ANALYSIS AND INTERPRETABILITY OF MACHINE LEARNING MODELS TO CLASSIFY THYROID DISEASE

Sumya Akter and Hossen A Mustafa

Abstract

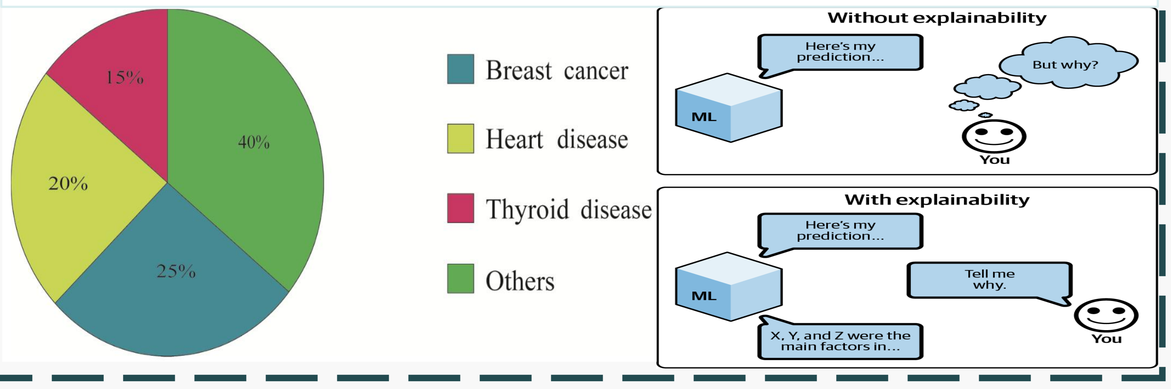
Thyroid disease is one of the most common diseases and increasing at an alarming rate daily. Accurate and early diagnosis is crucial for effective treatment and management. Machine learning (ML) models have shown promising results in diagnosing thyroid disease using patient data. But existing methods are less focused on dataset balancing and feature engineering part to detect thyroid disease. Our work proposes a cluster-wise data balancing technique to overcome the data imbalance problem. Some feature selection techniques are applied to improve the accuracy of detecting thyroid disease. After that, some machine learning algorithms are used, and their performance is compared with preprocessing techniques and without preprocessing techniques. Finally, some explainable artificial intelligence (XAI) tools are applied to make the model more transparent and trustworthy because of the machine learning black-box nature. Our results demonstrate the effectiveness of ML models in diagnosing thyroid disease and provide insights into the decision-making process of these models.

Background

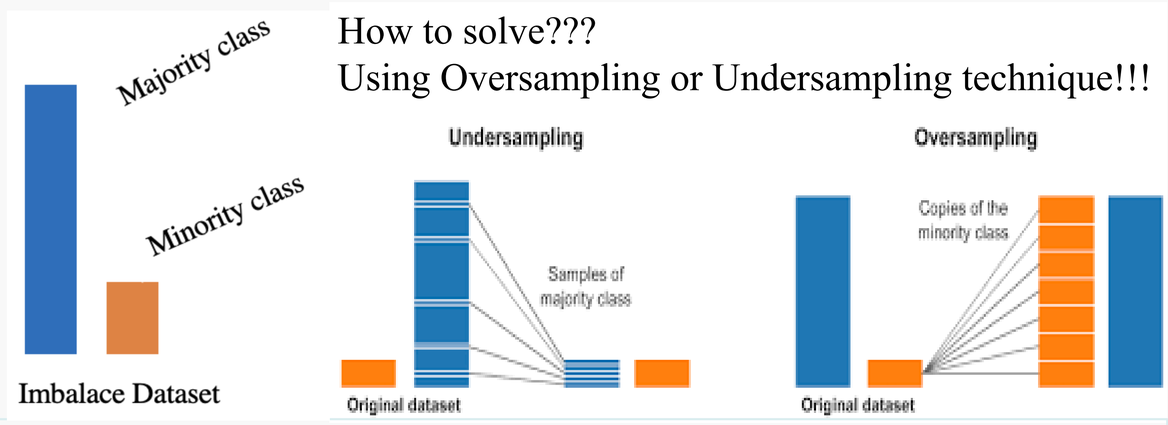
- In Alyas et. al's(2022) study, four machine learning algorithms namely Decision Tree (DT), Random Forest (RF), K-Nearest Neighbor (KNN), and Artificial Neural Network (ANN) were all used where RF provides the highest accuracy.
- In 2022, Chaganti et al's presented five ML models and three deep learning (DL) models in addition to four feature selection strategies.
- Again in 2022, Alsaadawi et al's used different ensemble methods after using Synthetic Minority Over-sampling Technique (SMOTE) to balance the data and the Recursive Feature Elimination technique (RFE) to pick the features.
- Arjaria et al's (2022) conducted an explainability approach to the thyroid dataset based on the logistic regression model but no preprocessing techniques were used.

Motivation

✓Among various diseases, thyroid disease has received comparatively less attention according to Rekha et al's study.
✓Previous works are less focused on data balancing and feature engineering part.
✓A comprehensive analysis of numerous ML models is now necessary to determine which performs the best.
✓Even though **honesty**, **transparency**, and **trust** are the major concerns in the medical field as a whole, a few publications have focused on using XAI to study thyroid disease.



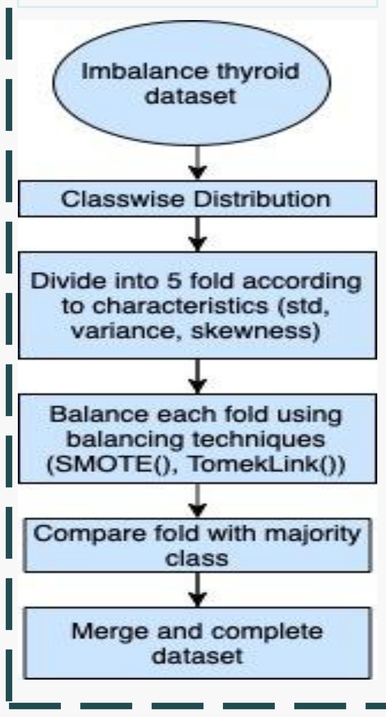
Proposed Idea and Methodology



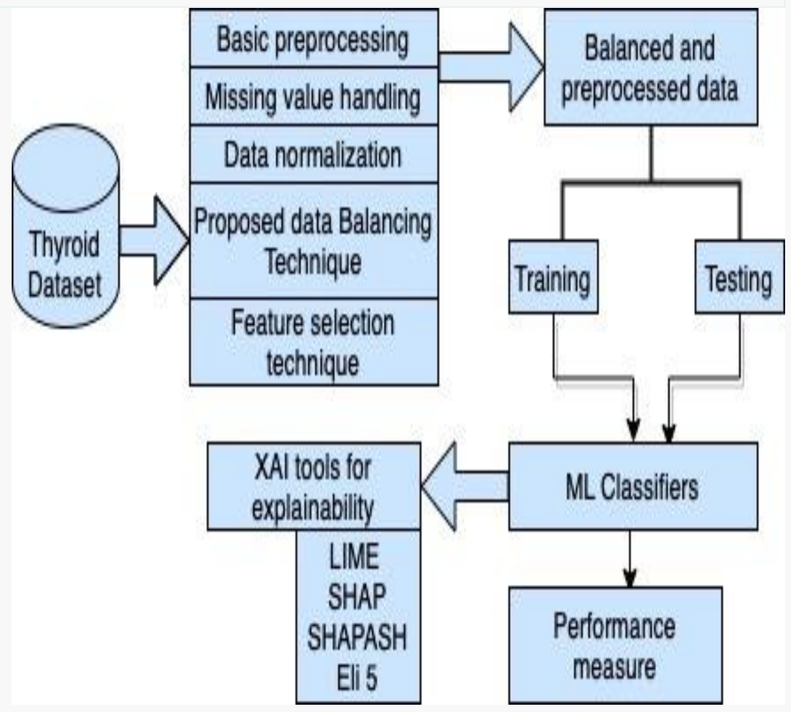
❖ Problem of oversampling: might occur **Overfitting**.
❖ Problem of undersampling: could **discard potentially useful data**.

To overcomne this problem, a combination of an oversampling technique (SMOTE) and an undersampling technique (TomekLinks) is used in our work after dividing the dataset clusterwise.

Proposed data balancing technique



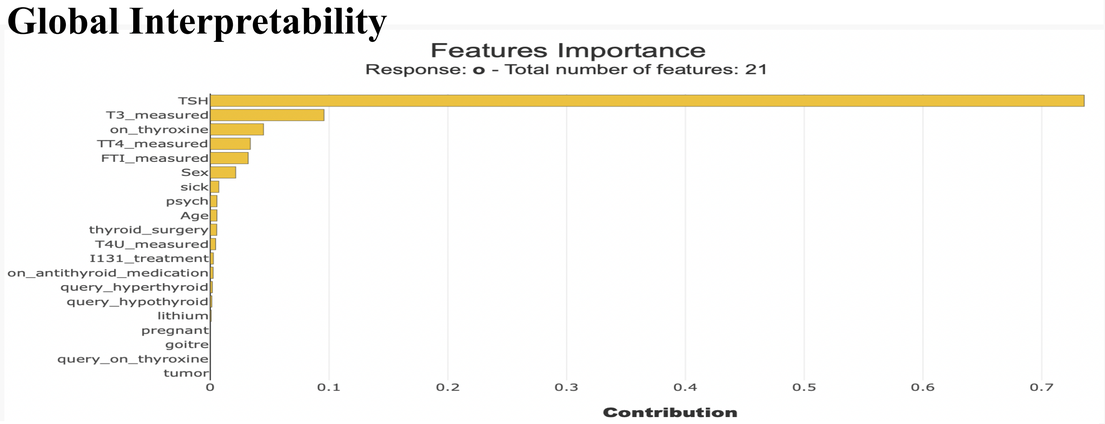
Proposed method



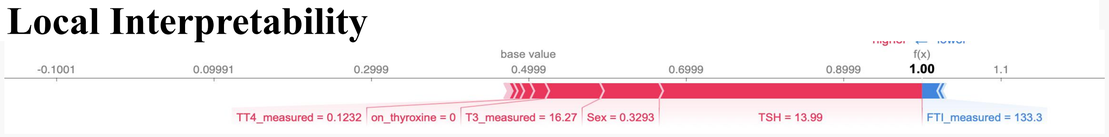
Results

Machine Learning Classifiers	Accuracy
Logistic Regression	0.78
Decision Tree	0.993
Random Forest	**0.996**
Multinomial Naive Bayes	0.73
K Nearest Neighbour	0.923
Gradient Boosting	0.995
AdaBoost	0.992
Stochastic Gradient Descent	0.73
XGboost	0.995

Global Interpretability



Local Interpretability



Skin Cancer Detection and Classification using Ensemble CNN and MLP

Nuzhat Tabassum Promi, Dr. Md. Liakot Ali

Abstract

Skin cancer is one of the most widespread and dangerous diseases because of its high mortality rate. Dermatologists conduct a preliminary clinical screening, followed by dermoscopic analysis, and histopathological examination to detect skin cancer which is a time-consuming process. An automated system can help both patients and dermatologists save time by performing early diagnosis of skin cancer. This research proposes an improved ensemble skin cancer detection and classification (SCDC) method that utilizes both dermoscopic images and clinical information. Combining clinical information such as patient age, sex, affected area location, cancer history, etc. with images can help us increase recall, precision, and accuracy. The proposed ensemble model includes CNN(ResNet, Xception, U-net) and MLP that is trained on the dataset publicly available from ISIC (The International Skin Imaging Collaboration). The main advantages of using CNN is that it can extract many informative features from the images that will strengthen our model.

Background & Motivation

- The number of new malignant melanoma diagnosis has annually increased by 27 percent in the past decade.
- Even though melanoma causes the most damage, ill-timed treatment of non-melanocytic cancers can also pose a danger.

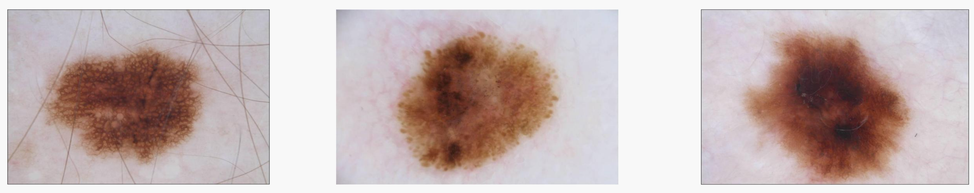


Figure 1: Some examples of Skin Cancer Images

- Many researchers are using machine learning techniques, e.g., SVM Random Forest, CNN, etc. However, skin cancer diagnosis still did not achieve good recall and precision. Without this, an automated system will pose risk.
- While many studies focus on skin cancer detection using available dermoscopic images, other clinical information such as patient age, sex, affected area location, cancer history, etc. are being ignored.
- Skin cancer is the most preventable cancer. When detected early, the five-year survival rate for melanoma is 99 percent. Proper Awareness and easily accessible detection system can go a long to help preventing skin cancer.

