2024 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)
October 14-18, 2024. Abu Dhabi, UAE

On Predicting Terrain Changes Induced by Mobile Robot Traversal

Miloš Prágr* Jan Bayer* Jan Faigl

***Abstract*— Mobile robots operating in convoys have a limited view of the terrain to be traversed if it is occluded by the preceding vehicle. Furthermore, the preceding vehicle might change the terrain geometry and eventually significantly alter its traversability by driving over the terrain. When the following vehicles do not consider such changes, they can use spurious terrain appearance and geometry to decide whether to follow in the tracks of the previous vehicle or to avoid them since the preceding vehicle's tracks can make the terrain untraversable. We propose to predict the terrain changes induced by the robot traversal on the traversed terrain and thus support the decision-making of the following vehicles. The developed model projects the robot wheel footprint along the planned robot path and combines the projection with the terrain appearance and prior terrain elevation. The coupled model is used in a convolutional neural network that predicts the elevation after traversal. The footprint projection component is designed so that learned networks can be transferred to vehicles with different wheel footprints without relearning the model. The proposed model is verified using a dataset captured using a real, one-ton, six-wheel robot traversing rigid roads and vegetated fields.**

I. Introduction

Terrain traversability assessment encodes mobile robots' experience over the traversed terrains. It is crucial in robots' autonomous decision-making to navigate safe, fast, and energetically efficient paths. Since traversability assessment is used in path planning, the terrain needs to be evaluated from a distance before passing through. Therefore, robots assess traversability based on the terrain's geometry, visual appearance, or overhead imagery. The assessment is limited by the range and modality of the sensors used.

The terrain assessment may also be influenced by changes to terrain caused by preceding vehicles in convoys. Furthermore, a window where the robot can observe the terrain traversed by the preceding vehicles might not be sufficient to adjust its path promptly and thus avoid hard-to-traverse segments such as deep mud tracks imprinted by the leading vehicle [1]. Hence, the robot should employ a traversability assessment model that predicts the changes imprinted onto the terrains by the traversal of other units in the convoy.

In this paper, we report our early results on predicting terrain changes induced by mobile robot traversal along the planned robot path. The model comprises a combination of a wheel footprint projector, designed to adapt the model to vehicles with different footprints, and a convolutional neural network that operates over the elevation and feature description along the vehicle path. The model is evaluated in a set of ablation tests on a dataset of a six-wheeled robot shown in Fig. 1 traversing different rigid roads and non-rigid vegetation that react differently to the robot traversal.

Fig. 1. Tracks imprinted on vegetation by the six-wheeled robot used in the model verification.

The rest of the paper is organized as follows. An overview of the related work on traversability through compliant terrains is presented in Section II. The addressed problem of modeling terrain changes induced by mobile robot traversal is formally introduced in Section III. The proposed model is detailed in Section IV, and evaluation results are reported in Section V. Finally, the paper is concluded in Section VI.

II. Terrain Traversability Assessment Approaches

Mobile robots avoid areas that would bar them from carrying out their mission, be it through damage to the robot or impeding further motion. Hence, robots reason about their traversability over the local terrains and adjust their decision-making accordingly. In this section, we briefly review the state-of-the-art traversability assessment approaches, with a particular focus on traversability through soft and compliant terrains.

Papadakis [2] presents a taxonomy of traversability assessment that divides the state-of-the-art approaches into appearance-based, geometry-based, proprioceptive, and hybrid. The exteroceptive geometry and appearance-based approaches assess terrains from a distance. Hence, these approaches are particularly suitable for path planning. Proprioceptive approaches are based on the robot's experience, such as awareness of movement and position, and range from robot velocity and energy consumption [3] to methods that utilize stability or vibration [4]. Unlike exteroception, robot proprioception cannot be applied directly in remote

The authors are with the Faculty of Electrical Engineering, Czech Technical University, Technicka 2, 166 27, Prague, Czechia, {pragrmi1|bayerja1|faiglj}@fel.cvut.cz. The work has been supported by the European Union under the project ROBOPROX - Robotics and advanced industrial production (reg. No. CZ.02.01.01/00/22_008/0004590). *Miloš Prágr and Jan Bayer are co-first authors. (Corresponding author: Miloš Prágr.)

terrain assessment. Instead, hybrid near-to-far methods learn to infer the robot experience near-observed traversability from descriptors observable from afar. Examples of hybrid methods are prediction of vehicle vibration from imagery data [5] or using overhead imagery to predict traversal cost defined from LiDAR measurements [6].

Terrain geometry offers a range of interpretable descriptors such as slope, step height, or roughness that can be directly tied to robot properties, albeit individual authors compute these descriptors differently, as demonstrated in [7], [8], [9]. Recently, both geometric and visual data have been applied in top-down Bird's Eye View (BEV) projections, which are robust concerning changing robot viewpoint and afford straightforward application for inpainting convolutional networks that fill in unobserved areas [10].

