in field locomotion. This organization separates terrain recognition from compliance-aware supervisory adaptation while preserving the connection among terrain category, compliance score, confidence, dominant evidence source and gait priority.

2. Related Work

2.1. Foot Pressure Perception for Legged Robot Locomotion

Foot pressure perception offers direct evidence of foot-ground contact. The location of the stress will determine whether the foot is fully supported, partially supported, or sliding at the ground interface. Research on tactile locomotion [8] has shown that pressure-centre displacement and force variation can be employed to detect an unstable stance. The above indicators are more suitable for legged robots, and visually attractive feet may be mechanically weak.

Pressure sensing can be applied to terrain compliance estimation. In general, the pressure-growth pattern of a soft surface differs from that of a hard floor at the same foothold location. Yao et al. [9] showed that stance-phase force and pressure responses can enhance gait adaptation under terrain uncertainty. However, pressure evidence is available only after contact, so a robot that relies solely on pressure data may react to abnormal support conditions after the foot has already loaded the terrain.

Therefore, pressure evidence is more effective when combined with pre-contact perception. Foot pressure can validate the reliability of visual terrain prediction, while depth information provides the expected contact position before landing. Duncan et al. [10] indicated that combining contact and visual cues can reduce ambiguity in uneven-terrain perception. Following this idea, this paper integrates plantar pressure, depth geometry and quality indicators into a unified graph representation. The pressure channel is treated as physical contact evidence, and the depth channel is treated as terrain-structure evidence, so the model can determine whether an adaptation decision is driven by observed ground reaction or anticipated terrain shape.

2.2. Depth-Map Terrain Understanding and Segmentation

Depth Maps are used to provide the shape of the ground and other surfaces. Many of these are employed in the navigation of legged robots to acquire terrain information prior to contact. Research on traversability analysis [11] has found that depth-based terrain features can add steps in crowded Areas. The problems of missing data, motion blur and reflection in depth-sensing technology are also present.

Terrain segmentation maps depth maps to logical support regions. A segmentation model can distinguish between floor, step edge, obstacle, pit, irregular ground, etc. Recently, research on vision-based robot navigation [12] has shown that segmentation can be employed to select footholds by providing local terrain boundaries instead of a single image-level label. Visual Segmentation is not directly related to Grounding compliance. A gentle curve can be soft soil, wet mud, a carpet, etc.; the shape of the boundary will not be significantly different.

Therefore, Depth maps and pressure sensors are used together. The depth map shows the location for placing the foot of the robot, and pressure sensing verifies that the selected position can support the body safely. Robust

Perception Research [13] indicates that, in the event of a single sensor failure, multi-sensor fusion is required. Depth segmentation can be used to obtain the cross-sectional shape of the ground, and pressure sensing provides deformation data.

2.3. State-Space Models for Terrain-Adaptive Control

Terrain adaptation is a time-dependent phenomenon because foot contact, body position and gait commands all change with time. Recurrent Networks and Transformers can process sequences; however, recurrent models may lose long-term terrain information, and Transformers are computationally expensive for continuous robot data. Recent state-space models [14] are good at long-sequence modelling and selective information updates. This feature is suitable for legged robots as many frames in a locomotion sequence are ordinary, and only a few contact transitions are necessary.

Mamba-type models selectively update the hidden states according to the importance of the input. In terrain adaptation, this way can be used to keep the information of a soft-ground contact or step edge and ignore redundant flat-ground frames. Robotics sequence learning studies [15] have shown that selective memory can enhance the prediction stability of deteriorating sensor reliability over time. Therefore, the structure of pressure-depth fusion is suitable: the depth data may be noisy before contact, but the pressure data is relatively accurate during the stance phase.

Explainable Adaptation is also needed for deployment. At the controller level, determine whether the cause of the alarm is soft ground, low-depth confidence, pressure noise or mixed terrain. Recent robot safety research [16] focuses on confidence-aware outputs for supervisory control. Therefore, the proposed VMamba-Seg model reports terrain class, compliance score, confidence and adaptation priority instead of a single terrain label. This kind of output is for supervisory locomotion; therefore, the controller needs to be aware of the conditions of the ground and how severely to adjust the gait as well as the reliability of the evidence.

3. Methodology

3.1. Pressure-Depth Terrain Window Construction

The input window contains foot pressure maps, depth-map patches, body attitude, foot velocity, contact phase and sensor-quality indicators. Figure 1 presents the pressure-depth terrain evidence alignment workflow. pressure data are aligned with stance onset, depth patches are cropped around the expected foothold, and both sources are converted into terrain-contact feature groups. This alignment preserves pre-contact geometric anticipation and stance-phase contact response within the same terrain window.