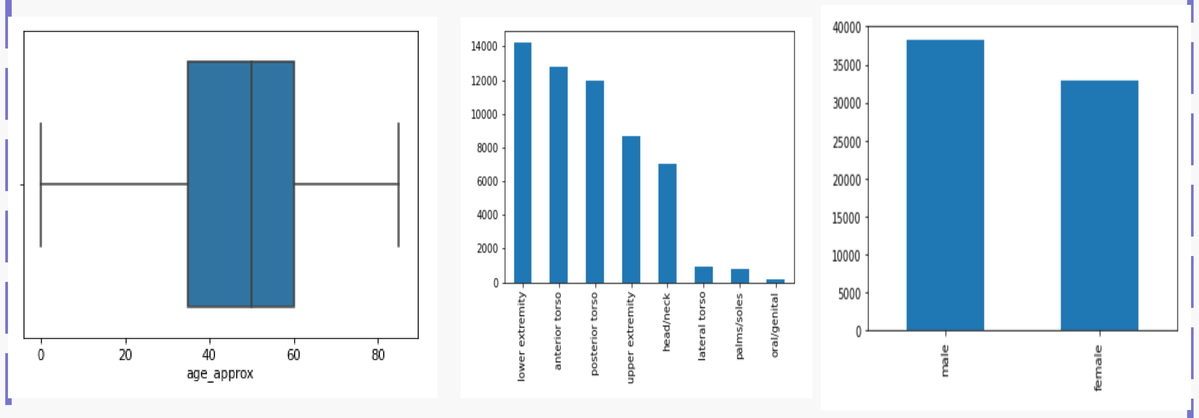


Figure 2: Clinical Features(Age, Location, Sex) from Dataset

Proposed Idea and Methodology

The simplified architecture of the proposed model is illustrated in Figure 3:

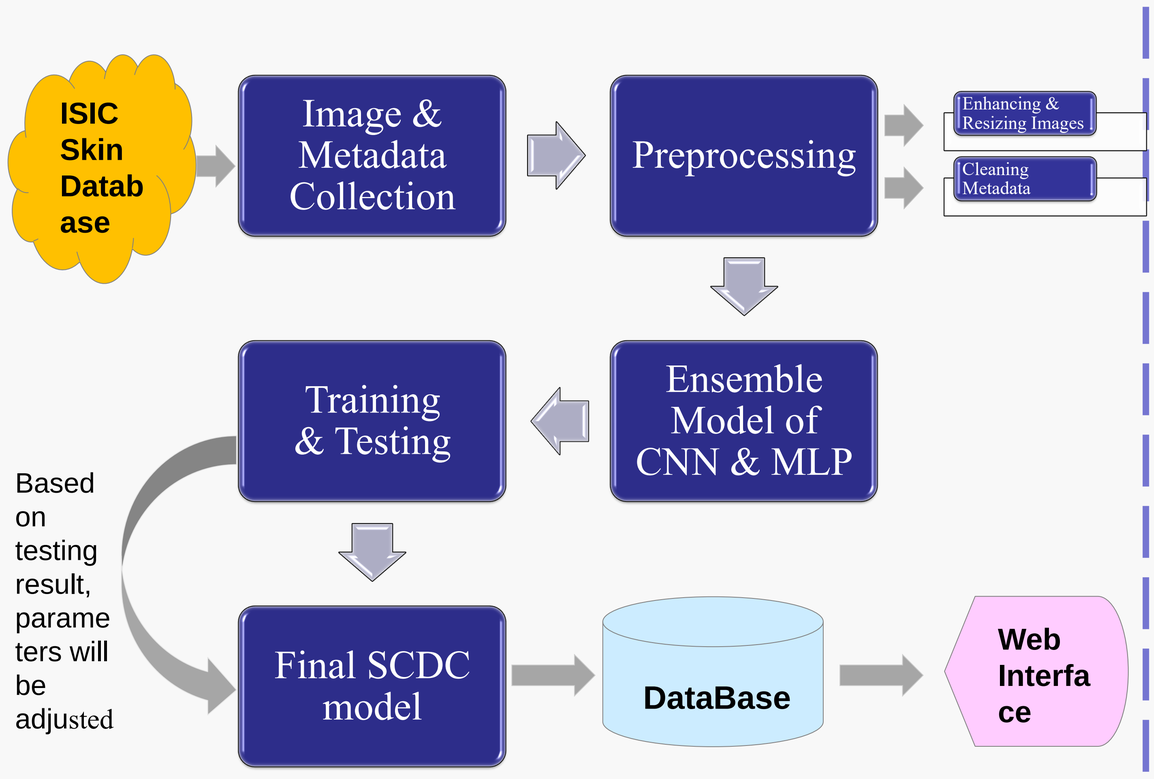


Figure 3: Simplified Conceptual Model of Proposed SCDC Model Framework

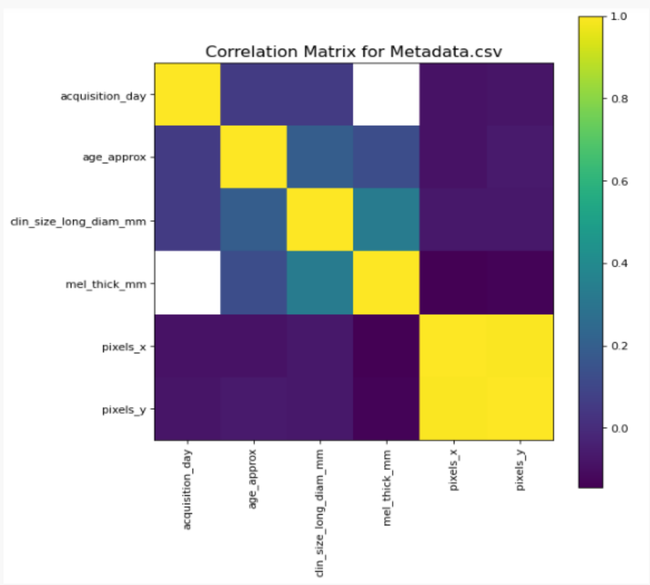
We will preprocess the images by resizing them to a uniform size and adjusting the color balance. Final SCDC Model will identify skin cancer and classify it into different groups. A web interface will be developed so that dermatologists and general people can easily upload an image and check by themselves. The proposed model will be implemented using Tensorflow and Keras libraries of python.

Results

❖We evaluated the performance of our skin cancer detection system on a test set of images. The train-test-spilt ratio is 0.2.

❖The system achieved an accuracy of 92%, sensitivity of 88%, and specificity of 91% in detecting skin cancer.

❖We are working on to tune the models, and parameters and improve accuracy, sensitivity, and specificity.



Developing a Real-time Hand-Gesture Recognition System for Wheelchair Control

Md Rafiul Huda, Md. Liakot Ali, Muhammad Sheikh Sadi

Abstract

Approximately 15% of the global population has a disability, who experience issues with the structure or function of their body, or challenges performing a task or action, worldwide. Many of them (131.8 million people currently) need to use wheelchairs to go around daily, yet they have been encountering many problems using traditional wheelchairs. This poster presents a prototype of an advanced algorithm for gesture recognition and decision-making for the control of a smart wheelchair, in line with contemporary technologies. Overall, hand and finger tracking using a robust hand-tracking solution and gesture recognition by a mathematical model are the main components of the proposed system. This method places a greater emphasis on users' flexibility by requiring less hand movement and environment independency. The experimental study shows that the proposed method outperforms existing methods.

Background & Motivation

- ❑ The number of wheelchair users is increasing day by day.
- ❑ Safe navigation, independent mobility and low cost are the key issues for wheelchair users, especially those suffering from mobility impairments.
- ❑ Disables people face many problems using traditional wheelchairs
- ❑ The advancement of technology has allowed wheelchairs to become intelligent by incorporating contemporary sensors, and AI.
- ❑ Gestures can be easily incorporated into wheelchair functionalities utilizing computer vision and AI.
- ❑ Several researches carried out i.e. joystick, depth camera or android application based, but not suitable for users having dexterity problems or face difficulties with movement of hand.
- ❑ Accuracy of existing systems is variable or not so high.
- ❑ Recent researchers used skin segmentation, KCF tracking, CNN model but couldn't not show good performance under daylight or if background matches with skin color.
- ❑ It's still a challenging issue to build an efficient gesture recognition system for wheelchair control that will consider users' flexibility, high accuracy, and environment independency.

eople

Proposed Idea and Methodology

- ❑ The image of hand is taken using a low cost RGB camera which will be attached to a wheelchair.
- ❑ The proposed methodology comprises 3 major steps :
  - ❖ Hand Detection and Hand Tracking by MediaPipe
  - ❖ Hand Landmarks Extraction
  - ❖ Gesture Recognition using Mathematical Model

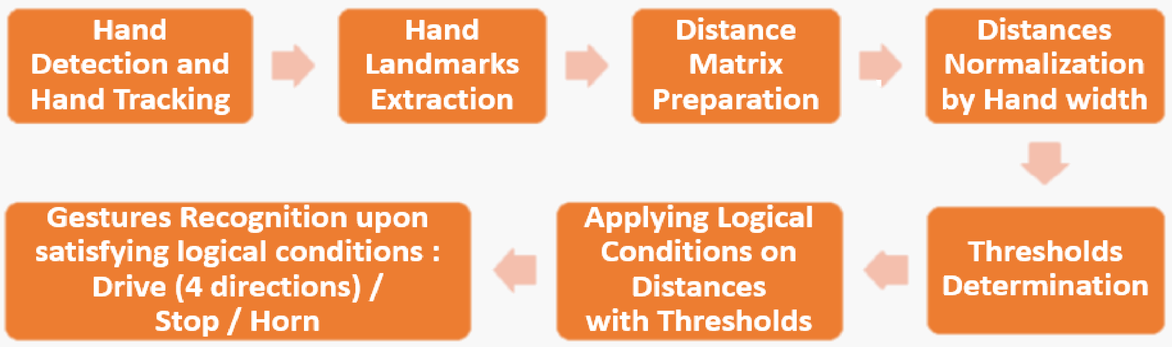


Figure 2. Architectural components of the proposed mathematical model

- ❑ The thresholds are determined at the start of hand detection, using the minimum and maximum distances among significant hand landmarks.

Drive: d(0,8) > Thld1 and d(0,12) < Thld2 and d(0,16) < Thld2 and d(0,20) < Thld2 and d(0,4) < Thld1

Stop: d(0,8) > Thld1 and d(0,12) > Thld1 and d(0,16) >Thld1 and d(0,20) >Thld1 and d(0,4) >Thld2

Horn: d(0,8) < Thld2 and d(0,12) < Thld2 and d(0,16) < Thld2 and d(0,20) < Thld2

Figure 3. Logical conditions for basic gestures recognition

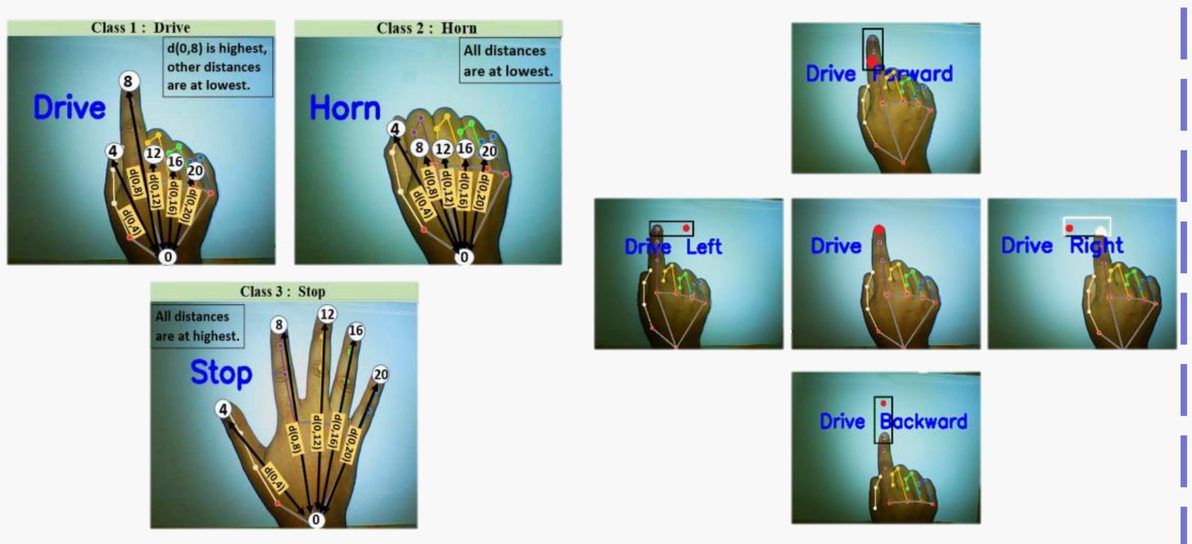


Figure 4. Classes of the projected hand gestures and concepts behind the proposed mathematical model for gesture recognition

Results

- ❑ The proposed system obtained a testing accuracy of 99.17% with a recognition rate of 21-27 FPS, tested by hands of different sizes.
- ❑ This system is invariant of skin color, indoor/outdoor environment and background.
- ❑ It's user friendly for the disabled persons, as it works only by short easy movement of fingers only. Hand raising is not required.