The traversability over soft and compliant terrains such as mud or vegetation is difficult to model since geometry without additional context cannot be used to assess such terrains. The authors of [11] learn the rigidity of obstacles encoded as the force required to pass through. The force is measured by a robot-mounted bumper, and a haptic antenna is used for a similar purpose in [12].

Visual appearance is beneficial when dealing with vegetation where geometry is ambiguous due to varying obstacle rigidity or when no geometry is available, such as in overhead terrain assessment [13]. Imagery in spectra outside of human visible ranges is suitable for detecting vegetation due to chlorophyll reflective properties. The authors of [14] review indices focused on highlighting vegetation. Furthermore, thermal sensing predicts the slip over potentially loose ground [15].

The rigidity of the supporting ground is a concern for wheeled robots that suffer from wheel sinkage. Approaches that address wheel sinkage rely on the identification of the terra-mechanics of the traversed soil. For instance, the authors of [16] propose to use the Kalman filter in the online estimation of the sinkage exponent, which is then used to enhance model predictive control in [17].

Since existing terra-mechanical models focus on vehicle-soil interaction, they are unsuitable for a broader class of terrains that might be covered by vegetation with unknown rigidity, which can spring up after traversal. Moreover, to the best of the authors' knowledge, no approach is explicitly focused on predicting the geometry of such terrain after robot traversal, which concerns mobile robots operating in convoys where the following vehicles have a limited view of the terrain to be traversed. Therefore, we propose to predict the terrain geometry change in BEV after robot traversal.

III. Problem Statement

The main goal of the proposed approach is to predict terrain changes that can be characterized by changes in terrain geometry, elevation in particular. Therefore, we consider the input to the predictor of the terrain changes to be the terrain elevation grid map $\mathcal{M}_{2.5D}$, in our particular case, with the cell size $\mathrm{d}_\nu = 0.1\,\mathrm{m}$. Each cell $\nu \in \mathcal{M}_{2.5D}$ is defined by its xy-position $\boldsymbol{p}(\nu) = (x, y)$ in the world reference frame and its elevation $z(\nu) \in \mathcal{R} \cup \varnothing$, where $\varnothing$ denotes that the elevation is missing as a result of occlusion. Besides, the additional predictor input is the planned path, along which it is assumed the robot would interact with the terrain. The path is represented as a sequence of poses

$$\begin{aligned}\Psi = \{&\mathbf{q}_1, \mathbf{q}_2, \ldots, \mathbf{q}_m; \mathbf{q}_i = (x_i, y_i, \theta_i);\\ &\forall i \in 1, 2, \ldots, m-1 : \mathbf{q_{i+1}} \in \text{8-nbh}(\mathbf{q}_i)\},\end{aligned} \tag{1}$$

where θ_i is the robot orientation, each position (x_i, y_i) of $\mathbf{q}_i$ corresponds with the position of the cell $\nu_i \in \mathcal{M}_{2.5D}$, and 8-nbh is the 8-neighborhood function on $\mathcal{M}_{2.5D}$.

The robot traverses the path Ψ, thus inducing changes on the terrain and changing the prior grid map $\mathcal{M}^{\text{prior}}_{2.5D}$ into the posterior grid map $\mathcal{M}^{\text{post}}_{2.5D}$. The changes induced on the terrain are affected by the robot configuration ρ. In particular, the robot wheel footprint is expressed as

$$\boldsymbol{W} = \begin{bmatrix} x_1 & x_2 & \ldots & x_n \\ y_1 & y_2 & \ldots & y_n \end{bmatrix} \tag{2}$$

that describes the relative positions of the robot's n wheels with respect to (w.r.t.) the robot position and heading.

The goal of the terrain change model $\mathcal{TC}$ is to predict the posterior elevation map $\widehat{\mathcal{M}}^{\text{post}}_{2.5D}$ after the robot traversal along the path Ψ, given the prior elevation map $\mathcal{M}^{\text{prior}}_{2.5D}$ and robot description $\rho = (\boldsymbol{W})$ as

$$\widehat{\mathcal{M}}^{\text{post}}_{2.5D} = \mathcal{TC}(\mathcal{M}^{\text{prior}}_{2.5D}, \rho). \tag{3}$$

Since the elevation grid map alone is not sufficient to discriminate between some individual terrains that might interact differently with the robot, such as concrete and mud, each cell carries an additional feature descriptor $f(\nu)$ selected w.r.t. the available sensor modalities and local surrounding environment. In this paper, we evaluate the proposed model over flat ground and vegetation with varying depths using a LiDAR-based dataset. Hence, without the loss of generality, the selected terrain feature is the min-max spread $\Delta(\nu)$ of the elevation values observed at the cell ν when constructing the grid map $\mathcal{M}_{2.5D}$ from localized LiDAR scans. The feature captures that LiDAR partially penetrates through vegetation and thus observes vegetation in one cell at different depths.