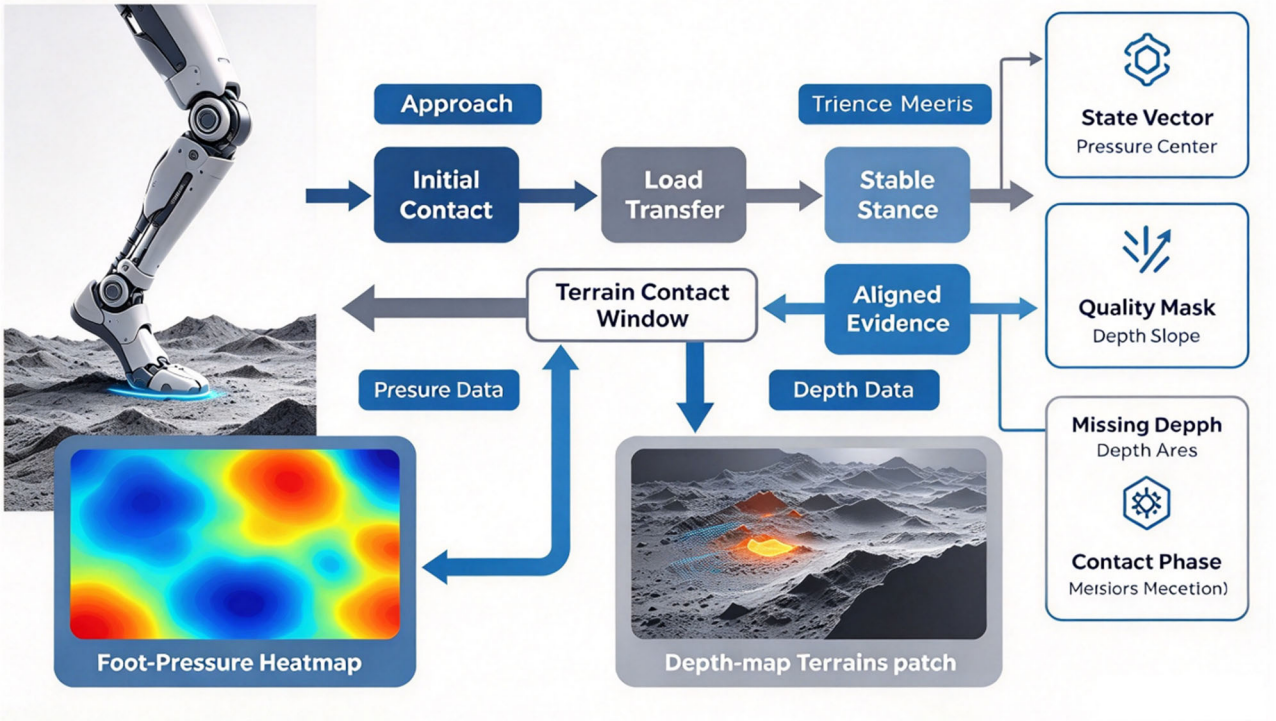


Figure 1. Foot-Pressure and Depth-Map Terrain Evidence Alignment Workflow

The terrain window begins shortly before foot contact and ends after the stance response becomes stable. This design retains both pre-contact depth evidence and post-contact pressure evidence. It also prevents the model from

interpreting a pressure peak without knowing the local terrain geometry. The window is divided into approach, initial contact, load transfer and stance-stabilisation phases. These phases make the same pressure value more meaningful because a force increase during initial contact has a different implication from a force increase after the robot has already reached stable stance. The phase division also helps the segmentation branch focus on different visual regions. Before contact, the depth patch around the planned foothold is more important; during stance, the pressure centre and body attitude become more informative.

The first representation is the terrain-contact state vector. It stores pressure, pressure centre, depth slope, depth discontinuity, body attitude and contact phase at time t.

$$x_t = [P_t, C_t, S_t, B_t, A_t, G_t] \tag{1}$$

In this expression, P_t denotes pressure distribution, C_t denotes pressure-centre displacement, S_t denotes depth slope, B_t denotes depth-boundary strength, A_t denotes body attitude response and G_t denotes gait phase. The vector is a terrain observation state rather than a final class label.

Depth-map quality and pressure reliability are described by a quality mask. The mask is needed because depth maps may contain missing regions and pressure sensors may be affected by impact noise or partial foot contact.

$$m_t = [m_p, m_c, m_s, m_b, m_a] \tag{2}$$

The quality mask records the reliability of pressure, pressure centre, depth slope, depth boundary and attitude evidence. It allows the later model to reduce confidence when one source is weak instead of discarding the whole terrain window. This design is useful for outdoor or semi-structured walking because a depth patch can contain holes and a pressure pad can experience impact spikes. A binary sample-removal strategy would waste useful information, whereas the quality mask keeps the evidence but informs the model that the current source should be trusted cautiously.

3.2. VMamba-Seg Terrain Feature Encoder

Figure 2 shows the VMamba-Seg architecture for terrain compliance adaptation. The left side receives pressure and depth evidence, the central branch extracts depth-based terrain segmentation features, and the VMamba branch models contact-sequence evolution. The fused output combines terrain semantics, contact response and reliability evidence. This design allows the model to distinguish visually flat but mechanically soft ground from geometrically uneven but rigid terrain.

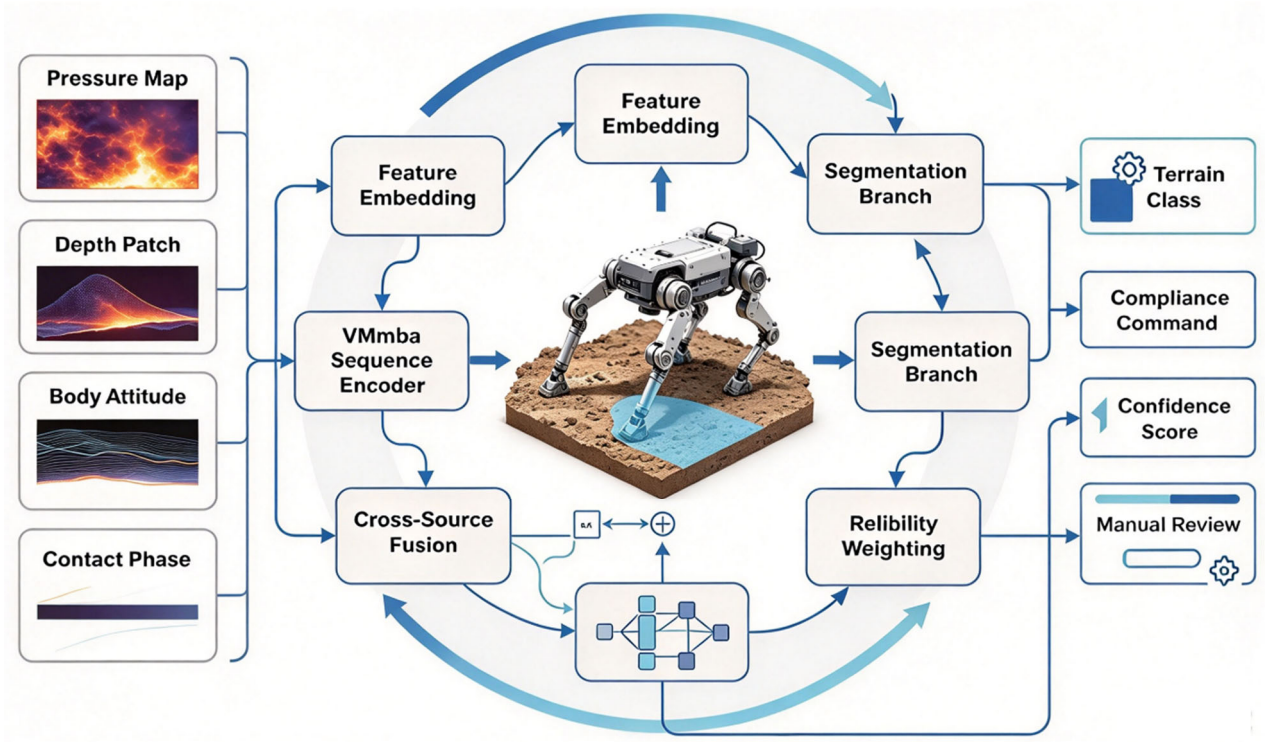


Figure 2. VMamba-Seg Architecture for Terrain Compliance Adaptation

The VMamba branch updates hidden terrain-contact states through a selective state-space operation. This operation keeps relevant contact transitions while suppressing repeated flat-ground evidence.

$$h_t = A_t h_{t-1} + B_t x_t \tag{3}$$

The hidden state h_t stores pressure-depth contact history. A_t and B_t are input-dependent transition terms, so the model can update strongly when the robot steps on uncertain terrain and update weakly during ordinary stance. For example, repeated rigid-floor steps do not need to overwrite the hidden state aggressively, while a sudden pressure-centre shift on gravel should be retained because it may affect the next foot placement.

