www.nature.com/scientificreports

scientific reports

Self-reconfigurable robot vision pipeline for safer adaptation to varying pavements width and surface conditions

Lim Yi[1], Braulio Félix Gómez[1], Balakrishnan Ramalingam[1]✉, Madan Mohan Rayguru[2], Mohan Rajesh Elara[1] & Abdullah Aamir Hayat[1]

This work presents the vision pipeline for our in-house developed autonomous reconfigurable pavement sweeping robot named Panthera. As the goal of Panthera is to be an autonomous self-reconfigurable robot, it has to understand the type of pavement it is moving in so that it can adapt smoothly to changing pavement width and perform cleaning operations more efficiently and safely. deep learning (DL) based vision pipeline is proposed for the Panthera robot to recognize pavement features, including pavement type identification, pavement surface condition prediction, and pavement width estimation. The DeepLabv3+ semantic segmentation algorithm was customized to identify the pavement type classification, an eight-layer CNN was proposed for pavement surface condition prediction. Furthermore, pavement width estimation was computed by fusing the segmented pavement region on the depth map. In the end, the fuzzy inference system was implemented by taking input as the pavement width and its conditions detected and output as the safe operational speed. The vision pipeline was trained using the DL provided with the custom pavement images dataset. The performance was evaluated using offline test and real-time field trial images captured through the reconfigurable robot Panthera stereo vision sensor. In the experimental analysis, the DL-based vision pipeline components scored 88.02% and 93.22% accuracy for pavement segmentation and pavement surface condition assessment, respectively, and took approximately 10 ms computation time to process the single image frame from the vision sensor using the onboard computer.

With urbanization, more infrastructure and pavements will be developed around the world. Due to this increase in pavements, more types of pavement are developed. These pavements generally vary in their properties, like surface condition, texture, and wet or waterlogged material, among others. In addition to the increasing number of pavement types, more pavements are required to be maintained daily to ensure a hygienic environment for social activities to occur. In Singapore, 200 km of sheltered pavements have recently been built, and it is expected to have more in the future[1]. A lot of manual labor and resources go into keeping these pavements clean for usage. With the rise of Industry 4.0 and automation, robots are researched extensively to reduce the amount of manual labor required for repetitive tasks, including the pavement cleaning industry.

In recent years many autonomous robots[2,3] and cleaning vehicles[4] were proposed for sweeping pavement tasks. However, those platforms have a lot of limitations, and is inefficient in covering the width of pavement, and are hard to use in the narrow pavements region. As a result, limited efficiency is achieved during the pavement cleaning tasks. Reconfigurable robots[5] are becoming a viable alternative for fixed morphology robots. These robots are developed with an inherent capability to autonomously change their kinematics to overcome difficulties in handling a given task and traversing the environment. Purposes of shape reconfiguration varies such as to perform another motion such as climbing[6], rolling[7], flying[8] and floating[9]. By considering the advantage of the reconfigurable mechanism, the pavement cleaning robot Panthera[10,11], as seen in Fig. 1, was developed with the reconfigurable mechanism. The robot moves through pavements with dynamically changing widths through a reconfigurable mechanism and performs efficient cleaning. To push the development of reconfigurable robot

[1]Engineering Product Development Pillar, Singapore University of Technology and Design (SUTD), Singapore 487372, Singapore. [2]Electrical Engineering, Delhi Technological University, New Delhi 110042, India. ✉email: balakrishnan@sutd.edu.sg

Related work	Algorithm	Application	Advantage	Limitation
Liang et al.[29]	D-UNet	Road surface condition prediction	Highest classification performance compared to ML algorithm	Due to pre-trained CNN classification, capacity is limited
Marcus et al.[30]	ResNet50 and InceptionV3	Road friction estimation	Classify six types of surface	Misclassification of wet asphalt and dirt as asphalt
Suryamurthy et al.[32]	Segnet	Terrain segmentation and roughness estimation	Real-time	Bias between flat surface and smooth boundaries

Table 1. Summary of related work.

Panthera towards self-configurable capability and safe autonomous locomotion with the instantaneous center of rotation[12], the kinematic control part of Panthera[13] needs a high-level vision pipeline that should have the capability to provide the information about changing pavement width, pavement types, and condition of the surface.

To our best of knowledge, Deep learning (DL) based vision pipeline for reconfigurable pavement cleaning robots is a new approach as it is not widely studied yet. This work proposed a DL-enabled vision pipeline for the Panthera robot, which provides essential information such as pavement type, pavement condition, and varying pavement width information for safe and efficient operation. The vision pipeline is built with Deep Convolutional Neural Networks (DCNN) based semantic segmentation algorithm and image classifier module. The DCNN modules are trained with large-sized pavement image datasets, collected from the pavements in Singapore and tested with offline and real-time environments.

Related work.

This section describe the existing study related to our work and its short summery is given in Table 1. Machine learning (ML)[14] and deep learning (DL)[15] are emerging techniques which are widely used for many indoor and outdoor robots[16], autonomous vehicle perceptual system design in recent years. These perceptual system vision pipelines data are widely used for[17], path planning[18], controlling , safe navigation[19] and efficient operation[20]. In ML techniques, support vector machine (SVM), K-nearest neighbors[21], Bayes classifier, and neural network (NN) are commonly used algorithms and perform surface condition assessment or classify the surface type from images or various sensors data.

In Khan et al. work[22], the authors proposed a terrain classification algorithm for mobile robot applications. The features extracted from the mobile robot collected images are used to classify terrain type using the Random Forest (RF) algorithm and scored 99.2% terrain classification accuracy. Omer et al.[23] investigate the feasibility of classifying winter road surface conditions as bare road, snowy road, and tracks. The authors use the support vector machine (SVM) algorithm, trained using 400 images each class collected through a vision system mounted on regular vehicles and have a classification accuracy of over 80%. Kawai et al.[24] propose a distinction method for road surface conditions at night. The author uses the differences in image features of dry, wet, and snowy roads under different light sources and combines the three features with color, brightness, and texture of road to classify the road surface condition. K-nearest neighbor algorithm is used for classification and reported 96.1%, 89.4%, and 95.6% classification accuracy, respectively.

In contrast with the ML technique, the DL scheme has a lot of advantages in perceptual system design, which has automatically extracted and learned the features from the bulk image datasets and performs classification and detection accuracy better than ML techniques. In literature, there are many DL based image classification framework such as ResNet[25], SqueezeNet[26], MobileNet and VGG16[27] which were trained and used for autonomous vehicle road pavement classification and condition detection task. In[28] Ramon et al. assess Neural Network's (NN) performance, ML, and DL for terrain classification and slip estimation. The authors performed the terrain classification estimation in the mobile robot Fitorobot and reported that deep learning models are optimal for solving terrain and ground robotics problems. In another study, Liang et al.[29] used the road surface status recognition system using a deep semantic segmentation framework. The author uses the D-UNet encoder–decoder framework for detecting the slippery road statuses caused by water, ice, and snow in the wintertime. Deep Convolutional Neural NetworK-based road friction estimation was proposed by Marcus et al.[30]. Here, the author trained and evaluated two pre-trained models, ResNet50 and InceptionV3[31], and reported that ResNet50 outperforms InceptionV3 for road friction estimation and classification tasks. Suryamurthy et al.[32] adopt the deep convolutional encoder–decoder framework in CENTAURO Robot for safe reconfiguration of leg joints and path planning application where the semantic segmentation framework "SegNet" was trained for terrain segmentation and roughness estimation task and obtained 64% classification accuracy.

The previous work[33], an RGB-D camera is used to estimate the vision feedback algorithm parameters for Panthera locomotion and reconfiguration using VGG16 semantic segmentation. Semantic segmentation used in[33] was unable to classify different pavement types in Singapore accurately. The parameters derived through the method directly go into Panthera kinematic control[13] and do not take into account the pavement types, conditions, and magnitude of the vision feedback algorithm parameters for safety. As pre-trained model VGG-16 pavement segmentation cannot perform accurately in Singapore pavements, safety of pavement users and the self-reconfigurable robot is compromised as it might make the robot move into non-pavements. On top of that, there is little or no work that is performed on a speed regulation on a self-reconfigurable robot based on pavement width changes and pavement conditions.