The quality of the predictions is quantified as the Root Mean Squared Error (RMSE) of the elevation prediction along the robot path Ψ w.r.t. the ground truth $\mathcal{M}^{\text{post}}_{2.5D}$ as

$$\begin{aligned}&\text{RMSE}(\widehat{\mathcal{M}}^{\text{post}}_{2.5D}, \mathcal{M}^{\text{post}}_{2.5D}) = \sqrt{\sum_{*} (z(\hat{\nu}) - z(\nu))^2},\\ &* = \forall \hat{\nu} \in \widehat{\mathcal{M}}^{\text{post}}_{2.5D}, \nu \in \mathcal{M}^{\text{post}}_{2.5D}, \exists \nu' \in \mathcal{M}^{\text{post}}_{2.5D} :\\ &\boldsymbol{p}(\hat{\nu}) = \boldsymbol{p}(\nu), z(\nu) \neq \varnothing, \nu' \in \Psi, \|\boldsymbol{p}(\nu'), \boldsymbol{p}(\nu)\| < \mathrm{d}^{\Psi},\end{aligned} \tag{4}$$

where the RMSE is computed over the cells within d^{Ψ}-distance of the path Ψ for which the ground truth posterior elevation can be observed with $z(\nu) \neq \varnothing$.

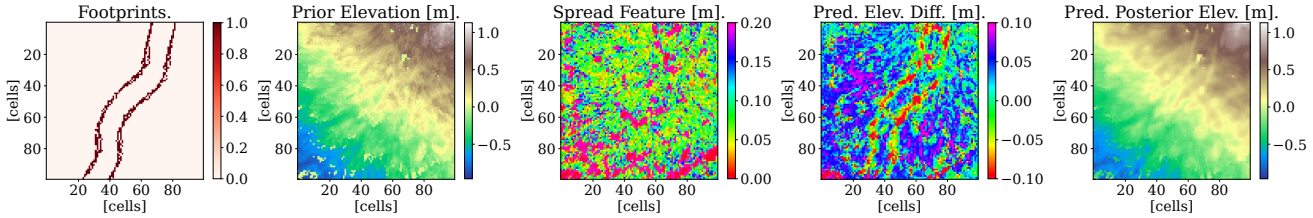


Fig. 2. The (a,b,c) inputs and (d,e) outputs of the convolutional neural network.

IV. Proposed Terrain Change Modeling

The proposed approach to predict terrain changes induced by mobile robot traversal comprises two parts. The first part is a footprint projector, which applies the robot description ρ to the elevation grid map $\mathcal{M}^{\text{prior}}_{2.5D}$. The second part is a convolutional neural network, which operates over the transformed inputs to predict the elevation after the robot traversal $\widehat{\mathcal{M}}^{\text{post}}_{2.5D}$. Each part is described in detail in the following paragraphs.

A. Wheel Footprint Projection

The wheel footprint projector applies the robot description ρ in the form of the wheel footprint $\boldsymbol{W}$ to the grid map $\mathcal{M}^{\text{prior}}_{2.5D}$ as follows. For each robot path pose $\mathbf{q} = (x, y, \theta) \in \Psi$, the wheel positions $\boldsymbol{W}'$ in the world reference frame can be expressed as

$$\begin{bmatrix} \boldsymbol{W}' \\ \mathbf{1}_n^T \end{bmatrix} = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & x \\ \sin(\theta) & \cos(\theta) & y \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \boldsymbol{W} \\ \mathbf{1}_n^T \end{bmatrix}, \tag{5}$$

where $\mathbf{1_n}$ is the n-length vector of ones $\mathbf{1}_n = (1, 1, \ldots, 1) \in \mathbb{R}^n$, where n is the number wheels in $\boldsymbol{W}$. Wheel width is left to be encoded in the neural network. The projected wheel positions are then applied to the gridmap so that $w(\nu) = 1$ if a wheel is projected to ν, and $w(\nu) = 0$ otherwise. An example of a wheel footprint is visualized in Fig. 2a.

B. Convolutional Neural Network

The convolutional neural network (CNN) is selected to infer the elevation change induced by the robot from the prior elevation extended by terrain feature description and wheel footprint projection along the robot path. The posterior elevation is inferred in the form of the elevation change as

$$\begin{aligned} \widehat{\mathcal{M}}^{\text{post}}_{2.5D} = \mathcal{TC}(\mathcal{M}^{\text{prior}}_{2.5D}) = \mathcal{M}^{\text{prior}}_{2.5D} + \widehat{\mathcal{M}}^{\text{diff}}_{2.5D} = \\ = \mathcal{M}^{\text{prior}}_{2.5D} + \text{CNN}(\mathcal{M}^{\text{prior}}_{2.5D}), \end{aligned} \tag{6}$$

following the intuition that it is preferable to infer zero elevation change for the inputs remote from the robot since the area can be sloped. The predicted change and elevation can be seen in Figs. 2d and 2e, respectively.