❑ **Performance Comparison among different methods:**

Features	System Types of various researches			
	Gao et al	***Oliver et al.***	***M. Repon Islam et al.***	***Proposed in this poster***
Functionality (Gesture Recognition)	Requires hand lifting, a Kinect Depth Camera and highly configured PC.	Requires wearing hand band, Joystick Manipulator, and accelerometer.	Requires movement of fingers, RGB Camera. It creates problems under daylight or if background matches with skin color.	Requires movement of fingers, RGB Camera. A mathematical model is used. Work well under daylight on any background.
Cost	High	Medium	Low	Low
Recognition Success rate	Depends on complexity (10–100%)	Actual accuracy is not measured	98.33%	99.17%

Enhancing Botnet Detection and Prevention using a Combination of Machine Learning Techniques and Real-time Model Updating

Md. Alamgir Hossain and Dr. Md. Saiful Islam

Abstract

One of the major security threats to computer systems and networks is caused by botnets. Due to the ongoing evolution of botnet characteristics, conventional approaches for detecting them are no longer reliable. Machine learning-based methods have become a potential botnet detection approach in recent years. The quality of a single model can decrease over time due to the ongoing evolution of botnets, hence no single learning technique can reliably detect all varieties of botnets. In this research, we offer a novel strategy for botnet detection that combines multiple learning techniques and incorporates real-time updating of the detection models to increase the efficiency of the detection of new botnets. Our experiments show that the proposed approach achieves higher accuracy and better detection rates than existing single-method approaches. This approach can be a valuable tool for enhancing the security of computer systems and networks against evolving botnets.

Background & Motivation

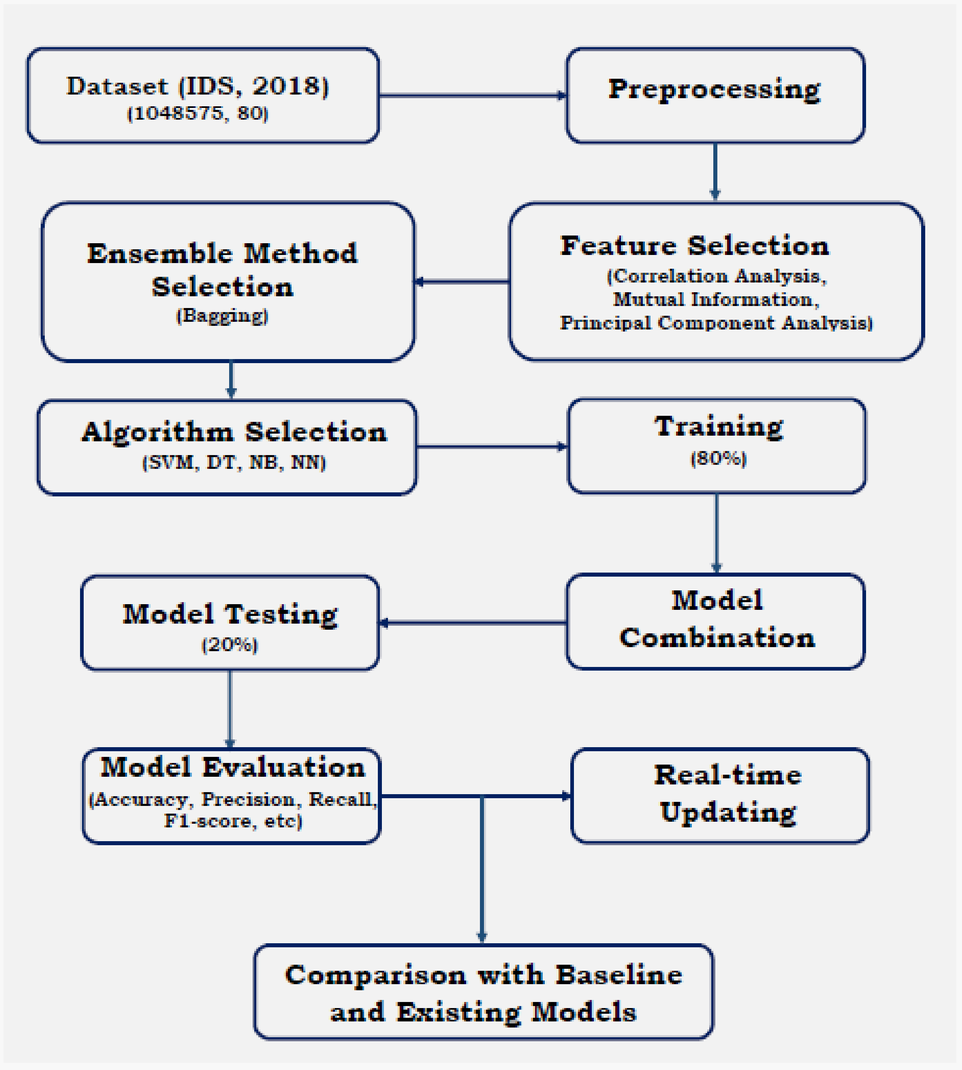
Botnets are collections of compromised devices that an attacker can remotely command to engage in a range of illegal actions like spamming, distributed denial of service attacks, identity theft, and eavesdropping. Botnets are a severe problem to computer systems and networks, affecting individuals and organizations all over the globe a lot of money and damage to their reputations. Because botnets are constantly evolving and employing advanced avoidance techniques, conventional approaches of botnet detection, such as signature-based detection, are losing effectiveness. As a result, there is a need for more reliable and flexible techniques of botnet detection.

Botnets are constantly evolving, making it difficult for current botnet detection techniques to keep up. Although machine learning-based approaches have produced promising outcomes in the detection of botnets, no single technique is capable of detecting all types of botnets. Furthermore, because botnets are dynamic, a single model's accuracy may decrease over time. As a result, a comprehensive strategy is required that combines various learning techniques and integrates real-time updating of the detection models to increase their efficiency in identifying new botnets. The goal of this research is to make computer systems and networks more secure against emerging botnets, which will ultimately protect individuals and companies from the financial and reputational harm that botnet attacks can bring.

Proposed Idea and Methodology

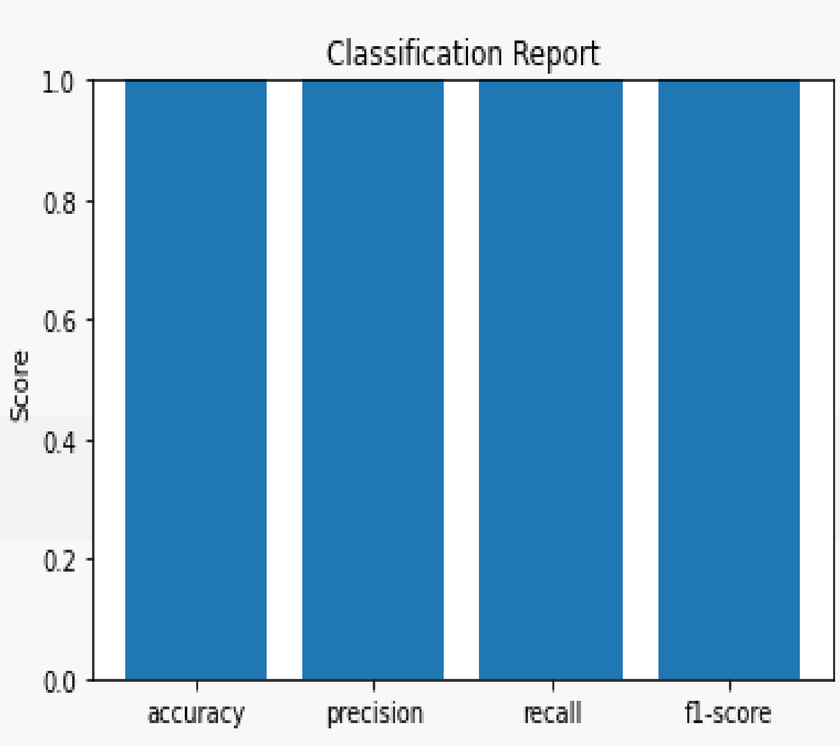
The proposed idea is to enhance botnet detection and prevention by combining multiple machine-learning techniques and real-time model updating. This involves investigating the current state-of-the-art in botnet detection methods, evaluating different machine learning algorithms, and developing a framework that combines multiple learning methods to improve the effectiveness of detecting new botnets.

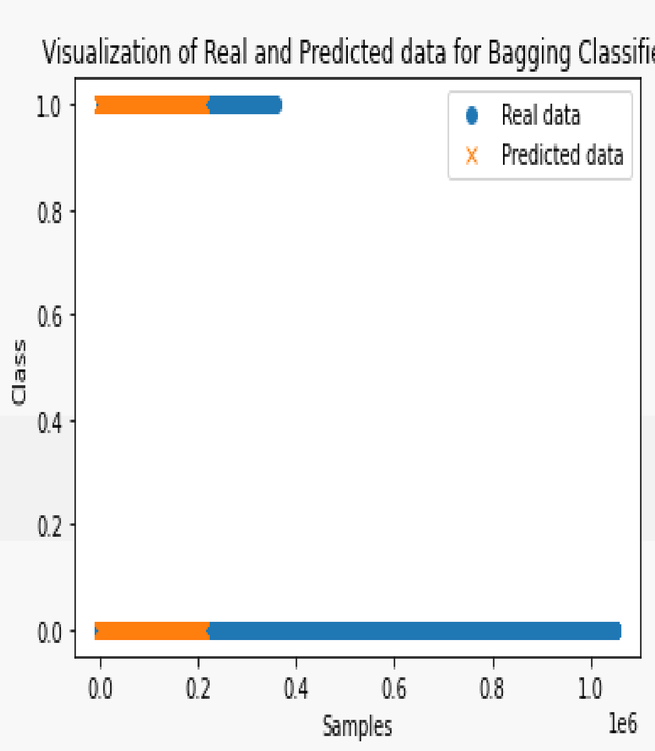
The methodology of this research is given below:

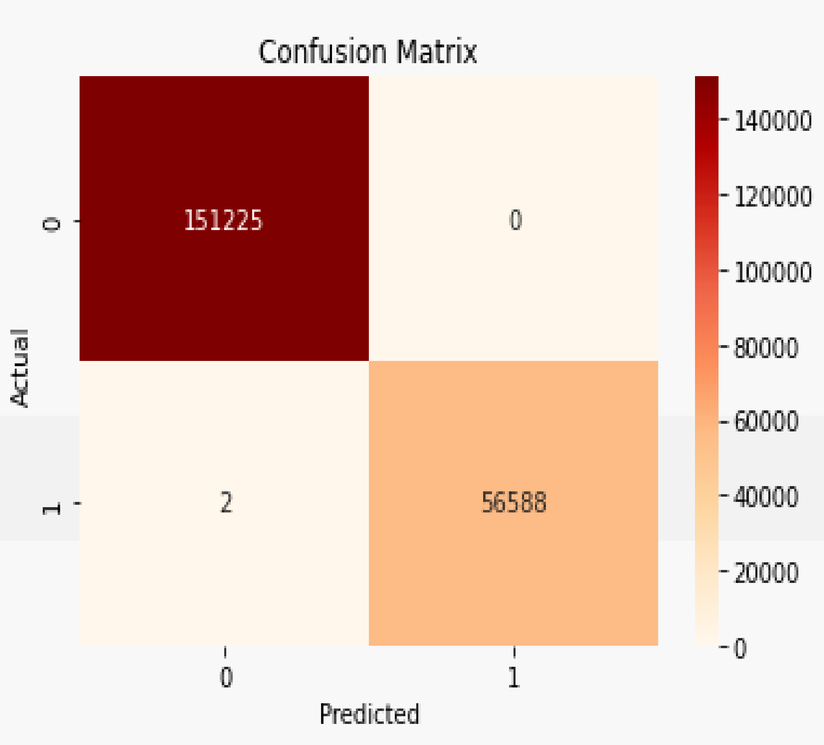


Results

- The proposed model, which uses 79 features, can accurately and successfully detect the botnet in all cases. The result of the various evaluation metrics is given below in figures.
- The accuracy of the proposed model is almost 100%, the error rate is 0.0000, and the accuracy of training and testing data is 100%. All the metrics are far better than the existing models, baseline models, and other ensemble methods like Boosting, Stacking, etc.







