An Intelligent Hybrid Ensemble Machine Learning Technique for Drowsiness Detection

*[1]Joshua B. Hassan, [2]Daniel D. Wisdom, [3] Taiwo D. Ajayi and Lateef G. Salaudeen
[1]Department of Computer Sciences, Federal University Oye Ekiti, Ekiti State, Nigeria
[2,4] Department of Computer Sciences, Chrisland University Abeokuta, Ogun State, Nigeria
[3]Department of Computer Sciences, Chrisland University Abeokuta, Ogun State, Nigeria
[1]joshua.hassan@fuoye.edu.ng | [2]ddaniel@chrislanduniversity.edu.ng | [3]lsalaudeen@chrislanduniversity.edu.ng

Received: 03-FEB-2025; Reviewed: 20-FEB-2025; Accepted: 28-MARCH-2025
https://dx.doi.org/10.4314/fuoyejet.v10i1.11

ORIGINAL RESEARCH

Abstract— Drowsiness is a major reason for road accidents especially for drivers on transit and in-cabin passengers resulting in mortality and Life threatening conditions. This paper investigate Drowsiness and it overall effects on Drivers and in-cabin passengers. Developed an Intelligent ensemble Random Forest machine learning (ML) technique for drowsiness detection. In order to accurately detect drowsiness, Dataset consisting of 1704 image was employed. The class distribution depicts 983 images for drowsy and 721 non-drowsy class. The paper developed a predictive model using deep learning algorithm of UNET and ResNet50 techniques were used during performance evaluation metrics based on confusion matrix. Data preprocessing, training, and testing of the models were conducted. Ensemble Random Forest ML techniques offered the best accuracy, F1-score, precision and recall of 97.8%, 94%, 96.2%, and 97% respectively. The results were compared with existing models, and the developed model outperformed the existing model with better efficiency in accurate detection of driver's drowsiness; with UNET outperforming others with a 96% F1 score, 96% accuracy, 96.5% precision, and 96% recall. Which may result in minimizing road accident and a significant reduction in mortality.

Keywords— Ensemble, Machine Learning, UNET, Resnet50, and Predictive-Model

1 Introduction

Drowsiness is a major reason for road accidents and workplace errors, resulting to increase yearly mortality rate, due to driver's lack of concentration. The term "drowsiness" is synonymous with sleepiness, which simply means an ability to fall asleep. There have been traditional approaches for detecting drowsiness which can be unreliable and inaccurate sleeping. Sleeping can be said to be one of the basic needs of the human. Lack of sleep can cause drowsiness. Drivers most times stays on the road for days without sleeping. For this reason, lack of sleep or driving for a long time causes driver's tiredness (Barua, *et al*., 2020). Resulting to drowsiness also known as Driver drowsiness. Driver drowsiness has been recorded to be one of the reasons why motor vehicles crashes. World Health Organization (WHO) has published reports that traffic accidents are one of the top 10 causes that lead to death in the world (world health organization). High rate of accidents in the world is caused by many factors, one of the common factors is related to driver fatigue. About 30% of all traffic accidents have been said to be caused by drowsiness (Abdusalomov, *et al*., 2023). The effects of both driver drowsiness and distraction are the same, i.e., driving performance is being decreased and a risk of crash involvement. (Haribabu, *et al*., 2022). Drowsiness detection systems been developed by manufacturers which is able to recognize signs of possible drowsiness (Albadawi, *et al*., 2022).

*Corresponding Author

Section B- ELECTRICAL/COMPUTER ENGINEERING &RELATED SCIENCES

Can be cited as:
Hassan J.B., Wisdom D.D., Ajayi T. D. and Salaudeen L.G. (2025). An Intelligent Hybrid Ensemble Machine Learning Technique for Drowsiness Detection. FUOYE Journal of Engineering and Technology (FUOYE**JET), 10(1), 66-75.** https://dx.doi.org/10.4314/fuoyejet.v10i1.11

Hence, there is a need for a smart drowsiness detection system that the automobiles in the industry can easily adapt. The numerous groundbreaking advances that the field of machine learning (ML) has, which different algorithms is used to train the model to be smart and autonomous (Ramalingam *et al*., 2020; Wisdom & Vincent 2024)

In the emerging digital era, ML has been employed in predicting Drowsiness. These developed ML Model are able to learn from dataset after training, if well trained with accurate dataset (Wisdom, *et al*., 2021). This may help in the development of drowsiness detection systems. Among the challenges which these systems frequently experience are high false positive rates, lack of adaptability to unique driving patterns, and inconsistent technical issues under different driving conditions (Haribabu *et al*., 2022). By applying deep learning techniques, this paper aim to build more effective Driver Drowsiness Detection System in order to address these issues.

Algorithm used by ML-based driver drowsiness detection system to evaluate a variety of real-time driving behavior indicators includes eye movements, facial expressions, and driving habits. With the help of training on drowsiness-related patterns, the paper was able to recognize indicators of driver sleepiness and send alarms or cautions in the case of an accident. A ML-based driver drowsiness detection system is developed, that continually tracks a driver's physiological and behavioral data, drowsiness detection systems grouped into three categories based on the measures that were used to detect the drowsiness signs (Albadawi *et al*., 2022), such as head motions, eye closure patterns, facial features, and steering behavior. The developed model learned and recognizes patterns that indicate fatigue or inattentiveness using ML techniques (Odunayo, *et al*., 2024). Employing these

© 2025 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)

indicators, the system can promptly issue warnings—such as visual or audio cues—to the driver, encouraging them to refocus or take a break, thereby minimizing the rate of accident and mortality with improved precision and accuracy in driver's drowsiness detection (Wisdom, *et al*., 2024). The paper employed facial recognition, computer vision, and steering behavior analysis to identify early signs of drowsiness and alert drivers in an accurate way. Improving accuracy while minimizing false positives, and building a system that can adapt to various driving conditions.