The network operates over the normalized $d^{\text{in}} \times h \times w$ input window $I \subset \mathcal{M}^{\text{prior}}_{2.5D}$, where $h = 100$ and $w = 100$ correspond to $10 \times 10\,\text{m}$ area and $d^{\text{in}} = 3$ is the input dimensionality that covers the normalized elevation map, projected robot footprint map, and spread Δ feature map[1]. The elevation is normalized, see Fig. 2b, in the way that the mean and unknown elevation values are set to 0

$$z'(\nu) = \begin{cases} z(\nu) - \frac{\sum_{\nu \in I, z(\nu) \neq \varnothing} z(\nu)}{|\nu \in I, z(\nu) \neq \varnothing|} & \text{if } z(\nu) \neq \varnothing \\ 0 & \text{otherwise} \end{cases}. \tag{7}$$

The spread feature is set to 0 for the unknown elevation cells

$$\Delta'(\nu) = \begin{cases} \Delta(\nu) & \text{if } z(\nu) \neq \varnothing \\ 0 & \text{otherwise} \end{cases}. \tag{8}$$

An example of the spread feature is depicted in Fig. 2c.

The network architecture comprises two convolutional layers and two deconvolutional layers, all using $d^{\text{nn}} \times d^{\text{nn}} \times k \times k$ sized kernels, where $k = 5$ is the kernel window size and $d^{\text{nn}} = 10$ is the network latent dimensionality. Besides, two additional convolutional layers are used as input and output projection, shaped as $d^{\text{in}} \times d^{\text{nn}} \times 1 \times 1$ and $d^{\text{nn}} \times d^{\text{out}} \times 1 \times 1$, respectively, where the output dimensionality is $d^{\text{out}} = 1$ since only the elevation difference is predicted. The padding of all the convolutional and deconvolutional layers is set up so that the network's latent layers are shaped as $d^{\text{nn}} \times h \times w$.

The network is learned in 250 iterations using the Adam [18] optimizer with 10^{-3} learning rate and 10^{-8} weight decay w.r.t. the mean squared error that is weighted by a path-centric mask over cells ν with known elevation in the ground truth $\mathcal{M}^{\text{post}}_{2.5D}$

$$\lambda(\nu) = \begin{cases} 0 & \text{if } z(\nu) = \varnothing, \\ 1 & \text{if } \exists \nu' \in \mathcal{M}^{\text{post}}_{2.5D} : \nu' \in \Psi, \|\boldsymbol{p}(\nu'), \boldsymbol{p}(\nu)\| < \text{d}^{\Psi}, \\ \frac{1}{\alpha^{\text{far}}} & \text{otherwise,} \end{cases} \tag{9}$$

where $\alpha^{\text{far}} = 5$ prioritizes predictions along the path while discouraging non-deterministic behavior in areas remote from the path.

V. Results

The proposed model has been verified using a dataset of flat roads and vegetated fields traversed by a six-wheeled one-ton robot with the footprint characterized as

$$\boldsymbol{W}_{\text{robot}} = \begin{bmatrix} 0.800 & 0.000 & -0.800 & 0.800 & 0.000 & -0.800 \\ -0.705 & 0.705 & -0.705 & 0.705 & -0.705 & 0.705 \end{bmatrix}. \tag{10}$$

The dataset comprises two trails through the environment, which are used as the training and testing set with 201 and

[1] In the ablation tests, the dimensionality may differ as for the considered modification of the network.

323 windows with the size $10 \times 10\,\mathrm{m}$ of the prior-posterior elevation, respectively. In the rest of the presented results, we report on the tests on the individual components of the proposed model and demonstrate its use in an online prediction within a Robot Operating System (ROS) [19] ecosystem. We discuss the properties of the model in Section V-F.

A. Dimensionality of the Latent Layers

We test the network architecture with the latent dimensionality $d^{\text{nn}} \in \{2, 5, 10, 20\}$. The results in Table I suggest that $d^{\text{nn}} \geq 10$ should be used, and no benefits have been observed when increasing the dimensionality to $d^{\text{nn}} = 20$. Hence, we select $d^{\text{nn}} = 10$ for the remaining experiments.

TABLE I
INFLUENCE OF LATENT NETWORK DIMENSIONALITY

d^{nn}	2	5	10	20
RMSE	5.56 ± 0.16	5.23 ± 0.21	$\mathbf{4.98} \pm 0.13$	5.00 ± 0.11

$[\mathrm{m} \times 10^{-2}]$, $\mu \pm \sigma$ of 9 learning runs.

