International Journal of
Science and Engineering Investigations
vol. 11, issue 125, June 2022
Received on June 8, 2022
ISSN: 2251-8843

Diagnosing Diabetes Using Wavelet-Based Machine Learning and Heart Rate Variability

Adeethyia Shankar
Brookfield High School
(shankarad@student.brookfieldps.org)

***Abstract*-** More than 400 million people around the world have diabetes mellitus so there is a need for a quick and accurate diagnosis of diabetes. Current diagnostic tests are lengthy and invasive, so in this research, we classify heart rate variability data from people who are healthy and people who have type 1 diabetes using a discrete wavelet transform (DWT)-based machine learning model that combines Random Forest and Support Vector Machine. Before inputting the data into the model, we transform the data using the DWT, utilizing components in both the wavelet domain and the original domain. Our combination model achieved a high accuracy of 93.64%, which was higher than the accuracy obtained by either Random Forest or Support Vector Machine alone. Also, inputting the components in the wavelet domain resulted in higher accuracy than inputting components in the original domain. In addition to high accuracy, our DWT-based Random Forest-Support Vector Machine model can be trained and produce a diagnosis in just minutes, promising a quick and accurate diagnosis of diabetes.

***Keywords*-** *Diabetes Diagnosis, Discrete Wavelet Transform, Machine Learning*

I. INTRODUCTION

More than 400 million people worldwide have diabetes mellitus (World Health Organization, 2021), a condition characterized by high blood sugar (What Is Diabetes?, 2020). Within this condition, two major types are type 1 diabetes, where the body does not produce insulin, and type 2 diabetes, where the body does not respond to insulin. Common complications of diabetes include hypoglycemia and diabetic ketoacidosis, which can be life-threatening (What Is Diabetes?, 2020). Moreover, Jinli Liu, after analyzing the incidence of diabetes from 1990 to 2017, estimated that almost 23 million people were newly diagnosed with diabetes in 2017, more than double the approximately 11 million people who were diagnosed in 1990 (Liu et al., 2020). With more and more people needing to be diagnosed with diabetes each year, the need for an accurate and efficient diabetes diagnostic test continues to grow. Currently, diabetes diagnostic tests include the fasting blood sugar test, random blood sugar test, and glycated hemoglobin, often called A1C, test. Table 1 describes these tests and their drawbacks.

TABLE I. CURRENT DIABETES DIAGNOSTIC TESTS

Test	Description	Drawbacks
Fasting blood sugar	Measures blood sugar level after at least 8 hours of fasting	Requires fasting before the test
Random blood sugar	Measures blood sugar level at an arbitrary time; convenient	Often must be repeated as blood sugar varies throughout the day
A1C	Measures the percentage of hemoglobin with sugar attached to it	Affected by hemoglobin variant, anemia, pregnancy, smoking

These tests are cumbersome. Specifically, the fasting blood sugar test requires fasting before the test, the random blood sugar test often must be repeated, and the A1C test is affected by numerous factors.

Nevertheless, diabetes diagnosis using heart rate variability (HRV) could resolve these issues. HRV measures the variation in RR intervals, where an RR interval is simply the time between consecutive heartbeats. An RR interval can be visualized as the time between two peaks on an electrocardiogram (ECG), such as in Fig. 1. For example, at a heart rate of 60 beats per minute, one RR interval is approximately one second; however, one RR interval may last 1.1 seconds while the next could be 0.9 seconds.

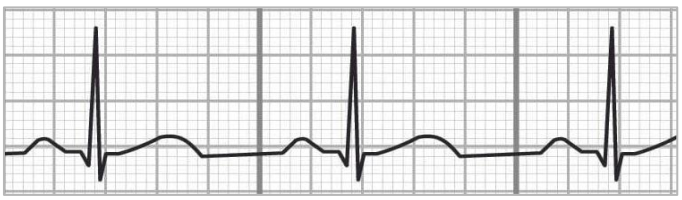

Figure 1. ECG graph: one RR interval is the time between two peaks (Sinus Rhythms1 Normal)

HRV is a useful measure for diabetes diagnosis because, over time, high blood sugar causes damage to the heart, blood vessels, and nerves (World Health Organization, 2021). In fact, Schroeder supports an association between HRV and diabetics, and that the heart becomes more damaged over time in diabetics (Schroeder et al., 2005). Although HRV can be

affected by heart conditions unrelated to diabetes, HRV is still a good measure in the absence of such conditions.

