ICIC Express Letters ICIC International ©2024 ISSN 1881-803X
Volume **18**, Number **1**, January **2024** pp. **79–86**

ROAD SURFACE CLASSIFICATION FOR INTELLIGENT VEHICLE PERCEPTION BASED ON INERTIAL SENSORS

Sakorn Mekruksavanich[1], Ponnipa Jantawong[1], Olarik Surinta[2] and Anuchit Jitpattanakul[3,4,*]

[1]Department of Computer Engineering
School of Information and Communication Technology
University of Phayao
19 Moo 2 Tambon Maeka, Amphur Muang, Phayao 56000, Thailand
sakorn.me@up.ac.th; ponnipa.jantawong@gmail.com

[2]Multi-Agent Intelligent Simulation Laboratory (MISL) Research Unit
Department of Information Technology
Faculty of Informatics
Mahasarakham University
Khamriang Sub-District, Kantarawichai District, Mahasarakham 44150, Thailand
olarik.s@msu.ac.th

[3]Department of Mathematics
Faculty of Applied Science
[4]Intelligent and Nonlinear Dynamic Innovations Research Center
Science and Technology Research Institute
King Mongkut's University of Technology North Bangkok
1518 Pracharat 1 Road, Wongsawang, Bangsue, Bangkok 10800, Thailand
[*]Corresponding author: anuchit.j@sci.kmutnb.ac.th

Received April 2023; accepted June 2023

Abstract. *In latest years, the necessity for several sources of situational information from the traffic environment has increased due to the growth of Intelligent Transport System (ITS) solutions, such as autonomous vehicles and enhanced driver support systems. Identifying Road Surface Type (RST) within this environmental information is essential and applicable throughout the ITS sector. The classification method must function successfully across various cars, driving behaviors, and situations in which a vehicle might operate. In this study, we use inertial sensors, such as accelerometers, gyroscopes, and magnetometers, which are reliable, non-polluting, and low-cost solutions appropriate for large-scale deployment, to develop a deep learning model that classifies road surface characteristics effectively. These sensor data were employed in three basic deep learning models, including our proposed RST-PyramidNet model: CNN-based, LSTM-based, and GRU-based models. A public benchmark dataset named Passive Vehicular Sensors (PVS) dataset based on the 5-fold cross-validation methodology is used to assess the effectiveness of these models. The experimental findings indicate that the proposed RST-PyramidNet surpasses previous benchmark deep learning models with an accuracy of 97.68% and an F1-score of 97.35%.*

Keywords: Road surface type, Deep learning, Classification model, Inertial sensor, Pyramidal residual network

1. **Introduction.** Inappropriate road surfaces could lead to dangerous and unpleasant travel and costly litigation and injury claims [1]. In these criteria, there are accidents directly impacted by the terrible state of the roadway, but drivers' conduct also causes problems in response to the poor condition of the roadway. Across the globe, the data about incidents that occurred due to poor road surface quality varies from low percentages

DOI: 10.24507/icicel.18.01.79

(high-income nations) to incredibly high figures representing low- and medium-income nations. Sustaining an appropriate level of effectiveness for the whole road infrastructure by implementing efficient road pavement management and repair programs is, thus, one of the most significant problems for Intelligent Transport Systems (ITS) throughout the globe [2, 3].

Conventional methods depend primarily on human perception, a global standard for all road inspection and professional practices [3, 4]. Nevertheless, the method is susceptible to subjectivity, time-consuming, and dangerous for drivers on the road. Modern techniques and high-performance approaches employ video and image analysis [5, 6, 7] to detect and categorize pavement surface distress more accurately [8, 9]; unfortunately, its present installation cost is considerable. For instance, the most prevalent tracking technique is automated technology such as ground penetrating radar, laser road imaging systems, and high-performance sensors paired with high-resolution cameras. Furthermore, the high costs connected with such technology are a vital barrier for road authorities, particularly small towns that could need more funding or specialized abilities to perform regular inspections.

