Scalable Computing: Practice and Experience, ISSN 1895-1767, http://www.scpe.org
© 2024 SCPE. Volume 25, Issues 4, pp. 2159–2174, DOI 10.12694/scpe.v25i4.2800

SPATIAL AND TEMPORAL CHARACTERISTIC ANALYSIS BASED LONG SHORT-TERM MEMORY FOR DETECTION OF SENSOR FAULT IN AUTONOMOUS VEHICLES

HONGWEI ZHANG[*], YANAN GAO[†], HUANXUE LIU[‡], AND YI CHEN[§]

Abstract. The artificial intelligence required to create self-directed automobiles relies heavily on the capability of precisely perceiving the environment around oneself. Most self-driving automobiles include several detectors, which work together to form a multi-source perception of the surroundings. Extended use of a system that drives autonomously will introduce a variety of worldwide and local failure indications due to the extreme sensitivity of the instruments involved to ambient or environmental situations. These failure indications pose significant risks to the technique's security. The paper presents a real-time information synthesis system incorporating techniques for identifying flaws and accepting faults. The compact connection can be recognized if the qualities mentioned above are provided, and the input information properties may be retrieved in real-time. One way to use the newly introduced method for assessing device reliability is to compute the detectors' worldwide and local degrees of trustworthiness. In order to ensure the precision and dependability of information combination, problem data is filtered out, and monitor duplication is used to assess both the worldwide and local assurance levels of data from sensors at the moment. The chronological and geographic association of data from sensors allows for this. Experimental findings show that the network's algorithms can outperform current techniques in terms of both rapidity and precision and can pinpoint the object's location even when specific sensors are blurry or broken. This research established that the proposed hybrid structure benefits autonomous vehicles' real-time reliability and speed.

Key words: Self-driving cars, Observation of objects, Combining statistics, Automatic failure analysis, Neural network with convolutions, Combining data, Robot on wheels.

1. Introduction. In the process of driving, unmanned vehicles need to rely on a large number of sensors to read and receive real-time road condition data from roadbed units, traffic signals, radar base stations and other vehicles. In order to achieve uninterrupted navigation with high reliability, multiple sensors are usually installed on the vehicle body at the same time to obtain sufficient redundant position information.Generally, unmanned vehicle network attacks are mainly divided into internal attacks and external attacks, external attacks mainly invade equipment other than unmanned vehicles, such as attacks and interference to infrastructure such as roadbed measurement units, traffic signals, road signs, communication equipment and other vehicle data; Internal attack is mainly to invade the system and sensor of the unmanned vehicle body, typically including the vehicle control system, navigation sensor, vehicle diagnosis system and other vehicle sensors.Among the many attack types, the attack of navigation sensor is the most direct and common And one of the most harmful types of attack it can be on people at all Noticeable cases of sudden or slow changes in the car's driving vehiclePath, travel speed and direction, the attack is extremely stealthy, giving people The psychology of creating a huge fear is the current industry forced to solve One of the key technical problems.

Recognizing obstacles on the road is crucial for entirely autonomous vehicles. The method provides self-driving automobiles with real-time data about many road elements, including other automobiles and humans [5]. Existing object detection technology uses cameras and other types of ambient awareness devices. The

[*]1. Department of automotive engineering, Hebei Petroleum University of Technology, Chengde, Hebei, 067000, China; 2. Hebei Engineering Research Center of Autonomous Driving Functional Safety, Hebei, 067000 China (HongweiZhang3@126.com)

[†]1. Department of automotive engineering, Hebei Petroleum University of Technology, Chengde, Hebei, 067000, China; 2. Hebei Engineering Research Center of Autonomous Driving Functional Safety, Hebei, 067000 China (Corresponding author: YananGao26@163.com)

[‡]1. Department of automotive engineering, Hebei Petroleum University of Technology, Chengde, Hebei, 067000, China; 2. Hebei Engineering Research Center of Autonomous Driving Functional Safety, Hebei, 067000 China (HuanxueLiu8@126.com)

[§]1. Department of automotive engineering, Hebei Petroleum University of Technology, Chengde, Hebei, 067000, China; 2. Hebei Engineering Research Center of Autonomous Driving Functional Safety, Hebei, 067000 China (YiChen685@163.com)

machine learning system can detect the object that interests it and decide its proper classification thanks to the camera's capacity to gather data about the surroundings, including its color and outlines [14].

In the 1990s, artificially generated characteristics combined with data mining were used for most road object detection tasks. HOG, LBP, and Haar characteristics are examples of synthetic characteristics that were popular then. A set of feature vectors will be generated when artificial characteristics are extracted from discrete areas of the image [9]. By classifying vectors of features using characteristic algorithms like SVM, it is feasible to ascertain both the location and classification of objects of interest.

This article will be the cutting edge of artificial intelligence The deep learning technology is introduced into the anomaly detection of unmanned vehicle navigation sensors In this paper, we propose a novel approach based on one-dimensional convolutional neural networks (1D Convolutional neural network, 1D-CNN) driverless car guide Intelligent diagnosis method and system framework of aero sensor anomaly. The model consists of Input layer, 1D convolution layer, 1D pooling layer, 1D global mean pooling Layer (1D Global average pooling, 1D-GAP) and Softmax Classifier composition.With the advent of machine learning, neural networks based on convolution have taken over the role of visual sensors in performing object identification. The AlexNet system reduced the error rates for the Top-5 categorization task to 15.7 percent [4]. ResNet divides the network into many blocks, each with its own output and input parameters, and connects them directly using a Shortcut architecture to solve the gradual vanishing problem. In order to accomplish object identification, the RCNN network proposed employs three phases: an extraction of features backbone, a network for proposing regions (RPN), and a region of interest (RoI)-Pooling. This addresses the time-consuming issue of typical algorithms for identifying objects repeatedly traversing sensor data. Influenced by RCNN, scientists are improving the effectiveness of identifying objects in certain circumstances, like traffic. Using the image frame difference data from traffic camera footage to teach the neural network and monitor the moving automobiles in the video in actual time improves the framework of RCNN. The goal was to improve the RCNN's motion detection capabilities. The investigation uses the Faster-RCNN algorithm combined with the selected search method to locate forward autos, which significantly helps with the automobile loss problem. In order to effectively detect detached moving vehicles and achieve a high degree of accuracy, employ CNN to process the data gathered by UAVs. To increase the subject's recognition rate to 25 frames per second (fps), we propose using a one-stage system to complete the bounding box and perform classification and regression simultaneously. This will fix the problem of the multiple-phase model's subpar performance in real-time [16].

Data fusion approaches are being developed for processing data from several sources on environmental perception. These techniques provide superior object identification results compared to using only one of the sensors available on autonomous cars. This will lessen the burden on the autonomous vehicle's processing and memory systems. Data fusion may be broadly classified into two approaches: before and after the fusion. The locations of fusion reactions provide the basis for these classifications. In the former approach, sensor data from the initial input layer is pooled, and the object identification network is designed to function with all the information [3].

The post-fusion method's conflict location and methodology are more malleable and dynamic than the pre-fusion method's. Feature extraction (FE) is often used to extract characteristics from the data collected by every detector, and then specialized fusion procedures are constructed following predetermined goals [10].