2 LITERATURE REVIEW

This section is structured based on driver's drowsiness, causes of driver drowsiness**.** *It presents the basis for driver's drowsiness with detection approaches. The subsection presents existing studies on drivers' drowsiness, courses of drivers' drowsiness various drivers' drowsiness approach.*

2.1 DRIVER'S DROWSINESS

The term "drowsiness" is synonymous with sleepiness, which simply means an ability to fall asleep. The term drowsiness typically refers to a state of sleepiness and apathy that leads to the tendency to fall asleep. Drowsiness is a process that occur due to lack of sleep (Chaabene, *et al*., 2021). Drowsiness causes a person to fall in sleep quietly or frequently. This is also the state of the human body where the body needs adequate sleep. Drowsiness can also be identified in number of ways, including self-reports, subjective ratings, eye closures, and reaction time measures. The inability of the eyes to be open for a short period of time is the main symptoms of being drowsy. Drivers are likely to experience drowsiness as shown below (IMAGE) due to lack of good sleep (Chandra, *et al*., 2021). Drowsiness is a huge threat for safe driving because it reduces the response time (Venkata *et al*., 2019). It is a state of reduced attention and vigilance towards any tasks being performed. Driver drowsiness mainly depends on: (i) the quality of the last sleep; (ii) the circadian rhythm (time of day) and (iii) the increase in the duration of the driving task.

2.2 CAUSES OF DRIVER'S DROWSINESS

The following criteria to identify accidents that are caused by drowsiness (Liu et al., 2009): 1. Blood alcohol level below the legal driving limit. 2. Vehicle ran off the road or onto the back of another vehicle. 3. No sign of brakes being applied. 4. Vehicle has no mechanical defect. 5. Good weather conditions and clear visibility.

There are multiple aspects that could contribute to drivers' drowsiness. Among the most important are fatigue (either mental or physical), lack of sleep, and drug consumption. The major contributing factors to drowsiness are long working hours, lack of sleep, and continuous driving (Farhangi, 2022). A report that drowsiness was one of the leading factors contributing to road accidents, among other factors including speeding and drinking (State of the Road Fatigue Fact Sheet, 2017). Drowsiness at the wheel is a major contributing factor in road accidents, causing a large number of deaths, serious injuries and financial losses (Traffic Safety Facts, 2017; State of the Road Fatigue Fact Sheet, 2017). Frequently by taking caffeine or different stimulants individuals keep on remaining conscious (Piyush *et al.*, 2019). Being physically unfit, by being either under or overweight, will cause weakness. Furthermore, being candidly focused will make the body get exhausted easily.

2.3 DETECTION APPROACHES

There are three approaches that are used: vehicle behavior- based or a driving-pattern-based approach, a driver's physiological signal-based approach, and a visual-based approach. There are various ways to carry out detection but a few are mentioned under this section such as: Sensing of physiological characteristics, Sensing of driver's eyes, Vehicle response and monitoring the response of driver at any moment (Flores-Monroy, *et al*., 2021). These techniques can be implemented in two ways: by accurately or practically measuring changes in physiological signals, such as eye blinking; and real time measuring physical changes such as leaning of the driver's head and the open/closed states of the eyes. Driver's operation and vehicle behaviors can be implemented in following ways, by monitoring the steering wheel movements, accelerometer or brake patterns, vehicle's speed, lateral acceleration caused and lateral displacements here. The final technique is by monitoring the response of the driver. The driver requires to consistently send the response but eventually this becomes tiresome and annoying for the driver. (Flores-Monroy, *et al*., 2021). In the first approach, the data sensed from the vehicle, such as velocity, steering angle, acceleration, and lane deviation, are used to detect a driver's abnormal condition (Flores-Monroy, et al., 2021). A number of metrics, including deviations from lane position, movement of the steering wheel, pressure on the acceleration pedal, etc., are constantly monitored and any change in these that crosses a specified threshold indicates a significantly increased probability that the driver is drowsy. The principal drawback of this approach is the variation of the detection performance depending on the road condition and the driver's driving ability. In the second approach, a driver's physiological signals, such as electrooculogram (EOG), electrocardiogram (ECG), electromyogram (EMG), and electroencephalogram (EEG), are used for this purpose. Although this approach provides better performance, it is invasive to the driver because the driver must use the wearable sensors during driving. Finally, in the visual-based approach, the driver's face image or video data are analyzed to determine the driver's drowsiness and distraction levels (Flores-Monroy *et al*. 2021; Phan *et al.*, 2021; Tamanani *et al*. 2021). A drowsy person displays a number of characteristic facial movements, including rapid and constant blinking, nodding or swinging their head, and frequent yawning Computerized, non-intrusive, behavioral approaches are widely used for determining the drowsiness level of drivers by measuring their abnormal behaviors. Different sensors have their own advantages and disadvantages. EEG is one of the most predictive and reliable techniques for drowsiness detection. However, EEG sensors are often expensive, time- consuming to setup, and intrusive. Camera-based drowsiness detection often does not require users to wear a device, therefore less intrusive. However, it is hard to

© 2025 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.

This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)
1http://dx.doi.org/10.46792/fuoyejet.v10i1.11 engineering.fuoye.edu.ng/journal

develop computer vision algorithms robust enough to detect faces and eyes of different skin colors, and under various weather and lighting conditions. Accelerator sensor is cheap but can only detect head movements.

2.4 Machine Learning

Since machine learning has grown so swiftly in recent years, many applications can now do cognitive tasks on their own (Greener *et al.*, 2022). Because it frequently enables systems to learn from and improve automatically, without being explicitly built, machine learning is sometimes regarded as the most well-liked technology in the fourth engineering uprising. Leverages state-of-the-art intelligent technology, such as machine learning automation, to continue automating both industrial and traditional production processes, such as data exploration. Learning algorithms are essential for analyzing these data critically and producing the necessary real-world applications. According to (Zhou *et al.*, 2020), supervised, reinforced, semi-supervised, and unsupervised learning are the four primary categories of learning algorithms. These categories are briefly addressed in Sect.

Some of the methods offered by ML algorithms to create a successful data-driven system include regression, data clustering, feature engineering (data preprocessing), classification analysis, dimensionality reduction, association rule learning, and reinforcement learning (Wisdom, *et al*., 2023).

2.5 Types of Machine Learning Techniques

Supervised learning, unsupervised learning, semi supervised learning, and reinforcement learning are the four main categories of machine learning algorithms (Zolanvari et al., 2019).

