Enhancing Health Data Assessment Performance: A Multi-Sensor Focused Study for Public Health

Palvi Soni[1], Gajendra Tandan[1]
[1]*Assistant Professor, Department of CS & IT, Kalinga University, Raipur, India.*
[2]*Research Scholar, Department of CS & IT, Kalinga University, Raipur, India*

KEYWORDS

Public Health, Wearable Technology, Multisensor, Horse Herd DiscriminantGenerative Adversarial Network (HHF-DGAN).

ABSTRACT

The term health data assessment performance described the process of assessing efficacy of acquiring and reporting health data. Multi sensor approach for predicting diabetes in public health scenarios, multi sensor for public health assimilates assorted data source and addresses community health challenges widely and efficiently. In this study, we proposed a Horse Herd Fused Discriminant Generative Adversarial Network (HHF-DGAN) algorithm for diabetes prediction. Sensor-based diabetes data is gathered from kaggle, and then data preprocessing using z-score normalization is inaccurate records from a dataset. To extract feature from the pre-processed data, Short-Time Fourier Transform (STFT) is used for dimensionality reduction. HHF-DGAN is a public health technique that addresses diabetes prevention and management while also including multi sensor assessment data. The proposed methods are to evaluate F1-score (86%), precision (92%), recall (93%), and accuracy (93%). The effectiveness of the suggested strategy in improving prediction reliability for diabetes detection in public health circumstances has been established by comparing it with existing methods. The study concludes field of healthcare assessments with a robust framework for leveraging multi sensor data and deep learning methods in disease prediction.

1. Introduction

Health data assessment performance by integrating multi-sensor technologies in public health outcomes using more comprehensive. Multi sensor systems in public health have become important for enhancing health data assessment and technological improvement. The combination of various sensors, such as wearable technology, biometric sensors, and environmental monitors, presents unmatched predictions for obtaining modern, high quality health data [1]. Technology assists in an exhaustive understanding of public health dynamics through enabling extensive tracking of contextual factors, behavioural patterns and physiological data [2]. Multi-sensor technologies might expressively increase the effectiveness of health assessments in the public health domain [6]. Timely actions are necessary to facilitate the treatment of health abnormalities [3]. Particularly crucial are the abilities to recognize infectious disease outbreaks, keep an eye out for environmental health risks, and manage chronic illnesses [12]. Multi sensor data combination includes wearable device, environmental sensor, and mobile application [4]. The diverse types of data are ranging from physiological parameters such as heart rate, and body temperature to environmental factors . The objective of this study is to predict diabetes disease in public health, and utilize multi-sensor data integration to improve accuracy and efficacy of health monitoring and interference strategies for improved health outcomes [15]. Remaining section are, portion 2 describes related works, portion 3 describes methodology, portion 4 describes result and portion 5 describes conclusion.

Related work

Aivaliotis et al [5] described an in-depth examination of effective smart health frameworks, which primarily require the employment of numerous energy resource inhibited strategies and sensor for internet of things (IoT) integration in the healthcare sector. The security required essential IoT modules, particularly the wireless connections. Xu et al [13] assimilated convolutional neural networks (CNNs) as rehabilitation assessment techniques that detect movements under various walking modes and provide an inclusive gait examination in the evaluation of movement-related diseases [9]. It enabled clinicians to analyse the effectiveness of rehabilitation efforts. The devices improved higher-limb endurance and body balance through rehabilitation. Hasan et al [7] address heart disease prediction challenges by focusing on inter-dataset performance discrepancies in machine learning (ML) models like random forest (RF) for feature selection. Their strategy improved the model's resilience and accuracy across a variety of datasets [11]. Wearable sensors were described by Ascioglu and Senol [8] as a way to enhance human activity recognition. CNNs and multi-sensor fusion were examples of deep learning (DL) techniques that improve accuracy and resilience under a variety of real-world scenarios. Nurmi and Lohan [14] examined the use of glucose, electrocardiogram (ECG), and accelerometer (ACC) sensors for diabetes prediction. In order to evaluate the predictive capacity of solo or combination sensors for diabetes, high prediction accuracy was achieved using the extreme gradient boosting (XGBoost) algorithm.