Vibration-based techniques might be a viable compromise between the high-cost monitoring procedure based on video or image interpretation and the necessity for frequent road network surveillance [10]. To avoid cost consumption, inertial sensor-based approaches (e.g., accelerometers) were used to produce a low-cost, dependable solution capable of real-time, flexible pavement inspection [11, 12]. The inertial sensor-based systems can monitor road surface irregularities through automobile vibrations, and the data are collected using vehicle-mounted accelerometers [8, 13]. In addition, the literature study revealed that the primary limitation of vibration-based techniques is the need for more ability to derive (from collected data) data regarding the kind of road surface concern. Furthermore, it was emphasized that this strategy could not evaluate road-surface damage in locations other than the vehicle wheel paths, with the result that the magnitude of pavement failure across the entire section of the road cannot be determined [14].

Over the past few years, there has been a lot of research focused on using inertial sensors to identify different types of road surfaces. Souza's study [15] utilized a flexible suction holder to secure a smartphone with sensors in a car, close to the dashboard. The data collected from the accelerometer (sampled at 100 Hz) and the GPS (for estimated velocity) were incorporated into a model (known as GPS). The most successful results in the report were obtained by combining the Longest Common Subsequence (LCSS) model with the Complexity Invariant Distance (CID). Overall, the model was able to accurately classify road surfaces as either asphalt/flexible pavement (98.28%), cobblestone streets (84.4%), or unpaved roads (78.64%), with an accuracy rate of 87.72%. In other studies, such as [16] and [17], accelerometer data from the suspension system and GPS speed were also used. To take account of the effects of the vehicle's suspension, the researchers applied the Quarter Car (QC) computational formula in the pre-processing stage and used the Fast Fourier Transform (FFT) to extract frequency domain features. The researchers were able to train a Support Vector Machine (SVM) to classify road surfaces as either asphalt (17.6%), concrete (99.6%), grass (74.9%), or gravel (85.3%) with an average accuracy of 69.4%.

Recent advancements in the classification of types of roads have resulted in the development of Deep Learning (DL) algorithms to increase identification results [18, 19]. In Varona et al.'s research [20], sensor data collected from accelerometers were fed into a Convolutional Neural Network (CNN) and a Long Short-Term Memory (LSTM) neural network to distinguish between four distinct categories of roads (concrete panels, cobblestones, asphalt, and dirt road). Like the study of Menegazzo and Wangenheim [21], they developed and tested three models for road surface type classification considering three stages – data collection, pre-processing and processing. This study also employed accelerometer data to train and evaluate DL models (CNN, LSTM, and Gated Recurrent Unit (GRU)). Training accuracy for the CNN-based model was an average of 93.04%, with a score of 98.60% for asphalt, 86.09% for cobblestone, and 90.78% for dirt. Moreover, current findings [22, 23] demonstrated an affordable and dependable system capable of real-time road pavement distress monitoring. Road surface categories were classified using various Machine Learning (ML) techniques (decision tree, k-nearest neighbor, and SVM).

Several studies have attempted to classify road surface categories using ML techniques. However, the interpretability of the classifiers was found to be hindered by the use of manually extracted features. In the same study of Souza [15], inertial sensors were used to determine road surface types, and the LCSS model was combined with CID to achieve an accuracy rate of 87.72%. Despite the successful results, the interpretation of the classifier was still limited by the use of manually extracted features. Similarly, in Basavaraju et al.'s study [24], ML models were trained to classify road surfaces using features extracted from accelerometer data and GPS speed. However, the classifier's interpretability was still compromised by the use of manually extracted features [25].

These findings highlight the need for alternative feature extraction methods and more advanced ML techniques to improve the interpretability of road surface classifiers. DL approaches, such as CNNs, may be promising for this purpose as they can automatically learn informative features from raw sensor data. By leveraging such methods, we can expect to achieve better classification results with improved interpretability, leading to more accurate and reliable road surface detection in real-world applications. Therefore, this study investigated road surface identification using inertial sensors and DL techniques that automatically generate distinguishing features. Models based on CNN, LSTM, and GRU are three fundamental deep-learning models designed to identify road surface variations automatically. Furthermore, a deep pyramidal residual model was established to categorize road surface classes effectively.

The remainder of the article is structured as follows. Section 2 describes the proposed RST-PyramidNet model in depth. Section 3 contains the findings of our experiments. Section 4 concludes this work with a discussion of demanding future research.