The detector will eventually start sending out failure indications after prolonged use. The safe functioning of self-driving automobiles depends on several elements, including the stability of the algorithm and the capacity for error diagnosis and tolerance. The authors describe a cutting-edge transfer learning technique for the identification of transmission failure with an adverse domain choice machine. When there is little data to work with, transfer learning of this kind is an excellent tool for tackling the real GFR problem [6]. As shown, the proposal of a faster and more accurate deep learning system for exact equipment fault identification using transfer teaching and its achievement of cutting-edge results for the primary mechanical records and transfer acquisition have been shown to allow and speed up the development of deep learning networks with excellent precision.

In this study, we provide a novel approach to recognizing objects, emphasizing addressing concerns that have been overlooked in other efforts.

To be more explicit, we provide a novel data integration framework that is compact in design and can do

real-time analysis of massive amounts of multi-modal data. To guarantee the precision and dependability of road object recognition, this research presents a fault diagnosis and avoidance (FDA) mechanism inside the data fusion architecture. It identifies mutual fault through the time and space relationship between sensor data. Because sensor faults might occur while driving, this is done to prepare for them.

The following is a summary of the main results of this investigation.

We choose a structure that is a lighter feature pyramid network (FPN) compared to the previously employed object recognition network to ensure the real-time efficiency of the data synthesis system while analyzing massive volumes of multi-modal data.

In order to guarantee reliable and precise identification results, it is recommended that the information integration framework involve a method called fault detection and isolation (FDA).

Finally, tests are conducted to verify the instantaneous precision and dependability of the fusion architecture in terms of its efficacy in 2D object detection and defect prevention using surrounding sensory information gathered in various circumstances. These tests are performed to ensure the accuracy of information on how people perceive their surroundings [21].

2. Related Works. The LSTM model is based on a variant of Recurent NeuralNetwork (RNN). RNN model is proposed to solve the problem of time series data dependence and is very sensitive to short-term input. However, due to the characteristics of the tanh layer inside RNN, it is possible to remember or forget all the memories, which will lead to the possibility of gradient dispersion and gradient explosion in long-term time dependence. Compared with ordinary RNN,LSTM model is a more efficient tool to obtain long time series information. In order to solve the long-term time dependence problem, LSTM adds the mechanism of gate and introduces the concept of cell state.To prevent situations that could be harmful, self-driving cars must have access to precise and trustworthy issue identification and monitoring technology. This piece outlines our plan for an integrated approach to defect identification and monitoring. First, condition faults in the self-driving car are identified by training the intersection curve between the secure and risky domains using the One-Class Support Vector Machine (SVM) method. This is done so that secure network users may quickly identify themselves as in the secure zone. Meanwhile, the linear kinematic vehicle bicycle model is used to inform the development of a Kalman filtering observer that can anticipate where the vehicle is. To assess whether the course of events has changed, the test known as Jarque-Bera is used to determine if the distribution of probabilities of the leftovers between forecast and observation is normally distributed. This is done to see whether there has been a change in direction. We use a fuzzy system to further categorize the faults that have been discovered. This architecture is based on a neural network variant, adding an affiliation layer following the input layer. Black box testing refines the fuzzy system's initial membership function, revealing the likely failure scenarios. This is because neutral networks have an impressive capacity for self-learning. The system's capacity to identify and diagnose issues was evaluated through a battery of tests on the real autonomous vehicle platform, Xinda. The trials' findings and the results of performance comparisons with other fault sensors proved the methods' worth. In this last section of the report, we will briefly overview our efforts. Our goal in this study is to develop a diagnosis and fault-finding system for autonomous vehicles [2].

Sensor fault pre-detection method The residual sequence obtained by the predicted value and the measured value is the key to analyze the sensor fault. When the sensor is in normal operation, the expectation of the obtained residual sequence should be infinitely close to 0. However, in the actual collection process of the field, it will be affected by complex environmental factors and the system itself, and the expectation of the residual sequence will have a certain deviation from the 0 value, but the residual sequence obtained from the normal operation data still conforms to the normal distribution.

Approaches to fault detection and diagnosis may be broadly classified into three types: model-based, signal-based, and knowledge-based.

Hybrid approaches are employed when considering practical uses for autonomous vehicles. To start, we utilize a One-Class SVM to check for physical flaws in the car's condition.

The distribution is then inferred by constructing a Kalman filter and acquiring the residuals between the predicted value and the measured value. This is done in order to figure out who gets what. This approach of finding errors validates the normality of the residual distribution and establishes whether or not the trajectory deviates in a checking period. Ultimately, a fuzzy system is built to examine the likelihood of each possible

component that might cause failures. Our neural network's implementation acquires fuzzy system membership functions, and fuzzy system inclusion function parameters are adjusted via black box testing and fitting.

One component of future work is identifying the source of the problem.We call the many potential defect-causing factors "Advanced Alarm" and "Steering Alarm" to indicate an abnormality in the longitudinal or lateral condition. Based on this classification, we may examine the part of the system that seems to be malfunctioning. If the alarm type is "Moving Alarm", for instance, the dynamic system of the car or the ABS may have been to blame, whereas if the alarm type is "Steering Alarm", the steering system may have been to blame. Assuring the safety of the autonomous vehicle requires isolating the faulty system, which can only be done if we know which subsystem caused the problems. It is also essential to examine and address ELM-based approaches, which are suitable ways to improve the training process of neural network-based methods and display the features of fast learning speed and strong adaptation capacity.

Uncrewed aerial vehicles (UAVs) rely heavily on fault diagnostics to ensure the security and dependability of their equipment. Problem-solving for UAV instruments is addressed, and a combination of feature modeling and a neural network-based approach is presented. Errant signals from failed sensors, such as the global locating network, inertial measurements (IMU), and air data system (ADS), were acquired. In this study, we employed STFT to convert the residual signal into a time-frequency representation. The malfunction diagnostic of the UAV instruments was then carried out, and a convolutional neural network (CNN) was utilized to extract the map's features. Finally, UAV flying tests are used to assess the effectiveness of the suggested technique. Compelling fault diagnostic logic between leftovers and good health status may be built from visualizing data using CNN to extract sensor defect knowledge. This research presents an advanced learning-based defect detection system for UAV sensors that combines a hybrid feature model. We used the short-time Fourier transform (STFT) to prepare for deep learning and produce nine time-frequency mappings from the EKF estimate residue. The input of CNN was formed from the two-dimensional pictures by creating a CNN architecture with eight subnetworks. Using multiple validations, we decided on the variables to use. Lastly, in-flight trials prove that the suggested strategy works. The suggested method's principal benefits were its hybrid nature, allowing fault identification and deep understanding. Model-based defect detection may also be used with other forms of machine learning, such as neural networks with recurrent architecture and layered auto-encoding [7]. CNN is one of the most representative deep learning algorithms in the field of artificial intelligence. It uses deep network structure to simulate the working mechanism of animal vision system, and can automatically extract key features directly from the original data, with strong feature extraction ability.Different from pattern classification tasks in computer vision and speech recognition, the unmanned vehicle network attack anomaly diagnosis task requires both high accuracy of the diagnosis model and fast real-time detection speed.

The primary goal of self-driving vehicle studies is to improve transportation accuracy and reliability. Despite significant advancements, even the most advanced algorithms might occasionally make errors; in some instances, the sensors are to blame. Refrain from failing to do so might have disastrous results.

This is why self-driving cars must be able to foresee and respond appropriately to problems as soon as they happen. It is possible to train data representations with the help of artificial intelligence techniques by combining real-world data with data gathered from correctly operating sensors and then artificially injecting a range of sensor flaws into the resulting dataset. This research proposes a novel paradigm for identifying faults, separation, recognition, and predictions (depending on observation) for multi-fault in multi-sensor systems, such as self-driving cars. This layout was designed with fault diagnosis, isolation, and identification. By combining the strengths of two distinct and very efficient profound neural network architectures, our recognition, verification, and separation solution delivers remarkable efficiency.

