Journal of Research Proceedings

JRP

Under the delegate of "Journal of Research Proceedings," we anchor a bimonthly electronic journal enclosing the diverse realms of the educational research field. JRP is providing a platform for the researchers, academicians, professionals, practitioners, and students to impart and share knowledge in the form of high quality empirical and theoretical research papers, case studies, literature reviews, and book reviews.

JRP Publications
www.i-jrp.com
journalrp.editor@gmail.com
9353189468

Machine Learning Approach for Prediction of Crop Yield

Ashwini V[1], Yogitha M[2], Prof.Veena M. Naik[3]
[1]Student ISE Department, Sri Krishna Institute of Technology, B'lore-560090, India
[2]Student ISE Department, Sri Krishna Institute of Technology, B'lore-560090, India
[3]Faculty ISE Department, Sri Krishna Institute of Technology, B'lore-560090, India

ABSTRACT

India is a global agriculture powerhouse. The average productivity of many crops in India is quite low and the current situation faced by farmers in India leads to increase in suicide rate over years, due to the impact of climate change in country. Crop productivity can be increased using Machine Learning (ML) methods and climate data. For weather forecasting, Machine Learning approaches like Recurrent Neural Network (RNN) are utilised, while Machine Learning categorization techniques like Decision Tree, Random Forest, and Nave Bayes are used to predict appropriate crops. Therefore, its necessary to build a model which takes into consideration of all the parameters for the better selection of crops which increases the crops yield.

KEYWORDS: Agriculture, Recurrent Neural Network (RNN), Decision Tree, Random Forest, Naïve Bayes.

I. INTRODUCTION

When it comes to the economic prosperity of a country like India, agriculture plays a critical role. Weather and soil parameters have the greatest impact on crop yield. It's critical to comprehend the effects of climate change on crop productivity. Weather has an impact on a plant's development and growth at every stage. Crop choosing using real-time contingencies depending on weather & soil parameters. Techniques that produce the highest yield while using the minimum amount of energy in constrained land areas must be employed. Precision Agriculture (PA) [1] is a popular term for this approach. Precision agriculture aims to improve crop yields and economics while reducing the amount of conventional inputs required to cultivate crops (land, water, fertilizer, herbicides, and insecticides). Several Machine Learning (ML) algorithms have been employed in agricultural in recent times for climate and groundwater quality forecasting and strategic thinking. Unsupervised and Supervised Machine Learning models are the two types of machine learning algorithms.

Supervised learning is a sort of artificial intelligence in which computers are trained with well-labeled training examples and then predict the outcome using that information. The labelled data indicates that some input data has already been tagged with the desired output. K-means grouping, KNN (K-nearest neighbours), and other unsupervised machine learning methods are examples. Supervised machine learning is employed throughout the classification methods used, like Decision tree, Random forest, and Nave Bayes [2-3].

II. LITERATURE SURVEY

Agriculture production is increased by using supervised Machine Learning technique. It gives information about how the crop should be selected so as to maximize the profit. It includes six major crops consisting of rice species like aus, aman, boro and other crops like potato, wheat and jute. The dataset is collected from Bangladesh Agricultural Research Council according to area. The algorithms used are Decision Tree Learning, ID3 and K-nearest neighbour algorithm. This method is proposed by Md. TahmidShakoor et.al. It is used to provide solution in selection of crop in more profitable manner [4]. Another method is presented by S. Veenadhri et.al designed machine on climatic parameter used in web application. It provides information about climatic condition which would affect the crop yield. The climatic terms that affect the crop yield in Madhya Pradesh was discovered which usesC4.5 Algorithm. The agro input parameters are not considered in this application as they differ with change in space and measure of individual field [5]. Nishit Jain et.al attempted to present a crop selection method based on various environmental factor using machine learning. This method helps the user to choose the crop which will maximize the crop yield by taking into consideration of all the parameters which affects the growth of crop [6-8]. This method also assists in predicting crop sequence as well as providing monitory benefits to farmers Economic and environmental parameters such as rainfall, temperature and type of soil plays a major role. Based upon all the parameters crop is cultivated based on seasons [9]. Igor Oliveira et al. present an advanced machine learning xsystem for pre-season agricultural yield forecasting in another approach. Agriculture stakeholder department yield prognosis, which can be obtained using machine learning modelling and obtaining additional data from external sources [10-14]. To create a pre-season projection using soil characteristics databases, satellite-derived precipitation, climate patterns forecast data using physical models, and other resources. The yield of soya beans or maize can be predicted without using data from the Normalized Difference Vegetation Index Information (NDVI). It yields useful results by obviating the need for elevated remote control visual representation and allowing farmers to plan for climate changes that affect harvest oscillation. Soya bean and maize