2. Methodology

In this section, initially diabetes data is gathered and data preprocessing is employed using z-score normalization, and STFT is used to extract the feature from pre-processed data, the HHF-DGAN is executed and explained in detail. Figure 1 illustrated the workflow of methodology.

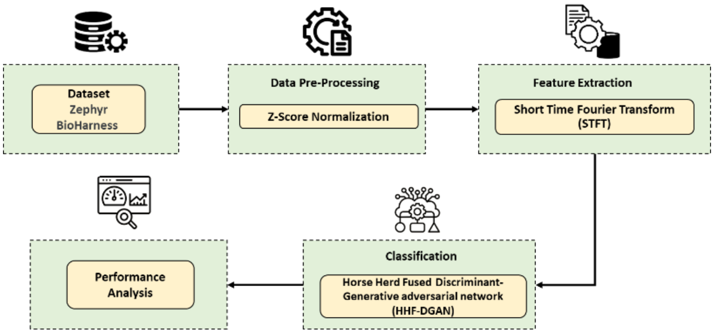

Figure 1: Workflow of methodology

Dataset

Data is gathered from kaggle (https://www.kaggle.com/datasets/sarabhian/d1namo-ecg-glucose-data) contains Zephyr BioHarness 3 wearable sensor [10]. The quality of data obtained from wearable sensors is typically inferior to that of medical-grade devices, which were obtained under highly controlled clinical settings. The dataset elaborated was developed on 20 healthy and 9 patients with diabetes. This is to identify hyperglycemia events by non-invasive ECG pattern analysis of breathing, glucose, and other dataset that can assist in the development of wearable device-based health care systems in non-clinical settings.

Data preprocessing using Z-score normalization

After gathering data, we utilize Z-score normalization, a data preparation technique that divides a feature's value by its standard deviation after subtracting the mean. To ensure that variables are comparable to prevent bias resulting in various scales, this approach is essential for predictive models are used to forecast diabetic disease. Z-score normalization involves varying the data by 0 and 1 standard deviation. This assures that the scales for the various parameters within the public health assessments are consistent. When it comes to performance measurements, user interaction statistics, or any other quantitative data, this is especially useful to enable a fair comparisons and equitable evaluations. The Z-score normalization process improves the dependability and efficiency of data driven insights, v_{th} denotes the value of specific variables in the data, and μ represents blood glucose stage of the feature's average value for each individual in the sample. The value of σ represents the degree of variance or dispersion from the mean in equation (1).

$$(Z - score)_T = \frac{(V_{th} - \mu)}{\sigma} \tag{1}$$

The formula above indicates that the maximum Z-score is to be obtained by dividing the characteristic mean by the standard deviation of the characteristic set. Standardizing variables to a conventional normal distribution is achieved through Z-score normalization, such as blood sugar levels. By providing accurate evaluation of risk factors and enhancing models for disease susceptibility prediction, it enables one to compare and forecast diabetes risk across various datasets.

Feature extraction STFT

The process of using STFT for diabetes prediction feature extraction involves breaking down time-domain data into minute windows, transforming each window using the Fourier transform, and then extracting frequency-domain features, such as spectral power or frequency bands, that assist in classification to be significant for the examination of repetitive behaviors that have a big impact on the prognosis of diabetes disease. However, STFT could be used to convert the initial data into frequency-domain characteristics. Where $STFT\{j(s)\}$ represents time frequency analysis, $W(m,\omega)$ denotes time shifted window function and $W(m,l)$ denotes the STFT can be characterized as follows in the discontinuous domains in equations (2) to (4) denote variables.