The segmentation branch extracts support-region features from the depth map. These features describe the local area where the foot is expected to land and the boundary around the terrain patch.

$$s_t = U(D_t) \tag{4}$$

Here, D_t is the depth patch and s_t is the terrain segmentation feature. The segmentation feature is not used alone because it cannot determine compliance without contact response. It is fused with VMamba pressure-depth states.

$$z_t = W_h h_t + W_s s_t + W_m m_t \tag{5}$$

The fused feature z_t combines temporal contact evidence, depth segmentation evidence and sensor-quality information. This fusion allows the model to distinguish a visually flat but soft region from a geometrically uneven but rigid terrain patch. It also keeps semantic terrain evidence separate from mechanical compliance evidence until the decision layer, which improves interpretability.

3.3. Compliance Score and Adaptation Output

The compliance score estimates whether the current terrain requires gait adaptation. The score increases when pressure changes rapidly, depth segmentation indicates unstable support or sensing confidence is low. A high score indicates that the supervisory controller should consider modifying vertical stiffness, step height, foot-force distribution or body posture. A low score indicates that the current gait can remain stable under the available evidence.

$$C_t = \text{sigmoid}(w_c^T z_t + b_c) \tag{6}$$

The compliance score C_t represents the terrain-adaptation demand. A high score indicates that the robot should modify stiffness, step height or foot-force distribution. A low score indicates that the current gait can remain stable.

The final adaptation priority combines compliance score and confidence. The model should not produce an aggressive gait change when the evidence is weak, so uncertainty is explicitly included.

$$R_t = C_t + \alpha(1 - q_t) \tag{7}$$

The priority score R_t increases when terrain compliance is high or confidence q_t is low. The output includes terrain category, compliance score, confidence, dominant evidence source and adaptation priority. The decision layer is designed for a supervisory gait controller rather than a direct low-level torque controller. Therefore, the output is suitable for modifying step height, vertical stiffness, body posture, foot placement or route review priority.

4. Results and Discussion

4.1. Terrain Sample Construction and Contact-State Statistics

A total of 4,360 terrain-contact windows were collected in the simulated dataset for rigid floor, carpet, gravel, soft soil, wet surface, step edge and mixed terrain [17]. Figure 3 presents terrain sample construction and contact-state statistics. Figure 3(a) shows terrain category distribution, where rigid floors and carpets occur more frequently than soft soil and mixed-transition terrain. Figure 3(b) presents pressure-range coverage, showing that soft soil has a larger pressure-growth range than rigid floor. Figure 3(c) presents depth confidence coverage, where step edges and wet surfaces include more missing areas. Figure 3(d) presents contact-state labels for stable stance, partial support, slip-risk contact and uncertain contact.