B. Effect of Input Types

Table II reports on tests investigating the effect of changing the network inputs. Three variants are considered. First, the proposed version that operates over the normalized elevation z, spread feature Δ, and footprint projection w, denoted as $z + \Delta + w$. The model corresponds to $d^{\text{nn}} = 10$ in Table I. Next, the wheel footprints are replaced by the robot path projection $\psi(\nu)$, which reports 1 if and only if $\nu \in \Psi$ and 0 otherwise. Hence, the network operates over elevation z, spread feature Δ, and path projection ψ, and the input set is denoted $z + \Delta + \psi$. Finally, the spread feature is omitted with the network operating over the elevation z and wheel footprint projection w, denoted as $z + w$.

TABLE II
INFLUENCE OF INPUT TYPES

Inputs	$z + \Delta + w$	$z + \Delta + \psi$	$z + w$
RMSE	$\mathbf{4.98} \pm 0.13$	5.17 ± 0.10	5.18 ± 0.09

$[\mathrm{m} \times 10^{-2}]$, $\mu \pm \sigma$ of 9 learning runs.

The proposed model $z + \Delta + w$ performs better than the alternative models. Moreover, a qualitative analysis suggests that although the difference in RMSE is relatively small when the spread is omitted, the network poorly distinguishes rigid terrains where tracks should not be imprinted, see Fig. 3.

Besides, since the wheel footprint is applied before the network, the proposed model supports prediction for vehicles with wheel footprints that are different from learning. Fig. 4 shows the predictions using narrow, standard, and wide wheelbase footprints.

C. Elevation Difference vs. Posterior Elevation Prediction

In (6), the CNN predicts the posterior elevation indirectly in the form of the elevation change, following the intuition that it is preferable to infer zero elevation change for the inputs remote from the robot. We investigate the effect of inferring the posterior elevation directly as

$$\widehat{\mathcal{M}}^{\text{post}}_{2.5D} = \mathcal{TC}(\mathcal{M}^{\text{prior}}_{2.5D}) = \text{CNN}(\mathcal{M}^{\text{prior}}_{2.5D}). \qquad (11)$$

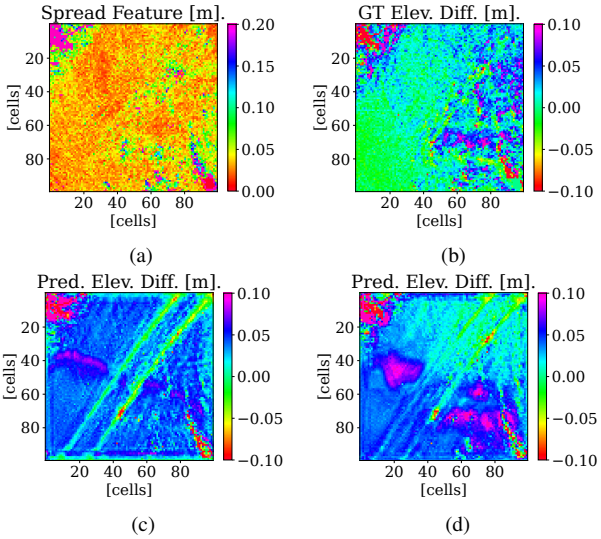
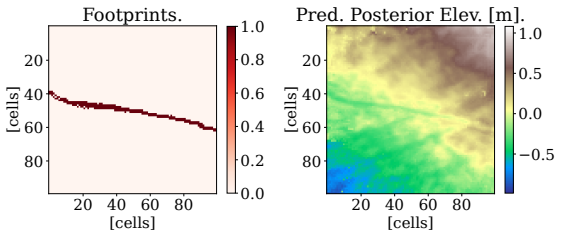
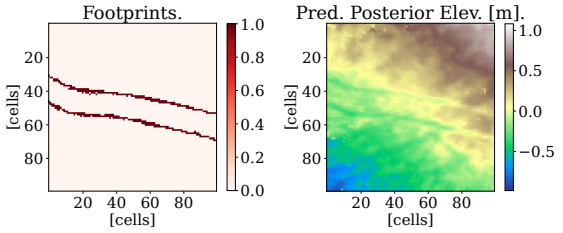


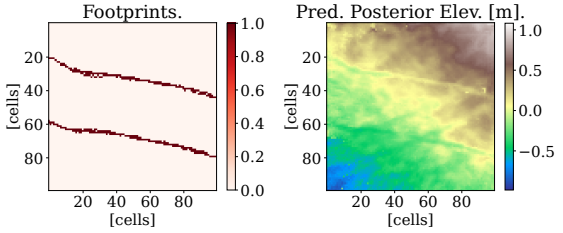
Fig. 3. (a) Influence of the spread Δ feature. (b) In the ground truth elevation change, faint shallow diagonal tire track is visible only around higher spread values, which correlate with tall grass where tracks tend to be imprinted. While both methods predict spurious deep tracks, these are much more pronounced (c) in the model missing the spread input. (d) In the proposed model with the spread feature, the deep tracks are limited to areas with higher spread values.