Taking into account the above facts, the following objectives of the present paper as:

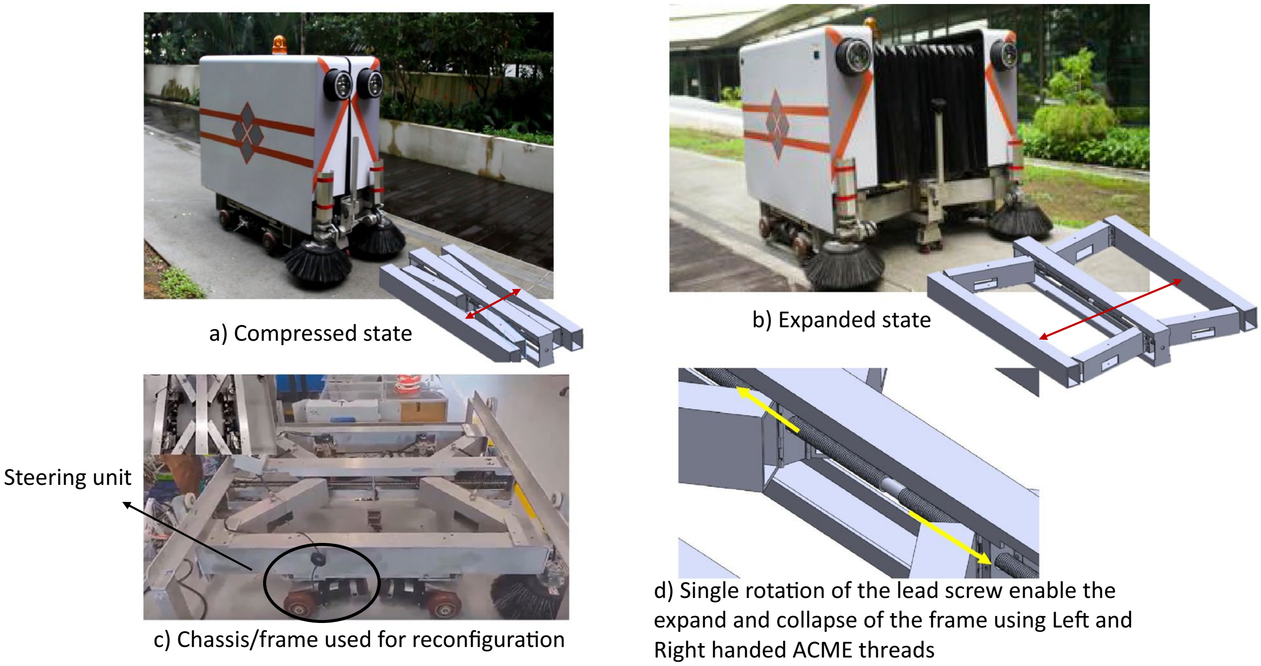

a) Compressed state
b) Expanded state
Steering unit
c) Chassis/frame used for reconfiguration
d) Single rotation of the lead screw enable the expand and collapse of the frame using Left and Right handed ACME threads

Figure 1. Self-reconfigurable pavement sweeping robot Panthera.

- Vision pipeline for semantic segmentation with pavement type classification (concrete, road, and paver block) and condition (bad, moderate, and good) and estimation of pavement width in Singapore.
- A fuzzy-based controller for safer adaptation based on pavement type condition classification and vision feedback control parameters.

The remaining of this paper is structured as follows. In "Panthera overview" section describe the proposed Panthera system and overview. In "Vision pipeline" section discuss the vision pipeline and its modules. In "Fuzzy inference system" section Section will discuss the fuzzy logic control system part. In "Experimental results" section will discuss the experimental results. In "Conclusions" section concludes the paper.

Panthera overview

Figure 1 shows the overview of the reconfigurable pavement sweeping robot Panthera. It was designed to reconfigure in shape during locomotion to adapt smoothly with respect to changing pavement widths[10,33–35]. The taxonomy of reconfigurable systems and its classification along with the sensor fusion are detailed in the respective works[5,36]. The detail of the mechanical system overview and control system architecture is briefly described next.

Mechanical overview. The Panthera core frame is made of an aluminum scissors mechanism, and the entire body is supported by four steering units attached to the aluminum frames. These four steering units are independent differential drives and consist of two wheels each. In total, Panthera has eight wheels with eight-wheel motors for locomotion. The mechanism responsible for the reconfiguration is the scissors mechanism that connects the central beam to two side beams. A double-threaded lead screw is connected to a motor which will drive the scissors mechanism to move as seen in Fig. 1d. The movement of the scissors mechanism enables the robot to expand and contract. As the steering units are connected to the aluminum scissors mechanism side beam as seen in 1c, the lead screw motor and steering units have to work in synchronization so that reconfiguration can be performed smoothly. Panthera core frame supports the robot's electronics, including the batteries, micro-controllers, relays, perspective sensors, and the industrial computer. The hardware and sensor components are sheltered by Panthera's two external aluminum covers, which are mounted on the two side beams of Panthera. The two external aluminum covers are connected by an artificial leather bellow, water resistance, and protect the internal electronics components from water and other foreign objects. The kinematics of Panthera can be found in our previous work[33].

Control system architecture. The Panthera central control system is built on an industrial computer with the operating system Ubuntu 16.04 and uses the middle-ware robot operating system (ROS) Kinetic version. ROS is capable of parallel information transfer between ROS nodes. It is used to publish and subscribe data within Panthera functional components. Figure 2 shows the hardware components of control system architecture. The industrial computer has a GPU, 8CPU core, and 16 GB RAM, performing high-level tasks including running ROS master, vision pipeline task, etc. The 24-V traction battery is used in the Panthera. The 24 V battery unit powers all of Panthera's motors, including eight-wheel motors, one lead screw motor, two brush motors, and a vacuum motor. It also powers the sensors and other low-level controllers in Panthera. The Panthera has been

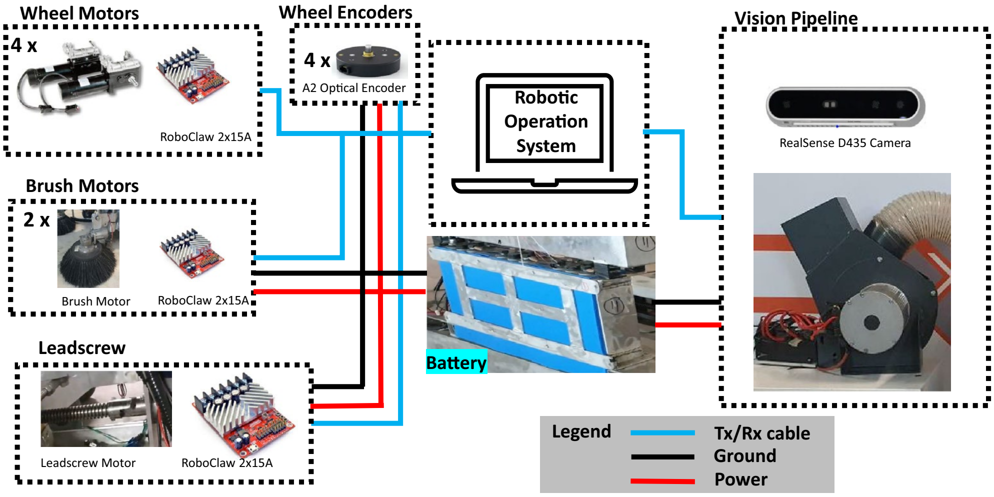


Figure 2. Hardware components and electrical layout.

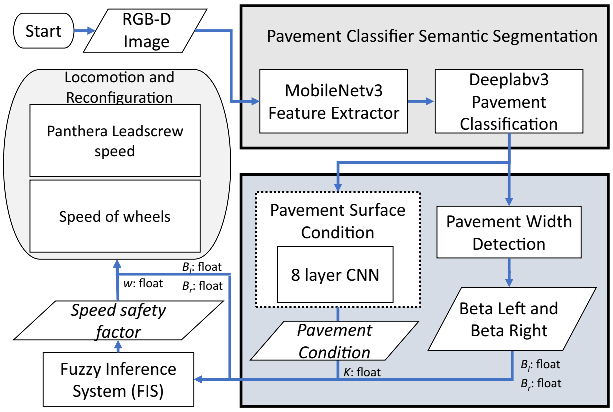


Figure 3. Vision pipeline.

built with three key sensors for safe locomotion operation: Digital Absolute Encoders, US Digital Incremental Encoders, and the RealSense D435 camera. The RealSense D435 camera is mounted at 95 cm in front of Panthera. The RealSense D435 RGB-D sensor (called perspective nodes) connects with the industrial PC through a USB 3.0 interface. It publishes RGB images and depth images in two topics as a sensor_msgs/Image message and is subscribed by the vision processing node. After subscribing to the information from the perspective nodes, the vision pipeline node will execute the pavement segmentation, pavement classification, surface condition prediction, and pavement width detection task.

Based on the type of pavement classified, surface condition predicted, and pavement width detection, the processing node will then publish three parameters: Pavement classification K factor, beta left, and beta right in a topic. These three parameters will be subscribed by the locomotion and reconfiguration node via a geometry_msgs/Twist message. Furthermore, the US Digital Absolute Encoders provide steering angle feedback for steering angle, while the US Digital Incremental Encoder provides the velocity feedback for the wheel's speed. Both steering angle feedback and velocity feedback are important for Panthera's control during locomotion and reconfiguration. The encoder module published geometry_msgs/Twist message and locomotion and reconfiguration node subscribed this topic for locomotion and reconfiguration operation. The inverse kinematics will determine the speed of the leadscrew motor and the wheel motors for reconfiguration during locomotion.