2. **Architecture of Sensor-Based RST Recognition.** The sensor-based RST classification framework, which is operated in this work, comprises four primary procedures: data acquisition, data pre-processing, data generation, and model training with evaluation, as shown in Figure 1.

2.1. **Passive vehicular sensors dataset.** We assessed our research using the Passive Vehicular Sensors (PVS) dataset [21], which is a benchmark that is accessible to the

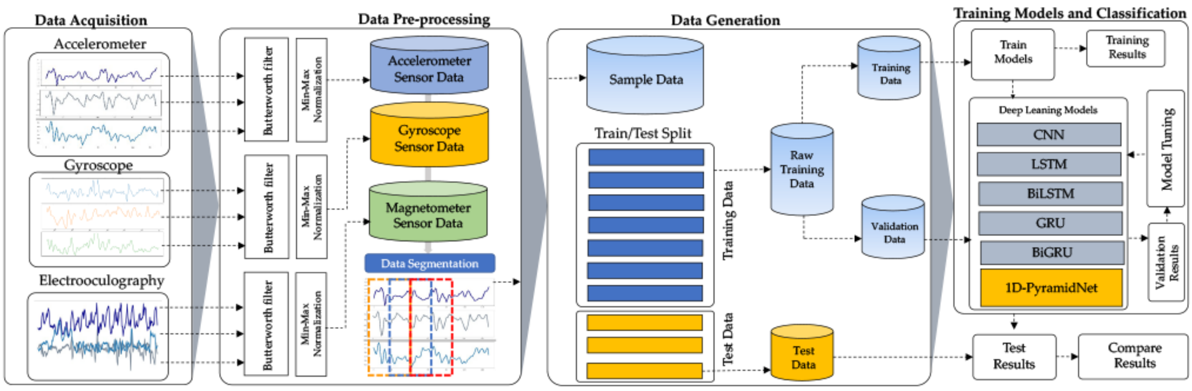

FIGURE 1. RST classification framework based on inertial sensors used in this work

public. Sensors were installed throughout the car to collect data, including a camera positioned on the vehicle's roof that recorded the external environment at a frequency of 30 frames per second. The in-dash GPS receiver sampled data at 1 Hz from an internal location. To account for information from multiple sources that could potentially affect the car's dependency attribute, six MPU-9250 units were strategically placed throughout the automobile. Three MPU-9250 were installed at each end of the front axle (right and left): one component was connected to the control arm, below and near the suspension; another component was located above and near the suspension, connected to the body immediately above the tire; and a third component was installed on the center console, inside of the cabin. In order to ensure that the MPU-9250's sample reference frame corresponds to the car's position and orientation, a controlled positioning method was utilized to install the components. To avoid signal saturation, the accelerometer and gyroscope were set to a full scale of 8 g and 1000 deg/s, respectively, and both were sampled at a frequency of 100 Hz.

The PVS dataset compiled information from various settings to provide a wide range of test cases for the model. The sensor mentioned above network was deployed across three vehicles (vehicular property), three drivers (driving property), and three situations (environmental property), wherein each situation provides three different surface types, among unpaved (dirt roads) and paved road segments (0-91.98 km/h) (asphalt or cobblestone roads).

2.2. **Data pre-processing.** The raw sensor data collected from the vehicle during the data collection process contained measurement noise and other unexpected noises, likely due to the sensor movements in the automobile. Noisy signals can cause data distortion, making it less reliable. As a result, it was critical to limit the impact of noise on signal processing so that useful information could be retrieved from the signal. Mean, low-pass, and wavelet filtering are some of the most frequently used techniques for filtration. Using a 3rd-order Butterworth filter, we de-noised all three dimensions of accelerometers, gyroscopes, and magnetometers using the 20 Hz cutoff frequency. At this pace, 99.9% of vehicle movements are captured, making it superior for recording motion.

It was necessary to alter the sensor data once it had been cleansed of unwanted noise. Each data point was transformed using a Min-Max normalization approach, which projects its values into the range $[0, 1]$. Having a way to balance the impacts of different dimensions might be beneficial for the learning processes. Normalized data from all sensors are split into equal-sized sections for model training using fixed-size sliding windows in the data segmentation stage of the process. To construct sensory data streams with a length, we employed a sliding window with a duration of 10 seconds in this study. The 10-second window is utilized for user identification because it is long enough to record crucial features of a thing's activities, such as numerous repeats of basic motions. Then, 2-second fixed-width sliding windows were applied with 50% overlap to the pre-processed sensor data.