We then utilize the sensor defect identification system's output to develop our gauge for the health index and feed it into the healthcare index prediction network's training procedure. This research looked at the underlying physical causes of sensor failures and classified them into four broad categories. The final goal was to develop a system for identifying faults, loneliness, proof of identity, and forecasting in-car electronics. Effectively handling many defects collected for various fault configurations that occurred in a multi-sensor scenario has been shown using a CNN-based sensor fault detector/classifier. This was demonstrated by its successful handling of the issues mentioned above. The proposed approach was evaluated and shown to be

99.95% accurate in finding errors. If a defect is detected, a signal is transmitted to the identity and isolation systems; if no malfunction is detected, the signal is forwarded to the detector's health prediction system. The outcome of the inspection for defects will inform this choice.

It has been shown that the proposed architecture, in conjunction with deep learning algorithms, may allow autonomous vehicles to perform fault isolation and identification. The signals will be sent to the system that evaluates and forecasts sensor health if the defect identification system finds the sensors functioning correctly. Three unique decay curves were used to help the detector identify defects in the system's output to get an HI value. This has been done so that we may speculate on the sensors' health. The HI metric is the information sent into the TFT network to foresee how the sensors will act and spot issues before they become critical.

This study employs a quantified reduction as a measure of effectiveness. Accordingly, obtainable losses of 0.0315, 0.0611, and 0.0299 at the 10th, 50th, and 90th percentiles have been reported. Audi, a German automotive manufacturer, has made public a dataset from the real world dubbed A2D2, which was used to validate the proposed method [22].

When coping with the challenging finding of damage challenges common in mechanical components, one possible option is to use a fault diagnostic strategy based on a combination of information from numerous sensors. Two potential problems with this approach are the challenge it may be to separate characteristics from many sources of sensory data and how challenging it can be to determine the optimal fusion level. Choosing the most appropriate feature or fusion level for a given fault diagnosis task can be challenging. This procedure requires a great deal of human work in addition to a deep understanding of the subject matter. We provide a multi-sensor data fusion approach to defect identification using deep convolutional neural networks that is both flexible and effective. To get around those two issues, we resort to this. The described method can automatically adapt to the requirements of each defect diagnosis task by learning features from raw data and maximizing the most practical combination of different fusion levels. In order to evaluate the effectiveness of the proposed technology, the global transmission test rig is used. In this study, we evaluate the performance of two traditional intelligent models, a back-propagation neural network and a support vector machine, in contrast to characteristics and integration levels that were created by hand, chosen by human judgment and applied to data from a single sense. Evidence from this trial suggests that the proposed technique outperforms all others in its ability to diagnose problems with the planetary gearbox quickly and precisely. To assess the condition of planetary gearboxes, this research presents a DCNN-based adaptive data fusion method. Information combination, a combination of features, selection combination, learning features, and detection of errors are adaptively merged into a single DCNN model. The proposed method utilizes the deep-layered configuration of DCNN to learn features from raw data while also adaptively fusing data, features, and choices. In addition, less time and effort from humans are needed to extract features and establish fusion levels when using this technique. The experimental results obtained using the planetary gearbox fault test apparatus serve as a gauge of the proposed method's efficacy. In this study, we put two widely used AI methods, BPNN and SVM, through their paces and compared their performance on various tasks, including feature-level fusion, decision-level combination, features created by hand, individual sensory information, and more. The experimental results provide persuasive proof that the proposed method works. When compared to other approaches utilized in the experiment, the suggested method achieves the best degree of testing precision.

In the future, we want to expand the range of mechanical items, failure modes, operating situations, and sensor types on which we conduct DCNN model-based learning of features and data fusion experiments. This will help us validate the techniques' efficacy and find further practical application recommendations. Also, due to the large number of variables in deep learning models, traditional parameter tuning can take a long time and result in a local optimum. At the same time, standard automated searching techniques can settle on a suboptimal solution quickly.

In order to adjust the parameters immediately, it is vital to investigate ways that are both more effective and faster. In summary, combining several neural network systems should enhance fault identification quality. Incorporating a self-encoding structure into the framework may increase its capability to learn more complex properties, making it particularly well-suited for forecasting potential problem situations. The model might be suited for such forecasts if the recurring structure was added [8].

Due to its importance, academic journals have extensively covered fault identification. To this end, several

machine learning algorithms have been created to better understand the diagnostic process and its constituent parts. In this study, we introduce a framework for learning termed a Hybrid Deep Belief Network (HDBN) for intelligent fault identification in motor-driven devices. The system combines data in various ways. The drive system of an automobile is an example of such a system. Based on our research, we provide three unique approaches to data fusion: data partnership, information arrival, and information hybridization. Looking at the fusion from the perspective of the energy contained in the signal might also shed light on its significance. More precise issue identification may be possible with the correct data fusion methods and formats suitable for model training purposes. In addition, a novel fusion method called mixed-precision training is employed to improve the model's overall performance. Our proposed model outperforms state-of-the-art methods, as shown by experiments performed on datasets provided by the modeling platform. As part of this body of work, we propose using data sources and their integration outcomes to build diagnostic models that extract depth information more efficiently. After establishing a DBN-based deep fault feature learning strategy, the HDBN diagnostic model was built using the hybrid data fusion technique. Second, three separate data fusion methods were developed in line with the nature and interpretation of the tangible form of the signal, rather than employing "information unification or information blend" as was done in the previous research.

The main contributions of this paper include the following three aspects: Firstly, a new scheme of network attack detection using the original location data of multi-sensor of unmanned vehicle is presented; Secondly, the deep learning technology is introduced into the field of unmanned vehicle network attack detection to solve the difficult problem that slow and small attacks are difficult to detect through its powerful feature extraction capability. Third, the existing 1D convolutional neural network algorithm is improved, and the 1D-GAP-CNN algorithm structure is designed, so that it is more in line with the fast real-time requirements of unmanned vehicle network attack detection.

The results of the experiments showed that the different data-fusing approaches were related to the true qualities of the data and that the learning accuracy could be improved by combining the correct data fusion methodology with the appropriate manner of dataset construction. Third, we applied a novel data fusion approach called mixed-precision learning to improve the model's overall accuracy significantly. The experimental findings showed that HDBM's artificial intelligence and information integration abilities significantly improved cognitive defect detection accuracy, far outperforming the other diagnostic systems. The author intends to continue investigating the problem of fault diagnostics in complicated environments. As part of this study, we will refine diagnostic procedures for usage in settings with a mix of loading and driving input interference [28].

3. Proposed Methodology.

3.1. Proposed Layout of Networks for Geographic Areas.

1. Feature Extraction Network (FEN)

This paper is about the network layer Number, size of convolution kernel, number of convolution kernel, activation function, pooling kernel The size and the number of pooling layers are adjusted repeatedly. The model consists of 10 network layers. There are five 1D convolution layers, two 1D pooling layers, and four RELUs Activation layer, 1 1D-GAP layer and Softmax output layer, all Both the convolution kernel and the pooling kernel are set to Padding= "Same", using Adam adaptive learning rate optimizer and mini-batch training method, 65 samples per batch, 210 rounds of training.Once a frame of a picture is acquired, it must be compressed, encoded, and run through an artificial brain to recover the structural data hidden in the spaces between the frames. The larger the perceived region of the resulting component map, the higher the FEN's complexity, but the more local detail is lost in the process [15]. The FEN presented here is based on the architecture of a neural network with residuals. This is done in order to tighten up the network's settings. The quantity of computational horsepower required by the system may be significantly lowered if, during the creation of the framework procedure, a priori characteristics are included to lower the size of map structures. The network's basic structure is laid out in Table 3.1.