forecasts are based on Brazil and America, which accounted for 44 percent of global food security in 2016 [15-17].

III. METHODOLOGY

Weather patterns estimation is important using a recurrent neural network (RNN). The random forest categorization technique is used to classify appropriate crops [18].

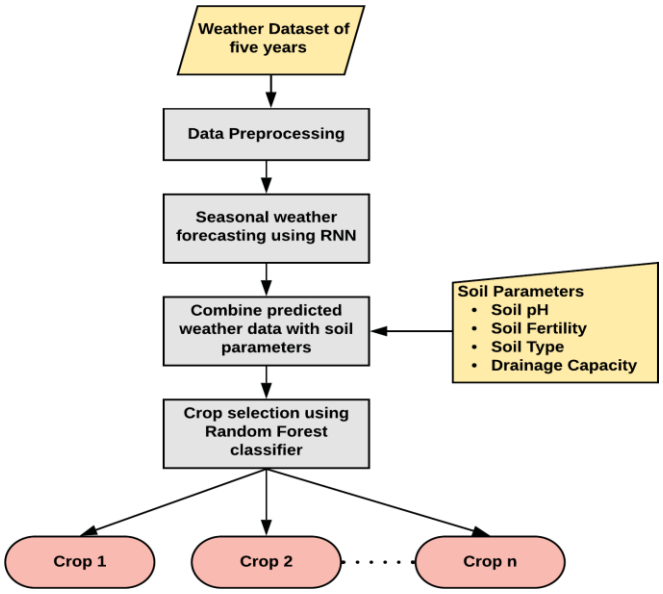


Figure-1: Schematic flow of proposed weather and soil forecasting based crop selection method

a) Weather Datasetoffiveyears: Atthisstage the dataset is collected, The dataset includes minimum and maximum temperature, PH, Humidity, and soil arduino board, and rainfall data is collected from Agri Ministry, IISc, and IIM.

b) Data Preprocessing : During this stage raw data is transformed into useful and efficient format. It involves cleaning of data, identifying the missing data, filling of missing values, ignoring the tuples.

c) Seasonal Weather Forecasting using RNN: At this stage RNN used in variant of Neural network. This Forecast predicts,

Weather at monthly interval upto 7 months and seasonal crop selection method suggests parameters. These data is collected through one or more crop suitable for a season.

d) Combine predicted weather data with soilparameters: Duringthis stagesoil parametersandtheyareasfollows,

Soil pH : It is a measures of acidity and basicity ofa soil. It influences the availability of essential nutrients. pH levels range from 0 to14.

Soil fertility: It increases the fertility soil we can use the method of recycling the organic waste.

Soil type:itdeterminesthetypeofsoil suitablefora particular cropsuch as sandy soil, clay soil etc.

Soildrainage: it determines how the water moves out of the soil as a result of force of gravity [19].

e) Crop selection using Random Forest Classifier : At this stage, it is suitable to predictcrop by considering all the parameters to have agoodyield.