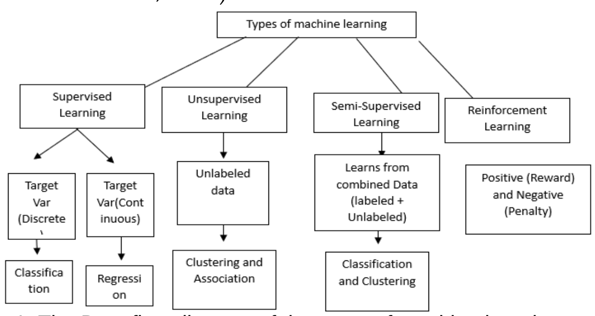


Fig 1. The Data flow diagram of the types of machine learning techniques

2.5.1 Supervised

In supervised learning, it is common practice to train functions that convert inputs into outputs using test input-output pairings. A function is inferred using labeled training data and a variety of training samples. A task-driven method or particular objectives that must be reached from a set of inputs must be used in order for supervised learning to be successful (*Greener et al*., 2022). Regression and classification are the two most common supervised tasks since they separate and fit the data, respectively. A good example of supervised learning is text categorization, which is the act of identifying the category or feeling connected to a text fragment, such a tweet or a product review.

2.5.2 Unsupervised

Unsupervised learning is frequently applied to experimental tasks, result aggregation, trend and structure detection, and feature extraction for generative features. Without the assistance of a person, unsupervised learning examines unlabeled data (Ray, 2019). Unsupervised learning tasks that are frequently performed include clustering, density estimation, feature learning, dimensionality reduction, association rule generation, anomaly detection, etc.

2.5.3 Semi-supervised

Semi-supervised learning, which employs both labeled and unlabeled data, can be conceptualized as a synthesis of the supervised and unsupervised techniques previously stated (Murphy, 2022). In light of this, it falls halfway between learning "with supervision" and "without supervision." Semi-supervised learning is advantageous since there are more unlabeled data sets in the actual world than labeled ones (Greener *et al*., 2022).

2.5.4 Reinforcement

Software agents and computers can increase productivity by using reinforcement learning, a form of machine learning (ML) method, which analyzes the optimal behavior in a given situation or environment automatically (Murphy, 2022). A well-trained AI system with an accurate dataset can use the knowledge gathered from environmental activists to take actions to either improve on accuracy or minimize hazards. However, developed AI Models are to be well trained tested several times before being deployed, which works well for creating AI models that can increase the automation or effectiveness of complex systems as robotics, self-driving cars, production, and logistics, especially in the emerging digital era of industry 4.0 and 5.0 advancement Technologies (Wisdom *et al*., 2025).

2.6 Model Evaluation

The F1 Score, Accuracy, recall, precision are the primary evaluation metrics considered in this work. These metrics, which included Logistic Regression, SVM, and Stacking helped determine the best algorithm for detecting intrusion.

2.6.1 F1 Score

The harmonic mean of recall and precision is used to generate the F1 score. To calculate recall, divide the total number of samples that should have been classified as positive by the total number of actual positive results. Furthermore, it is calculated by dividing the total number of positive outcomes, including those that were misclassified, by the total number of positive findings (Elsayed et al., 2020).

2.6.2 Accuracy

© 2025 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)

1http://dx.doi.org/10.46792/fuoyejet.v10i1.11 engineering.fuoye.edu.ng/journal

Accuracy refers to both how close measurement results are to the true value and how precise measurements are to a given value. It's both a true negative and a true positive in equal measure.

2.6.3 Recall

It is divided into two categories and shows how much of the news is genuinely pertinent. In contrast to specificity, which refers to the proportion of cases that were genuinely negative and were classified as such, sensitivity measures the proportion of cases that were actually positive and were classified as such (Saheed, 2022).

2.6.4 Precision

It is referred to as the news' pertinent portion. The capacity of a classifier to properly identify negative samples as such is referred to as precision of negative class. The ability of the classifier to properly identify positive samples as such is referred to as precision of positive class. Precision can range from a minimum of 0 to a maximum of 1 (Chicco & Jurman, 2020).

3.0 METHODOLOGY

This research was divided into 4 phases which include gathering of the dataset which is the first step in this study, preprocessing the dataset, making prediction using the deep learning algorithms (U-Net, and Resnet) and finally, evaluating the output of the model using some standard machine learning/deep learning algorithms (accuracy score, precision score, recall score, and lastly f1 score). The research system architecture is shown in figure 2.

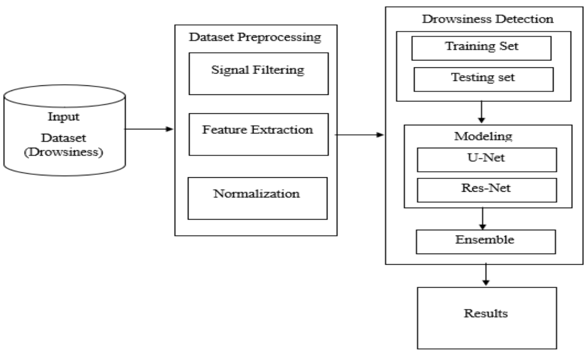


Fig 2. Proposed System Architecture

3.1 Dataset Gathering

The Kaggle platform provided the dataset for the drowsiness detection; it was particularly taken from the repository located at https://www.kaggle.com/datasets/hazemfahmy/openned-closed-eyes. In order to support research and development in drowsiness detection algorithms utilizing computer vision techniques, the dataset includes of photos showing people's eyes in both open and closed states. There are 5234 samples in all, each of which is an image of a person's eye in the dataset. For supervised learning tasks, ground truth labels were provided by labeling the images based on whether the eye is closed or open.