The computational capacity of the network may be used to roughly estimate the quantity of work that must be performed by the neural network's computation and memory components.

Computational energy within the system may be used as a proxy for the workload of the brain system's computation and memory components. The computational capability of a CNN level may be approximated

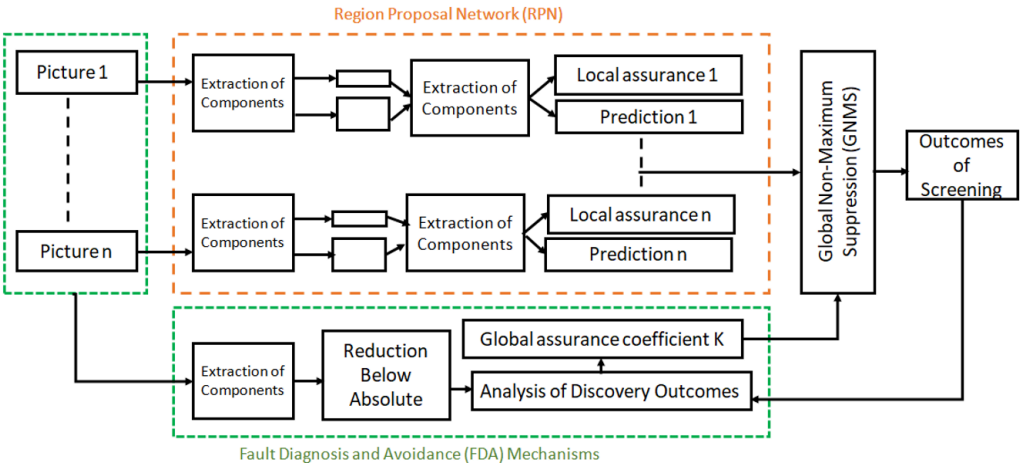


Fig. 3.1: Structured integration of information

Table 3.1: Feature extraction structure

	Category	Channel	Filter size	Outcome
	Input	4		416*416
	Ramification	32	3*3/2	104*104
2x	Ramification	64	3*3	52*52
	Ramification	64	3*3	
	Ramification	64	3*3	
	Remainder			52*52
2x	Ramification	64	3*3/2	26*26
	Ramification	64	3*3	
	Ramification	64	3*3	
	Remainder			26*26
2x	Ramification	128	3*3/2	13*13
	Ramification	128	3*3	
	Ramification	128	3*3	
	Remainder			13*13
	Required computing power: $0.69*10^9$ FLOPS			

using Eq. (3.1):

$$P = 2IX(C_{\text{in}}k^2 + 1)C_{\text{out}} \tag{3.1}$$

I and X stand for the length and breadth of the map of features, respectively; C_{in}, C_{out} stand for the number of input/output channels; and k stands for the length of the kernel.

The FEN uses a two-layer, 3x3 convolutional framework to reduce the size of the input picture. This takes 3.6 floating-point operations of processing power at the start. The technology strength for the system's Molecular weight layer is calculated to be 0.7*108 FLOPs, which is 1/27 of the calculating strength needed by DarkNet-53.

The compact FPN system uses a 26 x 26 and 13 x 13 feature map extracted from the core when building a characteristic structure. The bounding box extrapolation variables x, w, h as well as the desired category forecast probabilities Q(0), Q(1), and Q(2), are then computed to provide the output vector "x", "y", "w", "h", "C", and "Q(0)", "Q(1)", and "Q(2)". Remember that the i^{th} sensor will capture the predicted D_i area

after passing over the ultralight FPN connection. The GNMS component, discussed in more depth below, will synthesize data from all of the sensors to predict D_i and the related actual time probability. K_i, the RPN unit, determines where the goal D_{fusion} is in the 2D image and communicates this knowledge to the box-regression system, an element of the knowledge-merging structure [23].

2. Region Return on Investment Estimation

After the initial picture has been encoded and compressed by the FEN, numerous map files with different reduction levels have been generated. This is quite similar to the characteristic tower that the SIFT algorithm generates. Root regions are used to categorize the condensed and encrypted map of features in this piece of writing. The target's position and local probability data are reconstructed from these zones. Precisely, three roots of fixed dimensions are placed at each location of a feature chart in order to determine the object of interest using a grid-by-grid characteristic map. The quaternion vector "x, y, w, h" frequently expresses the anchoring point as a straightforward triangular boundary in both dimensions. The anchor coordinates and the enclosing box's dimensions combine to generate this data vector. The network can also predict the probabilities $P(0), P(1), \cdots$, and $P(n)$ that each Anchor's internal destination belongs to one of these categories. The output vector size of each item recognition level is determined by the system's capacity to distinguish between three unique goal classifications: people walking, cars, and bikes [18].

In order to establish a baseline for the Anchor information set, we first conduct K-mean clustering on the KITTI training set. With a more negligible size difference between the anchor and the target, the network can better pinpoint its location inside the local area divided by the Anchor. The final sizes of the six anchors obtained from clustering with $k = 6$ are (45, 42), (101, 61), (84, 168), (175, 98), (261, 160), and (377, 215). This study proposes two layers with features for the ultralight FPN system.

3. Loss Function

Following the completion of the system's framework, the impairment value must be used to direct the converged path of the network's settings in order to coax the system into continuous self-learning toward the specified quadratic value [19]. The coefficient of loss may be used to guide the development of the network's variables toward the target value. The system loss equation is divided into three subcomponents, each representing a distinct aspect of the algorithm's production vector: the boundaries box change loss, the leaping box object identification loss, and the objective categorization reduction. The entropy loss cross-product is utilized to calculate boundaries box change loss and the leaping box object identification loss since both the assurance of the box with boundaries and the expected value of the desired category are output as probabilities. However, the result of plugging in "x, y, w, and h" is a specific number, and this is what the square error loss function model uses to determine l_{bbox}. Each pixel in the final feature map is instructed to select the anchor with the highest Intersection-over-Union (IoU) score throughout testing. The standard setting of 1 has been assigned to these anchors, which equates to 1_{ij}; their anticipated value will be used for calculating l_{bbox} and l_{cls}. An anchor is considered empty and has no objects inside it if its IoU with any target is less than 0.5, and the associated 1_{ij} is set to 1. In addition, the anchor's expected confidence C will be used as a penalty term in the logarithmic shape to determine l_{conf}. This holds for anchors with a value of 0 in 1_{ij}.

Since the bulk of the information is made up of the anchor that does not have a target in the output, the value of grid2 i=0 A j=0 1_{ij} log C_{ij} will immediately affect the long calculation. A loss function of this kind makes no sense since the network consistently predicts C_{ij} to be very close to zero. We will manipulate the input data distribution using the weight coefficient 1_{ij} to achieve this objective. Similarly, pedestrian identification for autonomous driving focuses more on the accuracy of target position prediction than on more common tasks requiring object recognition [20].