Optical Properties Prediction of Negative Dispersion-Compensating Photonic Crystal Fiber Using Machine Learning

Md. Ibrahim Khalil and Dr. Md. Saiful Islam

Abstract

In this work, a novel and highly negative dispersion compensating photonic crystal fiber is structured, and then the study of machine learning approaches has been proposed to predict the output properties like effective refractive index, dispersion, confinement loss, effective area, and V-parameter by varying device dimensions to record different modal solution parameters. The proposed models take fewer computing resources and less time than COMSOL Multiphysics simulation. The machine learning models take milliseconds to train and less than one millisecond to test. Absolute percentage error of less than 0.05% in predicting an output has been obtained by Artificial Neural Network. The proposed PCF with negative dispersion characteristics has the potential for applicability in real high-rate optical communication. This study pave the step towards using machine learning based optimization techniques for such integrated silicon photonics devices.

Background & Motivation

- Microstructure PCFs are receiving particular interest because of their outstanding optical characteristics, which are unattainable with conventional optical fibers.

PCF gives an additional degree of flexibility in designing the guiding characteristics by altering the number, size, and orientation of the circular air holes [1,2].

- Efficient modeling, analysis, and simulation of PCF structures rely on numerical approaches such full vector Finite Element Method (FEM)[11].

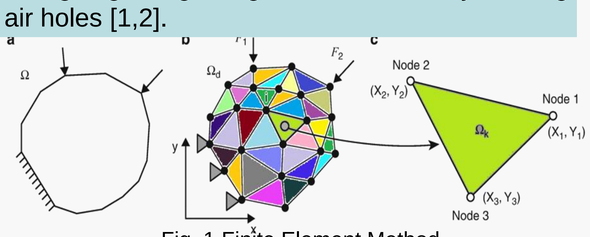


Fig. 1 Finite Element Method

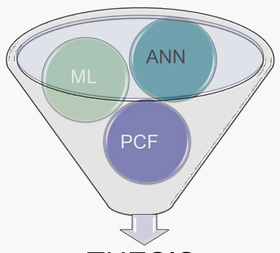
The first study applying a machine learning approach on only solid cores using Artificial Neural Network (ANN)

- Another work related to it combines solid, hollow, and multi-core PCF to predict the optical properties also using ANN [12, 13].

No other works related to this area have been published considering Negative Dispersion Compensating Photonic Crystal Fiber.

The main motivation was to

- Combine the finite element simulations, Artificial Neural Networks (ANN) for the quick and accurate computation.



THESIS

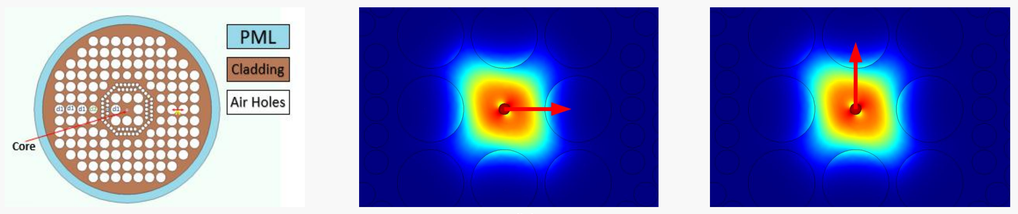
❖Designing and analyzing process requires a lot of computing resources, power, and time.
❖The optimization design also requires several iterative studies.
❖These procedures save valuable time, computing resources, and power.

Proposed Idea and Methodology

1. *Designing of the Devices*

- A total of six circular air cavities and one hexagonal air cavity.
- Pitch (Λ), diameter of air holes 'd', 'd_1 ' d_2', are on the μm scale.
- perfectly matched layer will be made with a thickness that is 9% less than the outer layer.

- Silica will be chosen as the background material because of its exceptional optical transmittance.



(a) (b) (c)

Fig. 2 (a) Cross section of proposed structure, Fundamental field distribution in (b) X-axis and (c) Y-axis

2. *PCF Modelling with ANN*

- The designed PCF is simulated to generate a finite and accurate dataset
- 80% Training data 10% Validation data, 10% Testing data samples are taken.
- 2 hidden layer with 50 nodes gives Best Model.

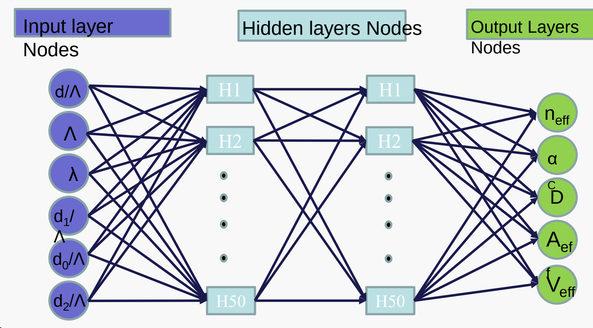


Fig. 3: ANN Used in our design

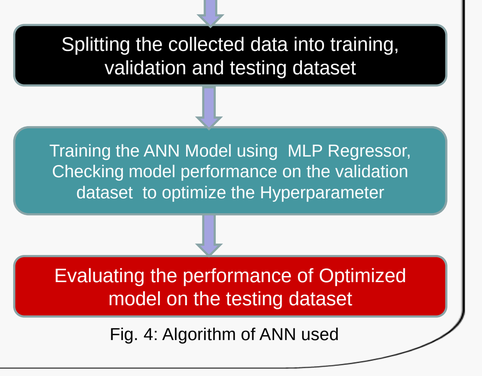
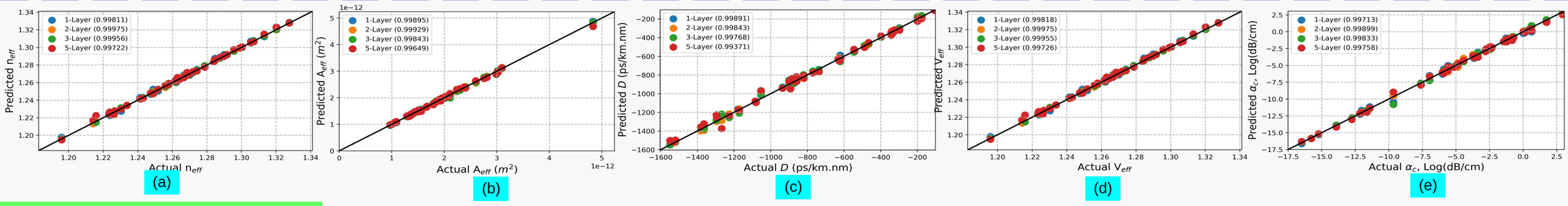


Fig. 4: Algorithm of ANN used

Results



. Inset shows the R squared (R2) obtained with epochs when training the ANN model.

Fig. 5: The scatter plot of training dataset produced by ANN for different number of hidden layer, comparing (a) n_{eff} (b) A_{eff} (c) D (d) V_{eff} (and (e) α_C values from the simulation (x-axis) and the ANN predictions (y-axis) along with the ideal linear model (y = x).

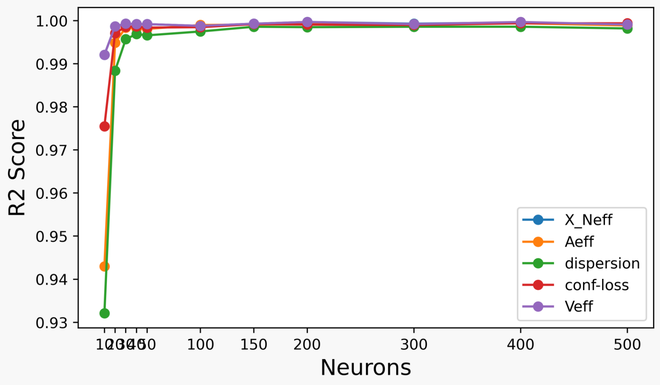


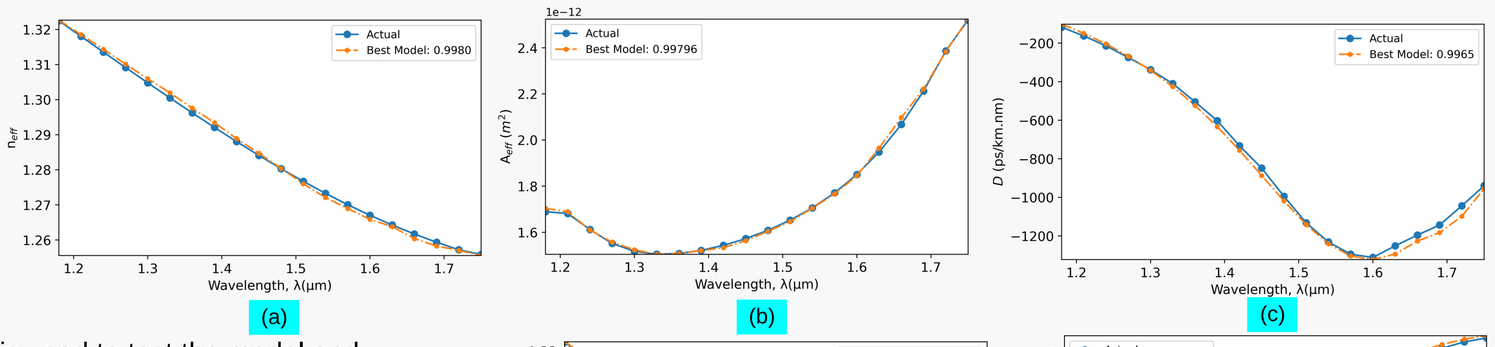
Fig. 3: Optimization of Number of Neurons

- The 300 neurons for first layer gives more accurate result for all parameters.

TABLE I. COMPARISON OF TIME TAKEN BY SIMULATION, ANN MODEL AND PROPOSED MODELS FOR A SINGLE PROPERTY.

Simulation Details	Time Taken	
COMSOL Multipysics (Normal Mesh)	21 Seconds	
COMSOL Multipysics (Finer Mesh)	1 minute 27 seconds	
COMSOL Multipysics (Extreme Fine Mesh)	2 minutes 33 seconds	
Models Details	**Training Time**	**Testing Time**
ANN model with 3 layers and 150 nodes[3]	20 seconds	5 milli-seconds
ANN model with 2 layers and 130 nodes [4]	19 seconds	5 milli-seconds
Proposed ANN models (2 layers, 50 nodes)	10 seconds	5 milli-seconds

- ✓ Less Time
- ✓ More Accurate
- ✓ Less resources
- ✓ Real-time
- ✓ Huge Future Scope



- Unknown dataset is used to test the model and compared with the Actual value.
- R Squared (R2) value is used as evaluation metric.
- All of the parameters gives cumulative R2 value around **0.9983**

Fig. 5: Comparing actual (simulation) and predicted (ANN model) comparing (a) n_{eff} (b) A_{eff} (c) D (d) V_{eff} (and (e) α_C for Best Model at an unknown parameters setting.

Future Scope:

The designed model can be applied to specific Bio-Sensor or such photonic devices to predict the properties within shortest time and almost without error.

Conclusion: Machine learning techniques have the potential to be used in real-time scenarios with the lowest computing resources in optical fiber communication

References:

[1]P. Russell, "Photonic crystal fibers," *science,* vol. 299, no. 5605, pp. 358-362, 2003.
[2]T. Matsui, J. Zhou, K. Nakajima, and I. Sankawa, "Dispersion-flattened photonic crystal fiber with large effective area and low confinement loss," *Journal of Lightwave Technology,* vol. 23, no. 12, pp. 4178-4183, 2005.
[3]S. Chugh, A. Gulistan, S. Ghosh, and B. Rahman, "Machine learning approach for computing optical properties of a photonic crystal fiber," *Optics express,* vol. 27, no. 25, pp. 36414-36425, 2019.
[4]H. Kumar, T. Jain, M. Sharma, and K. Kishor, "Neural network approach for faster optical properties predictions for different PCF designs," in *Journal of Physics: Conference Series*, 2021, vol. 2070, no. 1: IOP Publishing, p. 012001.

IoT BASED SMART WATER CONSUMPTION MONITORING SYSTEM FOR RESIDENTIAL COMPLEX

Imtiaz Ahamed, Md. Liakot Ali

Abstract

Environmental challenges are on the rise, and the population is expanding quickly, creating a huge demand for water supply. In turn, these problems have elevated water management and conservation to the level of a survival need for people. It is essential to use water effectively because of excessive worldwide water waste. Thus, one of the solutions that must be put into place to stop water waste is smart systems. In this poster, electronic sensors are used to monitor and regulate the flow of water, and LoRa technology is used to broadcast the sensor data to a central server via WiFi technology. The data traversed will be monitored by the central authorities continuously and the user will be updated on a real-time basis. The update will include on water usage volume. An application software will run and deliver these information straight to the user. Sensors (water flow/amount detecting sensor), a LoRa module, and a microcontroller based control unit are the basic construction components used in the process. Many useful reports on water usages can be prepared from the data gathered which will help the central authority to take a well decision to solve the water waste problem for the country.

Background & Motivation

Bangladesh is a densely populated country with limited freshwater resources. According to the World Bank, Bangladesh is one of the most water-stressed countries in the world, and its water resources are likely to become scarcer in the future due to climate change and population growth.