To facilitate the diagnosis process, in this research we used a DWT-based machine learning algorithm to classify HRV samples as either healthy or diabetic. Specifically, machine learning models are trained to classify a given set of data, and using their newly acquired knowledge, these models are then tested on a different set of data. Hence, the machine learning method is suitable for classifying HRV and performing a diagnosis. Furthermore, we used the DWT to remove outliers and white noise from HRV. We will explain DWT in more detail in the methodology section. Thus, our research question is, "How might a quick and accurate diagnosis of diabetes based on HRV be developed using the discrete wavelet transform and machine learning?" Our hypothesis is that inputting the wavelet domain components of the HRV data into a model that combines two machine learning models will yield a quick and accurate diabetes diagnosis.

II. Literature Review

The literature has generally agreed that there is an association between HRV and diabetes for at least 10 years; this association is necessary to establish that HRV is viable to use for diagnosing diabetes. In his 2002 study, Massimo Chessa investigated the association between HRV and diabetic autonomic neuropathy, which means diabetic nerve damage, by analyzing HRV from 50 children with type 1 diabetes and 30 healthy children. Specifically, Chessa extracted various statistical measurements such as pNN50, the proportion of consecutive RR intervals that differ by more than 50 milliseconds. The pNN50 parameter was significantly less in diabetic HRV, supporting an association between pNN50 and type 1 diabetes (Chessa et al., 2002). Since pNN50 was less in diabetic HRV, it must have significantly less variability than healthy HRV from one heartbeat to the next. Similarly, in 2006, Kudat investigated the relation between HRV and whether a person has either type 1 or type 2 diabetes by using pNN50. Of 30 healthy individuals and 31 diabetics, the diabetics had lower pNN50 and lower RR intervals than the healthy individuals, and among the diabetics, those with heart complications had lower pNN50 and lower HRV than those without heart complications (Kudat et al., 2006). In addition to Chessa's findings, Kudat supports that heart complications are associated with less HRV variability and faster RR intervals, posing heart complications as a confounding factor for HRV. Likewise, Schroeder analyzed longitudinal data from 6245 diabetics over nine years, as well as cross-sectional HRV data from 9940 prediabetics. In addition to an association between HRV and diabetes, Schroeder found that cardiac autonomic impairment, the inability of the involuntary nervous system to control heart rate, worsened over time in diabetics (Schroeder et al., 2005). Just as Kudat supports that having a heart complication confounds HRV, Schroeder supports that HRV worsens after years of diabetes, which may cause significant differences in HRV even among diabetic patients. Therefore, the literature has well established that there is an association between HRV and diabetes.

In addition, HRV is already used for the diagnosis of other illnesses. For example, it was used for the detection of cardiac autonomic neuropathy, a disease that damages the nerves controlling the heart (Bissinger, 2017). Likewise, Trivedi, through a review of many studies, found that HRV can be used to detect cardiac autonomic neuropathy even before the appearance of clinical symptoms. (Trivedi et al., 2019). Hence, due to diabetes results in cardiac autonomic impairment, HRV is a viable, as well as valid, measure to use to diagnose diabetes.

Previously, researchers have used statistical measures to differentiate between healthy and diabetic HRV, including pNN50, RMSSD (root mean square of the difference between consecutive RR intervals), SDNN (standard deviation of RR intervals), and NN50 (number of consecutive RR intervals that differ by more than 50 milliseconds). Chessa extracted pNN50, Kristiansen used pNN50 and rMSSD, Schroeder utilized rMSSD and SDNN, and Trivedi used SDNN, rMSSD, and NN50 (Chessa et al., 2002; Kristiansen et al., 2020; Schroeder et al., 2005; Trivedi et al., 2019). These statistical measures provide insight specifically into the variability of HRV.