Vision pipeline

The vision pipeline module executes three tasks: pavement type classification, pavement surface condition prediction, and pavement width estimation. The Fig. 3 shows the functional component of the vision pipeline. It comprises a DeepLabv3+ semantic segmentation framework, eight-layer CNN for pavement surface condition prediction, and pavement width estimation module. Here, the DeepLabv3+ is a critical component of the vision pipeline. The other two modules are built on top of the DeepLabv3+ framework. The vision pipeline ends with the output of the speed safety factor and the reconfiguration parameters[33] for Panthera locomotion. The details of each component and its integration are described in the subsections.

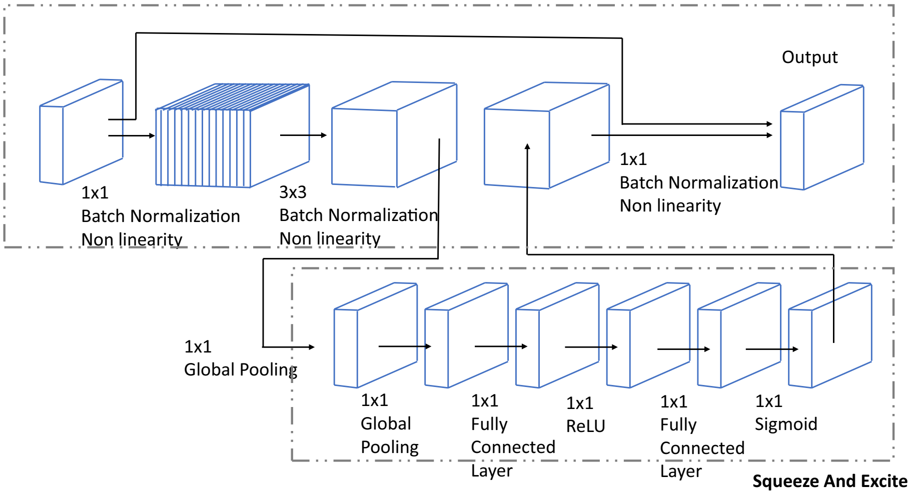


Figure 4. DeepLabv3+ semantic segmentation architecture.

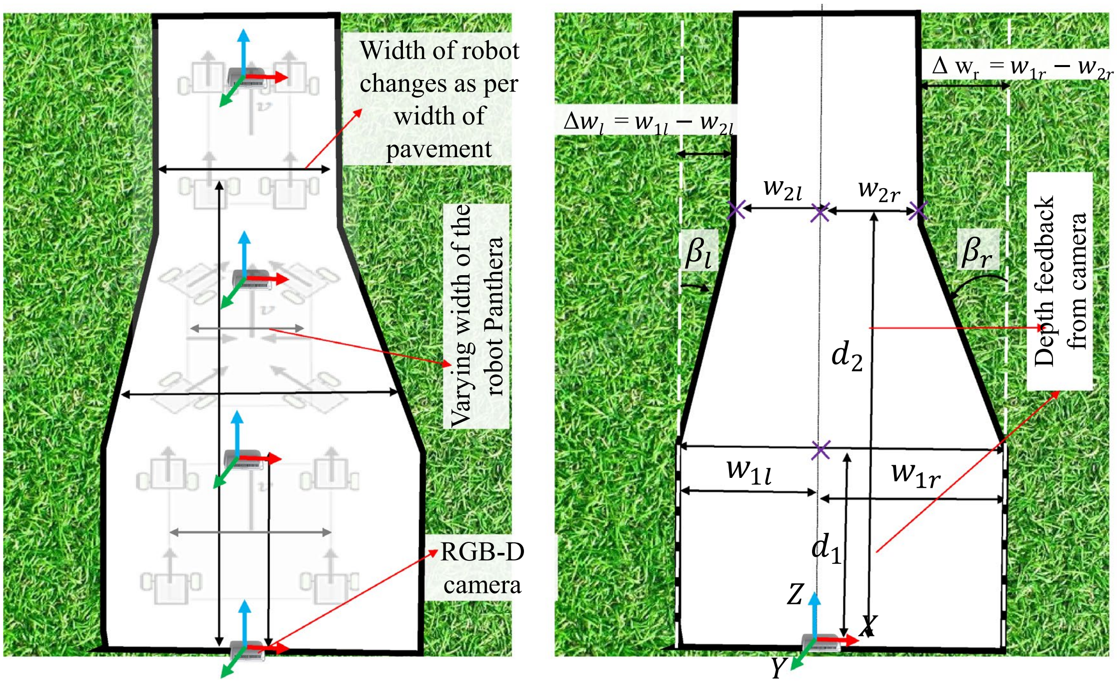
C Input	Operator	exp size	#out	Squeeze-And-Excite	Non-Linearity	Stride
$224^2 \times 3$	conv2d	–	16	No	h-swish	2
$112^2 \times 16$	bneck, 3×3	16	16	No	ReLU	1
$112^2 \times 16$	bneck, 3×3	64	24	No	ReLU	2
$56^2 \times 24$	bneck, 3×3	72	24	No	ReLU	1
$56^2 \times 24$	bneck, 5×5	72	24	Yes	ReLU	2
$28^2 \times 40$	bneck, 5×5	120	40	Yes	ReLU	1
$28^2 \times 40$	bneck, 5×5	120	40	Yes	ReLU	1
$28^2 \times 40$	bneck, 3×3	240	80	No	h-swish	2
$14^2 \times 80$	bneck, 3×3	200	80	No	h-swish	1
$14^2 \times 80$	bneck, 3×3	184	80	No	h-swish	1
$14^2 \times 80$	bneck, 3×3	184	80	No	h-swish	1
$14^2 \times 80$	bneck, 3×3	480	112	Yes	h-swish	1
$14^2 \times 112$	bneck, 3×3	672	112	Yes	h-swish	1
$14^2 \times 112$	bneck, 5×5	672	160	Yes	h-swish	2
$7^2 \times 160$	bneck, 5×5	960	160	Yes	h-swish	1
$7^2 \times 160$	bneck, 5×5	960	160	Yes	h-swish	1
$7^2 \times 160$	conv2d, 1×1	–	960	No	h-swish	1
$7^2 \times 960$	pool, 7×7	–	–	No	–	1
$1^2 \times 960$	conv2d, 1×1, No Batch Normalisation	–	1280	No	h-swish	1
$1^2 \times 1280$	conv2d, 1×1, No Batch Normalisation	–	k	No	–	1

Table 2. Specifications of MobileNetV3-Large.

DeepLabv3+.

Figure 4 shows the overview of the DeepLabv3+ semantic segmentation architecture. It comprised of the encoder–decoder function where the encoder function generates the feature map from the input images and decoder function gradually predict the object detail and spatial dimension of the objects.

Encoder. In DeepLabv3+, the encoder part consists of the backbone network, atrous separable convolution function, and atrous Spatial Pyramid Pooling (ASPP). In this work, MobileNetv3+ is configured as a backbone network. The layer detail of MobileNetV3+ is given in Table 2. In DeepLabv3+, the last convolution layer of MobileNetV3+ is replaced by an atrous separable convolution function to obtain the enlarged receptive field. Then, Atrous Spatial Pyramid Pooling (ASPP) function is applied on generated feature map, which applies four parallel convolution operations in feature map, including 1 × 1 convolution and three 3 × 3 convolutions dilation rates (6,12,18). In addition, ASPP applies Global Average Pooling (GAP) to the output features, a map from the last atrous block to obtained image-level features. In the end, the elements from all the branches are com-


Width of robot changes as per width of pavement
Varying width of the robot Panthera
RGB-D camera
$\Delta w_r = w_{1r} - w_{2r}$
$\Delta w_l = w_{1l} - w_{2l}$
w_{2l}
w_{2r}
β_l
β_r
Depth feedback from camera
d_2
w_{1l}
w_{1r}
d_1
Z
X
Y

a) Varying pavement width and Panthera reconfiguring ability

b) Top down view of the pavement and camera frame estimating width slopes on left and right side as β_l and β_r

Figure 5. Pavement width estimation using vision feedback for reconfiguration parameters.

bined into a single vector via concatenation. This output is then convoluted with another 1 × 1 kernel—using Batch Normalization (BN) and 256 filters.

Decoder. The decoder module applies the upsampling function to retain the details (specifically object boundaries) from the low dimension feature map. In the decoder side, the multi-scale feature map (extracted from ASPP) is bilinearly upsampled by a factor of 4 and then concatenated with the corresponding low-level feature map obtained from MobileNetV3+ last convolution layer. Before concatenation, 1 × 1 convolution is applied on a low-level feature map to reduce the number of channels. After the concatenation, a few 3 × 3 convolutions are applied to refine the features, followed by another simple bilinear upsampling of 4.

Pavement surface condition prediction. The pavement surface condition is computed by CNN based classifier algorithm. It was cascaded with pavement segmentation framework through preprocessing function and took the segmented pavement region (200 × 200 × 3) as input. The classifier comprises of eight CNN layers, flatten layer, and two fully connected layers. At the end of each convolutional layer, the ReLU activation function and max pooling function are applied, where the max-pooling function reduces the dimensional of the feature map at each stage. The last three layers of the classifier frameworks are the flatten layers and two fully connected layers. The flatten layer converts the multi-dimension tensor data into a single dimension tensor and feeds into two fully connected (FC) layers. In FC layers, the ReLU activation function is applied on the first layer, and the SoftMax function is used on the second layer, which generates the probabilities output of pavement condition, *K*, which is used to drive the fuzzy inference system (FIS).