In this context, the monitoring of water consumption in households can play an important role in ensuring the efficient use of water resources. By tracking water usage, residents can identify areas of waste and take steps to reduce consumption, ultimately helping to conserve water resources and reduce water bills.

Internet of Things (IoT) technology can provide an effective means of monitoring water usage in households. IoT devices can be installed on water pipes and connected to the internet, enabling real-time tracking of water consumption. This data can then be used to provide insights into patterns of water usage, identify leaks and inefficiencies, and enable residents to make informed decisions about their water consumption.

Figure 1. Water scarcity in near future due to lack of monitoring and wasting

Overall, an IoT-based water consumption monitoring system has the potential to promote sustainable water use, reduce water waste, and improve access to safe and clean water

Proposed Methodology

The proposed method is illustrated in Figure 2 in a nutshell.

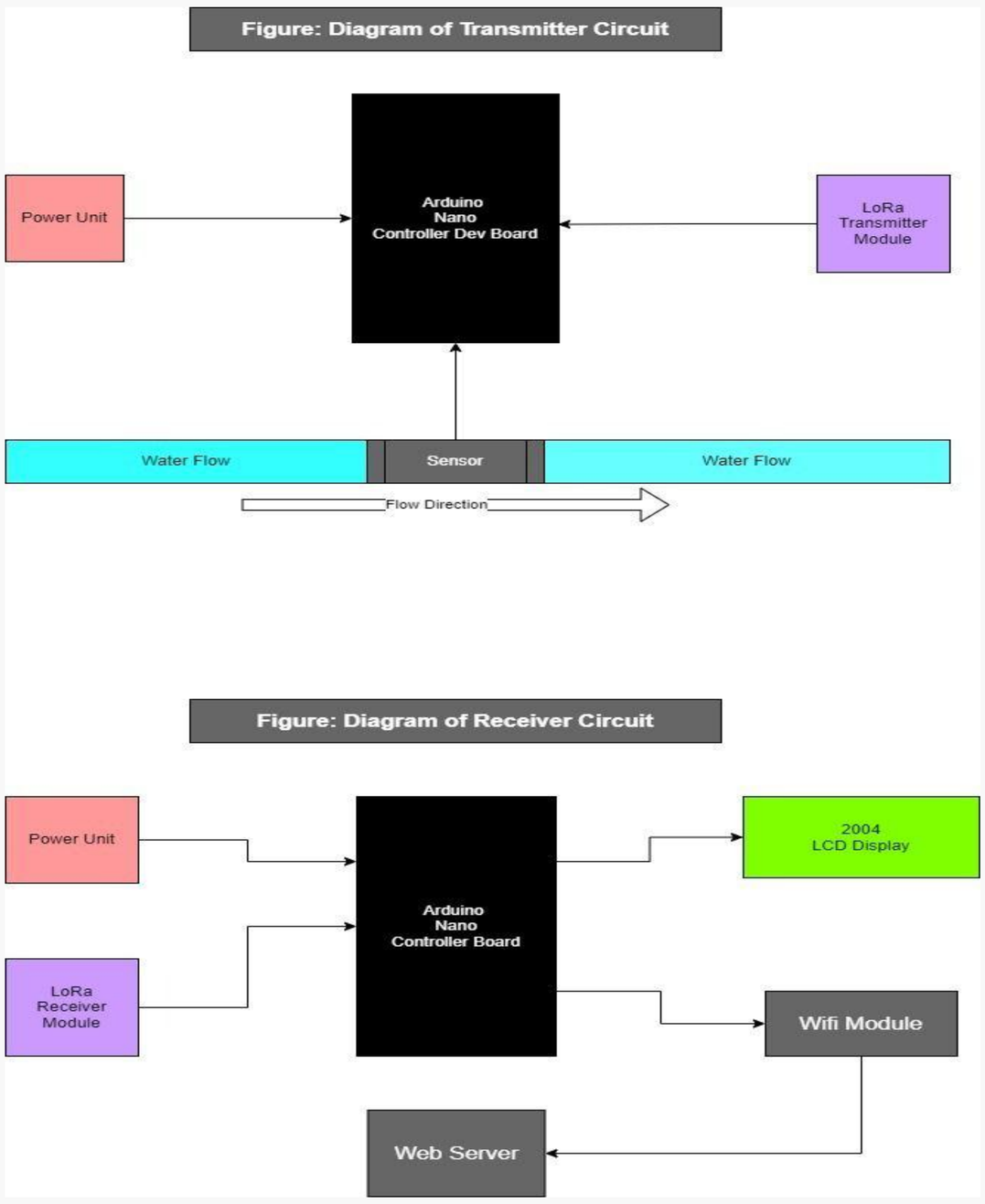


Figure 2. The flow-diagram of the proposed water flow measurement using LoRa

Results

- Real-time monitoring: With the implementation of IoT sensors, water usage can be monitored in real-time. This allows for better tracking of usage patterns and identification of any abnormalities or wastage.
- Reduction in water wastage: With the ability to monitor water usage in real-time, residents can be made aware of their consumption habits and take measures to reduce water wastage.
- Cost-effective: LoRa technology is a low-cost, low-power, and long-range wireless communication technology, making it a cost-effective option for implementing an IoT-based water consumption monitoring system.
- Improved sustainability: By reducing water wastage, the implementation of an IoT-based water consumption monitoring system can contribute to the overall sustainability of the residential area.
- Overall, IoT-based water consumption monitoring can help promote water conservation, reduce waste, and improve the efficiency of water management in residential areas.

❖ .

Automated Writing Evaluation Using Sentence by Sentence Scoring Model

Mehadi Hossain and Hossen A. Mustafa

Abstract

In this work, we propose a methodology for evaluating student essays using automated writing evaluation. The evaluation is done at the sentence level to overcome the ambiguity of evaluating essays as a whole. A publicly available sentence-wise dataset was collected and then evaluated manually to prepare scored dataset. Two models were developed: the first model is based on pre-trained models (BigBird, Longformer, and DeBERTa), while the second model is a newly designed neural network architecture based on multi-head-attention mechanisms. Models calculate the holistic score for each sentence. The final mark for each essay is obtained by summing up the marks for each sentence. Experimental results show that the model can score with high accuracy.

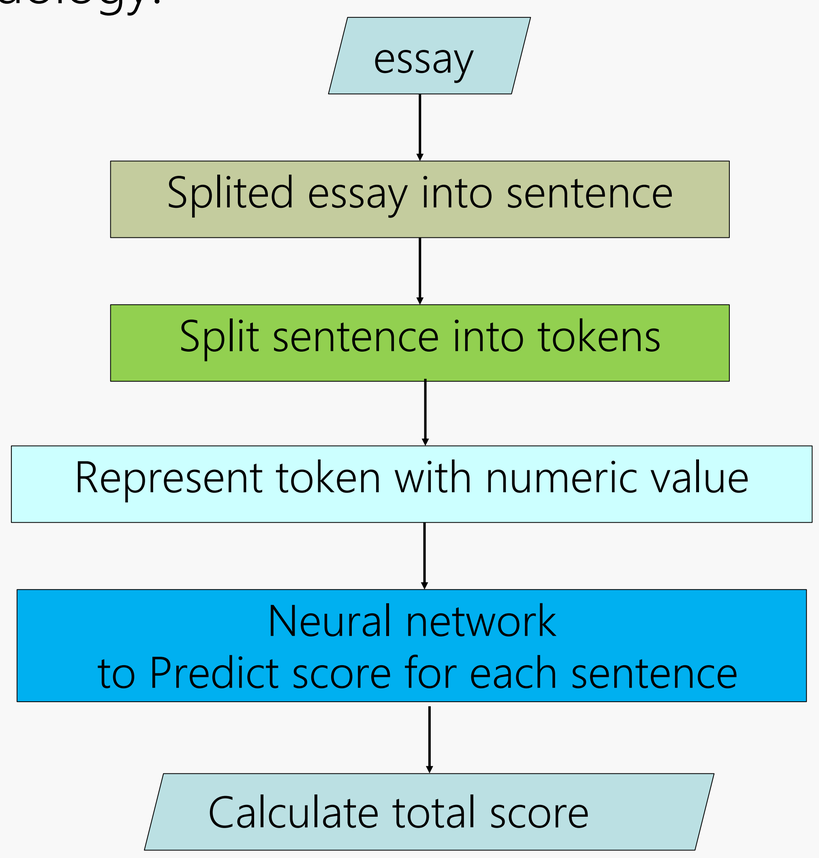
Background & Motivation

Automated writing evaluation is an automation process to evaluate writing with the help of computer programs. In educational settings, this process is used to give feedback on a piece of written work. Automated writing evaluation is a longstanding, active research area that attracts a lot of research efforts.

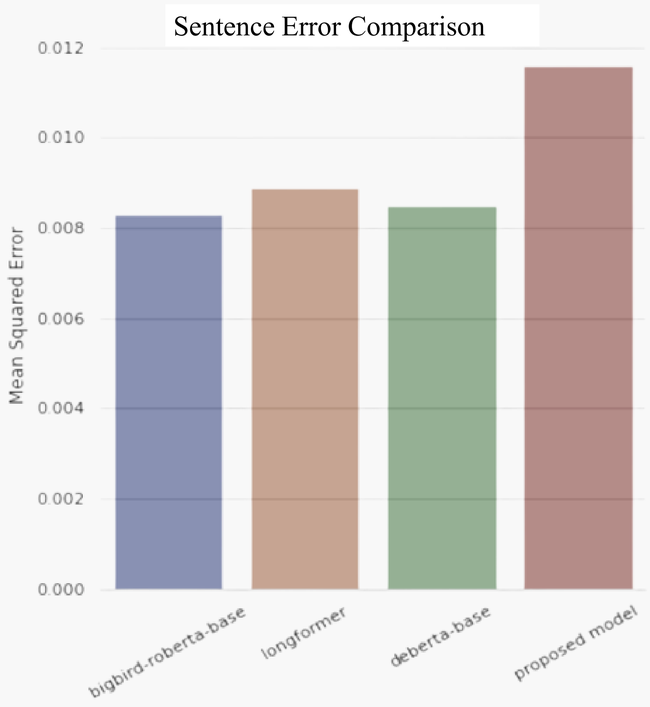
Generally, we see evaluations given based on a whole essay. But, a whole essay consists of many sentences. The ambiguity of sentences makes the learning model difficult to evaluate. For this reason, we propose a sentence-level evaluation system. In this case, each sentence will have a score and the total score will be calculated based on all sentence scores.

Proposed Idea and Methodology

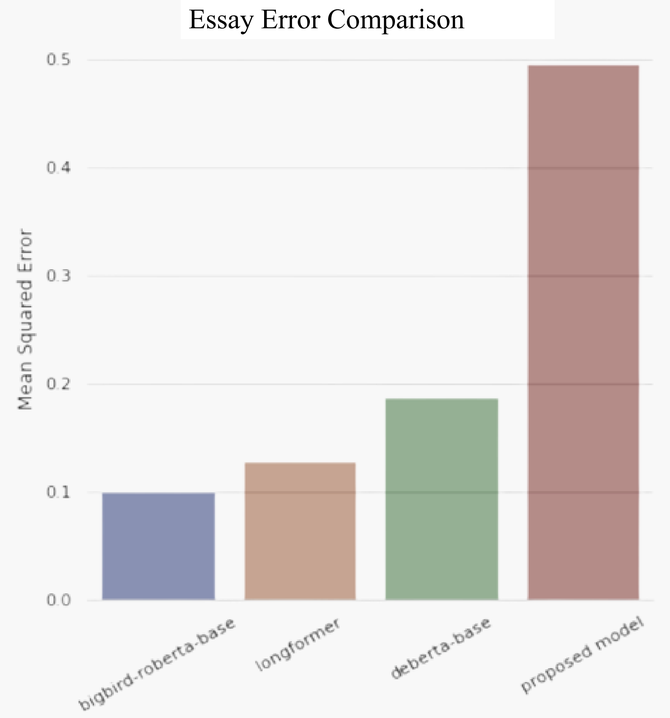
Our main focus of this research is to construct an intelligent system that will evaluate essay writing. Here is the outline of the methodology:



Results



- Pretrained models perform better than the proposed model (MSE: 0.4956).
- Deberta-base performs better in sentence-level scoring (MSE: 0.0086), bigbird-based performs better in essay-level scoring (MSE: 0.07).
- Bigbird-based model is best performer for full essay evaluation.
- The proposed model still viable for evaluating writing.



Towards Tolerating Soft Errors in Memristor-Based Memories

Md. Mehedy Hasan Sumon, Md. Liakot Ali, Muhammad Sheikh Sadi

Abstract

Memristor is creating revolutionary changes in memory technology due to its compact size, higher density, lower power consumption, faster operating time, and non-volatile property. However, an effective soft error tolerance method is required to enhance the reliability of memristor-based memory. This research proposes a new method to tolerate soft errors in memristor-based memories. It is the enhancement of SEC-DED (Single Error Correction and Double Error Detection) code. It includes horizontal and vertical check bits with SEC-DED and is defined for short the HV-SEC-DED method. This method can be applied to any size of dataword in the memristor-based memory. The complete set of equations for error correction are developed, and these are used to design the error corrector circuitry. A fault injector circuit has also been designed to inject a single, double, or triple-bit error in the dataword. MATLAB code and a Simulink model have been developed for implementing the proposed method in memristor-based memory. The proposed HV-SEC-DED (64) method has 24.48%, 26.56%, and 10.93% lower bit overhead than Golay, BCH, and HVD (64) respectively.