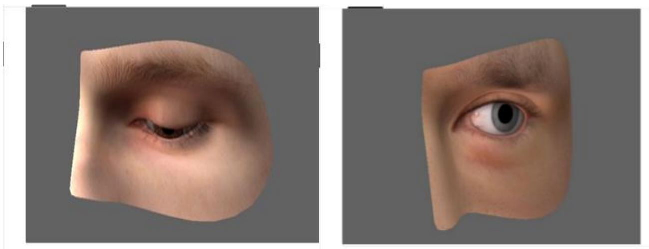


Fig 3. Sample of the image dataset

Fig. 3, depicts the sample of the image used in this study. The image at the left side shows that the person is drowsing while the image at the right side shows the person is not drowsing

3.2 Image Dataset Preprocessing

The process of preparing image datasets for drowsiness detection entails a number of procedures to get image data ready for machine learning models that are designed to identify drowsiness by analyzing visual clues from people's eyes. In order to enhance the functionality and resilience of the models, these preprocessing attempts to guarantee that the data is suitably formatted, cleansed, and enriched (Ma *et al.*, 2020). In this project, the following preprocessing techniques was used, signal filtering, feature extraction, and normalization. In order to train machine learning models and enable the accurate and reliable identification of weariness or drowsiness from eye pictures, these preprocessing techniques were utilized to effectively prepare image datasets for sleepiness detection. When developing and evaluating algorithms for drowsiness detection, these preprocessed datasets are helpful resources. The Fig. 4 shows the count or class distribution of the drowsiness dataset.

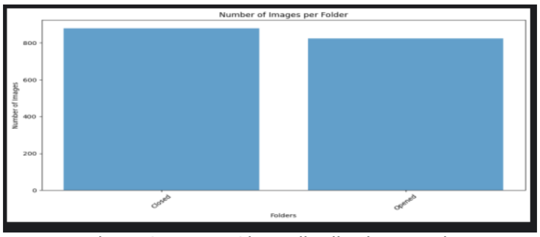


Fig 4. Count or Class distribution graph

After the collection of the dataset from the mention site, the images were imported into the environment used by used the matplotlib library in python. The library helped in accessing the images directly from the system OS. After the importing of the dataset, then further data preprocessing analysis was performed. In this project, the following are the preprocessing method used in the project, resizing and flattening of the images, after then the images was then converted into matrix for the model to train for it. And the Image work through is depicted in fig 5 respectively.

© 2025 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.

This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)
1http://dx.doi.org/10.46792/fuoyejet.v10i1.11 engineering.fuoye.edu.ng/journal

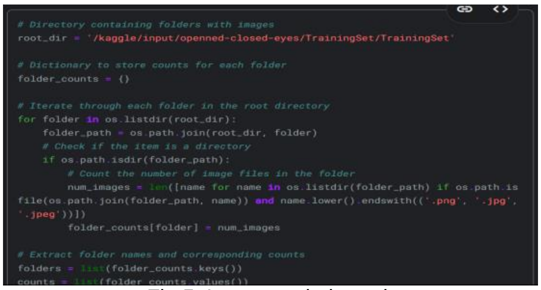


Fig 5. Image work through

Dataset Resizing

After the images was import into the environment, the images might have some unfixed size, which could cause a different dimension for the model, this was done to ensure the images present in the dataset has the same size and dimensions. This would make the model to process uniformly and make it easier to compare and analysis the images in the dataset. In this study, all the images are resized to (64, 64) the presentation of this is shown in figure 6 below.

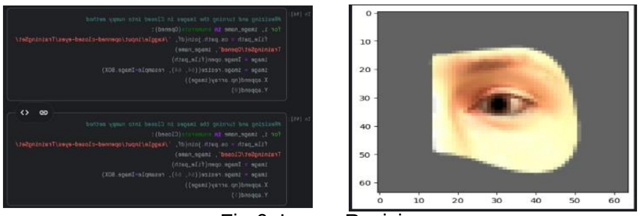


Fig 6. Image Resizing

Image Flattening

After resizing the images, one of the most crucial preparatory processes to apply in this project is picture model. Flattening works with algorithms, which need vector input data, because it turns multi-dimensional images into one-dimensional arrays. This step ensures uniform representation, streamlines the data structure, and reduces dimensional complexity to facilitate processing and manipulation of picture data. Flattening converts the 3D image matrices into 1D vectors, enabling effective feature extraction and classification, hence preparing the images for fully connected layers of neural networks. The result for this is shown at fig 7.

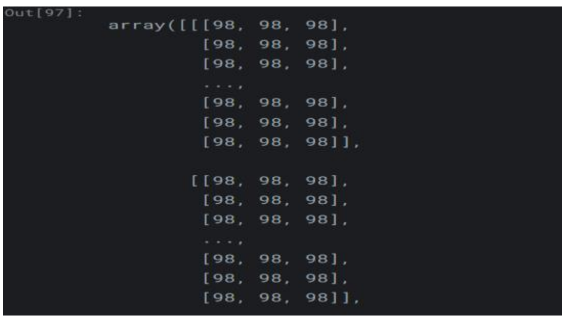


Fig 7. Factorized image

3.2.1 Signal Filtering

Digital signal processing techniques were used to filter physiological signals, to remove distortions and noise while retaining relevant information like drowsiness. For the purpose of detecting drowsiness, this process is called signal filtering (Chandra *et al.*, 2021). The processes in the process include comprehending the characteristics of the signal, choosing an appropriate filter type, establishing the filter's parameters, implementing the filter, evaluating its efficacy, and, if necessary, iteratively enhancing it. This is represented mathematically with convolutional operation between the input signal x(t) and the impulse response of the filter h(t) to obtain the filtered output signal y(t):

$$y(x) = x(t) * h(t) \qquad (1)$$

Where * donates the convolutional operation. The filter's impulse response h(t) is determined by its design parameter, such as cutoff frequencies and filter type, and the input signal x(t) represents the raw physiological data. Through the application of this mathematical procedure, the input signal can be made less noisy and distorted while retaining the necessary information for drowsiness detection in the filtered output signal. Noise reduction is optimized while signal integrity is preserved by iterative filter parameter revision based on performance evaluation.

3.2.2 Feature Extraction

The process of feature extraction for drowsiness detection is picking out informative traits from physiological data, in order to identify patterns linked to drowsiness. In this procedure, signal qualities are understood, pertinent features are chosen, extraction techniques are applied, and normalization and selection are optionally carried out. In mathematical terms, features are expressed as functions of the input signal, and processes related to subset selection are used to represent feature selection (Jabbar *et al.*, 2020). Machine learning algorithms are capable of accurately identifying tiredness by extracting and choosing pertinent features.