Pavement width estimation. The pavement width was estimated from the segmented pavement region. It was computed for each frame captured from the Realsense RGB-D sensor. The segmented pavement region from DeepLabv3+ is used as an input to pavement width estimation function. The width estimation function uses the left and right fences of the segmented pavement region and corresponding point cloud data to measure the width of the pavement. The distance between the leftmost and rightmost point of the pavement 3D point cloud data x-axis, y-axis, and z-axis data were used to compute the pavement width. Finally, the euclidean 3D distance function was applied to each leftmost and rightmost point to compute the width of the pavement. After estimating the pavement width, it will derive vision feedback parameters for Panthera to perform reconfiguration as seen in Fig. 5a. The derivation of the vision feedback parameters β_l and β_r and the reduction of noise due to robot vibrations through filtering can be found in the works of[37] and is visualized in Fig. 5b. The output *K*, and β_l, and β_r, which are the target heading angles of Panthera's wheels, will be passed on to the FIS to control Panthera speed safety factor, *w*.

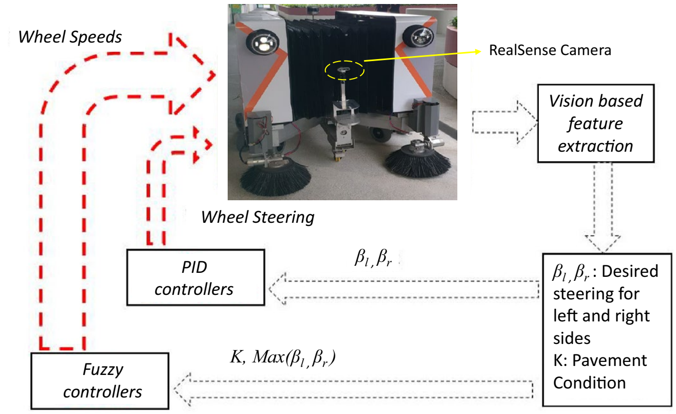

Wheel Speeds
RealSense Camera
Vision based feature extraction
Wheel Steering
PID controllers
β_l, β_r
β_l, β_r : Desired steering for left and right sides
K: Pavement Condition
Fuzzy controllers
$K, Max(\beta_l, \beta_r)$

Figure 6. Controller architecture.

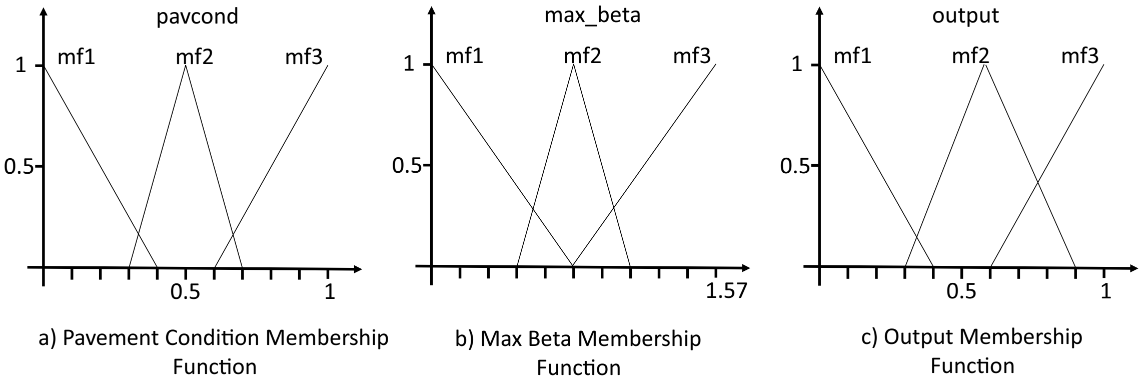


Figure 7. Membership functions.

Fuzzy inference system

Figure 6 presents the overall control architecture. Here, the fuzzy inference system (FIS) is a part of the Panthera control system. It receives the input from the vision pipeline, including pavement condition, K and locomotion and reconfiguration parameters, β_l, and β_r[37] to give the desired steering change for the wheels on the left and right side, respectively. Based on these factors, fuzzy logic is used to determine the safety factor, w, to apply to the robot speed without explicitly calculating the error variables. Given the desired headings, the conventional PID controllers control the steering. The robot's speed safety factor, w, depends on all three parameters, and a fuzzy controller is proposed to modulate the speed. The fuzzy logic inference engine works with a simple rule base and input variables, which are fuzzy in nature. It aims to exploit the fuzzy sets and fuzzy inference method to incorporate safety into the Panthera robot.

Fuzzy Controller: The controller exploits a Mamdani fuzzy inference engine. The pavement condition data received from the vision module is modeled as three Gaussian membership functions; bad (mf1), moderate (mf2), and good (mf3). The steering requirement is calculated from the pavement width data. The bigger steering requirement between β_l and β_r is taken as another input and modeled in the same manner as small, moderate, and large angles. The speed safety factor, w, is captured using a similar approach. The details of the membership functions and the set of nine rules are presented in Fig. 7 and Table 3 respectively.

Experimental results

This section describes the experimental setup procedure and outcome of the proposed vision pipeline system. The experimental setup procedure includes collecting the dataset image and labeling, training the model with labeled images, and evaluating the trained model using test images and real-time video stream.

Dataset preparation.

Panthera is developed for pavement sweeping tasks in public pavement spaces. Hence, the dataset images are collected from public pavement spaces in Singapore, including national parks, park connectors, residential parks, and school parks. The collected datasets are categorized into three classes such as concrete pavement, paver block pavement (a stone, brick, or block used for paving a surface), and roads to train the segmentation model. Furthermore, to train the pavement surface condition algorithm, the dataset were labelled into three category which include bad, moderate, and good as shown in Figs. 26, 27 and 28

Rule	pavcond	max_beta	speed safety factor
1	Bad	High	Slow
2	Bad	Medium	Slow
3	Bad	Small	Slow
4	Moderate	High	Slow
5	Moderate	Medium	Normal
6	Moderate	Small	Normal
7	Good	Small	High
8	Good	High	Normal
9	Good	Medium	High

Table 3. Fuzzy rules.

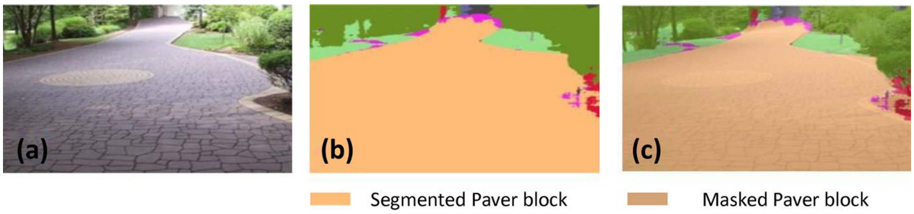


Figure 8. Offline results: paver block.

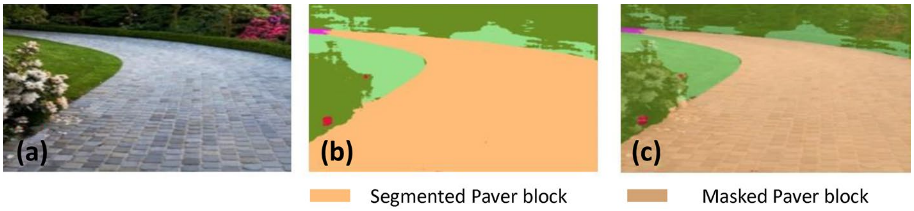


Figure 9. Offline results: paver block.

respectively. Intel RealSense D435 stereo vision sensor was used to collect dataset images. It was mounted on a bicycle to collect the images from different public pavement spaces at 95 cm height to get the same field of view as Panthera. After collecting the initial set of images of the three pavement types, the data augmentation process is applied to the ordered dataset, which involves adjusting the image orientation, varying brightness, adjusting the scale of the images, etc. This process will help control the over-fitting issue and make the model more robust. Then, the labeled image data were randomly divided into the train and test dataset according to the hierarchical sampling method.

Training. Transfer learning techniques were adopted to train the DeepLabv3+ pavement segmentation model, where cityscapes pre-trained weights files were used to fine-tune the model. The framework was trained with 2000 images in each class and used the following training parameter learning rate: 0.001, weight decay: 0.0005, momentum: 0.9, and batch size: 10. The eight-layer CNN framework was trained from scratch with Adam optimizer and used the segmented pavement region as a training dataset. The weight was initialized with the uniform method and applied the learning rate: 0.001, weight decay: 0.0005, momentum: 0.9, and batch size: 64. Both models were trained on NVIDIA RTX 3080 graphic processing unit enabled workstation and trained on GPU mode.