$$STFT\{j(s)\} = W(\tau,\omega) = \int_{-\infty}^{+\infty} j(s)g(s-\tau)f^{-i\omega s}cs \quad (2)$$

$$W(m,\omega) = \sum_{n=-\infty}^{+\infty} j(n)g(n-m)f^{-i\omega n} \quad (3)$$

$$W(m,l) = W(m,\omega)\big|_{\omega=\frac{2\pi}{M}l} = \sum_{n=-\infty}^{+\infty} j(n)g(n-m)-i^{\frac{2\pi}{M}ln} \quad (4)$$

Transforming time-domain signals into frequency-domain descriptions is the process of using STFT for predicting diabetes. To record spectrum modifications associated with diabetes, possibly improving accuracy in classification by disclosing frequencies patterns relevant to the start and progression of the disease.

Diabetics health data assessment framework

The proposed HHF-DGAN method is to predict diabetes using heterogeneous healthcare multi-sensor data for robust prediction. This approach enhances accuracy by including ECG. Glucose data refers to the development of wearable device-based health care systems in non-clinical conditions and personalized treatment of diabetes.

Discriminant Generative adversarial networks (DGAN)

To forecast diabetes using data from multiple sensors, feature extracted in the steps are given to Deep learning (DL) algorithms. DGAN has an influential class of artificial intelligence (AI) models designed to generate the real data circulations. Sensor configurations were assessed in multi-sensor prediction. The predicted rate using the most effective algorithms in those combinations, if the characteristics of two sensors glucose and ECG are combined, the combined data are modelled using the methods that improved the particular DGAN structures. DGAN for predicting diabetes using multi-sensor data is worked on by two networks that is generator and discriminator, the generator is combining accurate diabetes related multisensor data, while the discriminator calculates authenticity. By improving the generator's output through this adversarial process, sensor inputs can be used to anticipate outcomes more accurately. $\min_{H}\max_{C} U(C,H)$ Indicate the probability of certain diagnosis in health factor, $X(o_q, o_h)$ denotes diagnostic thresholds. Data includes determining C values potentially modifying H for the purpose to optimize the expected utility obtained from the diabetes detection diagnostic decisions in equations (5) to (6).

$$\min_{H}\max_{C} U(C,H) = \mathbb{E}_{w\sim o_{data}[\log C(w)]+\mathbb{E}_{y\sim o_{y(w)}}\left[\log\left(1-C(H(y))\right)\right]} \quad (5)$$

$$X(o_q, o_h) = \frac{1}{L}\sup_{\|e\|_K \le L} \mathbb{E}_{w\sim o_q[e(w)]} - \mathbb{E}_{w\sim o_h[e(w)]} \quad (6)$$

$$|e(w_1) - e(x_2)| \le L|w_1 - w_2| \quad (7)$$

Where equation (7) is $|e(w_1) - e(x_2)|$probability distribution, generator provides set of samples data.

DGAN method is predicting diabetes from multisensor data, by enhancing feature extraction and discrimination competences, DGANs optimize classification accuracy and DL techniques for managing diabetes with customized therapy.

Horse Herd Fused (HHO)

To classify the DGAN result to optimize HHO, we enhance predictive accuracy for diabetes by generating health assessment data and improving model training. HHO predictive capabilities could potentially detect diabetes more effectively. HHO is stimulated by the behaviour of horses in nature, with the leaders guiding herd towards optimal paths. To predict diabetes, HHO mimics mathematical models that mimic the herd's movements towards finding the best solution. It predicts diabetes risks by optimizing factors such as insulin sensitivity, blood glucose levels, and other health markers. Initialize a population of potential sensor configures based on the principles of horse herd behavior. Optimization is iteratively enhancing the solutions using the three main behaviours of HHO: exploration, exploitation, and abandonment. The best solution is found during the optimization process. Evaluation of each sensor configuration using HHO function to determine fitness and how well it predicts diabetes and equation (10) denotes the position of the horse$o_n^{iter,age}$, every horse's range is displayed based on age.