The single-sensor object detection network supplies data fusion's "preprocessing" connection. The success of the data fusion process depends on the precision with which this network determines the target's position. Therefore, a more excellent weighting factor box should be applied to l_{bbox} when calculating the total loss. This will increase the "penalty" of the bounding box's regression error on the network. Following is the formula for determining the ultimate network loss function:

$$l = l_{\text{conf}} + l_{\text{bbox}}\lambda_{\text{bbox}} + l_{\text{cls}} \tag{3.2}$$

3.2. Detection and Correction of Errors. Adverse climates, such as smog and precipitation, may be encountered during the functioning of the autonomous automobile. The communication on the data bus may

be disrupted, or the camera's optics could become soiled, rendering the device useless. These problems will create noise throughout the information utilized for contextual awareness, which may create issues like the failure to recognize or incorrectly identify goals, both of which pose severe threats to the security of self-driving cars. An error-finding and evasion method in the perception model has been developed to ensure redundant information [12]. With the help of redundant information for fault testing, the accuracy of sensor data can be evaluated right now, and the problem signal can be eliminated. This study uses The variable weight method to execute the FDA procedure. The information fusion architecture calculates the global degree of trust K and local credibility level C for data from sensors. Problems with both local and global flaws may be solved with the use of these confidence levels. During the fusing of the RoI region, the GNMS method filters away redundant and faulty signals to ensure accurate union outcomes are returned as the output of object detection. The K and C symbols represent the assurance factors used in this procedure.

1. Investigation of Flaws in a Sensor That Picks Up on Our Environment

The sensor's scanning architecture renders the image information susceptible to electromagnetic contamination from the surroundings. The lens is just like this. Dim lighting, blurry surrounding details, and shifts in viewpoint will impair the detectors' ability to see details [13]. The lens can roughly classify the background noise as "natural" or "human-made".

i. There is a fault line throughout the world. Poor lighting, a busy background, or a blurry lens will all contribute to an overall haze in the image.

ii. Localized failure. Information truncation, complete obstruction, and lens staining are just a few of the issues plaguing the photoreceptor chip. Because of these problems, the chip cannot do its job.

When a self-driving car gets motivated, the driver's perspective, the background, and the available light all shift continuously. Our method for recognizing defects uses an adaptive load strategy, which enables the score to be modified in real-time based on the data provided from all frames of a reputation to ensure the reliability of the sensory technology. In real-time, we employ a Kalman filtering approach to assess the reliability of every sensor's information and its impact on the overall outcome. We also fine-tune the sensor's general trust parameter as time passes. We correct worldwide sensor failures in this manner. To cope with local flaws in real-time, an assurance parameter for each local area of the image is calculated using RPN. The structure for combining data employs GNMS to combine local and global factors for reliability and accuracy in the final results of the information merging procedure.

2. Adjustment of International Reliability Rates

The detection zones of multiple cameras in a self-driving vehicle overlap; thus, the collected environmental information is redundant. By combining sensor measurements in a manner that accounts for local and worldwide reliability parameters, an information merging system produces potentially ideal outcomes for recognizing an object. As Root Truth, this information will be sent back into the network of fusion devices and used to fine-tune each sensor's broad reliability factor L [17].

The broader reliability factor L was calibrated using the Kuhn-Munkres (KM) method and the filter developed by Kalman in this research. In order to solve the problem of locating the most potent weight pairing in the divided chart, the KM approach was used. In this research, we know both the anticipated following mark S from the frame before it and the object recognition outcome C from the current frame. They form one of the two subgroups of the divided graph G. The prediction made in the prior frame by the sensor mark T is known. The KM method decides which of the two links is more reliable based on a preset set of principles and guarantees.

$$\sum_{j=0}^{n-1} N_{ij} \leq 1, \sum_{j=0}^{m-1} N_{ij} \leq 1 \tag{3.3}$$

Only one possible grade can be assigned for every test outcome. Before the KM method can be run, the Pearson correlation coefficient between the divided graph's nodes must be obtained. The value of this coefficient represents the degree of similarity between the indicator and the detection outcome. This article uses the IoU between the box boundaries as the association factor. Using the cross-section ratio, which comprehensively expresses the similarity of the position and the dimension of the two box boundaries, simplifies and improves the correlation coefficient calculation. The ratio of the points of intersection is also a helpful tool. Any

permutations where $T(i,j)$ is below 0.25 will be omitted once the value of the correlation matrix has been calculated.

These pairs of bounding boxes do not contain the same object because they are far apart. The absence of a correlation coefficient for the omitted combination shows no connection between the two points in the issue. In order to find the optimal combination of L and M, the KM technique repeatedly investigates the remaining connections. This is done so that the maximum number of correlation coefficients may be obtained between Tracker and Recognition by utilizing this specific conjunction:

$$\hat{N} = arg_{\max} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} N_{ij} T_{ij} \tag{3.4}$$

After obtaining the detection-tracking corresponding matrix T, we use the Kalman filter to update Tracker's monitoring of object data. It is possible to define the presence or absence of vehicles and people on the road based on the target's position and velocity relative to the world's geographic coordinates. The Kalman filter works on the premise that the target follows a Gaussian distribution in the kth structure, as this is the best approximation to the destination's natural state that can be made [1]. The mean of this distribution's values is sk, and its variance is s. Targeted velocity between successive frames may be roughly interpreted as consistent linear motion if the sample rate of the ambient awareness sensor is large enough. This happens when the ambient awareness sensor has a high enough sample speed.

Conversely, the law of object movement cannot be a perfectly uniform linear motion. The longer the time delay t, the more significant the discrepancy between the projected and actual values. This means that the measured data returned from the sensor must be used immediately to correct the predicted value.

In this article, we aggregate the data gathered from the sensors to learn about the target's location, size, and direction in a three-dimensional space. The sensor's speed data may be acquired by dividing the time gap between recognizing and the Tracker's locations by their magnitudes. Add the sensor readings for position and speed and write the result down as zk. Then, the Kalman filter merges the distributions of sensor readings P_{sensor} and the goal state prediction P_{pred}. After the two distributions have been reunited, they must be re-projected onto the initial target space of states. This distribution will be saved in the relevant Tracker so that a prediction can be made on the intended state of frame k+1. The SK file's location data will be used to fine-tune the four cameras' global weights. The following procedures constitute this process:

1. The 2D object recognition system predicts the perimeter of an object using the conventional NMS technique, which involves creating separate forecast data D_i for each image component.
2. Using the point of view projection alterations, read the sk file to determine the raster dimensions of the target's position pk and subsequently store the resulting ck file.
3. Frame k's worth of forecasts should be used as the Ground Truth to get the mean IoU value for B_i under cw.
   Find the ck in D_i corresponding to each bounding box's topographical center representing a prediction. First, using D_i as the appropriate ck, build a bounding box and then calculate the IoU value of the box. Using a weighting scheme based on the assurance parameter C of D_i, we can get the average IoU value of D_i under ck.
4. Update the sum confidence for each sensor to be consistent with its average IoU. The following is an illustration of a possible method for updating the K_i reliability factor for the sensing device i:

$$\alpha = \frac{1}{2} \ln \left(\frac{mean(IoU)}{1 - mean(IoU)} \right) \tag{3.5}$$

3. Non-maximal Inhibition on a Global Scale

At time t 1, the camera's universal assurance level Ki is modified, and the outcome is sent back to the GNMS component for incorporation into the next information frame. After that, images captured at time t are sent into a data fusion system. After the compact FPN network has finished manufacturing, an ROI area with $\{D\}_{i=1}^{M}$ is built [27].