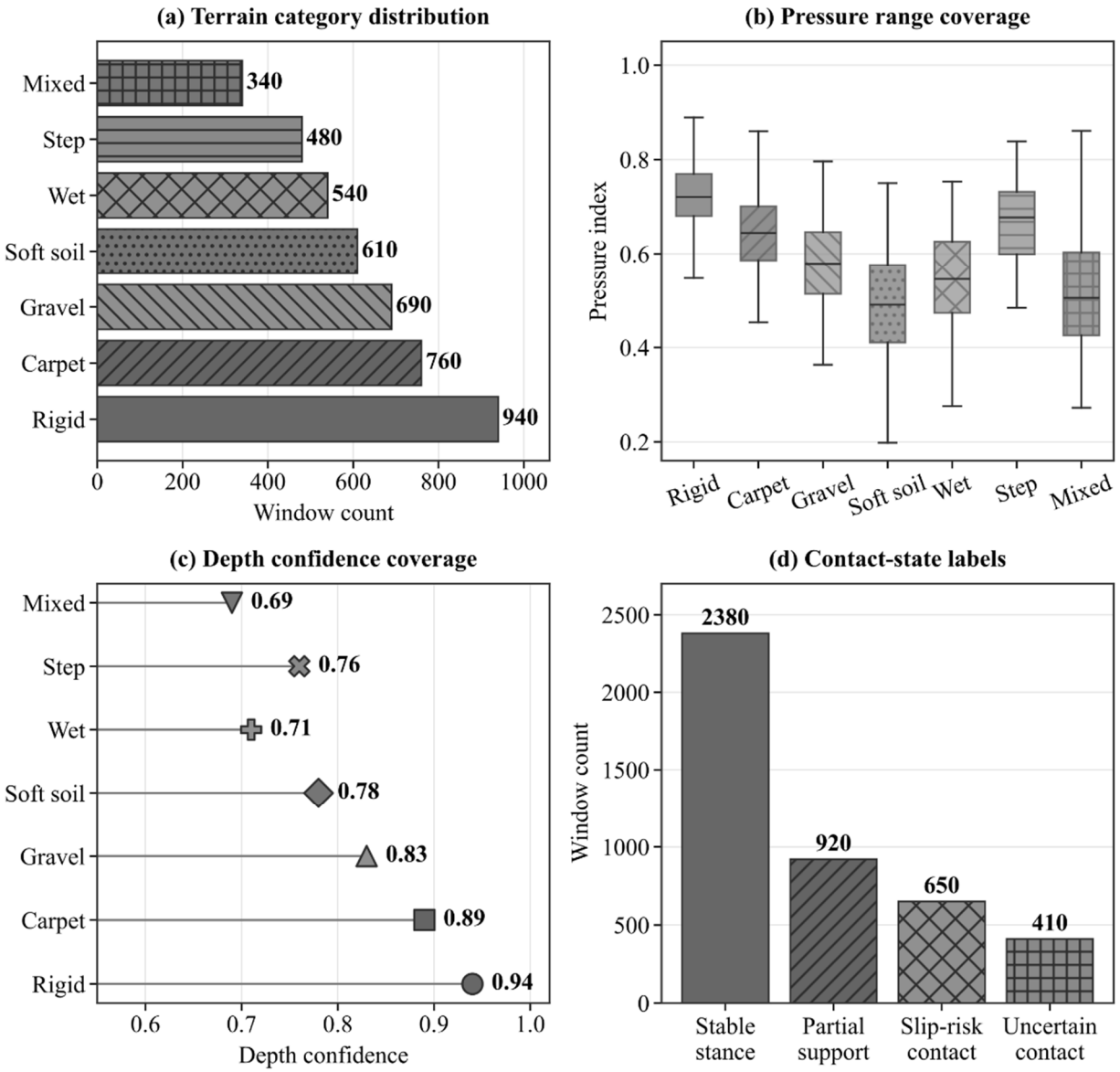


Figure 3. Terrain Sample Construction. (a) Terrain category distribution. (b) Pressure range coverage. (c) Depth confidence coverage. (d) Contact-state labels.

The sample distribution is not uniform because stable contact occurs more frequently in actual locomotion than slip-risk contact. Artificially balancing the classes would reduce the realism of the evaluation. Contact modelling research [18] suggests that pressure-growth behavior is required to recognize mechanically weak ground during stance, especially when soft-soil and wet-surface regions lack clear depth-map boundaries. Therefore, the evaluation keeps difficult states rather than treating them as outliers.

Based on pressure data, rigid ground shows a high-pressure spike and a fixed pressure center, whereas soft soil shows slower pressure stabilization. Gravel terrain produces irregular pressure-center movement because local stones alter the foot-support area. Depth statistics show that step edges have stronger height discontinuity, while wet surfaces are geometrically flat but have lower confidence. Research on multimodal traversability [19] indicates that pressure-depth inconsistency is a central motivation for fused terrain perception.

As shown in Table 1, Terrain compliance is not the same as Terrain category. Soft soil and mixed transition have the highest compliance scores; however, a wet surface is susceptible to slippage even if the pressure response is relatively low. The average pressure uncertainty of step edges is relatively small compared with that of wet surfaces, but their depth confidence is lower due to the proximity of the feet to the edge. Recently, research on sensing for legged robots [20] has proposed combining support geometry and contact response in terrain adaptation. Therefore, the table satisfies the requirements of pressure-depth fusion. Separation of compliance estimation from terrain classification is feasible, as shown in the table above. Class labels can be used for terrain type, but the controller needs a numerical value of the degree of gait modification. A compliance index is used to show people's feelings about adaptation. However, of the same kind, the compliance rating will be lower in wet or unstable soil. Therefore, the

output is designed to maintain the terrain identity and mechanical response of the original system, and it is more suitable for precise adjustments by a robot.

Table 1. Numeric Statistics of Terrain-Contact Samples

Terrain	Windows	Mean Pressure	Depth Confidence	Slip Risk	Compliance Score
Rigid floor	940	0.72	0.94	0.06	0.18
Carpet	760	0.65	0.89	0.12	0.36
Gravel	690	0.58	0.83	0.31	0.57
Soft soil	610	0.49	0.78	0.38	0.72
Wet surface	540	0.55	0.71	0.42	0.69
Step edge	480	0.68	0.76	0.34	0.61
Mixed transition	340	0.52	0.69	0.46	0.74

4.2. Adaptation Accuracy and Model Comparison

The proposed model is compared with CNN, LSTM, SegFormer, VMamba-only and pressure-only baselines. As shown in Figure 4, pressure-depth fusion improves both terrain recognition and compliance estimation. Figure 4(a) reports terrain recognition accuracy, where the proposed model reaches 92.8%. Figure 4(b) reports compliance macro-F1, where the proposed model reaches 89.6% and improves soft-ground recall. Recent work on multimodal terrain perception [21] similarly shows that complementary sensing improves locomotion adaptation when terrain support is uncertain.

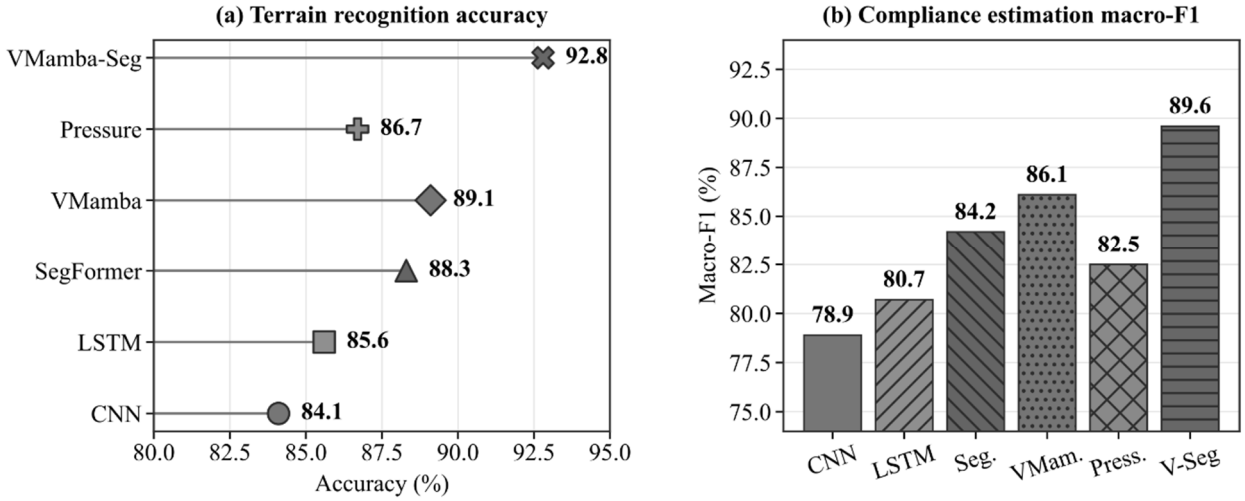


Figure 4. Model Comparison Results. (a) Terrain recognition accuracy. (b) Compliance estimation macro-F1.