Offline and real-time evaluation test. After training, the segmentation and detection framework performance was assessed through test images and real-time video streams. Totally 200 images were used for each class in test image dataset and its segmentation and classification accuracy was estimated through statistical measure parameter. To carry out the evaluation test, the trained model was loaded into NVIDIA RTX 3080 graphic processing unit enabled workstation and tested with collected pavement images. Figures 8, 9, 10, 11, 12, 13, 14, 15 and 16 shows the segmentation and classification results of test image dataset. Figures 17, 18, 19, 20, 21, 22, 23, 24 and 25 shows the online experimental results of pavement segmentation model. In all the Figures, (a) represents the Red Green Blue (RGB) image, (b) represents the segmented image and (c) represents the segmented masked image. The color schemed for paver block, concrete and road in (b) and (c) is consist-

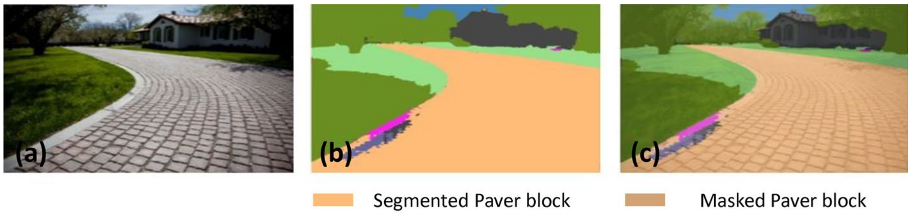


Figure 10. Offline results: paver block.

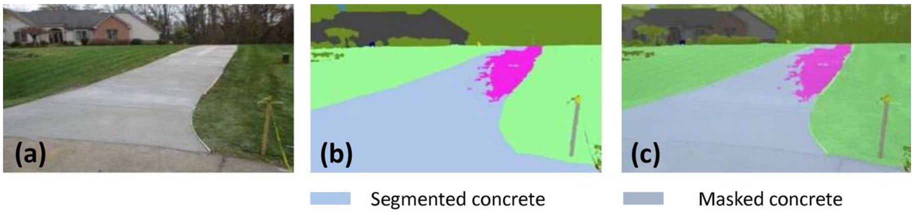


Figure 11. Offline results: concrete.

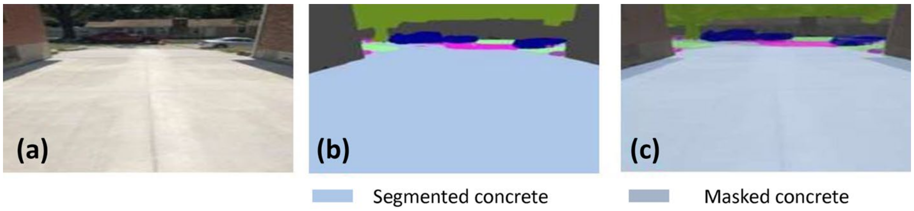


Figure 12. Offline results: concrete.

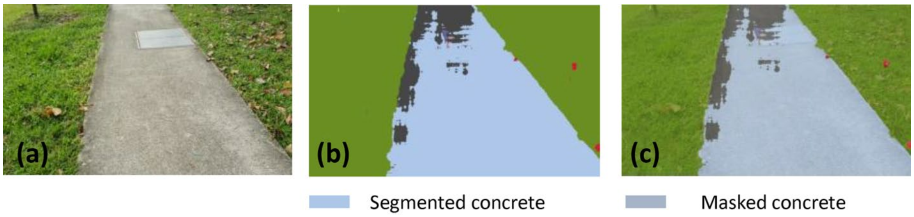


Figure 13. Offline results: concrete.

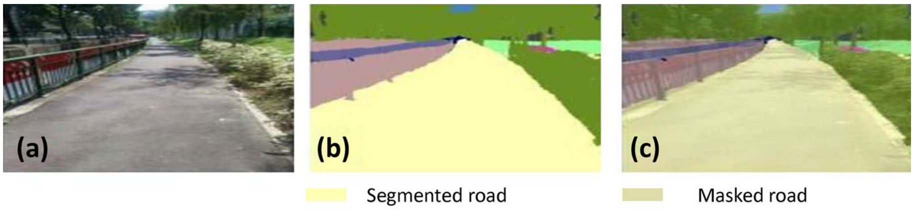


Figure 14. Offline results: road.

ent throughout the paper. Table 4 indicates the performance analysis report for offline and online test. In real time field trial , concrete, paver block and road pavement video streams are taken as the online input data to the model. In each class, a 100 m pavement was captured from Panthera robot perspective using RealSense D435 Camera. The image resolution of the RealSense D435 used is 640 × 480 where the processing node runs at about 20 frames per second.

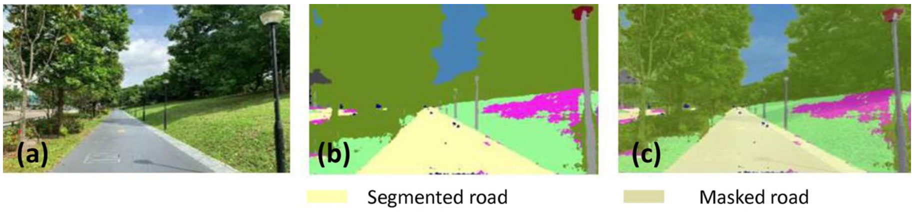


Figure 15. Offline results: road.

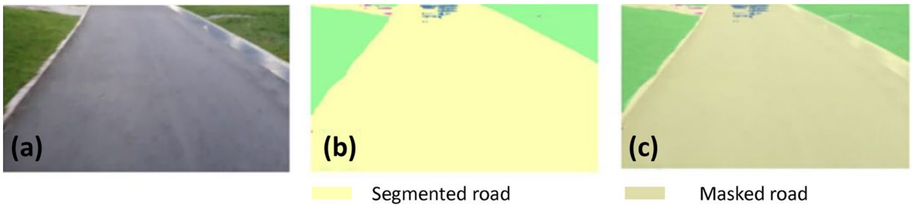


Figure 16. Offline results: road.

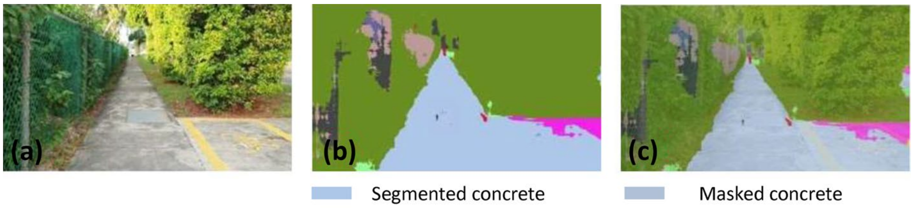


Figure 17. Online results: concrete.

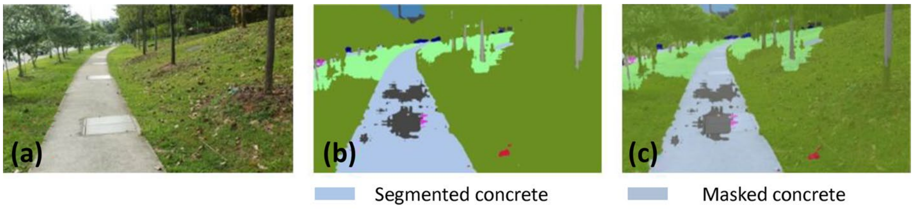


Figure 18. Online results: concrete.

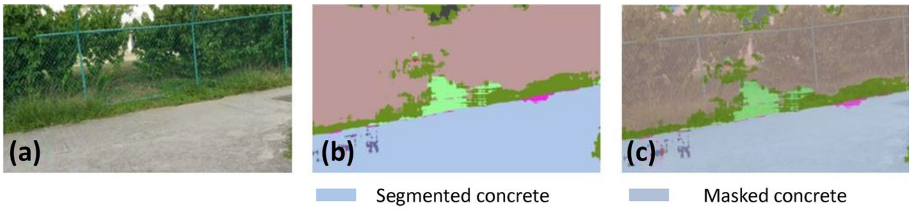


Figure 19. Online results: concrete.

The experiment results indicate that the segmentation algorithm accurately segmented the pavement and its boundary region and scored an average of 89.25% pixel classification accuracy for the test image dataset and 86.79% pixel classification accuracy for real-time collected pavement streaming video input which took approximately 10 ms to segment and classify each image. This analysis shows that the pixel classification accuracy for a real-time field trial is lower than the test image dataset. It is due to various environmental factors, such as jerks in locomotion, shadows, lighting conditions, etc.

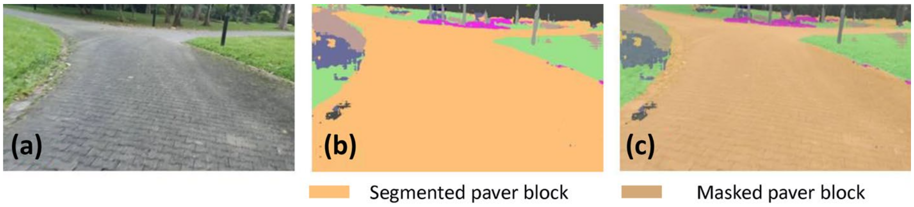


Figure 20. Online results: paver block.

