International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

Examining The Impact of Traffic Sampling on Methods for Network Intrusion Detection Based on Machine Learning

M. Yashwant Ravi Kumar ,Gajula Pavan Kumar ,Saroj Kumar Bharti
Dept. of Computer Science and Engineering
Bharath Institute of Higher Education and Research
Chennai, India
yaswanthravi11@gmail.com
gajulapavannaidu666@gmail.com

Challa Apoorva
Dept. of Computer Science and Engineering
Bharath Institute of Higher Education and Research Hyderabad, India
apoorvachalla14@gmail.com

Guide:
Mr.Krishnamoorthy
Assistant Professor Dept. of Computer science and Engineering, Bharath Institute of Higher
Education and Research, Krishnamoorthy.cse@bh arathuniv.ac.in

Abstract

Network intrusion detection is a key cybersecurity element that identifies and prevents unauthorized entry into computer networks Machine learning (ML) is an emerging solution for intrusion detection because it can process huge volumes of network traffic data and recognize sophisticated patterns of attacks. Yet, the success of ML-based intrusion detection systems (IDS) relies greatly on the input data quality and quantity.

This research evaluates the effect of traffic sampling on the performance of ML-based IDS. Two popular ML algorithms—Random Forest and Multi-Layer Perceptron—are tested using datasets with both unsampled and sampled network traffic data with varying sampling rates. IDS performance is measured in terms of accuracy, precision, recall, and F1-score.
The findings show that traffic sampling has a considerable impact on ML-based IDS, with
increased sampling rates tending to improve performance. Nevertheless, the best sampling rate depends on the ML algorithm and dataset employed.
The findings show that traffic sampling has a significant impact on ML-based IDS, with increasing sampling rates typically resulting in improved performance. The best sampling rate, however, depends on the particular ML algorithm and dataset utilized.

Decreasing the time and effort involved in GBM evaluation. The system not only increases diagnostic accuracy but also reduces the chances of human error, producing reproducible and consistent results irrespective of the different clinical environments. Furthermore, application of sophisticated deep learning models like Vision Transformers helps the model capture intricate spatial and contextual information from MRI scans and enhances the classification accuracy of genetic markers. The scalability of the framework enables it to be installed in numerous healthcare institutions, which makes it

International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 09 Issue: 04 | April - 2025 SJIF Rating: 8.586 ISSN: 2582-3930

an essential tool for large-scale GBM screening and research. Finally, this project hopes to eliminate the disconnect between medical imaging and genomics, enabling the creation of AI-enabled precision medicine for the diagnosis and treatment of brain tumours.

I.INTRODUCTION

Potholes are concave depressions on roads measuring as much as 10 inches in depth and result from road wear and weathering. Potholes emerge when the upper part of the road, most likely bitumen, wears out with heavy usage by vehicles, and the concrete below is revealed. Once a pothole is created, it can spread wider, particularly under rain conditions, making it one of the causes of motor accidents. Potholes are not just a major cause of damage to motor vehicles but life-threatening for motorcyclists too. They cause more danger at high speeds of travel, for drivers might miss the bumps within time. Even if a high-speed vehicle were to hit a pothole, the sudden impact would induce tire bursts, abrupt braking, rollover or loss of traction, resulting in accidents.

Because of these risks, detecting potholes during driving is the target of this study. The system proposed in this work tries to provide 3D pothole information and inform drivers of the distance of the pothole beforehand, enhancing road safety.

II. LITERATURE SURVEY

Detections of potholes have concerned researchers due to their effects on road safety as well as automobile maintenance. Traditional image processing approaches, as well as sophisticated machine learning techniques, have been developed for identifying and analysing potholes.

Initial research was aimed at applying image processing methods, in which road images were taken and processed using edge detection, texture analysis, and pattern recognition. These approaches depended greatly on hand-extracted features and were prone to being hindered by environmental conditions like light, shadows, and occlusions.

As sensor technology improved, scientists started using accelerometers and gyroscopes to identify potholes from the vibrations in vehicles. By monitoring differences in acceleration and angular velocity, these systems could identify road irregularities with fair accuracy. They did need to be highly calibrated for various vehicle types and road surfaces.