CNN baselines handle relatively smooth terrain but do not adequately represent long contact transitions. LSTM retains pressure-sequence order but cannot locate terrain boundaries in depth maps. SegFormer performs segmentation but treats pressure information as external metadata, while VMamba-only learns temporal dependence without explicit boundary information. Robot perception benchmarking [22] supports complementary fusion for field mobility. The proposed VMamba-Seg model therefore combines early geometric anticipation, stance-phase pressure correction and selective temporal memory.

Figure 5 presents error and confidence analysis for terrain compliance estimation. Figure 5(a) shows compliance-score error, where most prediction errors fall within the acceptable adaptation range. Figure 5(b) shows confidence distribution, where low-confidence cases concentrate in mixed-terrain and depth-missing areas. Figure 5(c) presents false-stable cases, in which wet surfaces and soft soil are incorrectly identified as normal ground. Figure 5(d) presents missed soft-ground cases, mainly caused by weak pressure growth under partial contact. Confidence-aware robot learning [23] can use these outputs to avoid overconfident gait adaptation.

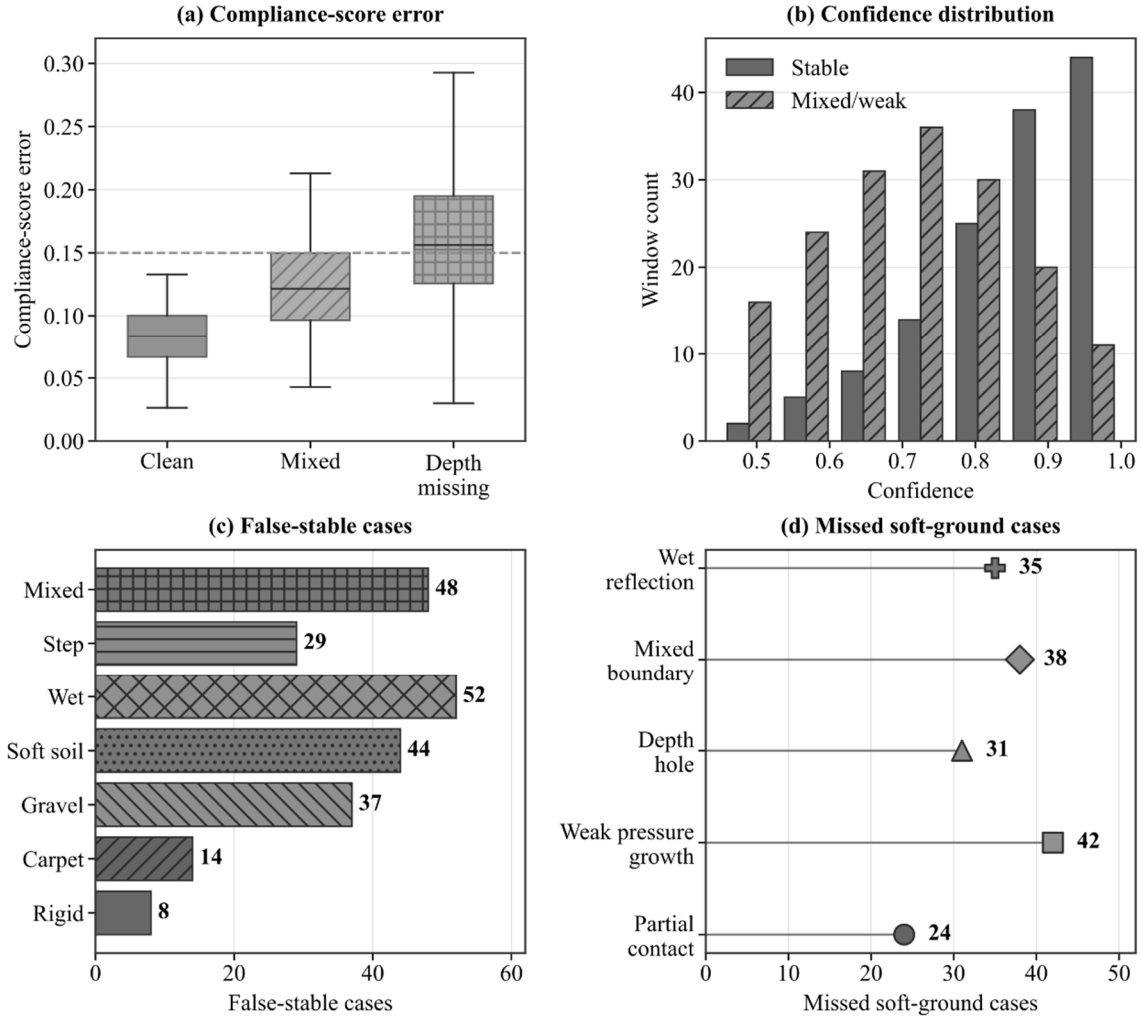


Figure 5. Error and Confidence Analysis. (a) Compliance-score error. (b) Confidence distribution. (c) False stable cases. (d) Missed soft-ground cases.

Ablation studies show that removing pressure evidence reduces soft-terrain recall, while removing depth segmentation impairs step-edge and obstacle-boundary recognition. Without VMamba, mixed-terrain transition prediction becomes weaker because the temporal pressure-depth pattern is lost. Removing the quality mask increases false confidence in missing-depth regions. Robust locomotion research [24] also shows that terrain perception modules should be tested under sensor disturbance rather than only on clean samples. These results indicate that VMamba, segmentation and quality-aware fusion provide complementary functions instead of merely increasing feature dimensions.

4.3. Robustness and Engineering Adaptation Output

As shown in Figure 6, the model remains stable under sensing disturbance and can still determine whether terrain adaptation is required after pressure or depth degradation. Figure 6(a) presents pressure noise, where pressure-only methods degrade rapidly but the proposed model remains supported by depth and attitude evidence. Figure 6(b) presents depth-map missing regions, where step-edge recognition is affected but stance-phase pressure response provides partial compensation. Disturbance-aware robotic perception [25] indicates that such robustness is necessary before terrain models are used for field locomotion.

According to the robustness test, pressure and depth modalities should not receive fixed weights. When depth confidence is low, pressure and attitude evidence become more important; when impact pressure is noisy, depth and body-attitude cues stabilize the prediction. Multimodal fusion research [26] supports reliability-aware weighting under different sensor failures. The relative contribution of each source therefore varies with terrain type, contact stage and sensor condition, which is essential for deformable surfaces, boundary terrain and flat but slippery ground.