3.2.3 Normalization

The training stability and performance of deep learning models can be enhanced by batch normalization, especially for tasks like drowsiness detection using physiological inputs. In order to stabilize the training process and quicken convergence, it entails normalizing the activations of each layer throughout training. In order to accomplish this normalization, batch statistics (mean and variance) are calculated for each mini-batch of data, and the activations are then modified appropriately. Population statistics are employed during inference to guarantee consistent behavior. Batch normalization can be easily incorporated into model architectures and used in conjunction with other methods to improve the precision and dependability of sleepiness detection (Nasri *et al.*, 2022).

Mathematical Implementation

© 2025 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)

1 http://dx.doi.org/10.46792/fuoyejet.v10i1.11 engineering.fuoye.edu.ng/journal

Now that we are clear on the requirement for Batch normalization, let's examine its operation. It involves two steps. The input is first normalized, and then rescaling and offsetting are carried out.

Normalization of the Input

The process of normalizing data involves changing its mean and standard deviation to 0 and 1, respectively. We need to get the mean of this concealed activation in this phase because we have our batch input from layer h.

$$\mu = \frac{1}{m}\sum h_i \qquad (2)$$

Here, m is the number of neurons at layer h.

Once we have meant at our end, the next step is to calculate the standard deviation of the hidden activations.

$$\sigma = \left[\frac{1}{m}\sum (h_i - \mu)^2\right]^{\frac{1}{2}} \qquad (3)$$

Additionally, since we are prepared with the mean and standard deviation. With the help of these data, we will normalize the hidden activations. In order to do this, we

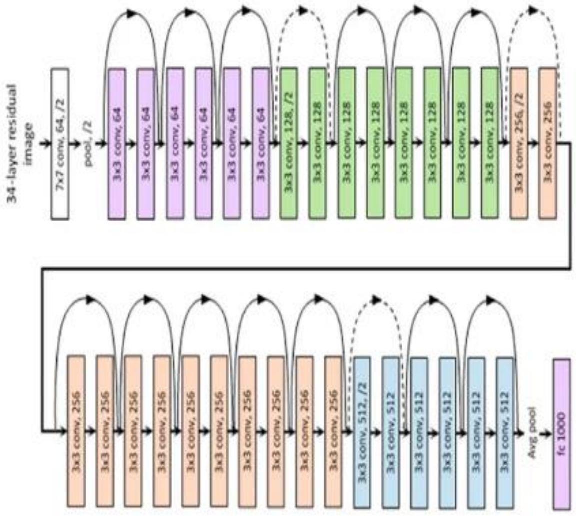


Fig 8. Resnet architecture

will take each input and subtract its mean before dividing the total by the smoothing term and standard deviation (ε).

The smoothing term (ε) assures numerical stability within the operation by stopping a division by a zero value.

$$h_{(norm)} = \frac{(h_i - \mu)}{\sigma + \varepsilon} \qquad (4)$$

Rescaling of Offsetting

In the final operation, the re-scaling and offsetting of the input take place. Here two components of the BN algorithm come into the picture, γ(gamma) and β (beta). These parameters are used for re-scaling (γ) and shifting(β) of the vector containing values from the previous operations.

$$h_i = \Upsilon\, h_{(norm)} + \beta \qquad (5)$$

These two are learnable parameters, during the training neural network ensures the optimal values of γ and β are used. That will enable the accurate normalization of each batch.

3.3 Drowsiness Detection

This is the next step for the prediction of the drowsiness system, in this phase, the preprocessed dataset was split into both training and testing set, and two deep learning algorithms was applied to it and eventually it was ensembled together to create robust, and improved model for the task of detecting a drowsiness in an image.

3.3.1 Training set and Testing set

Firstly, the sleepiness detection dataset is divided into training and testing sets. Usually, 80 percent of the data is set aside for training and the remaining percent is set aside for testing.

This division provides a distinct set of unseen data for assessing the models' performance and guarantees that the models are trained on a large enough dataset to discover patterns linked to drowsiness (Victoria, & Mary, 2021).

3.3.2 U-Net Model

Convolutional neural network architecture U-Net was first created for picture segmentation applications; however, it may be modified to detect drowsiness using physiological inputs. With the help of skip connections and an encoder-decoder architecture, it may extract hierarchical characteristics while preserving spatial information for accurate segmentation. U-Net gains the ability to translate input signals to segmentation maps that represent patterns of sleepiness during training. For Training, loss functions such as dice loss or binary cross-entropy are employed, and evaluation is carried out on different datasets (Chaabene *et al.*, 2021). For better performance, U-Net can be enhanced with methods like data augmentation or transfer learning. U-Net provides a strong framework for identifying patterns associated with sleepiness in physiological signals. Mathematically, it is represented as follows:

The mathematical formula for UNet are in several keys:

Convolutional Layers, Convolutional layers are used to extract features from the input signals. Mathematically, a convolution operation is performed between the input signal x and a set of learnable filters W, followed by an activation function σ:

$$z = \sigma(W * x + b) \qquad (6)$$

In this case, b is a bias term, z is the output feature map, and ∗ indicates the convolution operation.

Pooling Layers: In order to decrease spatial dimensions while maintaining significant features, pooling layers down sample the feature maps. Max pooling and average pooling are two common pooling methods.

Skip Connections: To preserve spatial information lost during downsampling, skip connections link related encoder and decoder layers. Skip connections add, concatenate, or mathematically combine the feature maps from the encoder layers with the decoder layers. **Upsampling Layers:** In order to fit the input size,

© 2025 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)
1http://dx.doi.org/10.46792/fuoyejet.v10i1.11 engineering.fuoye.edu.ng/journal

upsampling layers enlarge the spatial dimensions of feature maps. Transposed convolution and nearest-neighbor interpolation are two popular upsampling methods.

Loss Function: The difference between the ground truth and the anticipated segmentation map is measured by the loss function. Binary cross-entropy and dice loss are popular loss functions for binary segmentation tasks such as sleepiness detection.

3.3.3 Resnet50 Model

A deep convolutional neural network architecture called ResNet50, or Residual Network, was used for efficient training of extremely deep networks. It trains deep architectures without running into vanishing gradient issues by using residual blocks with skip connections to learn residual mappings (Magán *et al.,* 2022). ResNet can be used to learn hierarchical representations of physiological signals in the context of sleepiness detection, making it possible to identify patterns suggestive of drowsiness (Farhangi, 2022). Through the use of supervised learning, the network is trained to map input signals to predictions of drowsiness by minimizing a loss function. The model can be tested on fresh data after training to see how well it detects tiredness. All things considered, ResNet provides a strong foundation for effectively identifying sleepiness and capturing intricate correlations in physiological data. The developed resnet50 architecture is depicted in fig. 8.