(a) 0.1 m wheelbase.



(b) 1.4 m wheelbase (real robot).



(c) 3.4 m wheelbase.

Fig. 4. Projected footprints and posterior elevation predictions using different robot wheelbases, demonstrating (a) narrow, (b) normal, and (c) wide configuration. All predictions use the standard model learned using the (b) normal wheelbase.

Table III presents the influence of prediction through (6) and (11). The results suggest that the proposed indirect prediction through elevation difference results in lower RMSE.

TABLE III

INFLUENCE OF PREDICTION TYPE

Prediction	Through Elevation Difference	Directly as Elevation
RMSE	$\mathbf{4.98} \pm 0.13$	5.16 ± 0.26

$[\mathrm{m} \times 10^{-2}]$, $\mu \pm \sigma$ of 9 learning runs.

D. Effect of Path Distance Error Discount

The effect of the remote prediction discount used in learning is presented in Table IV. Moderate discount performs the best, which is the expected behavior since it is closer to the testing error (4) than $\alpha^{\text{far}} = 1$ while still incorporating areas without robot traversal. However, the difference between the values of α^{far} is negligible w.r.t. RMSE variance. Nonetheless, since $\alpha^{\text{far}} = 5$ performs the best without increasing the network complexity, we elect to use it to learn the proposed model.

TABLE IV

INFLUENCE OF PATH DISTANCE ERROR DISCOUNT

α^{far}	1	5	1000
RMSE	5.03 ± 0.11	$\mathbf{4.98} \pm 0.13$	5.04 ± 0.18

$[\mathrm{m} \times 10^{-2}]$, $\mu \pm \sigma$ of 9 learning runs.

E. Online Operation

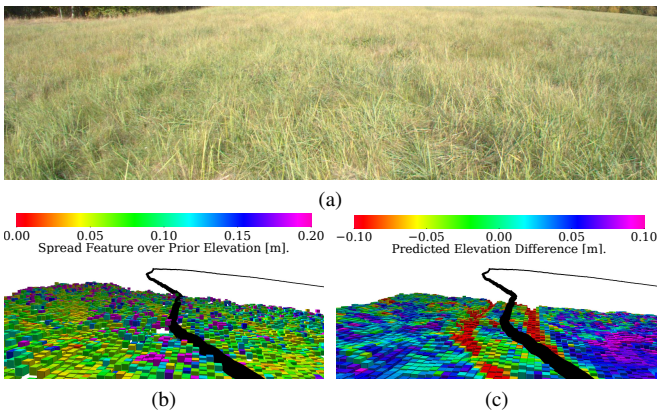


Fig. 5. Predictions over (a) tall grass field. (b) The spread feature (high values in purple) and the future robot path (in black). (c) Predicted elevation difference (jet color scheme, negative change in red, position change in blue/purple, no change in green). The red tracks are visible along the robot's path through the tall grass. The remaining areas along the path are predicted to have no elevation change, while the tall grass remote from the path, which is in general observed noisily in LiDAR, is predicted to rise slightly.

The proposed model has been verified in online operation when localized LiDAR data corresponding to the training and testing sets are played in a time-synchronized fashion from ROS bags (time-stamped archives of ROS data), and an online mapping system [20] is used to update the elevation grid map $\mathcal{M}_{2.5D}$. Hence, the model is implemented as a ROS node. The updated online grid is continuously fed to the proposed system along with the path through the environment, which the system uses to predict the robot elevation change. The achieved performance of the system

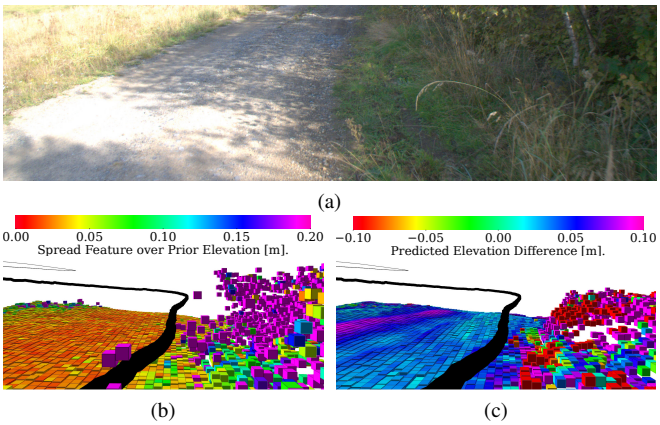


Fig. 6. Predictions over (a) a flat road. (b) The spread feature (high values in purple) and the future robot path (in black). (c) Predicted elevation difference (jet color scheme, negative change in red, position change in blue/purple, no change in green). No current tracks are visible along the robot's path through the rigid, flat ground.

is operation at 0.77 Hz using a workstation with the Intel i7-9700 4.7 GHz CPU, 32 GB memory, and NVIDIA GeForce RTX 3060 with 12 GB memory.