$$o_n^{iter,age} = Vel_n^{iter,age} + o_n^{(iter-1),age} \tag{9}$$

In the equation (10) denotes $vel_n^{iter,\alpha}$ is velocity of particular horse inmeals, stealing, and tolerance.

$$vel_n^{iter,\alpha} = Gra_n^{iter,\beta} + G_n^{iter,\beta} + Soc_n^{iter,\beta} + DefenseMec_n^{iter,\beta} \tag{10}$$

The motion vectors is computed using algorithms that effectively mimic the six motions of different horse groups. Based on the preceding behavioral patterns, Equations (11) and (12) could be regarded as the motion vectors of different aged horses during the entire cycle.

$$vel_n^{iter,\alpha} = Gra_n^{iter,\beta} + G_n^{iter,\beta} + Soc_n^{iter,\beta} + Imi_n^{\cdot iter,\gamma} + Roam_n^{iter,\gamma} + DefenseMec_n^{iter,\beta} \tag{11}$$

$$vel_n^{iter,\alpha} = Gra_n^{iter,\beta} + Imi_n^{\cdot iter,\gamma} + Roam_n^{iter,\gamma} \tag{12}$$

HHF-DGAN method is used to identify patterns symptomatic of diabetes inception and personalized procedures for prevention using multisensor. This method improves healthcare outcomes related to diabetes treatment by utilizing algorithms inspired by nature to increase predictive accuracy.

3. Results and discussion

This part describes the efficiency of proposed method HHF-DGAN, along with how we evaluated it using Python software. Parameters including f1-score, precision, accuracy, recall and existing methods are Naïve Bayes (NB) [11], OneR [11], and Sequential Minimal Optimization (SMO) [11]. Table 1 shows the numerical outcomes.

Table 1: Numerical outcomes of f1 score, recall, precision, accuracy

Methods	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)
NB [11]	85	91	85	83
OneR [11]	70	78	58	91
SMO [11]	89	89	92	77
HHF-DGAN [Proposed]	**93**	**92**	**93**	**86**

Accuracy: The accuracy of a measurement is the deviation from the true or acceptable value in public health evaluations. The comparison of accuracy among the other existing methods and proposed methods are shown in Figure 2. When compared to other existing methods, the proposed HHF-DGAN achieved 93%. The existing methods are NB, OneR, and SMO, which achieved values of (85%), (70%), and (89%). Finally, it shows the superior efficiency in predicting diabetes using HHF-DGAN.

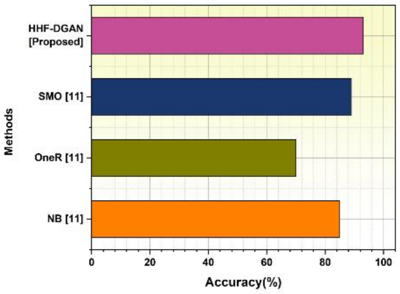

Figure 2: Result of accuracy

Precision: The percentage of true positive predictions made by those who are accurately classified as having diabetes among all positive predictions produced using multi-sensor data in diabetes prediction. Comparative analysis of precision graph shown in Figure 3, HHF-DGAN attained 92% and existing methods NB attained 91%, OneR attained 78%, and SMO attained 89%. It demonstrates superior efficiency in predicting diabetes using HHF-DGAN.

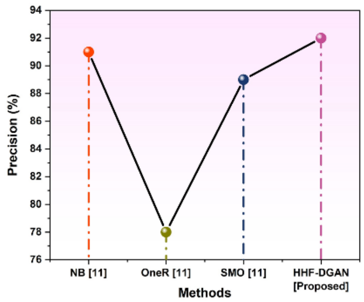

Figure 3: Comparison of precision

Recall: This metric calculates the ratio of true positive prediction to all positive cases. To predict diabetes using multisensor data, integrate inputs from multi sensor to enhance accuracy. It leverages assorted data source for robust prediction. The comparison evaluation of recall proposed and existing methods are illustrated in Figure 4. The proposed method has obtained 93% and existing methods such as NB obtains 85 %, OneR obtains 58%, and SMO obtains 92%. The result shows higher effectiveness in predicting diabetes using HHF-DGAN.