Background & Motivation

- Memories are widely used in systems where millions of transistors are integrated on a single chip.
- Due to the larger chip area of semiconductor memories, they suffer more cosmic radiation than other components.
- The impact of this radiation on semiconductor memories is huge because it can cause temporary malfunctions in memory cells which are defined as soft errors.
- The explosion of Ariane-5 missile system (as shown in Figure 1), is an example of the disastrous consequences of soft errors.
- Memristor is a promising candidate for replacing traditional semiconductor memories and has less sensitivity to soft errors.

Figure 1. The explosion of the Ariane 5: The most infamous and expensive accident in history due to soft-error.

- As far we have reviewed, negligible research on soft error tolerance in memristor-based memory is available.
- An efficient solution for soft errors in memristor-based memory is essentially required for adopting this advanced technology.

Proposed Methodology

The proposed method is illustrated in Figure 2 in a nutshell.

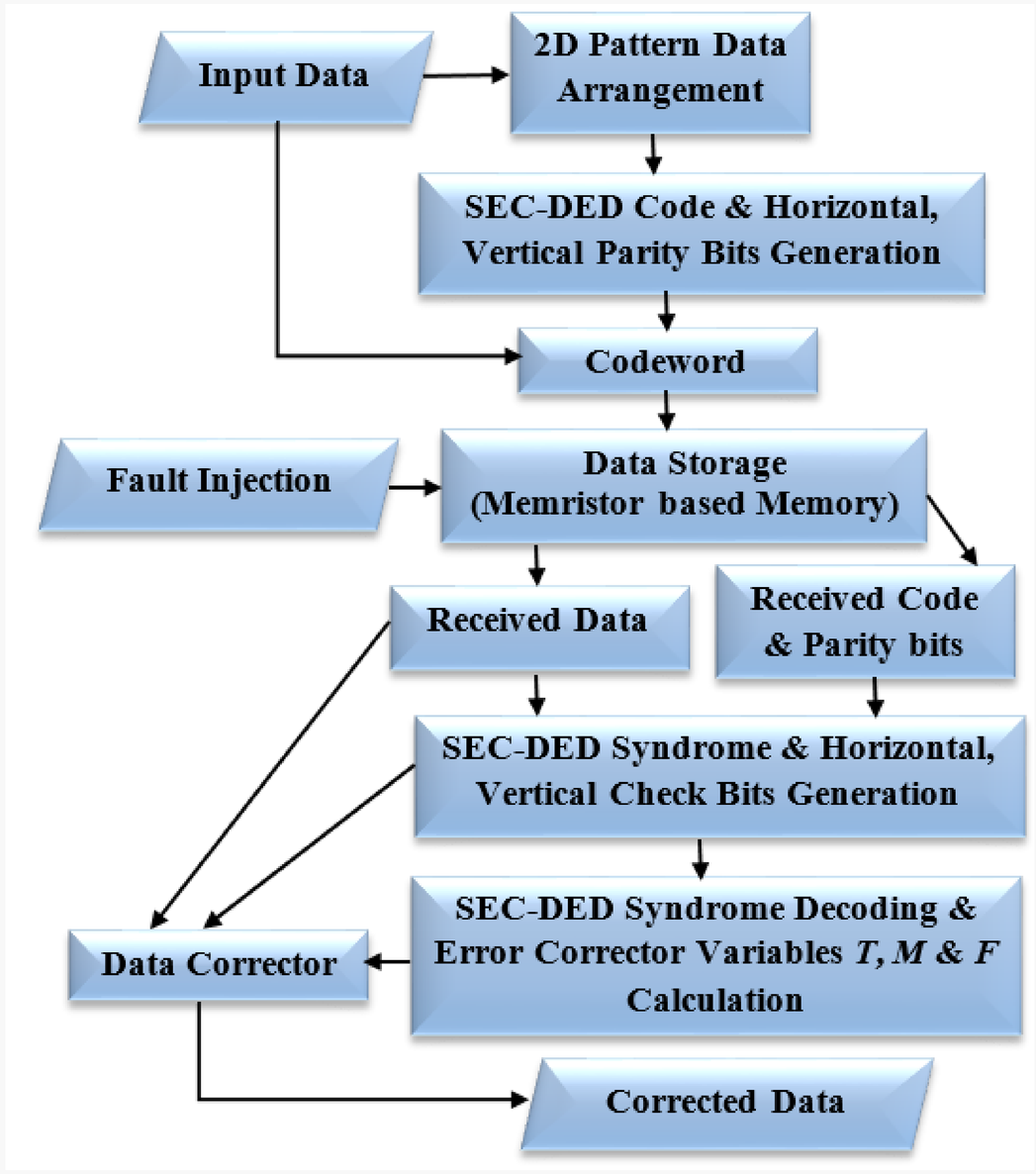
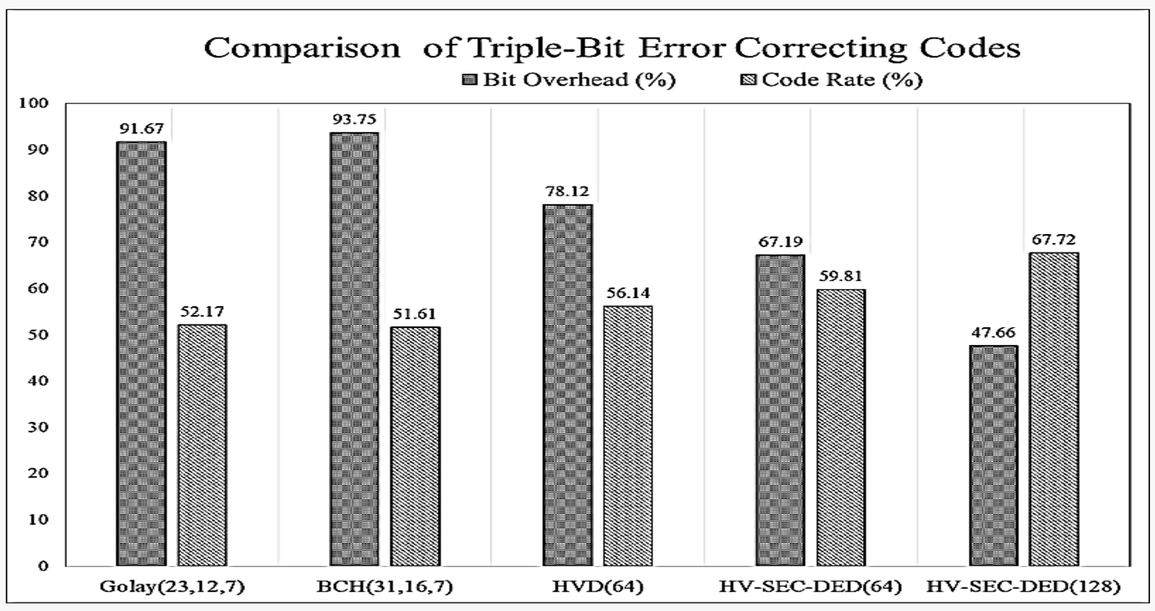


Figure 2. The flow-diagram of the proposed HV-SEC-DED method for 3-bit error correction.

Results

- The HV-SEC-DED (64) method has 24.48% lower bit overhead and 7.64% higher code rate compared to the Golay method. It has also 26.57% lower bit overhead and 8.2% higher code rate compared to the BCH method. And it has 10.94% lower bit overhead and 3.67% higher code rate than HVD.
- The HV-SEC-DED (128) method has 44.01% lower bit overhead and 15.55% higher code rate compared to the Golay method. It has also 46.09% lower bit overhead and 16.11% higher code rate compared to the BCH method. And it has 30.47% lower bit overhead and 11.58% higher code rate than HVD.



Real Time DDoS Attack Detection Using ML on SDN

Md. Nayem Khan and Dr. Md. Saiful Islam

Abstract: Today's world is getting largely dependent on internet communication and there is also happening various cyber-attacks too. Distributed Denial of Service (DDoS) is one of the most severe cyber-attacks that can prevent the legitimate users to access the wanted resources. With the use of emerging technologies, it is very easy and law cost to generate DDoS attack at very high range. DDoS attack is also changing its manner or types of attack vectors in such a sophisticated way that, it is very much difficult to trace out it as well. This project will be focused the way of DDoS detection and hance we use machine learning technique on SDN. The simulation process for making DDoS traffic, Software Defined Networking (SDN) is used. Here, proposed a network with sufficient hosts with server and attacker. For data collection, here used Wireshark network analyzer and CICFlowMeter-4 and from SDN Ryu controller and classified the normal and attack data. Then applied machine learning algorithms to classified data with previously available dataset. Here used information gain algorithm to calculate precision, recall and f1-score and plotting ROC curve to measure the accuracy of the applied ML algorithms. So, the proposed system will be able to detect DDoS traffic with high accuracy in real time.

Background & Motivation

Distributed Denial of Service (DDoS) attack is a malicious attempt to make a network resource unavailable to its intended users and affect the performance. In recent years there are several DDoS attacks. In Bangladesh Covid-19 vaccine website surokkha.gov.bd was under DDoS attack and vaccination service was hampered. Most of the researchers are trying to pay attention to detect and stop DDoS attack using machine learning approach. But there have some limitations to do so.

M. Aamir and etal. in their paper proposed a method of DDoS detection with feature engineering and machine learning. Their experiment accuracy was on average 92-95%. They proposed a solution using laboratory made dataset and their accuracy was good and time to detect the DDoS was moderate. Most of the researcher use only ML algorithm sand is not sufficient to detect DDoS effectively. Here, proposed methodology we will use dataset created by capturing the traffic. We also focus on secured environment setup, real time detection and high accuracy. We use Hybridge data collection system to meet the multidimensional traffic detection and accuracy.

There are several Machine Learning approaches are proposed like Artificial Neural Network (ANN), Support Vector Machine (SVM), Fuzzy Logic (FL), Bayesian Networks (BN), Decision Tree (DT), J48, Naïve Bayes (NB), Random Forest (RF) etc. On the other hand, Software-Defined Networking (SDN) is a networking approach that separates the control plane, which makes decisions on how data should be sent and received, from the data plane, which forwards the data itself. In traditional networking architectures, the control and data planes are tightly integrated, making it difficult to make changes or updates to the network's control logic. With SDN, the control logic is decoupled from the underlying hardware and implemented in software, making it more flexible and easier to manage. So, it is comparatively easy to generate/capture traffic with help of Ryu controller (OpenFlow Protocol, mininet) on this environment setup. Wireshark and CICFlowMeter-4 is also used to collect day to day use of real dataset. We will use various tools to generate DDoS traffic. Then we marge the dataset and level them into normal traffic and DDoS traffic as 0 & 1. In real world scenario it is very much tough to detect DDoS traffic by analyzing with the help of human & tools. On the same side, it is very much costly, tough, time and bandwidth consuming for the small business to manage cloud and hardware solution.