On the other hand, myriad machine learning models have been proposed to use HRV to diagnose diabetes. Yogender Aggarwal tested an Artificial Neural Network as well as Support Vector Machine (SVM) (Aggarwal et al., 2020). Moreover, Rajendra Acharya utilized the DWT to break down the HRV into multiple parts. Acharya then inputted these parts into a variety of machine learning models, such as Decision Tree, which has a flowchart-like structure that decides how to classify an input, and SVM, which constructs a boundary that separates data points by classification (Acharya et al., 2015). In addition, G. Swapna proposed a model that combines a convolutional neural network, long short-term memory, and SVM (Swapna et al., 2018). Furthermore, in our previous research, we tested Random Forest, which is a model that predicts the majority prediction of many Decision Trees (Shankar et al., 2022). A variety of statistical measures and machine learning models have been employed to distinguish between healthy and diabetic HRV, though the statistical measures did not specifically focus on diagnosis, and both the Random Forest and SVM models have yielded high diagnostic accuracy. While our previous work already used Random Forest, we combined Random Forest with SVM in this research to add the predictive power of both models as one, which should yield higher accuracy. Also, we utilize the DWT rather than statistical measures. However, while Acharya already used DWT to decompose the HRV into approximations and detail components, we went an extra step by denoising the HRV data, and we decomposed the HRV in novel ways.

III. Methodology

The purpose of this research was to create a quick and accurate diabetes diagnosis using HRV. We obtained the HRV data from the D1NAMO dataset. The D1NAMO dataset is a publicly available and anonymized dataset collected by Swiss researchers. This dataset contains data on ECG, HRV, breathing, walking, glucose measurements, and food eaten

from 20 healthy participants and 9 participants who had been diagnosed with type 1 diabetes (Dubosson et al., 2018). However, DICARDIA, a database that contains ECG data collected by researchers from Venezuela, also offered HRV data. Nevertheless, we did not use that data because many of those patients had cardiac autonomic neuropathy (Ledezma et al., 2014), which would confound a diabetes diagnosis. However, Dubosson acknowledges that some of the data is inaccurate and should be filtered. Inaccuracies in the HRV data were caused by occasions when poor conductivity between the device and the skin prevented the sensor from accurately recording a person's heartbeat (Dubosson et al., 2018). These issues caused some RR intervals, which are the times between consecutive heartbeats, to seem too long or too short due to failing to detect a heartbeat or falsely detecting a heartbeat, respectively. In fact, Fig. 2 illustrates an outlier RR interval that was recorded as more than eight seconds. Thus, we noted which of the RR intervals were either below 500 milliseconds or above 2000 milliseconds to identify outliers that are inaccurate.

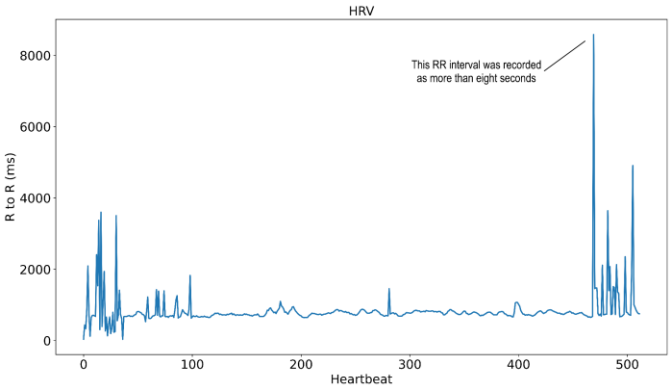

Figure 2. HRV with outliers

After identifying the outliers in the HRV data, we saved the data in between the outliers as samples of 512 consecutive RR intervals. we chose 512 data points to enable the usage of the DWT, which requires a data set with a length that is a multiple of four. Also, each group includes enough data to make a reasonably accurate diagnosis. At the same time, 512 RR intervals would only take about 10 minutes to record, so the diagnosis process would be quick. Thus, despite only having HRV data from 19 healthy and 9 diabetic people, we obtained many samples from the same person. Overall, we obtained 3003 healthy samples and 769 diabetic samples.

Nevertheless, the HRV data still had large peaks that were not filtered out by the initial outlier removal. We used DWT to remove lingering outliers, making the HRV data more accurate. In addition, we used DWT to break down the HRV into multiple parts, allowing us to determine which parts, when inputted into machine learning models, resulted in the most accurate diabetes diagnosis.