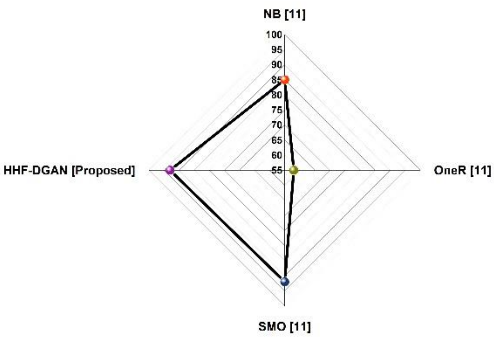

Figure 4: Result of Recall

F1 score: When predicting diabetes, a model's accuracy is determined using F1 score, which achieves the balance between recall that records all positive cases and precision accurately identifying positive cases. It gives a single number that, taking into account both false positives and false negatives, represents the model's performance in detecting diabetic cases. Figure 5 illustrates the f1 score outcomes. HHF-DGAN achieved 86% and the other existing models like NB achieved 83%, OneR achieved 91%, and SMO achieved 77%, respectively.

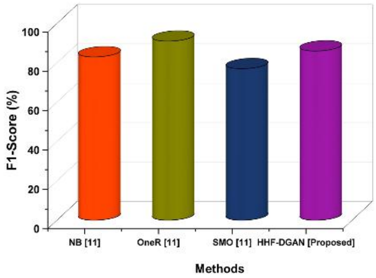

Figure 5: Comparison outcomes of F1 score

4. Conclusion and future scope

The study described a multi sensor approach for predicting diabetes in public health scenarios, and multi sensor data assimilation for addressing community health challenges widely and efficiently. HHF-DGAN diabetes prediction. To assess a suggested classification algorithm to predict diabetes using multisensor data was gathered from kaggle, and then effectively pre-processed by z-score normalization in accurate records from a dataset, features were extracted by STFT, which consistently extracted unwanted data. HHF-DGAN applied a method of public health that addressed diabetes prevention and management while also incorporating multi sensor assessment data. The proposed methods were to evaluate F1-score at 86%, precision at 92%, recall at 93%, and accuracy at 93%. The effectiveness of the suggested strategy in improving prediction reliability for diabetes detection in public health circumstances has been established by comparison with existing methods. The study concluded that the field of healthcare assessments was enhanced by a robust framework for leveraging multi sensor data and deep learning methods in disease prediction. While glucose serves as a major energy source for cells, the cellular energy currency, adenosine triphosphate, is not produced in sufficient amounts when glucose levels are low. The future research will predict diabetes using AI technology, various sensors, real time risk assessments the ability of wearables with integration for improving early detection and lifestyle management.

Reference

[1] J.H. Li, P.W. Yu, H. C. Wang, C.Y. Lin, Y. C. Lin, C.P. Liu, C.Y. Hsieh, and C.T. Chan, "Multi-Sensor Fusion Approach to Drinking Activity Identification for Improving Fluid Intake Monitoring". Applied Sciences, 14(11), p.4480, 2024. DOI: https://doi.org/10.3390/app14114480

[2] A. Ettefagh, and A. Roshan Fekr, "Enhancing automated lower limb rehabilitation exercise task recognition through multi-sensor data fusion in tele-rehabilitation". Bio Medical Engineering OnLine, 23(1), p.35, 2024. DOI: https://doi.org/10.1186/s12938-024-01228-w