The GNMS method performs fault diagnosis in line with the present frame's reliability coefficient (K, C). Reliable fusion outcomes may be achieved when unnecessary or inaccurate sensor data is discarded.

Table 4.1: KITTI information complexity classification

Degree of complexity	Minimal case e levation (Pixels)	Extreme obscuration	Most severe abbreviation (%)
Simple	45	Visible	20
Moderate	30	Partial	35
Hard	25	Not visible	45

Unlike traditional algorithms for environmental awareness, GNMS may focus on combining data operations and limit input from sensors in real-time using calculated regional and global reliability factors.

After the picture region of interest (ROI) set $\{D\}_{i=1}^{M}$ is muted by the global non-maximum value, junk data is purged, and gaps are filled in.

The results of object identification utilizing shared sensor information have finally been obtained.

4. Experimentation & Results.

4.1. Assessing the Efficiency of Networks.

1. Laboratory Environment

The following sections will illustrate how these setups are used to produce all of the results. The model is first pre-trained for 0.5 million cycles in a batch size of 8 using the COCO dataset. The Adam optimizer's training begins with a learning rate 103 and is further enhanced at steps 400 thousand and 500000 by multiplying by a hyper-parameter of 0.01. The pre-trained model is then utilized to train at an ongoing pace for an additional hundred iterations on the KITTI dataset. The weighting coefficient of the bounding box regression loss specified previously has been set to 1 to facilitate faster convergence. Two RTX 2300Ti GPUs are used in every training session [24].

2. Evaluation of Object Identification Efficiency

Scientists now have a fully developed model evaluation mechanism due to the KITTI information set. The program improves upon the PASCAL VOC-created computational method called average accuracy. The system cares little about anything that is too little to be seen by the camera or too far away from it to be of interest [25]. The strictness of the model's requirements is reduced. In addition, KITTI's model evaluation method categorizes the job into simple, mild, and complicated categories, depending on the size of the target boundaries and the degree of obstruction. Table 4.1 provides a detailed description of the difficulties.

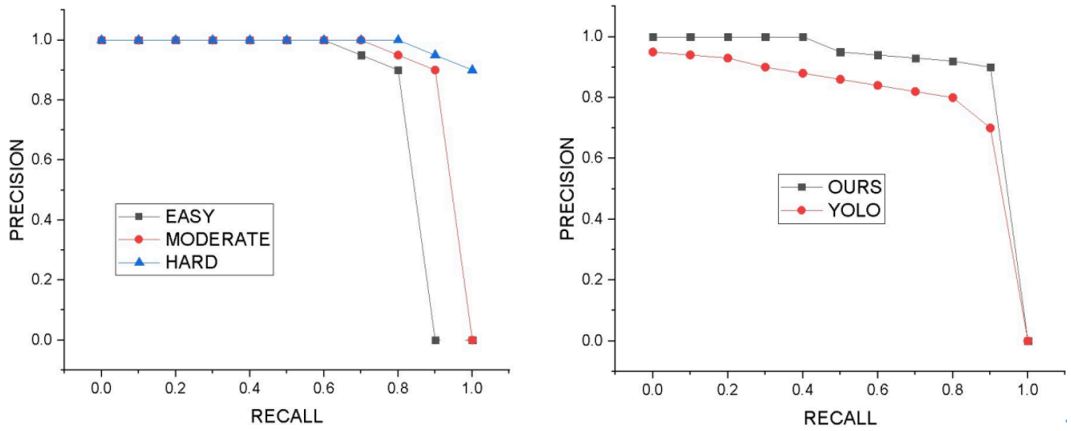
The items in the KITTI dataset are divided into three categories for evaluation. Vehicles, customers, and cyclists are all examples of these types. The AP70 measure should be used for vehicle goals, and the AP50 metric should be used for passenger and bicycle goals when assessing the efficacy of the network.

We chose to use the set-aside method to assess network performance because the test set that the KITTI provided lacked a clear name. The test set was built by selecting 2000 snapshots at random from the initial training set. Please remember that these 1,000 images are not part of the network's training [11]. Finally, we compile the evaluation findings of other notable networks on the Leader Board of the official KITTI website and use them to evaluate our network's efficiency. Table 4.2 displays our comparative findings.

In this article's network, the precision and recall (PR) curve is shown in Figure 4.1(a) for three different levels of complexity. During this process, the PR50 curve for YOLO_V3 is compared across all of the test sets, as shown in Figure 4.1(b). The object detection network must process four separate 2D images simultaneously. In order to reduce the number of network variables and the required amount of processing power, we have used several methods, including HOG feature extraction and residual networks. This was accomplished keeping the algorithm's real-time speed and memory size of the GPU in mind. Table 4.2 and Figure 4.1 show that even after extensive optimization of the network's settings and calculating velocity, our device identification system still ensures a high accuracy and recall rate, and its object detection capabilities can meet the current needs of self-driving cars in terms of the precision of their sense of the environment.

Table 4.2: Evaluating 2D image detection's effectiveness

Community	Existing Duration (s)	Simple (%)	Medium (%)	Difficult (%)
YOLO800	0.14	79.12	75.32	64.84
Fast-SSD	0.07	86.11	67.8	58.9
ResNet-RRC	0.07	92.43	86.34	75.25
MonoFENet	0.16	92.68	85.64	77.72
RefineNet	0.3	92.9	82.02	66.69
RetinaNet	0.3	94.3	83.74	69.38
Faster-RCNN	2.2	89.56	84.17	73.63
Proposed method	0.03	96.46	89.21	87.45



(a) PR curve for 2D image detection (b) Performance of 2D image detection compared with YOLO_V3

Fig. 4.1: Evaluating 2D image detection effectiveness

4.2. Effectiveness in Detecting and Preventing Problems. This paper's FDA architecture is developed primarily for two sorts of problems: worldwide problems and localized failures.

The study mimics the worldwide defects with the device's out-of-focus issue. When a photograph is down-sampled, its overall quality decreases, resulting in a blurrier, less distinguishable, and more difficult-to-find picture.

For the sensor's out-of-focus defect, we may utilize the down-sampling procedure as indicated in Eq. (10), where the input image matrix is I and the output fault image matrix is E.

$$E = I(i - (i \mod t), j - (j \mod t)) \tag{4.1}$$

where t is the number of rounds in the down-sampling process.

For regional issues, we employ the picture-shortening issue as a model, and we replicate the most severe jumbled impact by disrupting the vertical organization of the localized portion of the picture. This form of distorted coding additionally renders the system more vulnerable to identifying error issues, and the resulting impact is more noticeable in the test scenario since the error message is created on the foundation of the initial information [26]. Assuming the shuffling() method can cut and rearrange a stream of constant information, we can write an expression that represents the effect of regional garbling on the individuals:

$$E = I(\text{shuffle}(i), j) \tag{4.2}$$

Table 4.3: Evaluation of reliability under stress

Video routine	Out of focus (FDA)	OoF	Shortened (FDA)	Shortened
09-27-0010	0.8092	0.74	0.82	0.69
09-27-0018	0.8834	0.81	0.89	0.67
09-27-0052	0.7723	0.66	0.78	0.61
09-27-0057	0.8823	0.76	0.89	0.83
09-27-0058	0.8465	0.75	0.83	0.81

Table 4.4: Multi-sensor reliability assessment of efficiency

Video routine	Out of focus (FDA)	OoF	Shortened (FDA)	Shortened
09-27-0010	0.803	0.699	0.801	0.643
09-27-0018	0.882	0.743	0.872	0.612
09-27-0052	0.771	0.641	0.763	0.601
09-27-0057	0.880	0.740	0.873	0.822
09-27-0058	0.844	0.721	0.802	0.791

The full results are shown in Table 4.3. Table 4.4 shows the average IoU acquired by our intranet and the group serving as the control when two cameras experience both local and global faults at the same time.