Thealgorithmsusedinthismodelare:

1. DecisionTree
2. NaïveBayes
3. RandomForest

1] DecisionTree

Internal nodes represent properties of the dataset, branches indicate decision rules, and each node represents the result in a decision tree. The process starts with root node in a decision tree for estimating agricultural yield given a data set. The best property in a dataset that can be chosen using attribute selection methods is the root node of a tree[12-13]. The method analyzes the roots attribute to the actual dataset attributes, then continues the branch and goes towards the next node presents a comparison. It's done again and again until the leaf node is found [20-24].

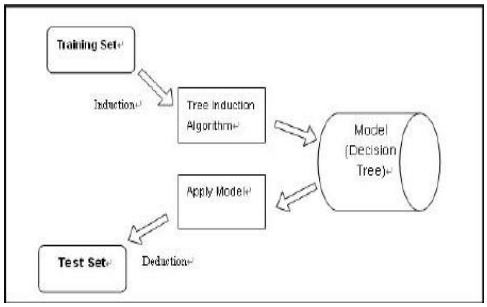


Figure-2: ArchitectureofDecisionTreeAlgorithm

2].NaivesBayes

The Nave Bayes classifier makes predictions based on an object's likelihood. It is based on the Bayes theorem premise. With prior knowledge, the Discrete wavelet transform (dwt is used to compute the likelihood of a hypothesis. The weather dataset is turned into a frequency distribution table utilizing the Nave Bayes classifier, and a probability table is constructed by calculating the probabilities for each parameter utilizing the Bayes theorem [24-27].

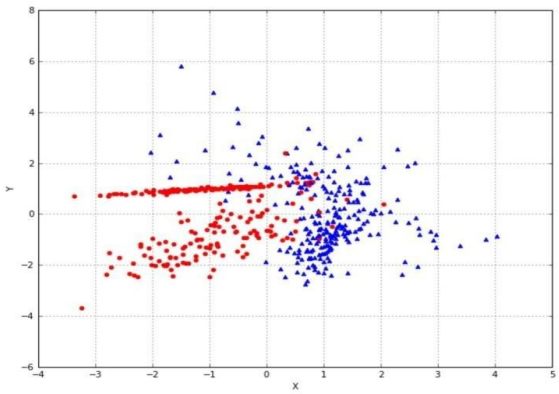
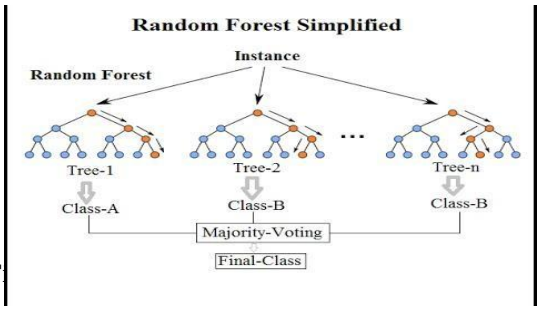


Figure-3:Graph Representation of Naïve Bayes Classification

3]RandomForest

Random forest classification is based on ensemblelearning. It is the process of combining multipleclassifiertoimprovetheperformanceofthemodel [28]. It has a large number of decision trees on different subsets of the supplied weather dataset and averages them to increase the predicted accuracy of the dataset, which is very useful. The higher the number of trees in the forest, the better the accuracy and the less chance of overfitting.



F

IV. MODELLING AND ANALYSIS

The dataset is divided into training and testing data in this study. The training set is a part of the data used to train a model, and the testing set is a subset of the data used to test the trained model [29-31]. The initial stage is to gather data, which includes variables like temperatures, humid, PH, rain, and other soil factors. These are dependent variables where as the predicted crop is an independent variable. Using proven classification methods, the final forecast crop is described as a simpler depiction of the complex link among weather and soil on only one hand, and crop quality (such as growth or yield) on another [32-34].