2.3. **DL models recognition.** After pre-processing the data, DL-based classification methods are utilized. This process involves constructing computational models with multiple processing layers to identify the data's representations at different levels of abstraction. As a result, layered representations are produced, expressed in terms of other layers, allowing for the creation of complex ideas from simpler ones. We developed deep neural networks based on LSTM, GRU, and CNN as fundamental models for DL. We employed the Adam optimizer in conjunction with the Categorical Cross Entropy loss function to construct all models.

This study proposed a one-dimensional pyramidal network named RST-PyramidNet, illustrated in Figure 2, for effectively classifying road surface categories using signal data

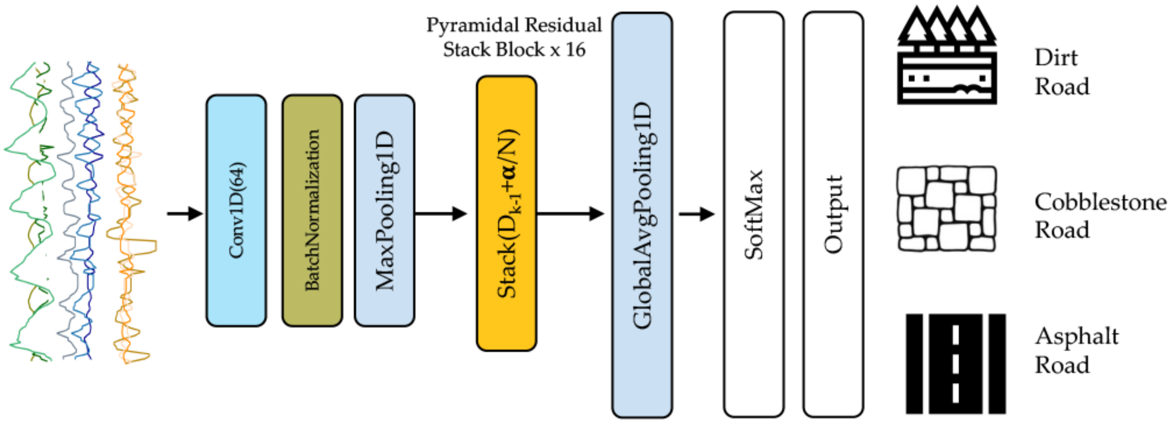

FIGURE 2. RST-PyramidNet architecture used in this work

from inertial sensors. The RST-PyramidNet is developed based on Han et al.'s deep pyramidal network [26].

The RST-PyramidNet is an improvement over the residual unit model based on the residual network. The main idea behind the pyramidal network is to gradually expand the feature map dimensions rather than abruptly increasing them at each residual unit through downsampling. Most deep CNN architectures enhance feature map dimensions significantly when the feature map scale decreases and increases feature map dimensions only when they approach a downsampling level. To address this issue, the network aims to progressively expand the feature map dimensions instead of increasing them at a single residual unit and evenly distribute the increasing feature map load. This approach results in a gradual increase in the number of channels as a function of layer depth, similar to a pyramid shape that gradually expands from top to bottom. In our research, the proposed RST-PyramidNet adopted the additive PyramidNet model, geometrically expanding the feature map dimension. The process of increasing the dimensionality of a feature map can be stated as follows:

$$D_k = \begin{cases} 64 & \text{if } k = 1 \\ \left| D_{k-1} + \dfrac{\alpha}{N} \right| & \text{if } 2 \leq k \leq N+1 \end{cases}$$

In the proposed network, the step factor for increasing dimensions, denoted by α, is set to 48. Here, N denotes the total number of residual units.

The network architecture also includes a combination of simple and residual networks, utilizing zero-padded identity-mapping shortcut connections when expanding the feature map dimension. ResNet investigated several types of shortcuts, including identity-mapping shortcuts. The identity-mapping shortcut is highly effective and has a lower risk of overfitting than other shortcuts due to the lack of parameters, resulting in improved generalization performance. Additionally, it can strictly pass through the gradient based on the unique mapping, providing more excellent stability during training. However, in our RST-PyramidNet, identity mapping alone cannot be used as a shortcut because the feature map dimension varies between residual units. Therefore, the zero-padded shortcut does not increase the risk of overfitting since there are no new parameters, and it exhibits better generalization performance than other shortcuts. As a result, the proposed zero-padded identity-mapping shortcut, as shown in Figure 3, can have a combined effect on the residual and plain networks, leading to significant improvement.