Figure 7 shows the terrain adaptation output used by the supervisory gait controller. Figure 7(a) presents step-height adjustment, where step-edge and gravel terrain require higher lifting. Figure 7(b) presents foot-force redistribution, where soft soil and wet surfaces require smaller impact force and softer stance. Figure 7(c) presents slip-risk priority, which is higher for wet surfaces and mixed transitions. Figure 7(d) presents the gait review record that stores terrain

class, confidence and dominant evidence source. Gait adaptation research [27] shows that this type of output is more useful for locomotion control than a single terrain label.

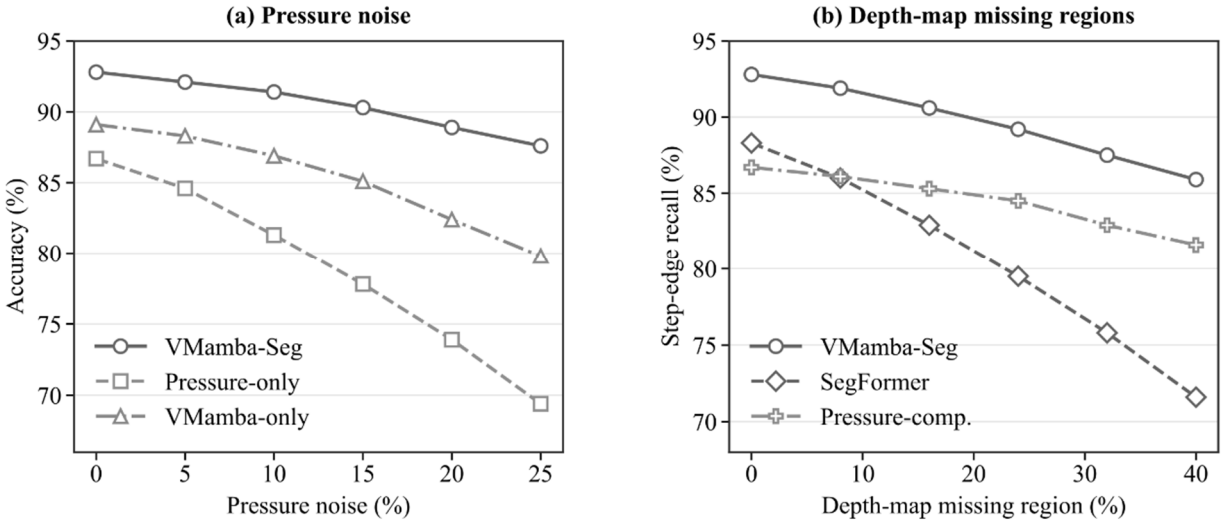


Figure 6. Robustness Under Sensing Disturbance. (a) Pressure noise. (b) Depth-map missing regions.

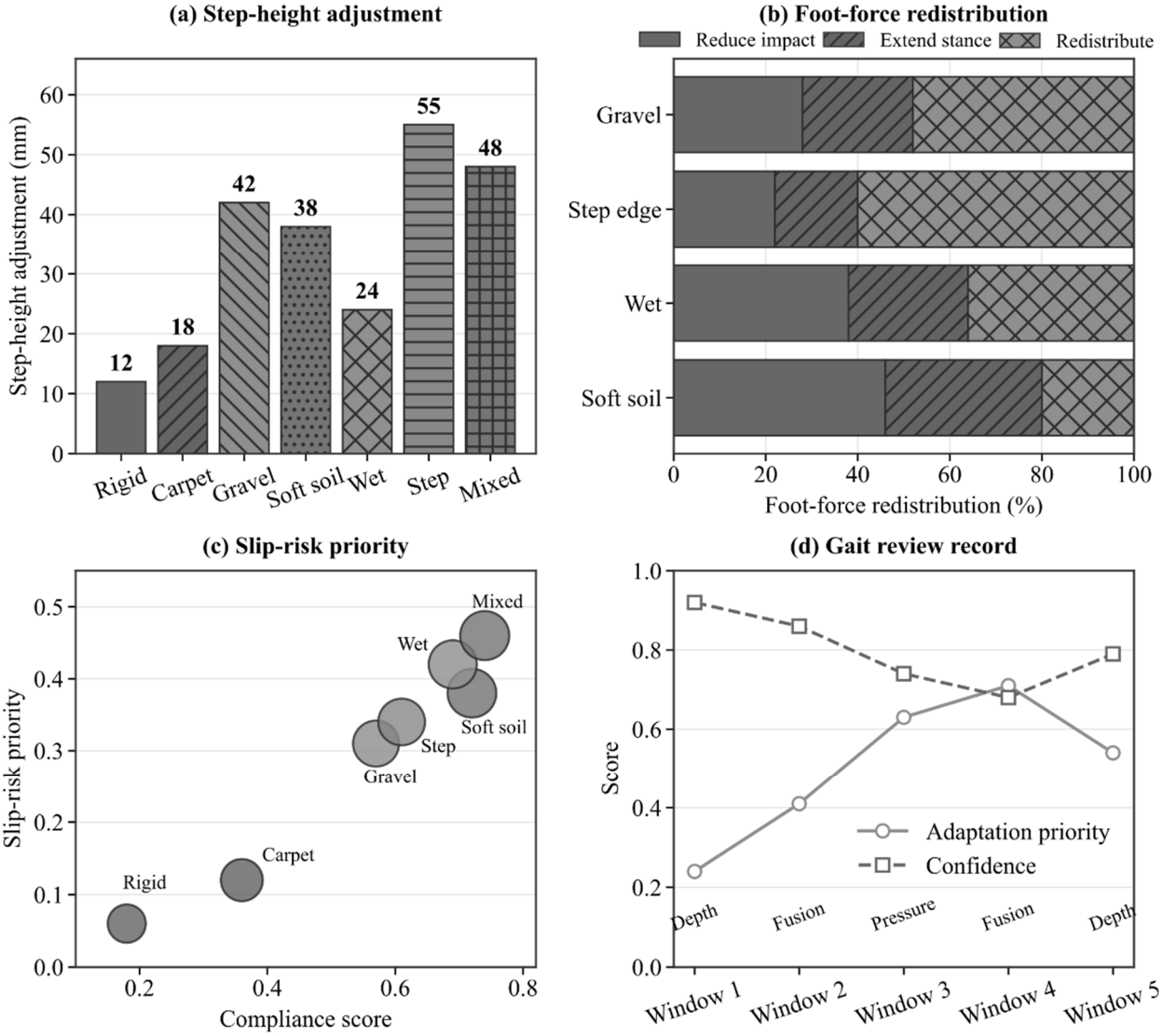


Figure 7. Terrain Adaptation Output. (a) Step-height adjustment. (b) Foot-force redistribution. (c) Slip-risk priority. (d) Gait review record.