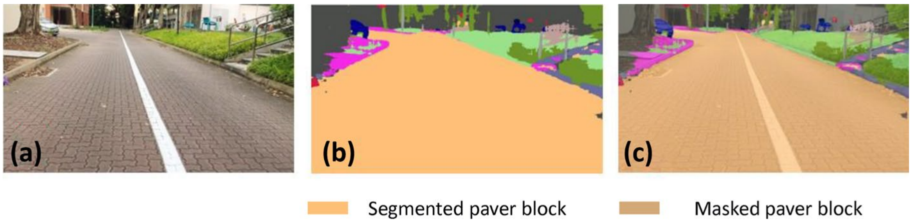


Figure 21. Online results: paver block.

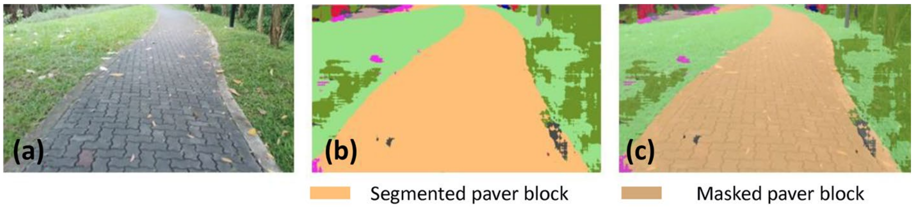


Figure 22. Online results: paver block.

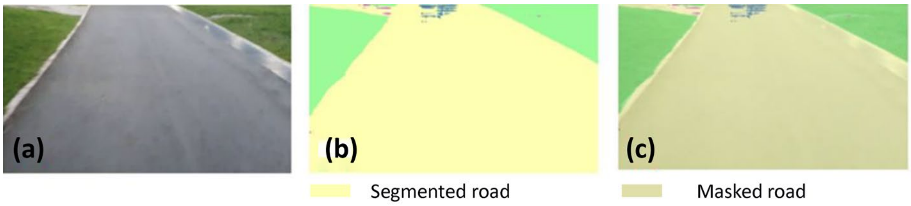


Figure 23. Online results: road.

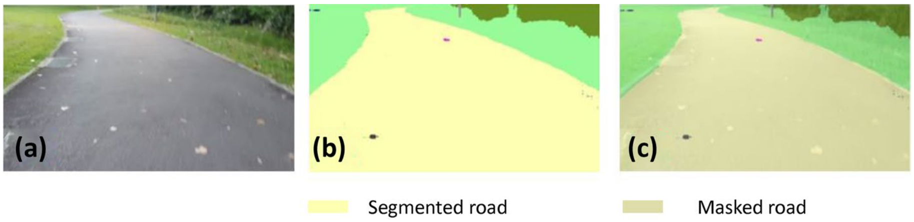


Figure 24. Online results: road.

Pavement surface condition and width estimation. The Fig. 26, 27 and 28 shows the pavement surface condition prediction result computed from segmented pavement region and Table 5 shows the statistical measures results for pavement surface condition model computed through confusion matrix parameters. In this experimental analysis, it observed that the surface condition classification model obtained average classification accuracy of 92.93%, and its prediction confidence score range is 87–94% respectively. Moderate class precision

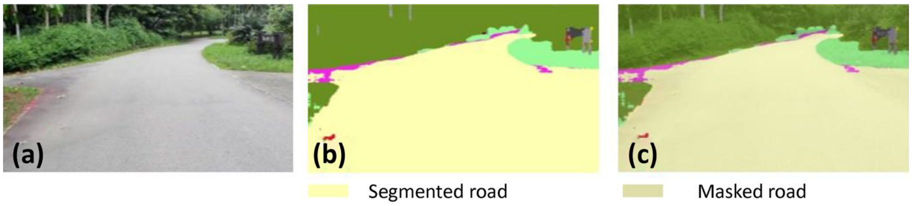


Figure 25. Online results: road.

Test model	Pixel accuracy (Average)	mean-IOU	Dice score
Offline test	89.25	88.22	88.27
Online test	86.79	85.54	86.12

Table 4. Performance analysis for pavement segmentation.

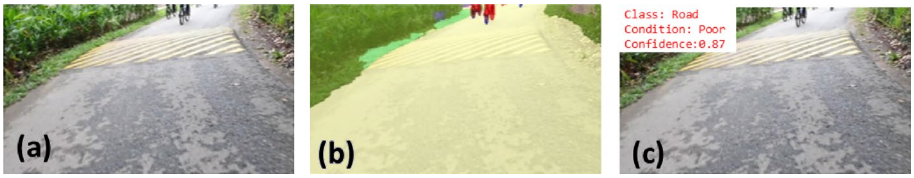


Figure 26. Pavement condition: bad.

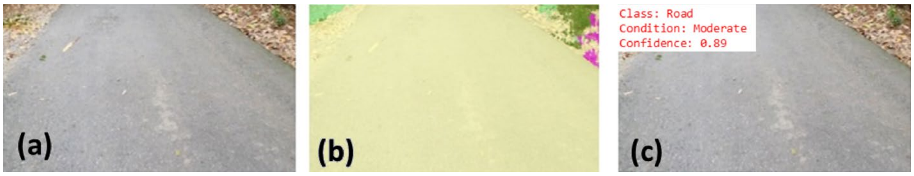


Figure 27. Pavement condition: moderate.

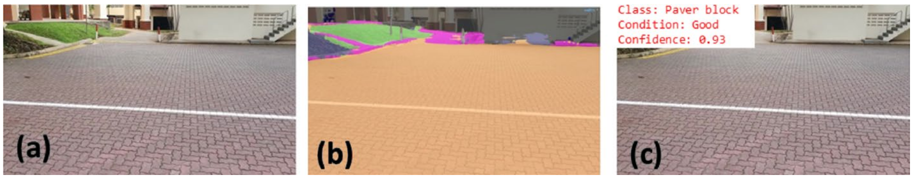


Figure 28. Pavement condition: good.

Class	Precision	Recall	F1	Accuracy
Good	93.44	92.89	92.65	93.17
Moderate	92.12	90.31	93.23	91.89
Poor	94.12	93.32	92.97	93.74

Table 5. Statistical measures analysis for pavement's surface condition.

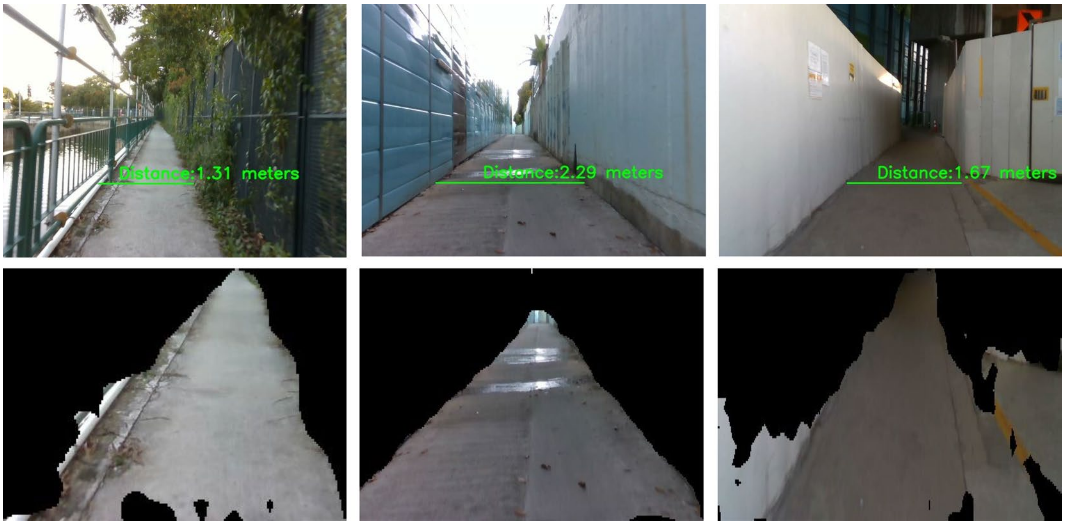


Figure 29. Offline results: pavement width estimation from pavement segmented region.

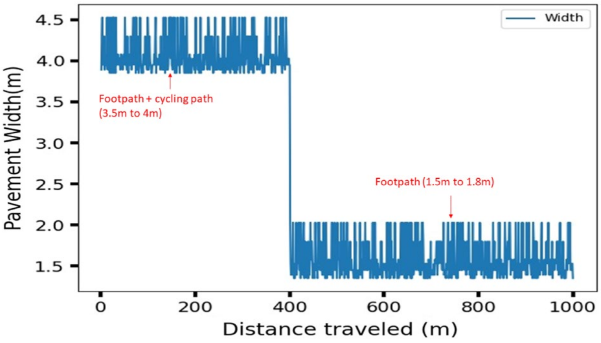


Figure 30. Pavement width estimation.

and accuracy is slightly lower than Good and Poor class as it has two boundaries for the classification whereas Good and Poor class only have one boundary.