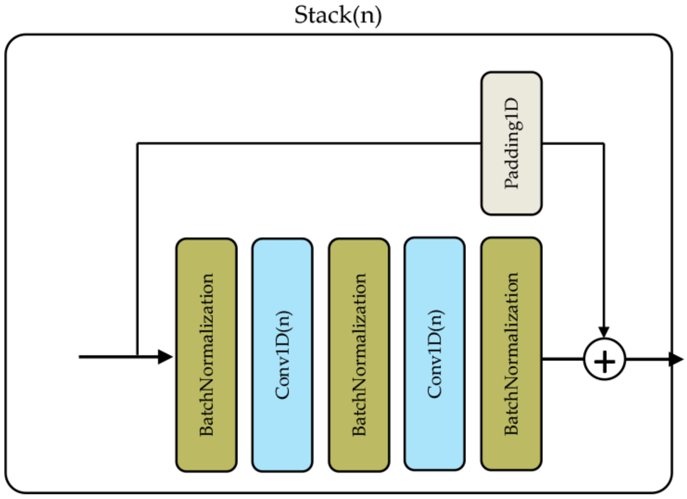

FIGURE 3. Structure of a stack block used in the proposed RST-PyramidNet

3. **Experimental Results and Discussion.** In this section, we explain the experimental setup and present the experimental findings required to accurately assess the proposed RST-PyramidNet model for classifying road types utilizing IMU sensor signal data.

This study conducted each experiment on the Google Colab Pro platform using a Tesla V100. Python 3.6.9, TensorFlow 2.2.0, Keras 2.3.1, Scikit-Learn, Numpy 1.18.5, and Pandas 1.0.5 libraries were utilized in addition to the Python programming language. Two experiments were conducted using sensor data from different locations to investigate the efficacy of DL methods.

- Case I involves using signal data collected from sensors positioned both above and below the left suspension of the vehicle.
- Case II involves using signal data collected from sensors positioned both above and below the right suspension on the right side of the vehicle.

Experiments were conducted to investigate the effectiveness of DL models based on signal data from IMU sensors. CNN, GRU, LSTM, and the proposed RST-PyramidNet all optimized their hyperparameters using the Bayesian technique. The experiments were assessed for their identification performance using various metrics such as accuracy, precision, recall, and F1-score.

Table 1 shows the accuracy and F1-score measures acquired from the different DL networks trained on the PVS dataset in Case I.

TABLE 1. Identification effectiveness of DL models conducted in Case I

Model	**Recognition performance**		
	Accuracy	**Loss**	**F1-score**
CNN	94.39% (±0.97%)	0.34 (±0.03)	93.68% (±1.06%)
LSTM	96.48% (±0.44%)	0.17 (±0.03)	95.98% (±0.50%)
BiLSTM	96.18% (±0.23%)	0.20 (±0.03)	95.65% (±0.28%)
GRU	96.62% (±0.27%)	0.20 (±0.04)	96.15% (±0.31%)
BiGRU	96.34% (±0.56%)	0.18 (±0.04)	95.83% (±0.66%)
RST-PyramidNet	97.54% (±0.49%)	0.11 (±0.01)	97.21% (±0.55%)

Table 1 demonstrates that the RST-PyramidNet network proposed in this study outperforms all other network models with an accuracy of 97.54% and an F1-score of 97.21%. This indicates that the RST-PyramidNet has better interpretability than the baseline deep learning models.

The outcomes of Case II, which used signal data from sensors positioned above and below the right suspension of the car, are presented in Table 2. The table shows that the proposed RST-PyramidNet network outperforms all other network models with an accuracy of 97.68% and an F1-score of 97.35%. The performance of the RST-PyramidNet is superior to that of the baseline deep learning models.