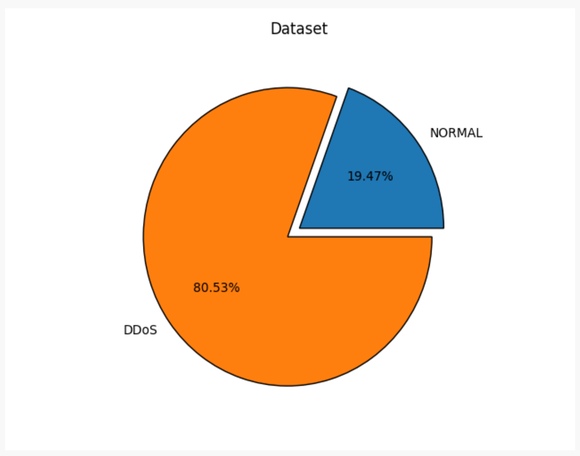
Proposed Idea and Methodology

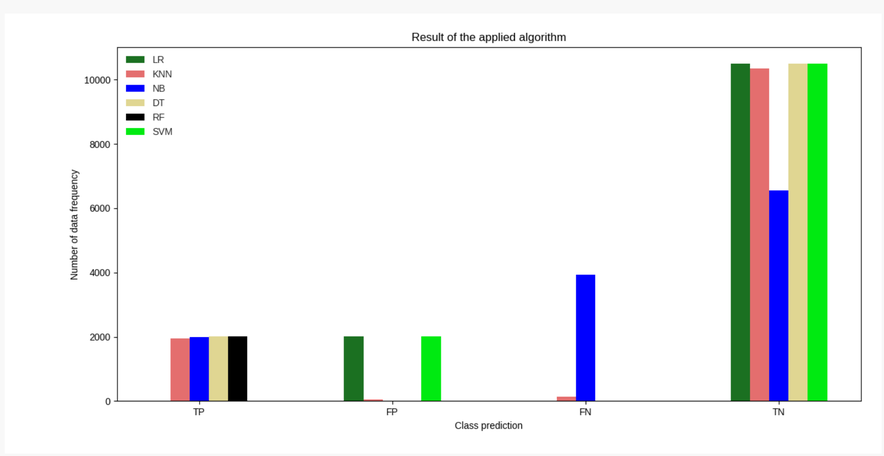
To get rid of this problem we proposed a machine learning approach to detect DDoS using various machine learning algorithm on SDN. Methodology is given below-

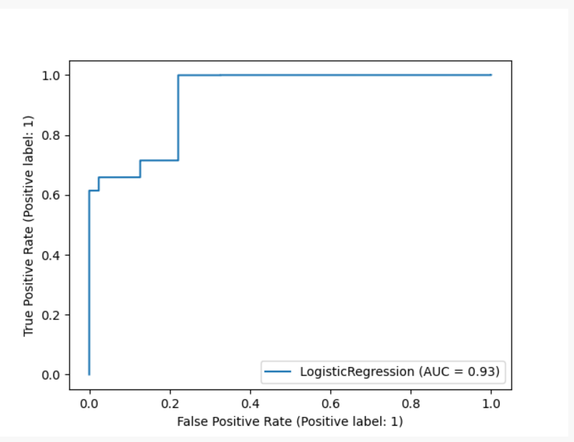
1. Various DDoS attack will be performed on SDN environment with the help of Ryu SDN Controller (mininet, OpenVSwitch) and may be set up in VMware or in Ubuntu 20.00 TLS.
2. Different types of DDoS attack will be executed using three or more DDoS attack generation tools. Hping3 will be used to generate TCP, UDP and ICMP based flood attack. Law Orbit Ion Canon (LOIC) or High Orbit Ion Canon (HOIC) is also used to HTTP/ Flood. Legitimate or Benign traffic will be collected from Lab environment or daily uses of internet traffic.
3. Dataset Size is more than 300MB+ and there will be more than 20m traffics are capture. DDoS are captured with 3 sec latency and normal traffic are captured with 10 sec periodical gap.
4. Wireshark a packet analyzer will be used to capture traffic. Also, CICFlowMeter-4 and developed script (if needed using python) will be used to fine tuned the dataset. A new dataset will be created using all the recorded traffic where every record will be like real world data.
5. The dataset will have the characteristics of relevance, representativeness, non-redundancy, scalability, and reusability. All the combination of possible DDoS volumetric attack in SDN will be incorporated by parameterizing the attack tools to make the dataset more realistic.
6. Detection system will be real time with proper information.
7. The performance will be evaluated by calculating precision, recall and f1-score and plotting ROC curve. Finally, to assess the effectiveness of our proposed model, a comparative analysis among the machine learning algorithms and the methods for DDoS attack detection will be compared to each other.

Results

1. A simple dataset containing different types of DDoS attack in SDN environment.
2. Finding appropriate Machine Learning algorithm to detect DDoS attack in real time.
2. Comparison of performance with other machine learning algorithms and to validate the obtained results.
3. Finally, Real-Time detection of DDoS Attack with sufficient warning message of affected Server/Clients.
4. Some Figures are given below-







Economic Denial of Sustainability attack detection and mitigation using machine learning

Md. Sharafat Hossain and Dr. Md. Saiful Islam

Abstract Though cloud computing saves the organization from investing a lot of money from buying and managing IT infrastructure it suffers from various security flaws. Economic Denial of Sustainability (EDOS) is one of which mainly abuse the pay-as-per-use and elasticity property of cloud computing. AI and Machine Learning are showing great success in the field of detecting malicious packet including distributed one. In this research we will implement an EDoS attack detection system using machine learning and also will implement an attack mitigation system.

Background & Motivation

The EDoS attack is a new type of cyberattack that targets cloud computing. It takes advantage of the auto-scaling feature to increase the number of resources assigned to the cloud customer and drive up the cost of the service.

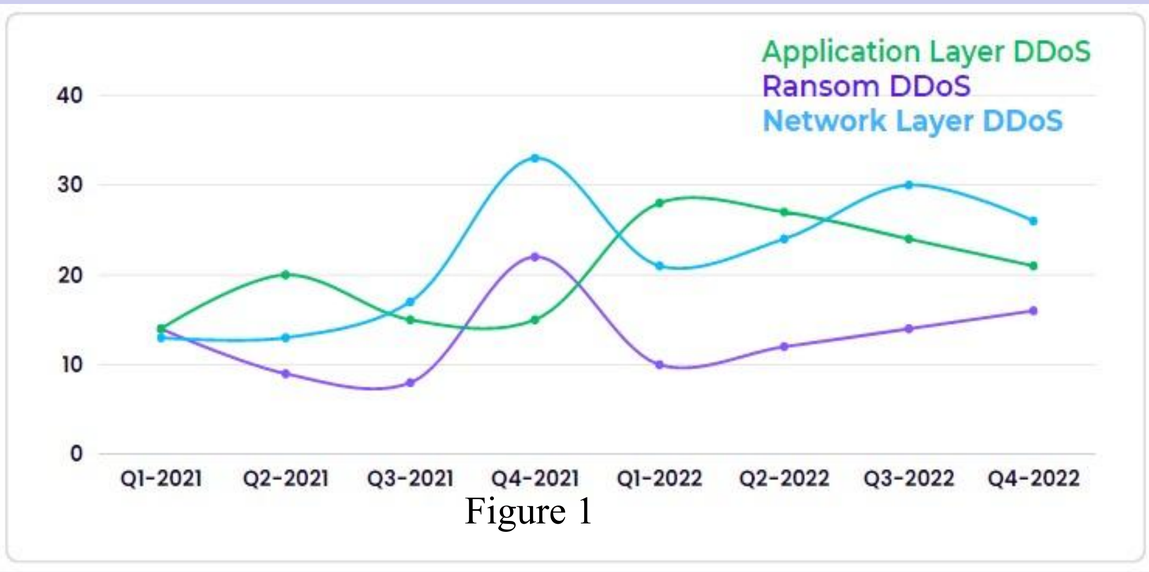


Figure 1

have experienced at least one DDoS attack.

- Financial loss due to an EDoS attack is growing rapidly and becoming a big threat to business organizations.
- Machine learning algorithms are doing well in the packet classification problem.

Proposed Idea and Methodology

First, for our proposed system, we build a machine learning model that can identify malicious packets. We will deploy the model once we have received satisfactory results from it.

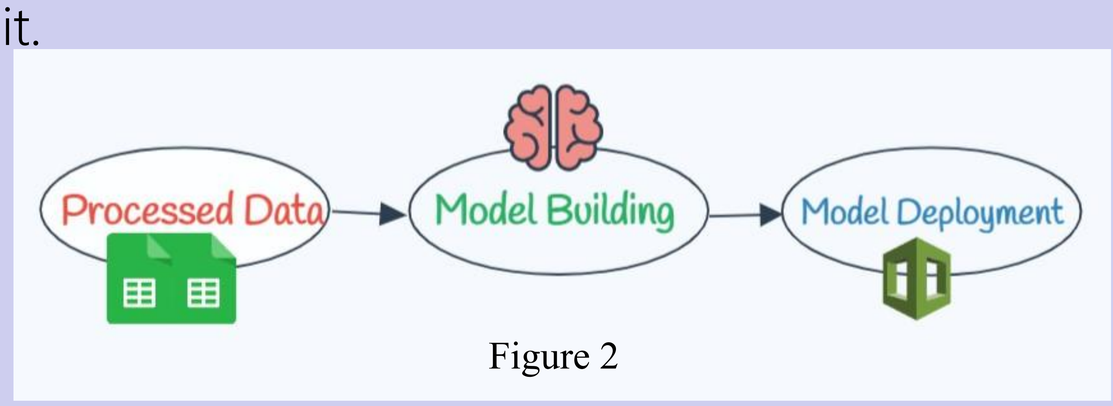


Figure 2

Next, after the deployment of the model, we will integrate the model into our attack mitigation system. The figure here describes the structure of the overall system.

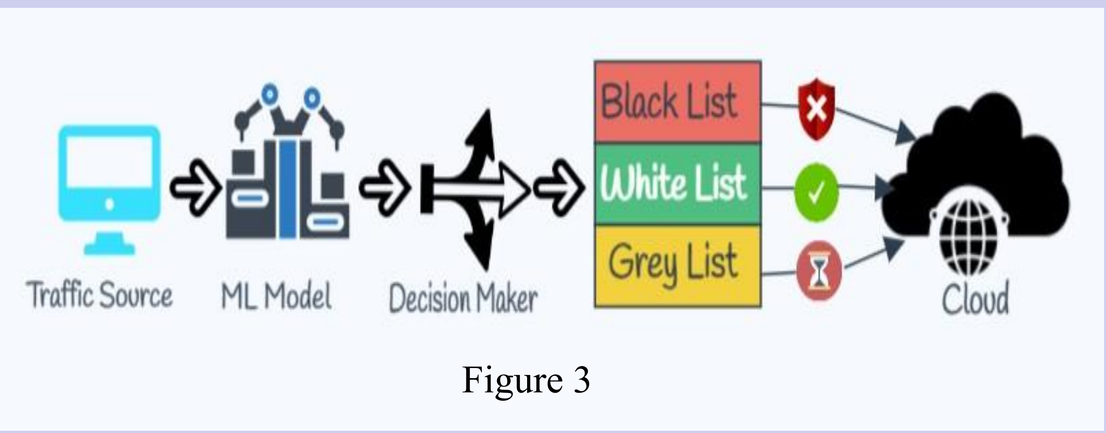


Figure 3

Expected Results

ML Model	SVM	DT	RF	XGB	E.Vote	NN	CNN	RNN
Curr. Acc.	92.55	93.73	95.16	95.01	95.03	97.78	94.03	-
Exp. Acc.	~95	~95	~96	~96	~96	>99	>99	>99

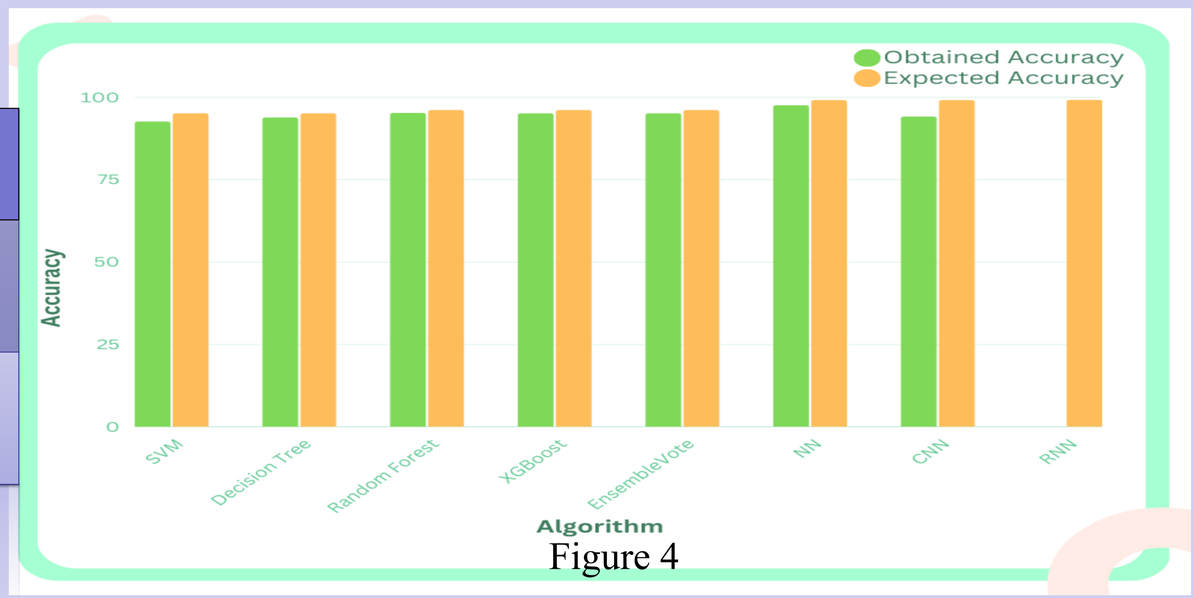


Figure 4

Towards Assisting the Visually Impaired to Detect Objects

Sheezanul Hassan, Md. Liakot Ali, Muhammad Sheikh Sadi

Abstract

Safe and independent mobility is one of the major daily challenges faced by the visually impaired. The visually impaired should be aware of obstructions and other interference to navigate properly in new or unfamiliar situations. However, establishing secure and safe mobility for the visually impaired is a difficult undertaking that must be done accurately and efficiently. Recognizing currency is another severe problem that the visually impaired confront because different notes, in our country, have similar surfaces and sizes. In this research, we propose an object and currency detection system that will benefit the visually impaired to detect objects and Bangladeshi currency notes and make the appropriate decision regarding their situation in both indoor and outdoor aspects. The model proposed in the system has been evaluated with real-world objects for evaluating the performance of the proposed method. The experimental analysis shows that the system has an average accuracy of 99.44%, a recall of 99.38%, a precision of 99.39%, and an F1 score of 99.37% after experimenting with 5281 images of real objects.