DWT is a mathematical data transformation technique that has the ability to both denoise and decompose data into different frequency components. First, our 4-Band DWT requires a 512 by 512 matrix that we constructed using filters in the 4-Band DWT filter bank. We denote this matrix by T. After constructing T, we applied 4-Band DWT to HRV sample S, a column vector of length 512, by multiplying it by T on its left as TS—still a 512-length column vector, the wavelet coordinates of S in the wavelet domain, and stored TS. Like S, TS has 512 entries. Notably, TS can be thought of as four groups of 128 entries each that are stacked on top of each other: the first 128 entries are a_1, the next 128 entries are d_1, the following 128 entries are d_2, and the final 128 entries are d_3. Equation (1) illustrates this stacking.

$$TS = \begin{bmatrix} a_1 \\ d_1 \\ d_2 \\ d_3 \end{bmatrix} \tag{1}$$

The result of this step of DWT is that a_1 can be used to approximate S with outliers removed, while d_1, d_2, and d_3 contain the removed outliers. However, d_1, d_2, and d_3 also contain many smaller noisy values that are not the removed outliers. These smaller values can be removed by setting all values with a magnitude less than a certain threshold value equal to zero. We used three separate threshold values for each of d_1, d_2, and d_3. We set each threshold to $\sigma_i\sqrt{2\,log(n\,/\,4)}$, where σ_i is the standard deviation of d_1 for $i = 1, 2, 3$. The subsequent step of DWT is to use the resulting threshold value to set the values that have a magnitude smaller than the threshold equal to zero. Setting these small values equal to zero results in a different column of values, which we refer to as (TS)*. At this point, we had only completed part of DWT and obtained the components in the wavelet domain: TS, a_1, d_1, d_2, d_3, and (TS)*. To complete the transformation, we multiply the transpose of T by (TS)* to obtain S*, the denoised version of the original vector S.

$$T^t(TS)^* = S^* \tag{2}$$

We can also multiply the transpose of T by TS to obtain S, which can be expressed as the sum of four orthogonal vectors, A_1, D_1, D_2, and D_3. These four vectors correspond to a_1, d_1, d_2, and d_3 transformed back out from the wavelet domain.

$$T^tTS = S = A_1 + D_1 + D_2 + D_3 \tag{3}$$

Our approach for using DWT is unique because we input DWT of S, its components in both Wavelet domain and original domain, denoised S and its DWT: TS, (TS)*, a_1, d_1, d_2, and d_3 in addition to S* and A_1 into machine learning models. In contrast, Acharya only input components in the original domain, namely A_1, D_1, D_2, and D_3. By inputting components in the wavelet domain, we assessed whether the denoising and decomposing ability of the first few steps of DWT are superior to applying the entire process by comparing diabetes diagnostic accuracies.

After applying DWT, the data imbalance needed to be resolved before inputting the data into a machine learning model. There were vastly greater numbers of healthy HRV samples at this point compared to diabetic samples, so even a model that always predicts a sample to be healthy would obtain high accuracy. Thus, we used SMOTE Tomek resampling, a

technique that can use existing samples to create more samples, to generate more diabetic samples based on the already existing diabetic samples. We avoided removing healthy samples as doing so would remove valuable information from the dataset. After correcting the data imbalance, we obtained about 3000 healthy samples and 3000 diabetic samples.

After producing approximately equal numbers of healthy and diabetic samples, the HRV data was ready to input into a machine learning model. This research implemented a novel machine learning model for diagnosing diabetes with HRV. Our model combined Random Forest and Support Vector Machine. Random Forest constructs many flowchart-like structures that each vote on how to classify the input. Random Forest then predicts the classification with the majority vote. On the other hand, SVM creates a boundary that separates data points by their classification. For example, in Fig. 3, the curve separates the blue points from the red points.

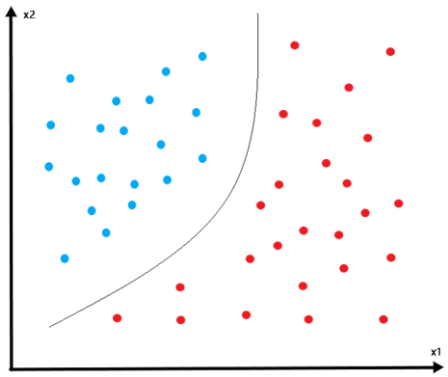

Figure 3. Visual representation of SVM

Random Forest obtained high accuracies in our previous research (Shankar et al., 2022), and SVM has also performed very well previously (Swapna et al., 2018; Aggarwal et al., 2020). Hence, combining the two models should yield even higher accuracy due to combining the predictive power of both models.