Figure 29 shows the pavement width estimation results computed from segmented pavement region part and Fig. 30 shows the pavement width graph for two different pavements computed for 1 km. To predict a one-meter pavement's width, 100 frames results were considered, computed based on robot operation speed of 700 m per hour. At the end of the vision pipeline, the generated surface condition and pavement width information are passed to the FIS to output the speed safety factor of the robot.

Fuzzy controller output.

Based on the membership function and rules, the output of the fuzzy controller is shown in a surface plot given in Fig. 31. Speed of Panthera is multiplied by the output of the fuzzy controller speed safety factor, w, to consider the vision pipeline parameters K, β_l, β_r where K is assigned based on the pavement classification. It can be observed that the robot's speed is maximum only when pavement condition is good and steering requirement is less. If the steering requirement is high, the speed is always minimum irrespective of pavement condition. This feature also suits the dynamic reconfiguration while avoiding various obstacles during motion. Similarly, the speed is minimum when the pavement condition is bad, irrespective of steering requirement. Panthera speed is adjusted for safety during bad pavement conditions and large steering requirements through the fuzzy controller.

Comparison analysis.

The DeepLabv3+ semantic Segmentation model was compared with the UNET Segmentation model. Here, the UNET model was trained with the same dataset till convergence and uses pixel classification accuracy and inference time as evaluation metrics. Figures 32, 33, 34, 35, 36 and 37 shows the segmentation result of DeepLabv3+ and UNET framework.

The comparison analysis indicates that DeepLabv3+ outperforms UNET in terms of segmentation accuracy and inference time. In UNET, the segmentation performance was relatively poor for night mode collected images and water puddles in the driveway. In this comparison analysis, DeepLabv3+ scored 92% classification accuracy and took 10 ms inference time. On the other hand, UNET scored 87% classification and took 22 ms for inference one image.

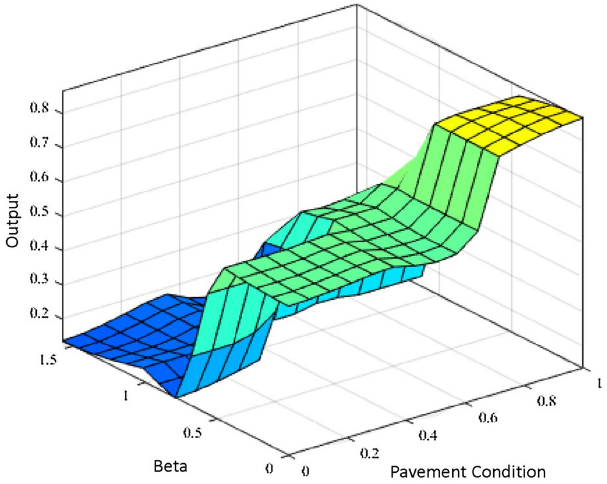


Figure 31. Surface plot for output speed safety factor.

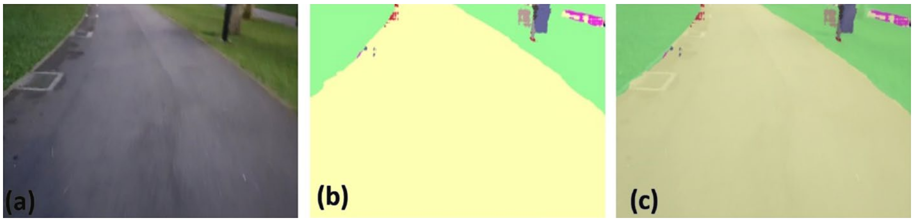


Figure 32. DeepLabv3+ segmentation results.

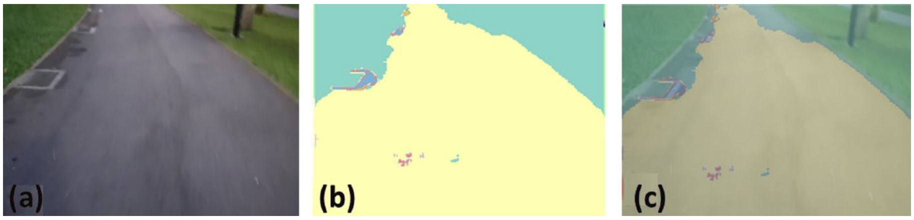


Figure 33. Unet segmentation results.

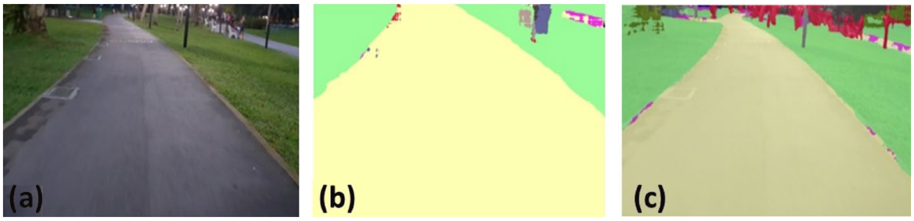


Figure 34. DeepLabv3+ segmentation results: pavement with night mode.

A short summary of the advantages and limitations for the proposed algorithm are as follows:

- The advantages:
  - Safer autonomous pavement sweeping self-reconfigurable robots which changes speed based on a speed safety factor, *w*, that adapts to varying reconfiguration parameters and pavement conditions.
  - High level of accuracy of pavement segmentation allows self-reconfigurable robot to more accurately identify the reconfiguration parameters β_l and β_r.
  - Inference time of the vision pipeline is low.

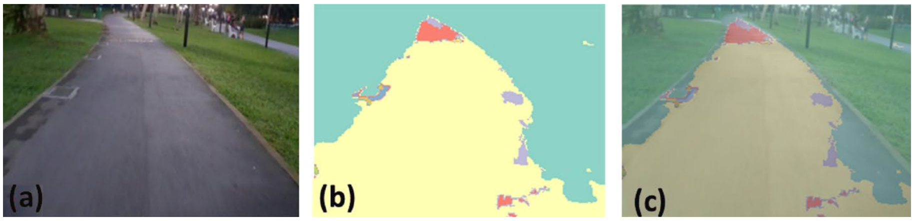


Figure 35. Unet segmentation results: pavement with night mode.

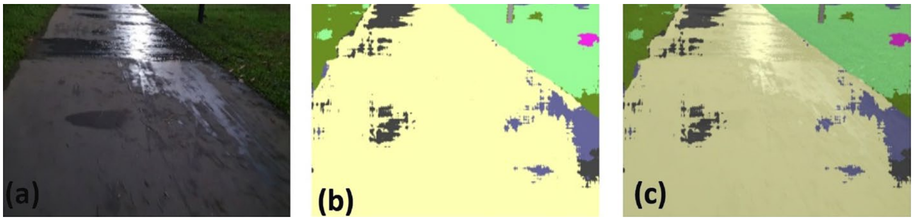


Figure 36. DeepLabv3+ segmentation results: pavement with water puddle.

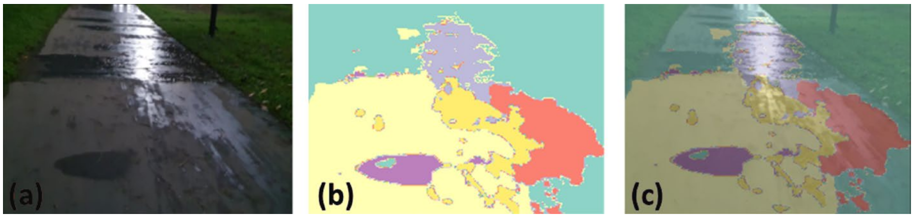


Figure 37. Unet segmentation results: pavement with water puddle.

- The limitations:
  - Models need to be retrained for different pavement types such as wood.
  - Accuracy might be affected by environmental changes such as heavy rain due to lack of visibility
  - Proposed algorithm focuses heavily on pavement segmentation and condition classification. Addition of other classes such as people, animals and vehicles might reduce the model accuracy.

Conclusions

This work proposed the deep learning-based vision pipeline for the reconfigurable pavement sweeping robot Panthera. Through a vision pipeline, the robot has been able to identify the pavement type, pavement condition, reconfiguration parameters to allow the robot to adapt to pavements of changing width with a safety factor based on fuzzy control. The efficiency of the deep learning-based vision pipeline was evaluated with a real pavement testbed, and its detection accuracy was estimated with standard performance metrics. The experimental results indicate that the vision pipeline classifies pavement segmentation and surface condition with 88% and 93% pixel-level classification accuracy. Furthermore, the proposed system was tested in varying lighting conditions and pavement types in Singapore and ensured that the model segmentation and detection accuracy were more stable with various pavement conditions. From the high level of accuracy in the vision pipeline, reconfiguration parameters and safety factors can be derived accurately for the safer operation of the Self-reconfigurable robot Panthera.

Data availability

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Received: 19 March 2022; Accepted: 2 August 2022
Published online: 25 August 2022

References

1. Lta. https://www.straitstimes.com/singapore/new-covered-walkways-to-hit-200km-mark. Accessed 18 March 2022.
2. Gaussian robotics. https://www.gaussianrobotics.com/. Accessed 18 March 2022.
3. Lionsbot. https://www.lionsbot.com/. Accessed 18 March 2022.