Table 2. Identification effectiveness of DL models conducted in Case II

Model	**Recognition performance**		
	Accuracy	**Loss**	**F1-score**
CNN	94.62% (±0.43%)	0.36 (±0.04)	93.91% (±0.46%)
LSTM	96.66% (±0.36%)	0.16 (±0.02)	96.21% (±0.40%)
BiLSTM	95.42% (±0.73%)	0.26 (±0.03)	94.75% (±0.83%)
GRU	96.39% (±0.56%)	0.20 (±0.04)	95.92% (±0.64%)
BiGRU	95.81% (±0.31%)	0.26 (±0.03)	95.22% (±0.37%)
RST-PyramidNet	97.68% (±0.33%)	0.11 (±0.02)	97.35% (±0.37%)

4. **Conclusion and Future Works.** In this study, we established classification models based on DL methodologies to handle signal data from inertial sensors. Considering its conservative approach, these sensors are harmless, non-polluting, and competitive, representing an appealing option for large-scale applications. These sensor data were then pre-processed and segmented in particular studies to assess the identification capabilities of DL networks. These experiments were conducted using CNN, LSTM, and GRU as baseline models for deep learning models. We proposed an RST-PyramidNet network to solve sensor-based RST employing sensor data from inertial sensors to enhance the classifier's performance. We have assessed the effectiveness of the RST-PyramidNet model using several indicators and an available PVS dataset. The findings show that the suggested deep pyramidal residual network surpasses the other baseline network by using the autonomously spatial-temporal feature extraction from raw sensor data with an average accuracy of 97.61% and F1-score of 97.28% for both scenarios.

In future studies, we aim to enhance the classification of road types by exploring several avenues. One such avenue includes optimizing the hyperparameters of the proposed models and conducting experiments with hybrid deep learning networks, such as LSTM-CNN, GRU, and ConvLSTM. Additionally, we plan to investigate other features that can be applied to analytical domains such as time, frequency, and time-frequency domains.

Acknowledgment. This research is supported by Thailand Science Research and Innovation Fund and University of Phayao (Grant No. FF66-UoE001).

REFERENCES

[1] J. Eriksson, L. Girod, B. Hull, R. Newton, S. Madden and H. Balakrishnan, The pothole patrol: Using a mobile sensor network for road surface monitoring, *Proc. of the 6th International Conference on Mobile Systems, Applications, and Services (Mobisys08)*, New York, USA, pp.29-39, 2008.

[2] S. Sousa, A. Santos, A. Costa, B. Dias, B. Ribeiro, F. Gonçalves, J. Macedo, M. J. Nicolau and Ó. Gama, A new approach on communications architectures for intelligent transportation systems, *Procedia Computer Science*, vol.110, pp.320-327, 2017.

[3] M. N. Tahir and U. Rashid, Demo: Intelligent transport system (ITS) assisted road weather & traffic services, *Proc. of the 2020 IEEE Vehicular Networking Conference (VNC)*, New York, USA, pp.1-2, 2020.

[4] L. C. González, R. Moreno, H. J. Escalante, F. Martínez and M. R. Carlos, Learning roadway surface disruption patterns using the bag of words representation, *IEEE Transactions on Intelligent Transportation Systems*, vol.18, no.11, pp.2916-2928, 2017.

[5] S. Gonwirat and O. Surinta, DeblurGAN-CNN: Effective image denoising and recognition for noisy handwritten characters, *IEEE Access*, vol.10, pp.90133-90148, 2022.