When serving as a comparison group, the network uses the same system variables as when actively fusing information. However, it does not utilize the GNMS module or the Kalman filter section, instead relying only on the weight-based method. The software sensor failure simulation is applied to the KITTI Raw-Data dataset, and we then run our fusion framework and the control group to achieve object identification results. Next, we compare the Mean-IoU value from out-of-focus and truncated fault data to see how they compare. The FDA effectiveness of these two networks may be compared more understandably by drawing the mean IoU curve of the system below the footage data, with the total number of pixels functioning as the x-axis and the IoU value of each picture functioning as the y-axis.

Figure 4.2 depicts the mean IoU curve for this network and the conventional data fusion network with the sensor down using the video data with the identifier 09-26-0056 as input.

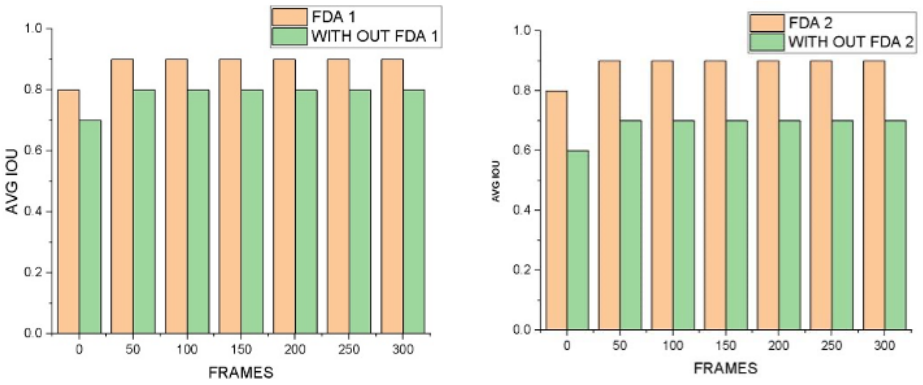
In order to evaluate the maximal defect capacity of the information integration architecture developed for this article, we inflict progressively severe global and local sensor malfunctions on the sensor data of numerous webcams and calculate the mean IoU value in the same manner. This enables us to learn the framework's maximum fault tolerance. Select the 07-19-0073 footage as the input and draw the IoU graph that occurs when both sensors fail simultaneously, as shown in Figure 4.3.

Figure 4.4(a) demonstrates that when the same defect is presented to 3 of the four recording devices, the problem-prevention mechanism of the network developed in this study malfunctions. The inability to recover has also led to a significant drop in the mean IoU gradient.

Difficulties arise because our system considers every finding outcome genuine and modifies the detector's load based on this information. The fault tolerance mechanism becomes useless due to the discrepancy between the first collision identification outcome and the actual reality. When subjected to the identical fault condition, every single camera will provide error findings that are comparable to one another. When garbled characters with different positions and ways of making them are sent to all three cameras, as shown in Figure 4.4(b), and the mean IoU curve is calculated from this data, the faults added to the sensor are correctly removed.

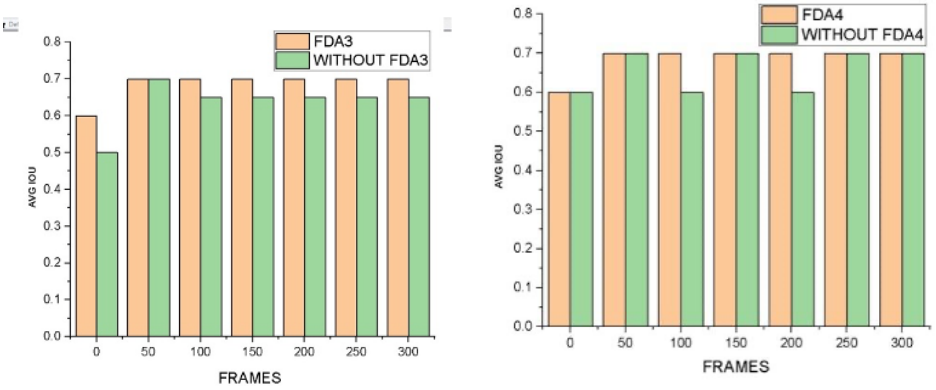
The effect of error messages on how well an object is recognized and how well our system works to get rid of problem detection findings can be shown more clearly by distributing the detection findings of the information combination structure described in this study and comparing them with the initial image information for the control group.

Tests show that when two or more cameras have accurate data, the fault-detection approach developed for



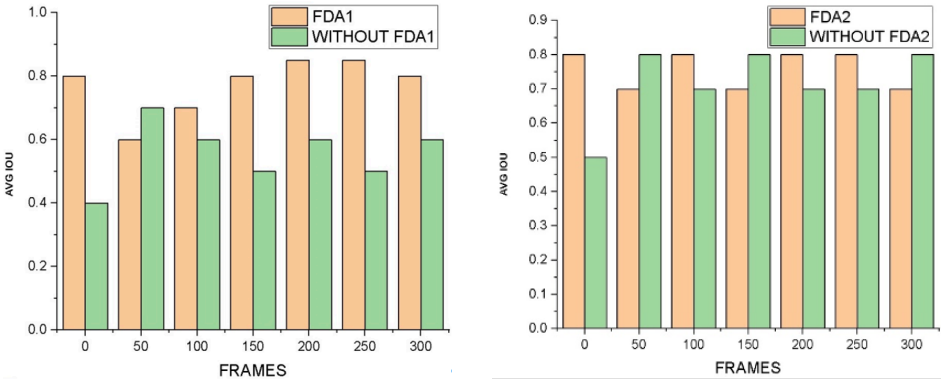
(a) Out of Focus Failure (b) Truncation Failure

Fig. 4.2: Performance of FDA



(a) Out of Focus Failure (b) Truncation Failure

Fig. 4.3: Evacuation of FDA



(a) Create a similar flaw] (b) Use a variety of failure scenarios

Fig. 4.4: Assertion from the Food and Drug Administration (three defective sensors)

this research can effectively filter out the distracting data caused by the sensor's global and local flaws. Global and local conviction procedures may eliminate perception mistakes during network functioning in time, and they can make full use of more precise data from sensors to provide the optimal combination impact even if the sensor in question does not have a substantial failure event. This is the case even if no significant failure occurs with the device.The results show that The proposed method can effectively diagnose the unmanned vehicle GPS sensor The same anomaly under cyber attacks, compared to mainstream intelligence The diagnostic algorithm has higher diagnostic accuracy and faster anomaly detection Measure the speed.

5. Conclusion. In this paper, an improved navigation sensor for unmanned vehicles is proposed to solve the problem of being vulnerable to hackers 1DGAP-CNN's new deep learning algorithm for detecting unmanned vehicles Driving a car's GPS navigation sensor due to physical failure or cyber hacking A variety of abnormal problems arising from intentional attacks.In this paper, one dimensional convolutional nerve, which is at the forefront of deep learning, is firstly introduced The network algorithm is introduced into the field of attack diagnosis of unmanned vehicle sensor network. Secondly, the real-time requirement and realization of unmanned vehicle network attack diagnosis Row 1D-CNN algorithm parameters too many insufficient, through the design of one A one-dimensional global mean pooling layer to replace the current 1D-CNN Flatten layer and the fully connected network part of Layers 2 to 3. The results show that The improved method can effectively reduce the number of model parameters and improve them And the diagnostic accuracy rate reached more than 99%.