Deep learning and machine learning techniques have helped enhance pothole detection rates considerably. New research has discussed the use of convolutional neural networks (CNNs) in image-based detection of potholes, whereby models are learnt using large road image datasets. These models have the ability to learn features automatically and detect potholes with high accuracy. Other research has combined LiDAR sensors and GPS information to build precise 3D models of road surfaces, enabling real-time pothole detection and localization.

While all these advancements exist, it remains challenging to ensure the robustness of pothole detection systems when applied in diverse driving scenarios and road surfaces. The quality of input data, efficiency in computations, and the ability to perform in real-world driving conditions determines the performance of such systems. Through this paper, we find inspiration in existing literature by integrating real-time alert and 3D pothole detection to enhance road maintenance and drivers' safety.

grading methods, their characteristic sets, classifiers, segmentation methods, and classification types.

III. PROBLEM STATEMENT

Potholes are a safety concern regarding roads, vehicle condition, and the longevity of infrastructure. Potholes lead to accidents, vehicle damage, and increased repair expenditures for the general public and municipal governments. Despite many attempts at preventing pothole formation, current pothole detection technology and warning technology are ineffective in most areas.

Conventional pothole detection techniques based on manual surveys and citizen reports are ineffective, time-consuming, and result in delayed road maintenance. Though certain automated techniques employ image processing or sensor-based techniques, they are limited by accuracy, environment robustness, and real-time usage.

This research would create a cost-effective pothole detection system, which uses new technologies to create 3D pothole information and give drivers real-time warning. Through the accurate detection of potholes and timely notification, the system will make travel on roads safer, minimize damage to vehicles, and help road maintenance agencies allocate repair priorities more effectively.

Potholes not only inconvenience drivers but also contribute to deadly road accidents, especially for motorcyclists and cyclists. When a car drives over a pothole at a high speed, it may result either in sudden jerks, tire punctures, suspension failure, or even loss of vehicle control, hence more opportunities for collision. Additionally, high-frequency roads with potholes may result in traffic congestion and reduced speeds, hence resulting in overall inefficiencies in transport networks. It can be addressed by a dependable and Existing pothole detection methods are beset by accuracy, flexibility, and cost issues. Road inspection is time-consuming and non-real-time, and as such, it is not straightforward for the authorities to respond in a timely manner. Image processing methods, while useful, are generally beset by illumination, weather, and road surface changes, leading to misclassifications. Similarly, accelerometer-based methods detect road undulations from the road vibration but must be properly calibrated for different vehicles and road surfaces and are hence impossible to use universally.

To overcome such problems, in this research we propose an intelligent pothole detection system using machine learning, sensor data, and real-time alerts. Based on computer vision techniques, deep learning models, and GPS-localization, the system will identify potholes and ascertain the severity. Using 3D pothole data generated, drivers will get real-time alerts for road obstacles ahead, so they can adopt precautionary measures and avoid accidents. The system can also prove to be beneficial for local administrations by marking the location of potholes and planning road maintenance more efficiently.

The creation of such a system demands meticulous focus on several factors such as hardware choice, data aggregation, and model optimization. The solution proposed will include aggregation of road data from various sources, training deep learning models for efficient pothole identification, and incorporating an easy-to-use interface for real-time notification. Bridging the current loopholes in pothole identification, this work will enhance road safety, reduce vehicle damage, and develop a general enhanced driving experience for drivers.

IV. METHODOLOGY

Existing System

Today, pothole detection and road maintenance rely on a chain of traditional and half-automatic methods. While these methods have been to some extent useful, they do not possess accuracy, efficiency, and real-time application. Traditional pothole detection systems can be categorized as follows: **Manual MRI Segment**

1. Manual Inspection and Citizen Reports

The most widespread practice among municipalities is manual road surveys, wherein the authorities inspect the roads personally to spot potholes. Second, governments and civic organizations provide citizens reporting devices in the form of helplines, websites, or mobile apps through which citizens can report potholes. All these practices, however, involve hard work and time. And inefficient, leading to delayed road repairs. Citizen reports may also be inconsistent, with some potholes going unnoticed or unreported.

2. Image Processing-Based Detection

Researchers have investigated the use of image processing algorithms in pothole detection from road images and videos. Edge detection, histogram equalization, and texture analysis algorithms are applied to detect potholes according to shape and contrast change. The algorithms are very sensitive to illumination, road textures, and camera directions and thus produce false alarms and wrong results in real-world usage.