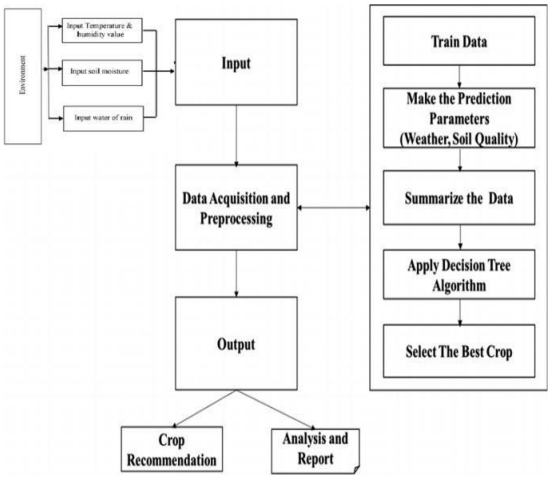


Figure-5:Flowchart of crop yield prediction

V. RESULT AND DISCUSSION

In this section, all experimental details and trail data sets output have been analysed. All experiments are performed using Arduino boards and windows OS installed on intel Core i5 CPU with 4GB RAM. We followed the guidelines which was mentioned in the Sonal Jain *et al* [14] using some of the data from the training set andsome from the testing set and feeding these sets in to the classification algorithms [35-40]. Further as mentioned in the Sonal Jainetal, we have taken the images of both schematic flow and rain fall graph as shown in fig-1 and andfig-7.

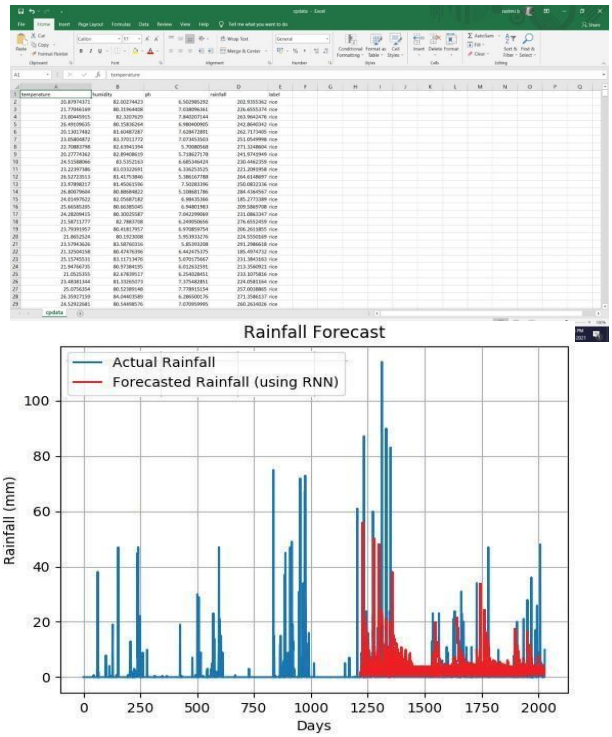


Figure-6: Sample Dataset for training classification model
Figure-7: Rain fall Forecast

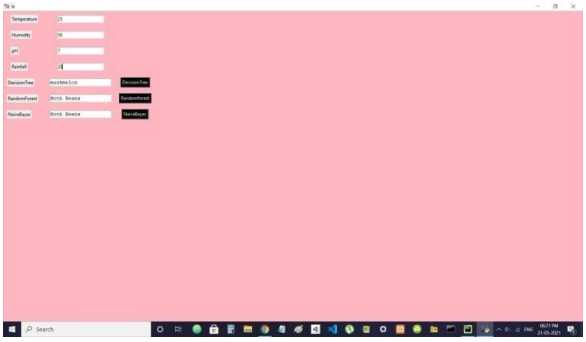


Figure-8:Predicted Output

VI. CONCLUSION

A result is the consequence of actions or events expressed qualitatively or quantitatively. Performance analysis is a set of basic quantitative relationship between the performance quantities. Although there are many techniques which
shows promising results in crop yield detection. As some times there might be sudden change in weather so prediction, so weather plays a prominent role. Therefore, in this paper we prepare a profitable model for crop prediction. Hence, our model is the best situatedfor the modern crop yield detection system when compared to the traditional and unconventional techniques.