The system operation over tall grass and road is depicted in Fig. 5 and Fig. 6, respectively. Besides, the operation is also shown in the Supplementary Video. The system predicts tracks to be imprinted in the tall grass but not on the rigid road, which is the expected behavior. Interestingly, for some flat areas where tracks are not imprinted, the system predicts that there should be a slight global elevation increase. The possible cause is discussed in the following paragraphs, where we also summarize the system's capabilities.

F. Discussion

The presented evaluation results demonstrate a system capable of learning the changes induced on terrain elevation by robot traversal. The capabilities of the system are demonstrated within an environment of flat, rigid roads and tall grass. Hence, we have elected to use a terrain descriptor suitable for describing the depth of the vegetation as one of the inputs of the network. In different environments, such as mud or forest, an operator should consider an alternative feature suitable for such environment, e.g., one capable of discriminating dead and living vegetation based on appearance in the near-red spectra.

The reported RMSE values exhibit relatively high variance. In the authors' opinion, the variance is linked to high prevalance of vegetation, which geometry might differ somewhat in the prior and GT posterior views in Figs. 7a and 7b, respectively. Besides, the system exhibits slight bias for predicting positive elevation in areas where it does not imprint robot tracks. In the authors' opinion, the behavior is rooted in the existence of such bias in the learning set; see Fig. 7. Two particular biases can be observed. First, the overall prior and posterior elevation is offset slightly in some areas. It is likely caused by an elevation or pitch error in the localization process since the prior and elevation grid must be observed from different locations on the map as the robot

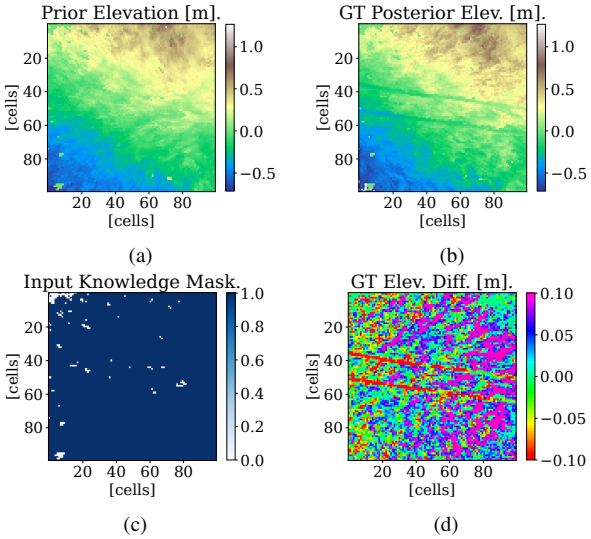


Fig. 7. Possible bias in the learning set. An example of (a) prior elevation, (b) ground truth elevation, (c) input knowledge-mask, and (d) ground truth elevation difference, which might be biased towards an overall positive change. Given the network architecture, the effect of footprints is limited to 17 cell windows around the projected footprints.

moves. Second, the model learns a particular transformation for such areas since the elevation is normalized so that unknown areas are set to the zero-mean elevation. However, in the authors' opinion, the first bias is dominant since using an additional input knowledge-mask $\kappa(\nu) = 1[z(\nu) \neq \varnothing]$ has a negligible effect on the prediction error as demonstrated in Table V.

TABLE V

INFLUENCE OF USING PRIOR ELEVATION KNOWLEDGE MASK

Inputs	$z + \Delta + w$	$z + \Delta + w + \kappa$
RMSE	$\mathbf{4.98} \pm 0.13$	5.07 ± 0.23

$[\mathrm{m} \times 10^{-2}]$, $\mu \pm \sigma$ of 9 learning runs.

VI. CONCLUSION

In this paper, we report on the study of inference of terrain change induced by robot traversal over compliant terrain. The proposed model comprises a module that projects wheel footprints along the planned robot path, coupled with normalized local elevation and terrain depth descriptor. These inputs are then passed to a convolutional neural network that predicts the elevation change along the robot path and, thus, the elevation after the robot traversal. The proposed model is verified on a dataset captured using a six-wheeled robot traversing flat roads and vegetated fields, where tires might be imprinted. The system can discriminate terrains where changes should be imprinted, and its operation is demonstrated in an online setup as part of the ROS ecosystem. Besides, we demonstrate that the system can be reparametrized to predict changes induced by a vehicle with different footprints while using one pre-learned model. Future extensions of the approach can be in using additional terrain descriptors suitable for various terrain types and extending the projected robot parametrization by adding per-wheel robot mass and tire size.

ACKNOWLEDGEMENTS

We would like to acknowledge Roman Adámek and Jan Nohel for the help with data collection and Jan Mazal for providing access to the experimental platform and area.