The purpose of the proposed model in engineering is to separate terrain category, compliance demand and confidence. High-confidence stiff ground permits normal gait, high-confidence soft terrain supports stiffness reduction

and longer support time, and low-confidence terrain windows should trigger cautious stepping or remeasurement before sudden gait switching. Supervisory control studies [28] indicate that separating prediction from controller action improves safety under uncertain robot operation. The proposed output therefore informs gait adaptation without replacing the low-level controller.

The model also produces route-level terrain records. If several low-confidence windows occur within one route segment, the robot can flag that area for terrain inspection. Frequent pressure evidence in visually flat areas may indicate mechanically weak ground, and repeated depth evidence near step edges can justify increased clearance before the next pass. Long-term field robotics research [29] shows that such logs improve deployment by linking perception output with later locomotion results.

The remaining limitation is that the current dataset is artificial and does not include the full range of outdoor disturbances. Sand, mud, grass, snow and loose debris may produce pressure-depth relationships that differ from the tested classes. Recent adaptive robot systems [30] recommend updating terrain models with field data after deployment. Future field deployment should therefore include continuous calibration, sensor-health monitoring and human review of uncertain terrain windows, with compliance thresholds adjusted after field confirmation.

5. Conclusion

This study proposed a VMamba-Seg model for terrain compliance adaptation of legged robots based on foot pressure and depth maps, and the overall conclusion is that pre-contact depth evidence, stance-phase pressure response, body attitude and sensor-quality indicators should be evaluated jointly for reliable supervisory gait adaptation. First, experiments on 4,360 terrain-contact windows show that the proposed model achieves 92.8% terrain recognition accuracy, 89.6% compliance macro-F1 and 84.7% soft-ground recall. Second, the comparative results show that pressure evidence is most informative for soft-ground compliance, while depth segmentation is more effective for step edges and obstacle boundaries. Third, the main limitation is that the dataset remains artificial and covers a limited set of terrain categories; therefore, future validation should include real terrain-contact data, more terrain materials, different robot speeds and real gait-control feedback.

The proposed output separates terrain class, compliance score, confidence, dominant evidence source and adaptation priority. This structure supports normal gait for reliable rigid ground, stiffness reduction for compliant surfaces, higher step clearance for step edges and cautious stepping for low-confidence terrain windows. The confidence score is especially important because sensor uncertainty should reduce controller authority rather than force an abrupt gait switch. Evidence records can also support route review and later calibration of terrain-specific thresholds.

Future work should extend the dataset with field terrain-contact logs, additional terrain materials, robot-speed variation and real gait-control feedback. Deployment studies should examine energy consumption, slip recovery, step replanning, body stability and locomotion safety under sand, mud, grass, snow and loose debris. These extensions will clarify how perception confidence should be connected with controller authority during extended monitoring and emergency-response locomotion in outdoor environments.

6. Declarations

6.1. Author Contributions

Maria Kuchta contributes to conceptualization, methodology, software, validation, analysis, investigation, data collection, draft preparation, manuscript editing, visualization, supervision, project administration, and funding acquisition. Natalia Kurzawa contributes to validation, analysis, investigation, data collection, draft preparation. All authors have read and agreed with the manuscript before its submission and publication.

6.2. Funding and Acknowledgments

This research received no specific financial support from any funding agency.

6.3. Institutional Review Board Statement

Not applicable.

6.4. Informed Consent Statement

Not applicable.

6.5. Data Availability

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

6.7. Declaration of generative AI and AI-assisted technologies

During the preparation of this manuscript, the authors utilized Grammarly for linguistic polishing, logical organization and draft revision. The generative AI tool was not involved in research design, data collection, data analysis, or the equation of core conclusions. All outputs generated by AI were reviewed, revised and verified manually by the authors. The authors take full responsibility for all content of this manuscript.

7. References

[1] Zhuang, H., Wang, J., Wang, N., Li, W., Li, N., Li, B., & Dong, L. (2024). A review of foot-terrain interaction mechanics for heavy-duty legged robots. Applied Sciences, 14(15), 6541. https://doi.org/10.3390/app14156541

[2] Wang, C., Zhang, R., Dong, W., Li, T., Jiang, L., Xu, W., Xu, P., Zhou, Y., & Zou, M. (2024). Trafficability anticipation for quadruped robot in field operation. Journal of Field Robotics, 41(4), 851-866. https://doi.org/10.1002/rob.22292

[3] Moreno, H. A., Moreno, L. A., Valentin-Coronado, L. M., & Flores, G. (2025). Soft paw sensor for tactile and force sensing in legged robots. Mechatronics, 111, 103407. https://doi.org/10.1016/j.mechatronics.2025.103407

[4] Carvalho, A. E., Ferreira, J. F., & Portugal, D. (2024). 3D traversability analysis and path planning based on mechanical effort for UGVs in forest environments. Robotics and Autonomous Systems, 171, 104560. https://doi.org/10.1016/j.robot.2023.104560

[5] Chen, Y., Yuan, H., Du, R., Song, S., et al. (2024). Terrain recognition in real-time for a legged robot based on ontology information. Procedia Computer Science, 250, 101-107. https://doi.org/10.1016/j.procs.2024.11.015

[6] Liu, Y., Tian, Y., Zhao, Y., Yu, H., Xie, L., Wang, Y., Ye, Q., Jiao, J., & Liu, Y. (2024). VMamba: Visual state space model. Advances in Neural Information Processing Systems, 37. https://doi.org/10.52202/079017-3273

[7] Cummins, L., Somers, A., Riza, T., et al. (2024). Explainable predictive maintenance: A survey of current methods, challenges and opportunities. IEEE Access, 12, 57574-57602. https://doi.org/10.1109/ACCESS.2024.3391130

[8] Shi, G., Yao, C., Liu, X., Zhao, Y., Zhu, Z., & Jia, Z. (2024). Foot Vision: A vision-based multi-functional sensorized foot for quadruped robots. IEEE Robotics and Automation Letters, 9(7), 6720-6727. https://doi.org/10.1109/LRA.2024.3408084