3. GPS-Based Pothole Mapping

Several studies have proposed GPS-based pothole mapping, where vehicles equipped with sensors record the location of road anomalies and share the data with a centralized database. This approach helps authorities track pothole-prone areas and plan maintenance schedules. However, the accuracy of GPS alone is limited, and the system does not provide real-time alerts to drivers.

Proposed System

To overcome the limitations of existing pothole detection systems, the system utilizes deep learning, sensor data, and real-time alarming to make efficient and accurate pothole detection pos

1.Data Collection Using Multi-Sensor Approach

The system integrates multiple sensors like cameras, accelerometers, and GPS to gather rich road condition information.

• **Cameras:** Record high-definition road images and videos to identify potholes through deep learning methods.

• **Accelerometers & Gyroscopes**: Measure vehicle vibrations and angular shifts to detect road irregularities.

• **GPS Module**: Logs the precise location of detected potholes, enabling real-time mapping.

2.Machine Learning-Based Pothole Detection

A high-precision pothole classification is obtained by a deep learning-based trained Convolutional Neural Network (CNN).

•The model is trained on a broad dataset of road images taken under different weather and lighting conditions to provide resilience.

• Preprocessing operations, including edge detection and contrast enhancement, enhance detection effectiveness.

•Pre-trained models such as Reset or Mobile Net are used with transfer learning to improve accuracy and decrease training time.

3.Real-Time Pothole Detection and Classification

The model, having been trained, is deployed in a vehicle-mounted embedded system that continuously observes road conditions in real time.

• Potholes are identified from live camera images, and the severity level (shallow, medium, deep) is classified.

• Accelerometer measurements supplement image-based detection to establish the existence and severity of potholes.

• The system minimizes false positives through multi-sensor validation.

IMAGE PROCESSING

Pics provide context, and every picture is different. It contains important figures, which may be beneficial in various ways. These figures can be accessed through a method known as image processing.
Image processing is one of the building blocks of computer vision, and it is extremely important in most applications in real-world scenarios, such as robotics, autonomous cars, and object detection. It enables us to process and convert thousands of images at a time and extract meaningful information from them. It has wide-ranging applications in industries around the globe.

Python is one of the widely used programming languages for image processing due to its remarkable libraries and tools that efficiently handle complex image-related tasks

What is Image Processing?

As the name implies, image processing is the process of carrying out operations on an image to meet a particular objective. This ultimate output may be in the form of an improved image or modified form to be utilized for additional analysis and decision-making.

An image can be represented as a function **F(x, y)**, where **x** and **y** are the spatial coordinates. The magnitude of **F** at any given point represents the intensity of the image at that location. If the values of **x, y**, and intensity are finite, the image is referred to as a **digital image**.

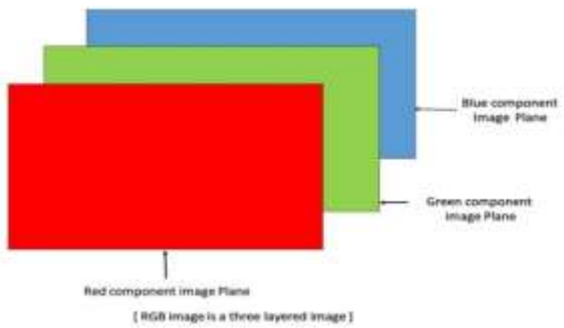
Types of Images:

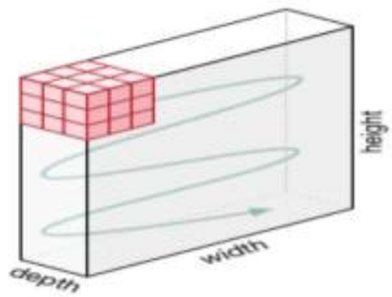
- **RGB Image**: Contains three layers of a 2D image: red, green, and blue (RGB) channels.
- **Grayscale Image**: Consists of shades of black and white and includes only one channel.

Convolutional Neural Networks (CNN) in Image Processing

CNNs are widely used for image recognition and analysis. A CNN typically consists of three primary layers:

1. **Convolution Layer (CONV)**:
This is the most basic form of a CNN, where the convolution is performed with a kernel (filter). The kernel moves over the image, shifting horizontally and vertically as per the value of stride until the image is fully covered.