3.3.4 Ensemble Technique

The paper employed a average weighted approach to incorporate predictions from both models to ensemble U-Net and ResNet50 for sleepiness detection. First, we use the sleepiness detection dataset to train U-Net and ResNet50. Next, we use both models to produce predictions for the test dataset. We can regulate each model's influence by giving its forecasts weights (Jasim, & Hassan, 2022). The final ensemble prediction is derived from the weighted average of the guesses. We can maximize the ensemble model's performance by tweaking the weights. By using the complementing advantages of ResNet50 and U-Net, this method may improve overall performance in sleepiness detection. It is flexible to fine-tune the ensemble strategy to get the best outcomes by adjusting the weights

3.4 Hardware and Software

The classification model engaged in this study was developed and experimented on a computer with the following specifications:

3.4.1 Hardware Requirement

The experimentation and coding process was carried out with a steady power supply and a notebook P.C. with the following key hardware configuration.

- Windows 7 Home Premium x64 (and above)
- 5GB RAM DDR3 1600MHz
- 320GB SSD HD
- Intel Core i5-3317U @ 1.70GHz

3.4.2 Software Requirements

The following software tools and systems were used for the development and experimentation of the proposed classifiers.

i. Operating system, required compiler, and libraries

ii. Python programming language (3.6.8)

iii. Python-based machine learning suites: Scikit Learn, Pandas, Numpy, Matplotib and Seaborn

3.4.2 Development Environment

The classifiers in this project were implemented using Python programming language. A high level, all-purpose programming languages. Python is a programming language used for developing websites, machine learning software, and other cutting-edge software (the python version used in this project is Python 3).

3.4.3 Libraries and Extensions

i) Numpy

A multidimensional, uniform collection of elements is referred to as a NumPy array. An array is defined by the kind of elements it contains as well as by the shape of the array. An array of shapes (M*N) comprising, for instance, floating points or complex integers can be used to represent a matrix. In contrast to matrices, NumPy arrays can have any number of elements.

ii) Pandas

Panda's Library, a statistical computing platform and database language in development since 2008, aims to bridge the gap between Python, a general-purpose systems and scientific computing language, and various domain-specific statistical computing platforms and language.

iii) Matplotlib

A portable 2D plotting and imaging tool called Matplotlib was created for the purpose of visualizing data in the fields of science, engineering, and finance. Using matplotib interactively from the Python shell, calling it from Python scripts, or including it in a graphical user interface (GUI) program (Tk, Wx, GTK, Windows) are all possible.

iv) Seaborn

A matplotlib-based Python data visualization tool called Seaborn is closely related to Pandas data structures.

v) Sci-kit learn (Sklearn)

Modern machine learning methods are combined in a Python module called Scikit-learn for medium-scale supervised and unsupervised situations. By using a general-purpose, high-level language, this package focusses on making machine learning accessible to non-

© 2025 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)
1http://dx.doi.org/10.46792/fuoyejet.v10i1.11 engineering.fuoye.edu.ng/journal

specialist. The efficiency, documentation, and consistency of the API are all given priority.

3.5 Performance Evaluation Metrics

The F1 Score, Accuracy, Specificity, and Sensitivity are the basic evaluation metrics considered in this research.

3.5.1 F1 Score

The precision and recall harmonics are averaged to determine the F1 score. It is determined by using the test's recall and precision, where precision is the proportion of "true positive" results to all "positive" results (including those that were misclassified), and recall is defined as the proportion of "true positive" results to all "positive" samples that actually should have been classified as such. It is represented mathematically as (Pham *et al.*, 2020).

$$F1 = \frac{2 \times precision \times recall}{precision + recall} \quad (7)$$

3.5.2 Accuracy

Accuracy is the degree to which measurement findings are close to the genuine value; it is the degree to which measurements are close to a certain value. It's the total of true negative and true positive values. It is represented mathematically (Tuan *et al.*, 2020);

$$Accuracy = \frac{\text{True positives} + \text{True Negatives}}{\text{True positives} + \text{False Negatives} + \text{False Positives} + \text{True Negatives}} \quad (8)$$

3.5.3 Recall

It's the percentage of relevant news that's retrieved, and it can be described in terms of either class.

Sensitivity refers to the percentage of cases that were truly positive but were labelled as negative, whereas specificity refers to the percentage of cases that were actually negative but were labelled as negative. (Mahmood *et al.*, 2022).

$$Recall = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (9)$$

3.5.4 Precision

It's defined as a percentage of relevant news. The capacity of the classifier to prevent misclassifying negative samples as positive is known as negative class accuracy. The ability of the classifier to prevent misclassifying positive samples as negative is referred to as positive class accuracy. Precision has the highest value of 1 and the worst value of 0. (Nandhini, & KS, 2020).

$$Precision = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (10)$$

4.1 RESULTS

This Section presents the details of the implementation of the proposed system using machine and deep learning algorithms. The experimental result and the analysis were also presented, as well as the libraries and extensions used to develop the classifiers.

4.1 Model Report

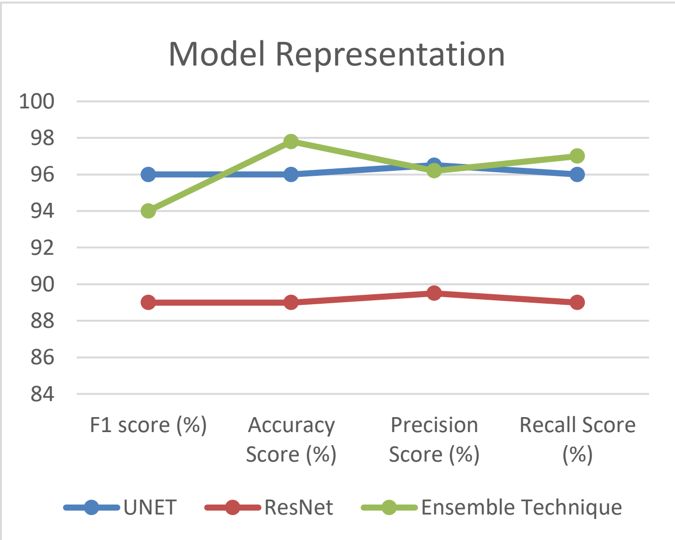
After the image preprocessing stage, the next thing that was done is the classification. In this stage, the image dataset which was a total number 1704 images was split into X and y which represent the features and the label of the image after extraction, and the images was divided into both training and testing set which was 1363 (80%) for training and 341 (20%) respectively. After the dataset was split, the 80% of the dataset was trained on the models used in this study and 20% was used to test them, and the evaluation of the both were taken using the standard machine learning metrics (Accuracy score, F1 score, Precision Score, and Recall Score). These metrics helped us to know how well the model performed on each measure. The table 1. below shows the information of how well each of the model used performed.