To combine Random Forest and SVM, we scrutinized the performances of Random Forest alone and SVM alone. We found that when SVM was tested to diagnose diabetes using d_3 HRV data, every sample that was predicted to have diabetes truly had diabetes. However, the model also predicted many healthy samples to be diabetic, resulting in poor performance overall. On the other hand, Random Forest, when tested to diagnose diabetes using (TS)* HRV data, performed well overall (Shankar et al., 2022).

Thus, we decided to combine Random Forest and SVM by generating a prediction for the HRV sample with both models, and then applying the below logic.

If: SVM predicts the patient has type 1 diabetes, predict type 1 diabetes.

Else: Predict Random Forest Prediction.

As our proposed model depends largely on the accuracy of the Random Forest model, we tested inputting the three DWT components among TS, (TS)*, a_1, d_1, d_2, and d_3 that, when inputted into Random Forest, obtained the highest mean accuracies. These three data types were (TS)*, TS, and a_1 (Shankar et al., 2022). In addition, we inputted No DWT, S*, and A_1 data as controls to assess the effect of DWT on the diagnosis accuracy of our model. We refer to the data inputted into Random Forest as Random Forest <DWT component>. We always inputted d_3 data into SVM because every sample the model predicted to have type 1 diabetes truly had type 1 diabetes; thus, we refer to SVM as SVM d_3. By relegating the diagnostic possibility of predicting type 1 diabetes for a type 1 diabetes sample to SVM, the model is more accurate than Random Forest alone.

Finally, we tested the model using 10-fold cross-validation, allowing us to obtain ten different results of how our model performed to avoid overfitting. For each individual result, we recorded the proportion of test samples that were true positives, true negatives, false positives, and false negatives. Using the mean of these proportions, we calculated accuracy, precision, recall, and F1. These metrics provide detailed insight into the performance of the model. Also, we compared these results with the results from Random Forest alone, SVM alone, and other diabetes diagnostic tests to determine if our proposed model is relatively accurate.

IV. Results

Below in Table 2 are the results we obtained after testing the Random Forest SVM combination model. We tested inputting no DWT, a_1, TS, (TS)*, S*, and A_1 data into the Random Forest; we inputted d_3 data into the SVM.

TABLE II. Diabetes mean diagnostic results for different inputs to Random Forest and D_3 data into SVM

Random Forest Input	Metrics			
	Accuracy (%)	*Precision (%)*	*Recall (%)*	*F1*
TS	93.64[a]	96.64	90.92	0.9370[a]
A_1	92.98	97.13[a]	88.85	0.9280
a_1	92.58	95.77	89.38	0.9247
No DWT	92.33	91.76	94.18[a]	0.9296
(TS)*	92.00	95.24	89.31	0.9218
S*	90.46	96.64	89.31	0.9234

a. Best obtained metrics

We display the results in the form of confusion matrices in Fig. 4. A confusion matrix is a matrix that records how a machine learning model classified different classes of samples. In each of our confusion matrices, there are four entries corresponding to the proportion of samples that were either healthy or diabetic and predicted to be either healthy or diabetic. Within each entry, there are two numbers: the mean proportion across the ten results as a percentage followed by the standard deviation of the ten proportions from each result.

To obtain the mean accuracy from each confusion matrix, add the mean proportions from the top-left entry and bottom-right entry, which represent the samples correctly predicted as healthy and correctly predicted as diabetic, respectively. To the right of each confusion matrix is a gradient of green representing the percentage of the total number of HRV samples. The color of each box corresponds to the mean proportion of the total number of HRV samples that were either healthy or diabetic and predicted to be either healthy or diabetic. The gradient ranges from 0% to 50% as our use of SMOTE Tomek resampling is expected to yield around equal numbers of healthy HRV samples and diabetic HRV samples.