V. CONCLUSION

Lastly, this study tested the impact of traffic sampling on machine learning-based network intrusion detection techniques. We tested the performance of four different machine learning techniques between sampled and unsampled network traffic data. From our findings, we observed that machine learning techniques for network intrusion detection were impacted significantly by sampling. Specifically, under heavy network loads, sampling tends to bring about a significant reduction in the performance of machine learning techniques. Conversely, under regular network loads, sampling can be an excellent way of streamlining machine learning techniques.

We further discovered that performance of the machine learning model depended on the sampling rate and using a greater sampling rate would ensure improved performance. From our work, we are suggesting the adoption of both the unsampled network traffic data as well as sampled network traffic data in a hybrid approach in order to achieve accurate network intrusion detection with high scalability and computational efficiency. Our suggestion is that model training can utilize unsampled data while sampled data can be utilized for testing as well as for model evaluation. In a nutshell, our paper emphasizes the value of diligent measurement of traffic sampling impact on machine learning-based network intrusion detection and presents an actionable hybrid approach to balance accuracy with efficiency under realistic conditions.

VI. FUTURE ENHANCEMENT Future Enhancements:

Over the coming years, the system will be optimized through the integration with advanced deep learning algorithms and real-time processing. Integrations like increased accuracy based on better datasets, cross-sensor integration to enhance detection, and adaptive algorithms for learning will enhance the performance of the system. Cloud deployments will even allow for largescale processing to enable broader availability and integration in the smart city system.

Besides, AI-driven automation in decision-making post-image processing output can also render the system more user-friendly. Utilization of real-time alerting via mobile apps or IoT devices can also enable instant alerting and proactive response for road safety or maintenance.

As AI and computing have evolved, it is now possible for the system to include self-learning mechanisms, where the models get better over time based on user feedback and changing environmental conditions. This will enhance the responsiveness of the system to changing real-world conditions.

Additionally, creating an intuitive dashboard with interactive charts will be capable of empowering decision-makers to efficiently analyse the data. Integration with GIS mapping services will enable authorities to dynamically monitor and schedule maintenance work.

improving the overall performance of the solution

1. **Integration of Multi-Omics Data:**
Incorporate genomic, transcriptomic, and proteomic data alongside MRI scans to improve the accuracy of molecular subtyping and provide a more comprehensive understanding of glioblastoma heterogeneity.
2. **Real-Time Clinical Deployment**: Optimize the model for real-time processing and integrate it into clinical workflows through cloud- based platforms, enabling seamless deployment in hospital settings.
3. **Incorporation of Explainable AI (XAI):**
Develop interpretable models that provide visual explanations and confidence scores for segmentation and classification outcomes, enhancing trust and transparency in clinical decisions.
4. **Adaptive Learning Models:**
Implement models that continuously improve by learning from new data, allowing for better.

VII. REFERENCES

1. Workshop on Mobile Computing System and Applications, 2021
2. R. Gass, J. Scott, C. Diot, “Measurements of In-Motion 802.11 Networking,” IEEE 3. X. Zhang, J.K. Kurose, B.N. Levine, D. Towsley, H. Zhang, “Study of a bus-based disruption-tolerant network: mobility modeling and impact on routing,” 13th annual ACM international conference, 2021
4. “http://dev.emcelettronica.com/datasheet/ st/LIS3L06AL,” Datasheet of ST LIS3L06AL accelerometer, JAN 2021
5. “http://www.its.dot.gov/vii,” RITA | ITS |
Vehicle Infrastructure Integration, JAN 2021
6. “http://www.gps.gov/,” Global Positioning System, JAN 2022
7. J.W. Byers, M. Luby, M. Mitzenmacher, “A Digital Fountain Approach to Reliable
Distribution of Bulk Data,” SIGCOMM, 2022
8. M. Mitzenmacher, “Digital fountains: a survey and look forward,” Information Theory Workshop, 2021. IEEE, 2021
9. “Pothole detection System using WiFi,”
MTech project report submitted by Shonil
Vijay, JUL 2021
10. “FireBird Reference manual,” Embedded and real-time systems lab, Computer Science and Engineering Department, IITB,
2021