[3] S. Neelima, Manoj Govindaraj, Dr.K. Subramani, Ahmed ALkhayyat, & Dr. Chippy Mohan. (2024). Factors Influencing Data Utilization and Performance of Health Management Information Systems: A Case Study. *Indian Journal of Information Sources and Services, 14*(2), 146–152. https://doi.org/10.51983/ijiss-2024.14.2.21

[4] L.Zhang, J.Yu, Z.Gao, and Q. Ni, "A multi-channel hybrid deep learning framework for multi-sensor fusion enabled human activity recognition". Alexandria Engineering Journal, 91, pp.472-485, 2024.DOI: https://doi.org/10.1016/j.aej.2024.01.030

[5] V.Aivaliotis, K. Tsantikidou, and N.Sklavos, "IoT-based multi-sensor healthcare architectures and a lightweight-based privacy scheme". Sensors, 22(11), p.4269, 2022. DOI: https://doi.org/10.3390/s22114269

[6] Kodric, Z., Vrhovec, S., & Jelovcan, L. (2021). Securing edge-enabled smart healthcare systems with blockchain: A systematic literature review. *Journal of Internet Services and Information Security*, *11*(4), 19-32.

[7] M.Hasan, M.A.Sahid, M.P, Uddin, M.A. Marjan, S.Kadry, and ,J.Kim, "Performance discrepancy mitigation in heart

disease prediction for multisensory inter-datasets". PeerJ Computer Science, 10, p.e1917,2024. DOI: https://doi.org/10.7717/peerj-cs.1917/fig-17

[8] G. Ascioglu, and Y. Senol, "Design of a wearable wireless multi-sensor monitoring system and application for activity recognition using deep learning". IEEE Access, 8, pp.169183-169195, 2020. DOI: https://doi.org/10.1109/ACCESS.2020.3024003

[9] Malathi, K., Shruthi, S.N., Madhumitha, N., Sreelakshmi, S., Sathya, U., & Sangeetha, P.M. (2024). Medical Data Integration and Interoperability through Remote Monitoring of Healthcare Devices. *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications (JoWUA), 15*(2), 60-72. https://doi.org/10.58346/JOWUA.2024.I2.005

[10] DINAMO ECG Glucose data is gathered from kaggle website (https://www.kaggle.com/datasets/sarabhian/d1namo-ecg-glucose-data)

[11] A.Rghioui, A. Naja, J.L.Mauri. andA.Oumnad, An IoT based diabetic patient monitoring system using machine learning and node MCU. In Journal of Physics: Conference Series (Vol. 1743, No. 1, p. 012035). IOP Publishing, 2021. DOI: 10.1088/1742-6596/1743/1/012035.

[12] B. S. Ronald, and R. Raman, "Multi-Sensor Fusion in Livestock with IoT for Biometric Sensing and MQTT for Enhanced Health Insights". In 2024 2nd International Conference on Disruptive Technologies (ICDT) pp. 935-940,2024.DOI: https://doi.org/10.1109/ICDT61202.2024.10489625

[13] Z.Xu,Z.Wu, L.Wang, Z.Ma, J.Deng, H.Sha, and H.Wang, "Research on Monitoring Assistive Devices for Rehabilitation of Movement Disorders through Multi-Sensor Analysis Combined with Deep Learning". Sensors, 24(13), p.4273, 2024.DOI: https://doi.org/10.3390/s24134273

[14] J. Nurmi, and E.S. Lohan, Machine-learning-based diabetes prediction using multi-sensor data. IEEE Sensors Journal 2023. https://doi.org/10.1109/JSEN.2023.3319360

[15] Bobir, A.O., Askariy, M., Otabek, Y.Y., Nodir, R.K., Rakhima, A., Zukhra, Z.Y., Sherzod, A.A. (2024). Utilizing Deep Learning and the Internet of Things to Monitor the Health of Aquatic Ecosystems to Conserve Biodiversity. *Natural and Engineering Sciences*, *9*(1), 72-83.