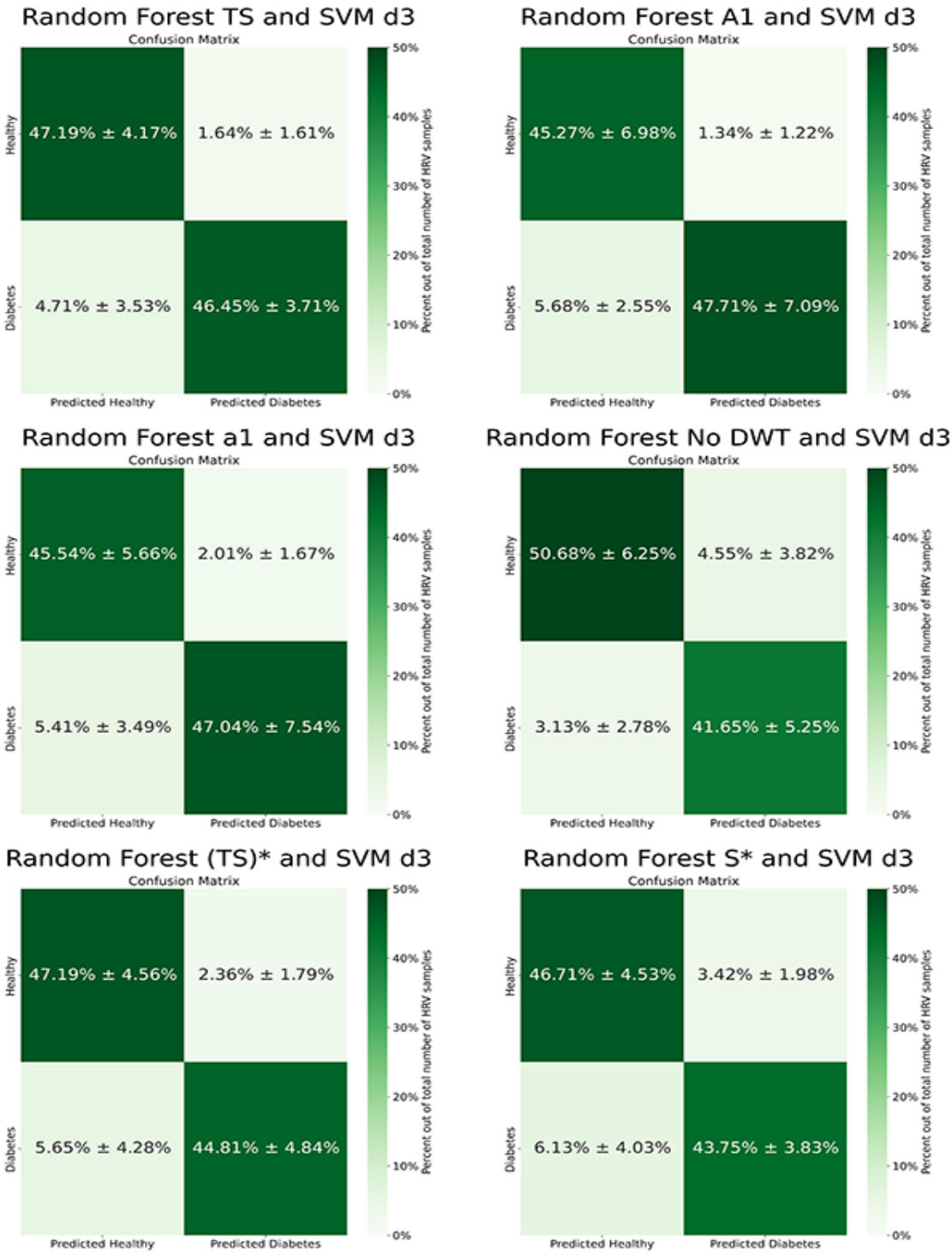

Figure 4. All six confusion matrices of results of Random Forest - SVM combination model

The highest mean accuracy of 93.64% was achieved with Random Forest TS and SVM d_3, which also achieved the highest F1 of 0.9370. With this model, only 1.64% of HRV samples were falsely diagnosed as diabetic, though 4.71% of samples were diabetic but undiagnosed with diabetes. Also, of the six tested models, results had the least variability with Random Forest TS and SVM d_3. But, performance was still substantially more variable than in our previous work, where standard deviations were less than 1% compared to greater than 3% in the current results.

Although not the model with the greatest mean accuracy, Random Forest No DWT and SVM d_3 achieved the highest recall of 94.18% and had only 3.13% of samples be false negatives, more than one percentage point less than with Random Forest TS. However, precision was only 91.76%, the lowest out of all the models, and 4.55% of samples were falsely diagnosed with diabetes, suggesting that this model minimizes false negatives by simply predicting diabetes more often.