REFERENCES

[1] Muhammad, L.J., Algehyne, E.A., Usman, S.S., Ahmad, A., Chakraborty, C. and Mohammed, I.A., 2021. Supervised machine learning models for prediction of COVID-19 infection using epidemiology dataset. SN computer science, 2(1), pp.1-13.

[2] Parameshachari, B.D., Panduranga, H.T. and liberata Ullo, S., 2020, September. Analysis and Computation of Encryption Technique to Enhance Security of Medical Images. In IOP Conference Series: Materials Science and Engineering (Vol. 925, No. 1, p. 012028). IOP Publishing.

[3] Ngo, T.D., Bui, T.T., Pham, T.M., Thai, H.T., Nguyen, G.L. and Nguyen, T.N., 2021. Image deconvolution for optical small satellite with deep learning and real-time GPU acceleration. Journal of Real-Time Image Processing, pp.1-14.

[4] Md. Tahmid Shakoor, Karishma Rahman, Sumaiya Nasrin Rayta, Amitabha Chakrabarty, "Agricultural Production Output Prediction Using Supervised Machine Learning Techniques", IEEE, 2017, pp:182-187.

[5] S. Veenadhari, Dr. Bharat Misra, Dr. CD Singh, "Machine learning approach for forecasting crop yield based on climatic parameters", International Conference on Computer Communication and Informatics (ICCCI - 2014), Jan. 03 -05, 2014, pp: 200-205.

[6] Prabu, S., Balamurugan, V. and Vengatesan, K., 2019. Design of cognitive image filters for suppression of noise level in medical images. Measurement, 141, pp.296-301.

[7] Dwivedi, R., Dey, S., Chakraborty, C. and Tiwari, S., 2021. Grape disease detection network based on multi-task learning and attention features. IEEE Sensors Journal.

[8] Lingappa, H., Suresh, H. and Manvi, S., 2018. Medical image segmentation based on extreme learning machine algorithm in kernel fuzzy c-means using artificial bee colony method. Int. J. Intell. Eng. Syst, 11, pp.128-136.

[9] Igor Oliveira, Renato L. F. Cunha, Bruno Silva, MarcoA. S. Netto, "A Scalable Machine Learning System for Pre-Season Agriculture Yield Forecast", 14th IEEE eScience, pp: 423-430.

[10] Nishit Jain, Amit Kumar, Sahil Garud, Vishal Pradhan, Prajakta Kulkarni, "Crop Selection Method Based on Various Environmental Factors Using Machine Learning", International Research Journal of Engineering and Technology (IRJET), Volume: 04 Issue: 02, February -2017, pp : 551-558.

[11] Babu, N.V., 2014. Design and performance analysis of clustering mechanisms for wireless sensor networks.

[12] Vu, D.L., Nguyen, T.K., Nguyen, T.V., Nguyen, T.N., Massacci, F. and Phung, P.H., 2020. HIT4Mal: Hybrid image transformation for malware classification. Transactions on Emerging Telecommunications Technologies, 31(11), p.e3789.

[13] Babu, R.G., Maheswari, K.U., Zarro, C., Parameshachari, B.D. and Ullo, S.L., 2020. Land-Use and Land-Cover Classification Using a Human Group-Based Particle Swarm Optimization Algorithm with an LSTM Classifier on Hybrid Pre-Processing Remote-Sensing Images. Remote Sensing, 12(24), p.4135.

[14] Sonal Jain, Dharavath Ramesh of "Machine Learning Convergence for Weather based Crop Selection, IEEE, 2020, pp: 1-6.