[6] S. Noppitak and O. Surinta, dropCyclic: Snapshot ensemble convolutional neural network based on a new learning rate schedule for land use classification, *IEEE Access*, vol.10, pp.60725-60737, 2022.
[7] K. Fithriasari and U. S. Nuraini, Face identification using multi-layer perceptron and convolutional neural network, *ICIC Express Letters*, vol.15, no.2, pp.157-164, 2021.
[8] H. Dong, K. Song, Y. Wang, Y. Yan and P. Jiang, Automatic inspection and evaluation system for pavement distress, *IEEE Transactions on Intelligent Transportation Systems*, vol.23, no.8, pp.12377-12387, 2022.
[9] S. Luo, J. Yao, J. Hu, Y. Wang and S. Chen, Using deep learning-based defect detection and 3D quantitative assessment for steel deck pavement maintenance, *IEEE Transactions on Intelligent Transportation Systems*, pp.1-10, 2022.
[10] B. Yu and X. Yu, Vibration-based system for pavement condition evaluation, *Proc. of the 9th International Conference on Applications of Advanced Technology in Transportation (AATT)*, Chicago, IL, USA, pp.183-189, 2006.
[11] S. Mekruksavanich, A. Jitpattanakul, K. Sitthithakerngkiet, P. Youplao and P. Yupapin, ResNet-SE: Channel attention-based deep residual network for complex activity recognition using wrist-worn wearable sensors, *IEEE Access*, vol.10, pp.51142-51154, 2022.
[12] S. Mekruksavanich and A. Jitpattanakul, Deep residual network for smartwatch-based user identification through complex hand movements, *Sensors*, vol.22, no.8, 2022.
[13] A. Susanto and G. P. Kusuma, Road condition detection based on accelerometer and gyroscope sensors using deep learning LSTM, *ICIC Express Letters*, vol.16, no.8, pp.803-810, 2022.
[14] T. Lee, C. Chun and S.-K. Ryu, Detection of road-surface anomalies using a smartphone camera and accelerometer, *Sensors*, vol.21, no.2, 2021.
[15] V. M. A. Souza, Asphalt pavement classification using smartphone accelerometer and complexity invariant distance, *Engineering Applications of Artificial Intelligence*, vol.74, pp.198-211, 2018.
[16] S. Wang, S. Kodagoda, L. Shi and X. Dai, Two-stage road terrain identification approach for land vehicles using feature-based and Markov Random Field algorithm, *IEEE Intelligent Systems*, vol.33, no.1, pp.29-39, 2018.
[17] S. Wang, S. Kodagoda, L. Shi and H. Wang, Road-terrain classification for land vehicles: Employing an acceleration-based approach, *IEEE Vehicular Technology Magazine*, vol.12, no.3, pp.34-41, 2017.
[18] J. Menegazzo and A. von Wangenheim, Road surface type classification based on inertial sensors and machine learning, *Computing*, vol.103, pp.2143-2170, 2021.
[19] L. Guo, R. Li and B. Jiang, A road surface damage detection method using YOLOv4 with PID optimizer, *International Journal of Innovative Computing, Information and Control*, vol.17, no.5, pp.1763-1774, 2021.
[20] B. Varona, A. Monteserin and A. Teyseyre, A deep learning approach to automatic road surface monitoring and pothole detection, *Personal Ubiquitous Comput.*, vol.24, no.4, pp.519-534, 2019.
[21] J. Menegazzo and A. von Wangenheim, Multi-contextual and multi-aspect analysis for road surface type classification through inertial sensors and deep learning, *Proc. of the 2020 X Brazilian Symposium on Computing Systems Engineering (SBESC)*, Florianopolis, Brazil, pp.1-8, 2020.
[22] E. Ranyal, A. Sadhu and K. Jain, Road condition monitoring using smart sensing and artificial intelligence: A review, *Sensors*, vol.22, no.8, 2022.
[23] R. Bibi, Y. Saeed, A. Zeb, T. Ghazal, R. Said, S. Abbas, M. Ahmad and M. Khan, Edge AI-based automated detection and classification of road anomalies in VANET using deep learning, *Computational Intelligence and Neuroscience*, vol.2021, pp.1-16, 2021.
[24] A. Basavaraju, J. Du, F. Zhou and J. Ji, A machine learning approach to road surface anomaly assessment using smartphone sensors, *IEEE Sensors Journal*, vol.20, no.5, pp.2635-2647, 2020.
[25] J. Liu and S. S. Chung, Automatic feature extraction and selection for machine learning based intrusion detection, *Proc. of the 2019 IEEE SmartWorld, Ubiquitous Intelligence & Computing, Advanced & Trusted Computing, Scalable Computing & Communications, Cloud & Big Data Computing, Internet of People and Smart City Innovation (SmartWorld/SCALCOM/UIC/ATC/CBDCom/IOP/SCI)*, Leicester, UK, pp.1400-1405, 2019.
[26] D. Han, J. Kim and J. Kim, Deep pyramidal residual networks, *Proc. of the 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Honolulu, HI, USA, pp.6307-6315, 2017.