On the other hand, regarding the Random Forest a_1 and SVM d_3, the Random Forest (TS)* and SVM d_3, and the Random Forest S* and SVM d_3 models, their mean accuracies were more than one percentage point less than the highest mean accuracy of 93.64% and the models had higher rates of false positives and false negatives.

It should be noted that the Random Forest A_1 and SVM d_3 model obtained the highest precision of 97.13%. However, the mean accuracy of this model was more than half a percentage point lower than the highest mean accuracy of 93.64%, and the recall and F1 metrics for this model were lower than those metrics for the Random Forest TS and SVM d_3 model.

V. Discussion

Overall, inputting TS data into Random Forest and d_3 data into SVM yielded the highest mean accuracy of 93.64%. Although the model's performance had a relatively high variation compared to our previous work, the model surpassed the highest mean accuracy of Random Forest alone, which was only 91.9%, and the highest mean accuracy of SVM alone, which was only 77.5% (Shankar et al., 2022). Combining Random Forest and SVM is more effective at diagnosing diabetes than just Random Forest. Also, the use of the TS component of DWT is more effective for diagnosing diabetes than the use of components in the original domain. The combination model performed with higher accuracy than the A1C test and fasting blood sugar test did in an Iranian population, with 90% and 80% accuracy, respectively (Ghazanfari et al., 2010). Our proposed model is also more accurate than the random blood sugar test, which obtained 92.2% accuracy in early pregnancy (Adefisan et al., 2020). Although the accuracy of the proposed model is less than the 95.7% accuracy obtained by Swapna with their combination model (Swapna et al., 2018), Random Forest and SVM only require minutes to be trained, whereas the neural networks used by Swapna can take upwards of hours for training. In the future, the Random Forest - SVM model can be tested for diagnosing HRV samples from different demographics, which would be facilitated by quick training times. As age and sex are important determinants of HRV (Tegegne et al., 2018), such an investigation could yield a diabetes diagnosis that is more relevant to a specific age range and sex.

However, a significant obstacle was the small sample size of only 29 patients in the D1NAMO dataset, even despite extracting multiple HRV samples from each patient. Hence, before clinical use, the model should be trained with HRV from a large number of people. In addition, to make the model's diagnosis more relevant to people who have not yet been diagnosed with diabetes, the model should specifically be trained with HRV from just-diagnosed diabetics. Moreover, we used SMOTE Tomek to generate more than 2000 extra diabetic HRV samples from just 769 samples. But, these extra diabetic HRV samples may not actually be plausible diabetic HRV samples. Nevertheless, collecting HRV from equal and large numbers of healthy and diabetic people can resolve this issue. Overall, the Random Forest - SVM combination model proposed in this research is preferable to other popular diagnostic tests because the combination model obtained a higher accuracy and performed quicker than the other diagnostic tests.

Just as this research combined the Random Forest and SVM models, future work could combine the tested Random Forest - SVM models to minimize false positives and false negatives. Moreover, HRV from many different people is essential in the future for results that are representative of large populations. With more diverse and abundant HRV data, the Random Forest SVM model can be used in clinics to diagnose diabetes accurately and quickly.

Acknowledgment

Adeethyia Shankar would like to thank Dr. Xiaodi Wang for advising throughout this research, as well as providing valuable revision.

References

[1] Acharya, U. R., Vidya, K. S., Ghista, D. N., Lim, W. J. E., Molinari, F., & Sankaranarayanan, M. (2015). Computer-aided diagnosis of diabetic subjects by heart rate variability signals using discrete wavelet transform method. Knowledge-Based Systems, 81(June), 56-64. https://doi.org/10.1016/j.knosys.2015.02.005

[2] Adefisan, A., Olagbuji, B., Adeniyi, A., Ade-ojo, I., Ghazalli, S., & Olofinbiyi, B. (2020). Diagnostic accuracy of random plasma glucose and random blood capillary glucose in detecting international association of diabetes and pregnancy study groups- defined hyperglycemia in early pregnancy. Nigerian Journal of Clinical Practice, 23(8), 1086. https://doi.org/10.4103/njcp.njcp_221_19