[15] L. Tan, K. Yu, A. K. Bashir, X. Cheng, F. Ming, L. Zhao, X. Zhou, "Towards Real-time and Efficient Cardiovascular Monitoring for COVID-19 Patients by 5G-Enabled Wearable Medical Devices: A Deep Learning Approach", **Neural Computing and Applications**, 2021, https://doi.org/10.1007/s00521-021-06219-9

[16] L. Tan, K. Yu, F. Ming, X. Cheng, G. Srivastava, "Secure and Resilient Artificial Intelligence of Things: a HoneyNet Approach for Threat Detection and Situational Awareness", **IEEE Consumer Electronics Magazine**, 2021, doi: 10.1109/MCE.2021.3081874.

[17] L. Tan, N. Shi, K. Yu, M. Aloqaily, Y. Jararweh, "A Blockchain-Empowered Access Control Framework for Smart Devices in Green Internet of Things", **ACM Transactions on Internet Technology**, vol. 21, no. 3, pp. 1-20, 2021,https://doi.org/10.1145/3433542.

[18] Hemalatha, K. L., SUNILKUMAR MANVI, and HN SURESH. "ADAPTIVE WEIGHTED-COVARIANCE REGULARIZED KERNEL FUZZY C MEANS ALGORITHM FOR MEDICAL IMAGE SEGMENTATION." Journal of Theoretical & Applied Information Technology 95, no. 14 (2017).

[19] Hemalatha, K. L., S. M. Ashitha, and S. R. Meghana. "Design and implementation of modified FCM in the detection of brain tumor." Int. J. Adv. Sci. Res. Eng 3, no. 4 (2017): 2850-2858.

[20] Subramani, Prabu, K. Srinivas, R. Sujatha, and B. D. Parameshachari. "Prediction of muscular paralysis disease based on hybrid feature extraction with machine learning technique for COVID-19 and post-COVID-19 patients." Personal and Ubiquitous Computing (2021): 1-14.

[21] Pham, Dung V., Giang L. Nguyen, Tu N. Nguyen, Canh V. Pham, and Anh V. Nguyen. "Multi-topic misinformation blocking with budget constraint on online social networks." IEEE Access 8 (2020): 78879-78889.

[22] Rajendran, Ganesh B., Uma M. Kumarasamy, Chiara Zarro, Parameshachari B. Divakarachari, and Silvia L. Ullo. "Land-use and land-cover classification using a human group-based particle swarm optimization algorithm with an LSTM Classifier on hybrid pre-processing remote-sensing images." Remote Sensing 12, no. 24 (2020): 4135.

[23] Do, Dinh-Thuan, Tu Anh Le, Tu N. Nguyen, Xingwang Li, and Khaled M. Rabie. "Joint impacts of imperfect CSI and imperfect SIC in cognitive radio-assisted NOMA-V2X communications." IEEE Access 8 (2020): 128629-128645.

[24] Parameshachari, B. D., H. T. Panduranga, and S. K. Naveenkumar. "Partial encryption of medical images by dual DNA addition using DNA encoding." In 2017 international conference on recent innovations in signal processing and embedded systems (RISE), pp. 310-314. IEEE, 2017.

[25] Nguyen, Tu N., Bing-Hong Liu, Nam P. Nguyen, and Jung-Te Chou. "Cyber security of smart grid: attacks and defenses." In ICC 2020-2020 IEEE International Conference on Communications (ICC), pp. 1-6. IEEE, 2020.

[26] Parameshachari, B. D., Rashmi P. Kiran, P. Rashmi, M. C. Supriya, Rajashekarappa, and H. T. Panduranga. "Controlled partial image encryption based on LSIC and chaotic map." In ICCSP, pp. 60-63. 2019.

[27] Y. Gong, L. Zhang, R. Liu, K. Yu and G. Srivastava, "Nonlinear MIMO for Industrial Internet of Things in Cyber–Physical Systems," **IEEE Transactions on Industrial Informatics**, vol. 17, no. 8, pp. 5533-5541, Aug. 2021, doi: 10.1109/TII.2020.3024631.

[28] C. Feng, **K. Yu**, M. Aloqaily, M. Alazab, Z. Lv and S. Mumtaz, "Attribute-Based Encryption with Parallel Outsourced Decryption for Edge Intelligent IoV," **IEEE Transactions on Vehicular Technology**, vol. 69, no. 11, pp. 13784-13795, Nov. 2020, doi: 10.1109/TVT.2020.3027568.