REFERENCES

[1] J. P. Gray and V. V. Vantsevich, "An Estimation of multi-wheel drive vehicle convoy mobility in stochastic terrain condition," in *International Conference of the International Society for Terrain-Vehicle Systems (ISTVS)*, 2014, pp. 1–12.

[2] P. Papadakis, "Terrain traversability analysis methods for unmanned ground vehicles: A Survey," *Engineering Applications of Artificial Intelligence*, vol. 26, no. 4, pp. 1373–1385, 2013.

[3] N. Kottege, C. Parkinson, P. Moghadam, A. Elfes, and S. P. N. Singh, "Energetics-informed hexapod gait transitions across terrains," in *IEEE International Conference on Robotics and Automation (ICRA)*, 2015, pp. 5140–5147.

[4] K. Otsu, M. Ono, T. J. Fuchs, I. Baldwin, and T. Kubota, "Autonomous terrain classification with co- and self-training approach," *Robotics and Automation Letters*, vol. 1, no. 2, pp. 1–6, 2016.

[5] M. A. Bekhti, "Traversability cost prediction of outdoor terrains for mobile robot using image features," Ph.D. dissertation, Shizuoka University, Sept. 2020.

[6] B. Sofman, E. Lin, J. A. Bagnell, J. Cole, N. Vandapel, and A. Stentz, "Improving robot navigation through self-supervised online learning," *Journal of Field Robotics*, vol. 23, no. 11-12, pp. 1059–1075, 2006.

[7] M. Wermelinger, P. Fankhauser, R. Diethelm, P. Krüsi, R. Siegwart, and M. Hutter, "Navigation planning for legged robots in challenging terrain," in *IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, 2016, pp. 1184–1189.

[8] P. Krüsi, M. Bosse, and R. Siegwart, "Driving on point clouds: Motion planning, trajectory optimization, and terrain assessment in generic nonplanar environments," *Journal of Field Robotics*, vol. 34, no. 5, pp. 940–984, 2016.

[9] T. Homberger, M. Bjelonic, N. Kottege, and P. V. K. Borges, "Terrain-dependant control of hexapod robots using vision," in *International Symposium on Experimental Robotics (ISER)*, 2016, pp. 92–102.

[10] A. Shaban, X. Meng, J. Lee, B. Boots, and D. Fox, "Semantic terrain classification for off-road autonomous driving," in *Conference on Robot Learning*, vol. 164, 2022, pp. 619–629.

[11] M. Prágr, J. Bayer, and J. Faigl, "Autonomous exploration with online learning of traversable yet visually rigid obstacles," *Autonomous Robots*, vol. 47, pp. 161–180, 2023.

[12] J. Baleia, P. Santana, and J. Barata, "On exploiting haptic cues for self-supervised learning of depth-based robot navigation affordances," *Journal of Intelligent & Robotic Systems*, vol. 80, no. 3–4, pp. 455–474, 2015.

[13] M. Eder, R. Prinz, F. Schöggl, and G. Steinbauer-Wagner, "Traversability analysis for off-road environments using locomotion experiments and earth observation data," *Robotics and Autonomous Systems*, vol. 168, p. 104494, 2023.

[14] C. Ünsalan and K. L. Boyer, "Linearized vegetation indices based on a formal statistical framework," *Transactions on Geoscience and Remote Sensing*, vol. 42, no. 7, pp. 1575–1585, 2004.

[15] C. Cunningham, I. A. Nesnas, and W. L. Whittaker, "Improving slip prediction on Mars using thermal inertia measurements," *Autonomous Robots*, vol. 43, no. 2, pp. 503–521, 2019.

[16] J. Dallas, K. Jain, Z. Dong, L. Sapronov, M. P. Cole, P. Jayakumar, and T. Ersal, "Online terrain estimation for autonomous vehicles on deformable terrains," *Journal of Terramechanics*, vol. 91, pp. 11–22, 2020.

[17] J. Dallas, M. P. Cole, P. Jayakumar, and T. Ersal, "Terrain adaptive trajectory planning and tracking on deformable terrains," *IEEE Transactions on Vehicular Technology*, vol. 70, no. 11, pp. 11 255–11 268, 2021.

[18] D. P. Kingma and J. Ba, "Adam: A Method for stochastic optimization," *arXiv*, p. 1412.6980, 2017.

[19] M. Quigley, K. Conley, B. P. Gerkey, J. Faust, T. Foote, J. Leibs, R. Wheeler, and A. Y. Ng, "ROS: An open-source Robot Operating System," in *ICRA Workshop on Open Source Software*, 2009, pp. 1–6.

[20] J. Bayer and J. Faigl, "Speeded up elevation map for exploration of large-scale subterranean environments," in *2019 Modelling and Simulation for Autonomous Systems (MESAS)*, 2020, pp. 190–202.