1. An information integration architecture with a defect diagnostic method has been suggested to address the problem of identifying objects in self-driving situations. This structure eliminates the background data provided by the instruments during external information gathering. It allows for the reciprocal detection of defects amongst perception devices by establishing global and local reliability values.
2. Experiments with the KITTI collection show that the information integration system proposed in this study significantly reduces the load on the computing and storage units of the multi-source environmental monitoring technology without affecting the accuracy of item detection.
3. If the image sensor has an extensive global or local flaw, the data integration system would be able to erase malfunctioning data in a timely way using the FDA technique proposed. Doing so will guarantee trustworthiness and precision in the merged dataset.

REFERENCES

[1] W. Bao, B. Xu, and Z. Chen, *Monofenet: Monocular 3d object detection with feature enhancement networks*, IEEE Transactions on Image Processing, 29 (2019), pp. 2753–2765.

[2] Y. Fang, H. Min, W. Wang, Z. Xu, and X. Zhao, *A fault detection and diagnosis system for autonomous vehicles based on hybrid approaches*, IEEE Sensors Journal, 20 (2020), pp. 9359–9371.

[3] J. Fritsch, T. Kuehnl, and A. Geiger, *A new performance measure and evaluation benchmark for road detection algorithms*, in Proceedings of the 16th International IEEE Conference on Intelligent Transportation Systems, The Hague, Netherlands, 2013, IEEE, pp. 1693–1700.

[4] Z. Gao, C. Cecati, and S. X. Ding, *A survey of fault diagnosis and fault-tolerant techniques—part i: Fault diagnosis with model-based and signal-based approaches*, IEEE Transactions on Industrial Electronics, 62 (2015), pp. 3757–3767.

[5] R. Girshick, *Fast r-cnn*, in Proceedings of the IEEE international conference on computer vision, Santiago, Chile, 2015, pp. 1440–1448.

[6] R. Girshick, J. Donahue, T. Darrell, and J. Malik, *Rich feature hierarchies for accurate object detection and semantic segmentation*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, pp. 580–587.

[7] D. Guo, M. Zhong, H. Ji, Y. Liu, and R. Yang, *A hybrid feature model and deep learning based fault diagnosis for unmanned aerial vehicle sensors*, Neurocomputing, 319 (2018), pp. 155–163.

[8] L. Jing, T. Wang, M. Zhao, and P. Wang, *An adaptive multi-sensor data fusion method based on deep convolutional neural networks for fault diagnosis of planetary gearbox*, Sensors, 17 (2017), p. 414.

[9] E. Khalastchi and M. Kalech, *Fault detection and diagnosis in multi-robot systems: A survey*, Sensors, 19 (2019), p. 4019.

[10] S. Li, L. Ding, H. Gao, Y.-J. Liu, N. Li, and Z. Deng, *Reinforcement learning neural network-based adaptive control for state and input time-delayed wheeled mobile robots*, IEEE Transactions on Systems, Man, and Cybernetics: Systems, 50 (2018), pp. 4171–4182.

[11] S. Li, H. Wang, L. Song, P. Wang, L. Cui, and T. Lin, *An adaptive data fusion strategy for fault diagnosis based on the convolutional neural network*, Measurement, 165 (2020), p. 108122.

[12] T.-Y. Lin, P. Dollár, R. Girshick, K. He, B. Hariharan, and S. Belongie, *Feature pyramid networks for object detection*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA,

2017, pp. 2117–2125.
[13] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, *Focal loss for dense object detection*, in Proceedings of the IEEE International Conference on Computer Vision, Venice,Italy, 2017, pp. 2980–2988.
[14] T.-Y. Lin, M. Maire, S. Belongie, J. Hays, P. Perona, D. Ramanan, P. Dollár, and C. L. Zitnick, *Microsoft coco: Common objects in context*, in Proceedings of the 13th European Conference, Zurich, Switzerland, 2014, Springer, pp. 740–755.
[15] X. Liu, W. Liu, T. Mei, and H. Ma, *A deep learning-based approach to progressive vehicle re-identification for urban surveillance*, in Proceedings of the 14th European Conference, Amsterdam, The Netherlands, 2016, Springer, pp. 869–884.
[16] Y. Qu, L. Jiang, and X. Guo, *Moving vehicle detection with convolutional networks in uav videos*, in Proceedings of the 2nd International Conference on Control, Automation and Robotics, Hong Kong, China, 2016, IEEE, pp. 225–229.
[17] R. N. Rajaram, E. Ohn-Bar, and M. M. Trivedi, *Refinenet: Refining object detectors for autonomous driving*, IEEE Transactions on Intelligent Vehicles, 1 (2016), pp. 358–368.
[18] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, *You only look once: Unified, real-time object detection*, in Proceedings of the IEEE Conference on Computer vVsion and Pattern Recognition, Las Vegas, NV, USA, 2016, pp. 779–788.
[19] J. Redmon and A. Farhadi, *Yolo9000: better, faster, stronger*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 2017, pp. 7263–7271.
[20] ———, *Yolov3: An incremental improvement*, arXiv preprint arXiv:1804.02767, (2018).
[21] S. Ren, K. He, R. Girshick, and J. Sun, *Faster r-cnn: Towards real-time object detection with region proposal networks*, Advances in Neural Information Processing Systems, 28 (2015).
[22] S. Safavi, M. A. Safavi, H. Hamid, and S. Fallah, *Multi-sensor fault detection, identification, isolation and health forecasting for autonomous vehicles*, Sensors, 21 (2021), p. 2547.
[23] K. Shi, H. Bao, and N. Ma, *Forward vehicle detection based on incremental learning and fast r-cnn*, in Proceedings of the 13th International Conference on Computational Intelligence and Security, Hong Kong, China, 2017, IEEE, pp. 73–76.
[24] D. Stavrou, D. G. Eliades, C. G. Panayiotou, and M. M. Polycarpou, *Fault detection for service mobile robots using model-based method*, Autonomous Robots, 40 (2016), pp. 383–394.
[25] B. Wang, Y. Lei, N. Li, and W. Wang, *Multiscale convolutional attention network for predicting remaining useful life of machinery*, IEEE Transactions on Industrial Electronics, 68 (2020), pp. 7496–7504.
[26] B. Wang, Y. Lei, N. Li, and T. Yan, *Deep separable convolutional network for remaining useful life prediction of machinery*, Mechanical Systems and Signal Processing, 134 (2019), p. 106330.
[27] H. Yang and S. Yin, *Actuator and sensor fault estimation for time-delay markov jump systems with application to wheeled mobile manipulators*, IEEE Transactions on Industrial Informatics, 16 (2019), pp. 3222–3232.
[28] T. Zhang, Z. Li, Z. Deng, and B. Hu, *Hybrid data fusion dbn for intelligent fault diagnosis of vehicle reducers*, Sensors, 19 (2019), p. 2504.

Edited by: B. Nagaraj M.E.
Special issue on: Deep Learning-Based Advanced Research Trends in Scalable Computing
Received: Nov 6, 2023
Accepted: Jan 25, 2024