[29] S. Chen, L. Zhang, Y. Tang, C. Shen, R. Kumar, **K. Yu**, U. Tariq, and A. K. Bashir, “Indoor Temperature Monitoring Using Wireless Sensor Networks: A SMAC Application in Smart Cities”, **Sustainable Cities and Society**, vol. 61, p. 102333, July 2020.

[30] Subramani, Prabu, Fadi Al-Turjman, Rajagopal Kumar, Anusha Kannan, and Anand Loganthan. "Improving medical communication process using recurrent networks and wearable antenna s11 variation with harmonic suppressions." Personal and Ubiquitous Computing (2021): 1-13.

[31] Bhuvaneswary, N., S. Prabu, K. Tamilselvan, and K. G. Parthiban. "Efficient Implementation of Multiply Accumulate Operation Unit Using an Interlaced Partition Multiplier." Journal of Computational and Theoretical Nanoscience 18, no. 4 (2021): 1321-1326.

[32] Nguyen, Ngoc-Tu, Ming C. Leu, and Xiaoqing Frank Liu. "RTEthernet: Real-time communication for manufacturing cyberphysical systems." Transactions on Emerging Telecommunications Technologies 29, no. 7 (2018): e3433.

[33] Rajendrakumar, Shiny, and V. K. Parvati. "Automation of irrigation system through embedded computing technology." In Proceedings of the 3rd International Conference on Cryptography, Security and Privacy, pp. 289-293. 2019.

[34] Nguyen, Tu N., Bing-Hong Liu, and Shih-Yuan Wang. "On new approaches of maximum weighted target coverage and sensor connectivity: Hardness and approximation." IEEE Transactions on Network Science and Engineering 7, no. 3 (2019): 1736-1751.

[35] Manjanaik, N., B. D. Parameshachari, S. N. Hanumanthappa, and Reshma Banu. "Intra Frame Coding In Advanced Video Coding Standard (H. 264) to Obtain Consistent PSNR and Reduce Bit Rate for Diagonal Down Left Mode Using Gaussian Pulse." In IOP Conference Series: Materials Science and Engineering, vol. 225, no. 1, p. 012209. IOP Publishing, 2017.

[36] Le, Ngoc Tuyen, Jing-Wein Wang, Duc Huy Le, Chih-Chiang Wang, and Tu N. Nguyen. "Fingerprint enhancement based on tensor of wavelet subbands for classification." IEEE Access 8 (2020): 6602-6615.

[37] Parameshachari, B. D., H. T. Panduranga, and Silvia liberata Ullo. "Analysis and computation of encryption technique to enhance security of medical images." In IOP Conference Series: Materials Science and Engineering, vol. 925, no. 1, p. 012028. IOP Publishing, 2020.

[38] K. Yu, L. Tan, M. Aloqaily, H. Yang, and Y. Jararweh, "Blockchain-Enhanced Data Sharing with Traceable and Direct Revocation in IIoT", **IEEE Transactions on Industrial Informatics**, doi: 10.1109/TII.2021.3049141.

[39] K. Yu, L. Lin, M. Alazab, L. Tan, B. Gu, "Deep Learning-Based Traffic Safety Solution for a Mixture of Autonomous and Manual Vehicles in a 5G-Enabled Intelligent Transportation System", **IEEE Transactions on Intelligent Transportation Systems**, doi: 10.1109/TITS.2020.3042504.

[40] K. Yu, M. Arifuzzaman, Z. Wen, D. Zhang and T. Sato, "A Key Management Scheme for Secure Communications of Information Centric Advanced Metering Infrastructure in Smart Grid," **IEEE Transactions on Instrumentation and Measurement**, vol. 64, no. 8, pp. 2072-2085, August 2015.
https://ieeexplore.ieee.org/document/7138617