Table 1. Models Result

Models/performance	F1 score (%)	Accuracy Score (%)	Precision Score (%)	Recall Score (%)
UNET	96	96	96.5	96
ResNet	89	89	89.5	89
Ensemble Technique	94	97.8	96.2	97

The table 1. Shows the comparisons of drowsiness detection algorithms which shows that UNET outperforms the others with a 96% F1 score, 96% accuracy, 96.5% precision, and 96% recall. ResNet has a lower-than-average total score of about 89% across all dimensions. The model representation is depicted in fig 9. By recognizing drowsiness, the Ensemble Technique consistently outperforms other methods in terms of reliability and accuracy, with the maximum accuracy of 97.8%, recall of 97%, 94% F1 score, and 96.2% precision.

Figure 9. Graphical representation of the models



5. CONCLUSION

This section presents the summary and conclusion as well as limitation of the study as follows:

5.1 Summary

This paper examined the practicality of implementing machine-learning-based drowsiness detection system, the paper skillfully employed an ensemble algorithm

© 2025 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)

approaches to develop An intelligent hybrid ensemble machine learning technique for drowsiness detection, that can detect drowsiness in drivers actions. To identify the best strategy for a machine learning-based drowsiness detection system, the research employed UNET, RESNET50 and derived some equations used to analyze and test the developed scheme's performance using the variant dataset obtained from kaggle. Data preprocessing, training, and testing of the models were conducted. Ensemble Random Forest ML techniques offered the best accuracy, F1-score, precision and recall of 97.8%, 94%, 96.2%, and 97% respectively. The results were compared with existing models, and the developed model performed better than the existing model with better efficiency in accurate detection of driver's drowsiness.

5.2 Conclusion

Drowsiness during driving may be handled by a drowsiness detection system on machine learning (ML). An intelligent hybrid ensemble machine learning technique for drowsiness detection is developed, the ensemble algorithm was used to integrate the two drowsiness detection algorithms, UNET and RESNET. The model was evaluated using a public available dataset on Kaggle, which is perfect for the system. The results demonstrated that the suggested strategy achieves a high Accuracy, indicating a stronger overall performance in terms of detection. Based on experimental findings, it was determined that employing two drowsiness detection algorithms for an effective system is realistic and feasible. When compared to standalone approaches, using two techniques exhibits remarkable accuracy and can solve the problem of labeled data in detecting drowsiness. In the validation dataset, Precision, F1, Recall, Accuracy, and AUC scores of the proposed models all result to 100%. Which shows the comparisons of drowsiness detection algorithms that UNET outperforms the others with a 96% F1 score, 96% accuracy, 96.5% precision, and 96% recall. ResNet has a lower-than-average total score of about 89% across all dimensions. By recognizing drowsiness, the Ensemble Technique consistently outperforms other methods in terms of reliability and accuracy, with the maximum accuracy of 97.8%, recall of 97%, 94% F1 score, and 96.2% precision.

5.3 Limitations of the study

This paper introduced a hybridized technique for the driver's drowsiness detection system. Therefore, additional research into the issue based on real-time data and power time optimization is necessary and more techniques could be added to enhance the performance of the model in the future work.

Acknowledgement

The authors wishes to appreciate the reviewers for their constructive comments that have improved this paper. However, this research was not funded.

REFERENCES

Abdusalomov, A. B., Nasimov, R., & Cho, Y. I. (2023). "Real-Time Deep Learning-Based Drowsiness Detection: Leveraging Computer-Vision and Eye-Blink Analyses for Enhanced Road Safety." Sensors, 23(14), 6459.

Barua, S., Ahmed, M. U., Ahlström, C., & Begum, S. (2020). "Automatic driver sleepiness detection using EEG, EOG, and contextual information." Expert Systems with Applications, 115, 121-135.

Haribabu, J., Navya, T., Praveena, P. V., Pavithra, K., & Sravani, K. (2022). "Driver Drowsiness Detection Using Machine Learning." Journal of Engineering Science, 13(06)

Albadawi, Y., Takruri, M., & Awad, M. (2022). A Review of Recent Developments in Driver Drowsiness Detection Systems. *Sensors*, *22*(5), 2069. https://doi.org/10.3390/s22052069.

Ramalingam, V., Shivani, & Aditya. (2020). Driver Drowsiness Detection System using Machine Learning Algorithms. *International Journal of Recent Technology and Engineering (IJRTE)*, *8*(6), 990–993. https://doi.org/10.35940/ijrte.f7514.038620

Chaabene, S., Bouaziz, B., Boudaya, A., Hökelmann, A., Ammar, A., & Chaari, L. (2021). Convolutional neural network for drowsiness detection using EEG signals. *Sensors*, *21*(5), 1734.

Chandra, B. V. B., Naveen, C., Kumar, M. M. S., Bhargav, M. S. S., Poorna, S. S., & Anuraj, K. (2021). A comparative study of drowsiness detection from Eeg signals using pretrained CNN models. In *2021 12th international conference on computing communication and networking technologies (ICCCNT)* (pp. 1-3). IEEE.

Farhangi, F. (2022). Investigating the role of data preprocessing, hyperparameters tuning, and type of machine learning algorithm in the improvement of drowsy EEG signal modeling. *Intelligent Systems with Applications*, *15*, 200100.