[9] Yao, C., Shi, G., Xu, P., Lyu, S., Qiang, Z., Zhu, Z., Ding, L., & Jia, Z. (2024). STAF: Interaction-based design and evaluation of sensorized terrain-adaptive foot for legged robot traversing on soft slopes. IEEE/ASME Transactions on Mechatronics, 29(6), 4039-4050. https://doi.org/10.1109/TMECH.2024.3350183

[10] Duncan, J. A., Alambeigi, F., & Pryor, M. W. (2024). A survey of multimodal perception methods for human-robot interaction in social environments. ACM Transactions on Human-Robot Interaction, 13(4), 1-50. https://doi.org/10.1145/3657030

[11] Lee, K., & Lee, K. (2024). Terrain-aware path planning via semantic segmentation and uncertainty rejection filter with adversarial noise for mobile robots. Journal of Field Robotics, 42, 287-301. https://doi.org/10.1002/rob.22411

[12] Yu, J., Wang, D., Chen, Z., Chen, C., Wu, S., Wang, Y., & Xiong, R. (2026). Foothold-based planning for legged robot autonomous navigation over uneven terrain. IEEE/ASME Transactions on Mechatronics, 31(1), 93-104. https://doi.org/10.1109/TMECH.2025.3588355

[13] Park, K., Kim, Y., Kim, D., & Choi, J. W. (2025). Resilient sensor fusion under adverse sensor failures via multi-modal expert fusion. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 6720-6729. https://doi.org/10.1109/CVPR52734.2025.00630

[14] Qu, H., Ning, L., An, R., Fan, W., Derr, T., Liu, H., Xu, X., & Li, Q. (2026). A survey of Mamba. ACM Transactions on Intelligent Systems and Technology, 17(5), 1-43. https://doi.org/10.1145/3819582

[15] Hur, M., Kim, D., & Min, M. (2026). SensitiveMamba: Selective state space modeling for robust time series anomaly detection. Proceedings of the 41st ACM/SIGAPP Symposium on Applied Computing, 1156-1163. https://doi.org/10.1145/3748522.3779727

[16] Tsuchiya, R., & Kobayashi, Y. (2024). Evaluation of confidence-aware prediction of uneven terrain risk for self-supervised learning by a mobile robot. 2024 IEEE/SICE International Symposium on System Integration, 786-791.

https://doi.org/10.1109/SII58957.2024.10417249

[17] Chen, W., Yang, K., Yu, Z., Shi, Y., & Chen, C. L. P. (2024). A survey on imbalanced learning: Latest research, applications and future directions. Artificial Intelligence Review, 57(6), 137. https://doi.org/10.1007/s10462-024-10759-6

[18] Wang, T. K., & Lin, P. C. (2024). Legged robot locomotion on viscoelastic terrain using a reduced-order dynamic model with rolling contact and energy input and dissipation. Journal of Mechanics, 40, 571-585. https://doi.org/10.1093/jom/ufae048

[19] Lee, J., Bjelonic, M., Reske, A., Wellhausen, L., Miki, T., & Hutter, M. (2024). Learning robust autonomous navigation and locomotion for wheeled-legged robots. Science Robotics, 9(89), eadi9641. https://doi.org/10.1126/scirobotics.adi9641

[20] Liu, S., Pan, Z., Zhou, S., Niu, Z., & Wang, R. (2024). Sensorless ground reaction force observation with disturbance compensation in heavy-legged robots. IEEE/ASME Transactions on Mechatronics, 29, 3672-3683. https://doi.org/10.1109/TMECH.2024.3354989

[21] Yu, J., Wang, D., Chen, Z., Chen, C., Wu, S., Wang, Y., & Xiong, R. (2024). A fast motion and foothold planning framework for legged robots on discrete terrain. 2024 IEEE/RSJ International Conference on Intelligent Robots and Systems, 10678-10685. https://doi.org/10.1109/IROS58592.2024.10802606

[22] Viktor, P., & Kiss, G. (2026). Multimodal sensor fusion in autonomous vehicles: Technologies, architectures, and open challenges. Sensors, 26(11), 3528. https://doi.org/10.3390/s26113528

[23] Gao, Z., Jiang, X., Xu, X., Shen, F., Li, Y., & Shen, H. T. (2024). Embracing unimodal aleatoric uncertainty for robust multimodal fusion. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 26866-26875. https://doi.org/10.1109/CVPR52733.2024.02538

[24] Fan, D. D., Otsu, K., Kubo, Y., Dixit, A., Burdick, J., & Agha-Mohammadi, A. A. (2025). STEP: Stochastic traversability evaluation and planning for risk-aware navigation; Results from the DARPA Subterranean Challenge. IEEE Transactions on Field Robotics. https://doi.org/10.1109/TFR.2024.3512433

[25] Cai, X., Queeney, J., Xu, T., Datar, A., Pan, C., Miller, M., Flather, A., Osteen, P. R., Roy, N., Xiao, X., & How, J. P. (2025). PIETRA: Physics-informed evidential learning for traversing out-of-distribution terrain. IEEE Robotics and Automation Letters, 10(3), 2359-2366. https://doi.org/10.1109/LRA.2025.3527285

[26] Zhao, F., Zhang, C., & Geng, B. (2024). Deep multimodal data fusion. ACM Computing Surveys, 56(9), 1-36. https://doi.org/10.1145/3649447

[27] Chen, G., Zhou, Z., Lei, X., Lodewijks, G., Zhang, R., Zou, M., & Ji, A. (2025). Biomimetic transition gait for quadruped robot creeping from level to slope surfaces. Journal of Field Robotics, 42(8), 4090-4100. https://doi.org/10.1002/rob.70007

[28] Li, P., Yin, B., Zhang, L., & Zhao, Y. (2025). Adaptive control algorithm for quadruped robots in unknown high-slope terrain. Journal of Engineering Research, 13(3), 1771-1784. https://doi.org/10.1016/j.jer.2024.05.018

[29] Prados, C., Hernando, M., & Gambao, E. (2025). Path and footfall planning for N-legged and climbing robots: A model predictive control approach. Robotics and Autonomous Systems, 194, 105119. https://doi.org/10.1016/j.robot.2025.105119

[30] Zhang, Y., Huang, B., Hong, M., Huang, C., Wang, G., & Guo, M. (2025). A terrain classification method for quadruped robots with proprioception. Electronics, 14(6), 1231. https://doi.org/10.3390/electronics14061231