Background & Motivation

- Around 43 million people are blind and around 295 million people in the world have moderate to severe vision impairment
- Visually impaired people fail to detect and avoid obstacles in their path, thus causing them emotional suffering, undercutting their independence.
- Identifying currency-note is another major issue for the visually impaired.
- Smart Cane, Ultra-Cane, Ray Mobility Aid are some of the commercially available devices for visually impaired people.

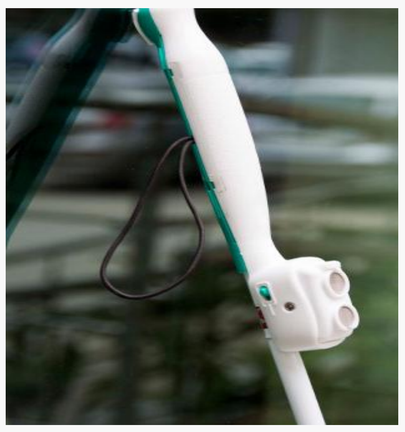


Figure 1. Smart Cane

Figure 2. Ultra Cane

- The acceptance rate of these devices is relatively low due to the high cost, inaccuracy for detecting obstacle types, the identification of different objects in both daylight and dark environments, and the detection of currency banknotes
- There is room for research to create a stand-alone system that can handle the objects and currency detection system for the visually impaired.

Proposed Methodology

- The proposed methodology can be summarized in the following steps

Figure 3. Steps of the proposed method

- The proposed system uses EfficientNetB0 CNN architecture for object and currency detection and transfer learning. The model for object detection is illustrated in Figure 4.

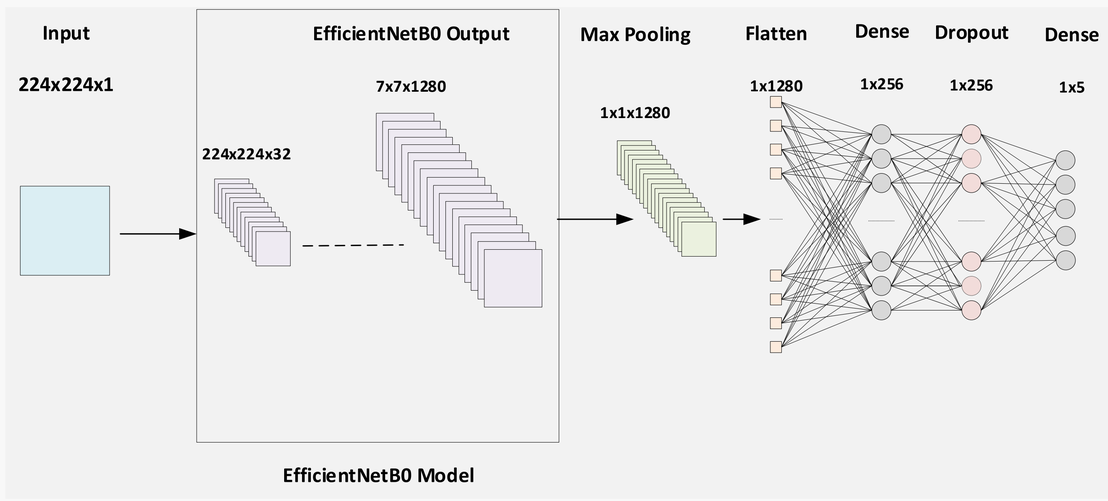


Figure 4. Summary of the proposed model for object detection

Results

- Weighted average, precision, recall, and F1 score for all five classes for object detection are 99.826%, 99.825%, and 99.825% respectively on 1144 samples.
- Overall values of precision, recall, F1 score, and accuracy of all of the fifteen classes for currency detection are 99.39%, 99.38%, 99.37%, and 99.44% respectively
- Comparison of performance between the proposed system and the existing system are shown in Table 1.

Table 1. Comparison of proposed system with existing systems

Method based on	Coverage Area	Class	Obstacles	Currency detection	Accuracy
Convolutional Neural Network (CNN), Recurrent Neural Network (RNN) (Ahmed, F. et al. 2019)	Indoor	5	Yes	No	95%
Faster region-convolutional neural networks (Faster R-CNN) with Resnet50 (Dahiya et al., 2020)	Indoor, outdoor	4	Yes	No	92%
Convolutional Neural Network (CNN) with Resnet50 (Afif et al., 2020)	Indoor, outdoor	5	Yes	No	99.80%
YOLO (You Only Look Once) based algorithm and IOU (Intersection of Union) (Kumar and Jain, 2021)	Indoor, outdoor	25	Yes	No	96.14%
Convolutional Neural Network (CNN) with EfficientNetB0 (our proposed method)	Indoor, outdoor	15	Yes	Yes	99.44%

UNSUPERVISED RECYCLED FPGA DETECTION USING SYMMETRY ANALYSIS

Tanvir Ahmad Tarique, Foisal Ahmed, Md. Liakot Ali

Abstract

Recently, recycled field-programmable gate arrays (FPGAs) pose a significant hardware security problem due to the expansion of the semiconductor supply chain. Ring oscillator (RO) based frequency analyzing technique is one of the popular methods, where most studies used the known fresh FPGAs (KFFs) in machine learning-based detection, which is not a realistic approach. In this study, we propose a novel recycled FPGA detection method based on an unsupervised anomaly detection scheme by analyzing the symmetry information of the RO frequency. As the RO frequencies in some neighboring logic blocks on an FPGA are similar because of the symmetrical array structure of the FPGA, our method compares the RO frequencies only of those blocks as a similarity analysis. The proposed method efficiently classifies recycled FPGAs through outlier detection using direct density ratio estimation. Experiments using Xilinx Artix-7 FPGAs demonstrate that the proposed method successfully distinguishes recycled FPGAs from 10 fresh FPGAs by x% fewer computations compared with the conventional recycled FPGA detection method.

Background & Motivation

- Nowadays, recycled field-programmable gate arrays (FPGAs) are a significant counterfeit issue in the IC supply chain due to the increased number of third-party IC vendors.
- More than 80% counterfeit components are recycled, and the recycled FPGAs have reliability risks and trustworthiness issues due to aging-induced performance degradation.
- Global FPGA market is increasing day by day with multitude of applications such as self-driving vehicles, neural network accelerator, cloud computing, 5G communication, aerospace and satellite communication, medical instruments etc.
- If untrusted FPGAs infiltrate mission-critical systems, the system's reliability may suffer, causing significant incidents which are global threat for both the government and industries.
- This is the concern to detect the recycled FPGA with low-cost and high accuracy.

Figure 1. Some applications of FPGAs. From top-left in clockwise direction, self-driving vehicles, cloud computing, aerospace and satellite communication, and medical devices.

- Currently, the techniques used to detect recycled FPGAs require very large number of datasets for the supervised ML approaches, and they require a vast amount of different types of calculations, which is very costly to accomplish.
- Also, the dataset of recycled FPGA is unavailable which led to use the unsupervised ML approach with K-means++ clustering algorithm.

Proposed Methodology

The proposed method is illustrated in Figure 2.

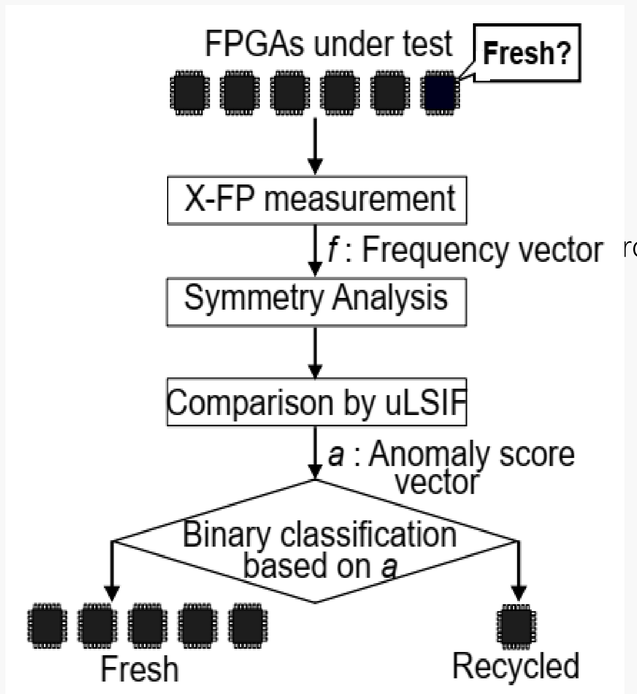


Figure 2(a). Flow chart of the proposed method.

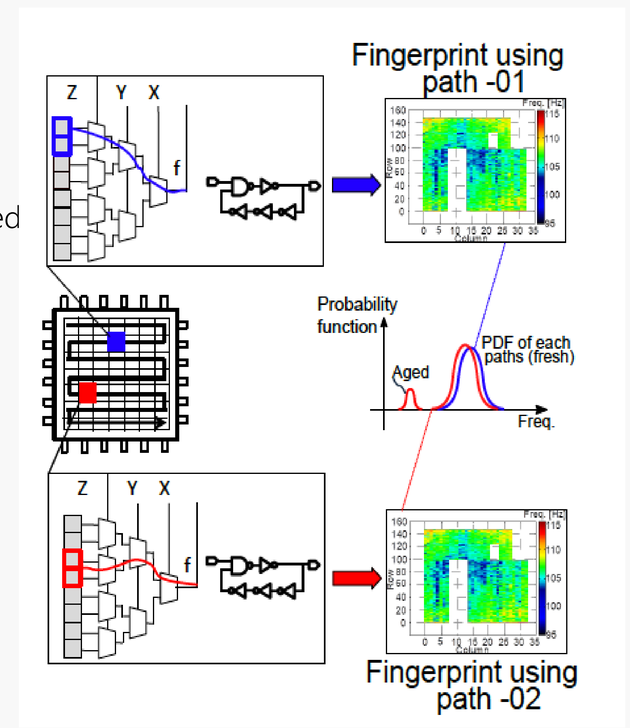


Figure 2(b). Basic Idea of the proposed method

Results

- The proposed method successfully demonstrates the use of symmetry analysis in case of FPGAs which requires around 40% less computations.
- This method achieves more than 92% accuracy within a very short period of time.
- It detects all the 3 recycled FPGAs accurately which were used for the purpose of this research.

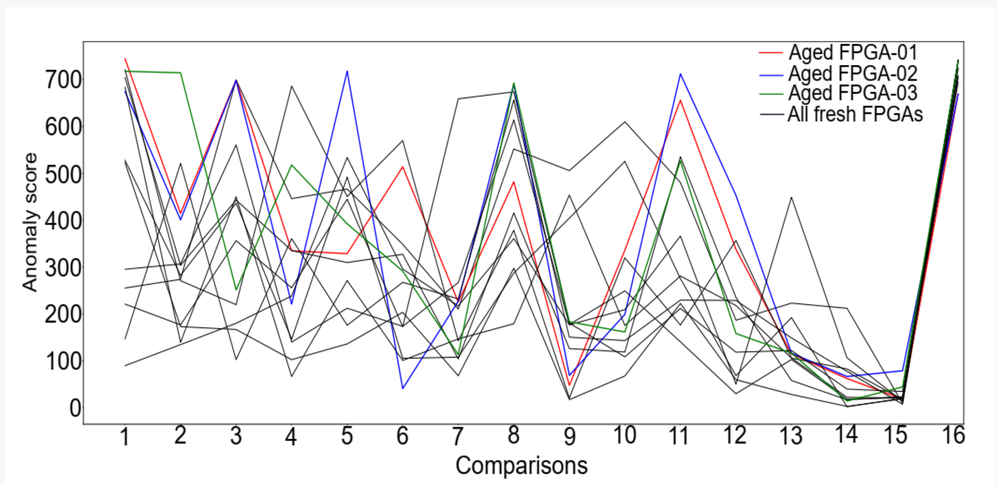


Figure 3. Maximum anomaly scores of different comparison paths using the proposed approach where used 10 fresh FPGAs and three artificially aged or recycled FPGAs

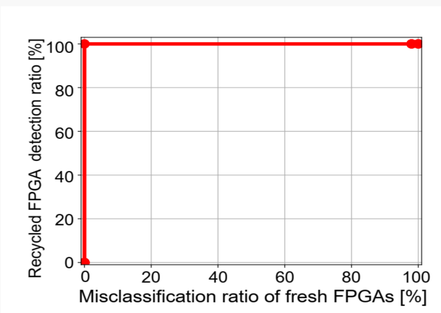


Figure 4. ROC curve of the proposed recycled FPGA detection method using 10 fresh FPGAs and 3 aged FPGAs