Flores-Monroy, J.; Nakano-Miyatake, M.; Perez-Meana, H.; Escamilla-Hernandez, E.; Sanchez-Perez, G. A (2022). CNN-based driver's drowsiness and distraction detection system. In Proceedings of the 14th Mexican Conference on Pattern Recognition, Chihuahua, Mexico, 22–25 June 2022.

Flores-Monroy, J.; Nakano-Miyatake, M.; Perez-Meana, H.; Sanchez-Perez, G. (2021). Visual-based real time driver drowsiness detection system using CNN. In Proceedings of the International Conference on Electrical Engineering, Computing Science and Automatic Control, IEEE, Mexico City, Mexico, 10 November.

Tamanani, R.; Muresan, R.; Al-Dweik, A.(2021) Estimation of driver vigilance status using real-time

© 2025 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)

facial expression and deep learning. IEEE Sens. Lett. 5, 6000904.

Wisdom, D. D., & Vincent, O. R. (2024). Cybersecurity concerns and risks in emerging healthcare systems. Cybersecurity in Emerging Healthcare Systems, Institute of Engineering and Technology (IET), UK. https://doi.org/10.1049/PBHE064E_ch1

Greener, J. G., Kandathil, S. M., Moffat, L., & Jones, D. T. (2022). A guide to machine learning for biologists. *Nature Reviews Molecular Cell Biology*, *23*(1), 40-55.

Wisdom, D. D., Igulu, K., Esther, O. O., Baba, G. A., Ahmad, A., & Sidi, A. (2021). A Review of Best Practices and Methods of Mitigating Cybersecurity Risks in Healthcare Systems and a Newly Proposed Algorithm. Computology: Journal of Applied Computer Science and Intelligent Technologies, 1 (2), 27-36. DOI: 10.17492/computology. v1i2.2104.

Jabbar, R., Shinoy, M., Kharbeche, M., Al-Khalifa, K., Krichen, M., & Barkaoui, K. (2020). Driver drowsiness detection model using convolutional neural networks techniques for android application. In *2020 IEEE International Conference on Informatics, IoT, and Enabling Technologies (ICIoT)* (pp. 237-242). IEEE.

Jasim, S. S., & Hassan, A. K. A. (2022). Modern drowsiness detection techniques: A review. *International Journal of Electrical and Computer Engineering*, *12*(3), 2986-2995.

Wisdom, D. D., Vincent, O. R., Igulu, K. T., Arowolo, M. O., Hyacinth, E. A., Christian, A. U., Oduntan, O. E., Baba, G. A., (2023). Mitigating Cyber Threats in Healthcare Systems: The Role of Artificial Intelligence and Machine Learning, Artificial Intelligence and Blockchain Technology in Modern Telehealth Systems. Institute of Engineering and Technology (IET), UK.

Magán, E., Sesmero, M. P., Alonso-Weber, J. M., & Sanchis, A. (2022). Driver drowsiness detection by applying deep learning techniques to sequences of images. *Applied Sciences*, *12*(3), 1145.

Odunayo, E. O., Hassan, J. B., Wisdom, D.D. Ugwunna, C. O., Isaac, S., & Falayi, C. F. (2024). Artificial intelligence & blockchain technology in modern telehealth systems. KASU Journal of Computer Science (KJCS). Vol 3(1), p 607-625.

Nandhini, S., & KS, J. M. (2020). Performance evaluation of machine learning algorithms for email spam detection. In *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)* (pp. 1-4). IEEE, February, 2020.

Nasri, I., Karrouchi, M., Snoussi, H., Kassmi, K., & Messaoudi, A. (2022). Detection and prediction of driver drowsiness for the prevention of road accidents using deep neural networks techniques. In *WITS 2020: Proceedings of the 6th International Conference on Wireless Technologies, Embedded, and Intelligent Systems* (pp. 57-64). Springer Singapore.

Wisdom, D. D., Ajayi, T. D., Balogun, O. A., Ayetuoma, I. O., Vincent, O. R., & Garba, A. B. (2024). A hybrid artificial intelligence system for the visually impaired. NIGERCON 2024:2024 IEEE International Conference on Electro-Computing for Sustainable Development, Afe Babalola University, Ado Ekiti, Ekiti State, Nigeria, November 26-28.

Pham, B. T., Jaafari, A., Avand, M., Al-Ansari, N., Dinh Du, T., Yen, H. P. H., ... & Tuyen, T. T. (2020). Performance evaluation of machine learning methods for forest fire modeling and prediction. *Symmetry*, *12*(6), 1022.

Phan, A.-C.; Nguyen, N.-H.-Q.;Trieu, T.-N.; Phan, T.-C. (2021) an Efficient Approach for Detecting Driver Drowsiness Based on Deep Learning. Appl. 11, 8441.

Wisdom, D. D., Vincent, O. R., Igulu, K., Christian, A. U., Hyacinth, E. A., Baba, G. A., Esther, O. O. (2025). The Protection of Industry 4.0 and 5.0: Cyber security Strategies and Innovations, CI-Industry-4.0: Computational Intelligence in Industry 4.0 and 5.0 Applications. Taylor and Francis group.

Tuan, T. A., Long, H. V., Son, L. H., Kumar, R., Priyadarshini, I., & Son, N. T. K. (2020). Performance evaluation of Botnet DDoS attack detection using machine learning. *Evolutionary Intelligence*, *13*(2), 283-294.

Victoria, D. R. S., & Mary, D. G. R. (2021). Driver drowsiness monitoring using convolutional neural networks. In *2021 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS)* (pp. 1055-1059). IEEE.

World Health Organization The top ten causes of death. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/thetop-10-causes-of-death

Zhou, Y., Cheng, G., Jiang, S., & Dai, M. (2020). Building an efficient intrusion detection system based on feature selection and ensemble classifier. *Computer networks*, *174*, 107247.

Zolanvari, M., Teixeira, M. A., Gupta, L., Khan, K. M., & Jain, R. (2019). Machine learning-based network vulnerability analysis of industrial Internet of Things. *IEEE Internet of Things Journal*, *6*(4), 6822-6834.

© 2025 The Author(s). Published by Faculty of Engineering, Federal University Oye-Ekiti.
This is an open access article under the CC BY NC license. (https://creativecommons.org/licenses/by-nc/4.0/)

1http://dx.doi.org/10.46792/fuoyejet.v10i1.11 engineering.fuoye.edu.ng/journal