4. Johnston sweepers. https://www.johnstonsweepers.com/wp-content/uploads/2017/12/c201-e6-brochure-31506.pdf. Accessed 18 March 2022.
5. Tan, N., Hayat, A. A., Elara, M. R. & Wood, K. L. A framework for taxonomy and evaluation of self-reconfigurable robotic systems. *IEEE Access* **8**, 13969–13986. https://doi.org/10.1109/ACCESS.2020.2965327 (2020).
6. Rus, D. Self-reconfiguring robots. *IEEE Intell. Syst. Appl.* **13**, 2–4. https://doi.org/10.1109/5254.708423 (1998).
7. Tan, N., Mohan, R. E. & Elangovan, K. Scorpio: A biomimetic reconfigurable rolling–crawling robot. *Int. J. Adv. Robot. Syst.* **13**, 1729881416658180. https://doi.org/10.1177/1729881416658180 (2016).
8. Mulgaonkar, Y. *et al.* The flying monkey: A mesoscale robot that can run, fly, and grasp. In *2016 IEEE International Conference on Robotics and Automation (ICRA)* 4672–4679. https://doi.org/10.1109/ICRA.2016.7487667 (2016).
9. Georgiades, C. *et al.* Aqua: An aquatic walking robot. In *2004 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS) (IEEE Cat. No.04CH37566)*, vol. 4, 3525–3531. https://doi.org/10.1109/IROS.2004.1389962 (2004).
10. Hayat, A. A., Parween, R., Elara, M. R., Parsuraman, K. & Kandasamy, P. S. Panthera: Design of a reconfigurable pavement sweeping robot. In *2019 International Conference on Robotics and Automation (ICRA)* 7346–7352. https://doi.org/10.1109/ICRA.2019.8794268 (IEEE, 2019).
11. Rayguru, M. M. *et al.* An output feedback based robust saturated controller design for pavement sweeping self-reconfigurable robot. *IEEE/ASME Trans. Mechatron.* **26**, 1236–1247. https://doi.org/10.1109/TMECH.2021.3063886 (2021).
12. Yi, L. *et al.* Anti-collision static rotation local planner for four independent steering drive self-reconfigurable robots. In *2021 International Conference on Robotics and Automation (ICRA)* (IEEE, 2022).
13. Hayat, A. A., Yi, L., Kalimuthu, M., Elara, M. R. & Wood, K. L. Reconfigurable robotic system design with application to cleaning and maintenance. *J. Mech. Des.*https://doi.org/10.1115/1.4053631 *(2022).*
14. Alpaydin, E. *Machine Learning* (MIT Press, Berlin, 2021).
15. LeCun, Y., Bengio, Y. & Hinton, G. Deep learning. *Nature* **521**, 436–444. https://doi.org/10.1038/nature14539 (2015).
16. Kunze, L., Hawes, N., Duckett, T., Hanheide, M. & Krajník, T. Artificial intelligence for long-term robot autonomy: A survey. *IEEE Robot. Autom. Lett.* **3**, 4023–4030. https://doi.org/10.1109/LRA.2018.2860628 (2018).
17. Zhang, X. & Chen, H. Independent wheel drive and fuzzy control of mobile pipeline robot with vision. In *IECON'03. 29th Annual Conference of the IEEE Industrial Electronics Society (IEEE Cat. No.03CH37468)*, vol. 3, 2526–2530. https://doi.org/10.1109/IECON.2003.1280643 (2003).
18. Arkin, R. C. Path planning for a vision-based autonomous robot. In *Mobile Robots I* Vol. 0727 (eds Marquina, N. & Wolfe, W. J.) 240–250 (International Society for Optics and Photonics (SPIE), Berlin, 1987). https://doi.org/10.1117/12.937802.
19. Murarka, A., Modayil, J. & Kuipers, B. Building local safety maps for a wheelchair robot using vision and lasers. In *The 3rd Canadian Conference on Computer and Robot Vision (CRV'06)* 25–25. https://doi.org/10.1109/CRV.2006.20 (2006).
20. Delmerico, J., Mueggler, E., Nitsch, J. & Scaramuzza, D. Active autonomous aerial exploration for ground robot path planning. *IEEE Robot. Autom. Lett.* **2**, 664–671. https://doi.org/10.1109/LRA.2017.2651163 (2017).
21. Kim, J.-H., Soo Kim, B. & Savarese, S. Comparing image classification methods: K-nearest-neighbor and support-vector-machines (2012).
22. Khan, Y. N., Komma, P. & Zell, A. High resolution visual terrain classification for outdoor robots. In *2011 IEEE International Conference on Computer Vision Workshops (ICCV Workshops)* 1014–1021. https://doi.org/10.1109/ICCVW.2011.6130362 (2011).
23. Omer, R. & Fu, L. An automatic image recognition system for winter road surface condition classification. In *13th International IEEE Conference on Intelligent Transportation Systems* 1375–1379. https://doi.org/10.1109/ITSC.2010.5625290 (2010).
24. Kawai, S., Takeuchi, K., Shibata, K. & Horita, Y. A method to distinguish road surface conditions for car-mounted camera images at night-time. In *2012 12th International Conference on ITS Telecommunications* 668–672. https://doi.org/10.1109/ITST.2012.6425265 (2012).
25. He, K., Zhang, X., Ren, S. & Sun, J. Deep residual learning for image recognition. 770–778. https://doi.org/10.1109/CVPR.2016.90 (2016).
26. Iandola, F. *et al.* Squeezenet: Alexnet-level accuracy with 50x fewer parameters and < 0.5mb model size (2016).
27. Simonyan, K. & Zisserman, A. Very deep convolutional networks for large-scale image recognition. arXiv:1409.1556 (2014).
28. González, R. & Iagnemma, K. Deepterramechanics: Terrain classification and slip estimation for ground robots via deep learning. arXiv:abs/1806.07379 (2018).
29. Liang, C. *et al.* Winter road surface status recognition using deep semantic segmentation network. In *Proceedings of the International Workshop on Atmospheric Icing of Structures (IWAIS 2019), Reykjavik, Iceland* 23–28 (2019).
30. Nolte, M., Kister, N. & Maurer, M. Assessment of deep convolutional neural networks for road surface classification. In *2018 21st International Conference on Intelligent Transportation Systems (ITSC)*, 381–386. https://doi.org/10.1109/ITSC.2018.8569396 (2018).
31. Melvin, L. M. J. *et al.* Remote drain inspection framework using the convolutional neural network and re-configurable robot raptor. *Sci. Rep.* **11**, 1–15. https://doi.org/10.1038/s41598-021-01170-0 (2021).
32. Suryamurthy, V., Raghavan, V. S., Laurenzi, A., Tsagarakis, N. G. & Kanoulas, D. Terrain segmentation and roughness estimation using RGB data: Path planning application on the centauro robot. In *2019 IEEE-RAS 19th International Conference on Humanoid Robots (Humanoids)* 1–8. https://doi.org/10.1109/Humanoids43949.2019.9035009 (IEEE, 2019).
33. Yi, L. *et al.* Reconfiguration during locomotion by pavement sweeping robot with feedback control from vision system. *IEEE Access* **8**, 113355–113370. https://doi.org/10.1109/ACCESS.2020.3003376 (2020).
34. Le, A. V., Hayat, A. A., Elara, M. R., Nhan, N. H. K. & Prathap, K. Reconfigurable pavement sweeping robot and pedestrian cohabitant framework by vision techniques. *IEEE Access* **7**, 159402–159414. https://doi.org/10.1109/ACCESS.2019.2950675 (2019).
35. Lim, Y. *et al.* Anti-collision static rotation local planner for four independent steering drive self-reconfigurable robots. In *2022 International Conference on Robotics and Automation (ICRA)* 1–8 (IEEE, 2022).
36. Povendhan, A. *et al.* Multi-sensor fusion incorporating adaptive transformation for reconfigurable pavement sweeping robot. In *2021 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)* 300–306 (IEEE, 2021).
37. Yi, L. *et al.* Locomotion with pedestrian aware from perception sensor by pavement sweeping reconfigurable robot. *Sensors*https://doi.org/10.3390/s21051745 *(2021).*

Author contributions

Conceptualization, B.K.R. and M.R.E.; methodology, M.R.E., L.Y. and B.K.R.; software, B.F.G. and B.K.R; validation, L.Y. and A.A.H.; formal analysis, L.Y., B.K.R. and A.A.H.; investigation, M.R.E. and B.K.R.; resources, M.R.E. and B.K.R.; data, B.K.R., L.Y., M.M.R, B.F.G and A.A.H.; writing–original draft preparation, B.K.R., A.A.H. and M.R.E.; supervision, M.R.E.,and B.K.R.; project administration, M.R.E.; funding acquisition, M.R.E. All authors reviewed the manuscript.

Funding

This research is supported by the National Robotics Programme under its Robotics Enabling Capabilities and Technologies (Funding Agency Project No. 192 25 00051), National Robotics Programme under its Robot

Domain Specific (Funding Agency Project No. 192 22 00058) and administered by the Agency for Science, Technology and Research. National Robotics Programme(Grant No. W1922d0110).

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to B.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.