[3] Aggarwal, Y., Das, J., Mazumder, P. M., Kumar, R., & Sinha, R. K. (2020). Heart rate variability features from nonlinear cardiac dynamics in identification of diabetes using artificial neural network and support vector machine. Biocybernetics and Biomedical Engineering, 40(3), 1002-1009. https://doi.org/10.1016/j.bbe.2020.05.001

[4] Bissinger, A. (2017). Cardiac autonomic neuropathy: Why should cardiologists care about that? Journal of Diabetes Research, 2017. https://doi.org/10.1155/2017/5374176

[5] Bonora, E., & Tuomilehto, J. (2011). The pros and cons of diagnosing diabetes with a1c. Diabetes Care, 34(Supplement_2), S184-S190. https://doi.org/10.2337/dc11-s216

[6] Chessa, M., Butera, G., Lanza, G. A., Bossone, E., Delogu, A., De Rosa, G., Marietti, G., Rosti, L., & Carminati, M. (2002). Role of heart rate variability in the early diagnosis of diabetic autonomic neuropathy in children. Herz, 27(8), 785-790. https://doi.org/10.1007/s00059-002-2340-4

[7] Dubosson, F., Ranvier, J.-E., Bromuri, S., Calbimonte, J.-P., Ruiz, J., & Schumacher, M. (2018). The open d1namo dataset: A multi-modal dataset for research on non-invasive type 1 diabetes management. Informatics in Medicine Unlocked, 13, 92-100. https://doi.org/10.1016/j.imu.2018.09.003

[8] Ghazanfari, Z., Haghdoost, A. A., Alizadeh, S. M., Atapour, J., & Zolala, F. (2010). A comparison of HbA1c and fasting blood sugar tests in general population. International Journal of Preventive Medicine, 1(3), 187-194. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3075530/

[9] Kristiansen, E., Wanby, P., Åkesson, K., Blomstrand, P., Brudin, L., & Thegerström, J. (2020). Assessing heart rate variability in type 1 diabetes mellitus—Psychosocial stress a possible confounder. Annals of Noninvasive Electrocardiology, 25(5). https://doi.org/10.1111/anec.12760

[10] Kudat, H., Akkaya, V., Sozen, A., Salman, S., Demirel, S., Ozcan, M., Atilgan, D., Yilmaz, M., & Guven, O. (2006). Heart rate variability in diabetes patients. Journal of International Medical Research, 34(3), 291-296. https://doi.org/10.1177/147323000603400308

[11] Ledezma, C. A., Severeyn, E., Perpinan, G., Altuve, M., & Wong, S. (2014, August 26-30). A new on-line electrocardiographic records database and computer routines for data analysis [Paper presentation]. 2014 36th Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Chicago, Illinois, The United States of America. https://ieeexplore.ieee.org/abstract/document/6944189

[12] Liu, J., Ren, Z.-H., Qiang, H., Wu, J., Shen, M., Zhang, L., & Lyu, J. (2020). Trends in the incidence of diabetes mellitus: Results from the global burden of disease study 2017 and implications for diabetes mellitus prevention. BMC Public Health, 20(1). https://doi.org/10.1186/s12889-020-09502-x

[13] Mayo Clinic Staff. (2021, March 27). Type 1 diabetes [Fact sheet]. Mayo Clinic. Retrieved September 25, 2021, from https://www.mayoclinic.org/diseases-conditions/type-1-diabetes/diagnosis-treatment/drc-20353017

[14] Schroeder, E. B., Chambless, L. E., Liao, D., Prineas, R. J., Evans, G. W., Rosamond, W. D., & Heiss, G. (2005). Diabetes, glucose, insulin, and heart rate variability: The atherosclerosis risk in communities (ARIC) study. Diabetes Care, 28(3), 668-674. https://doi.org/10.2337/diacare.28.3.668

[15] Shankar, A., Chang, S., Wang, X., & Zhao, Y. (2022). Wavelet based machine learning approaches towards precision medicine in diabetes mellitus. Proceedings of the 15th International Joint Conference on Biomedical Engineering Systems and Technologies, 4, 290-297. https://doi.org/10.5220/0010993100003123

[16] Sinus rhythms1 normal [Photograph]. (n.d.). https://cdn.shopify.com/s/files/1/2630/9452/files/sinus-rhythms1-normal.jpg?v=1616525205

[17] Swapna, G., Vinayakumar, R., & Soman, K.p. (2018). Diabetes detection using deep learning algorithms. ICT Express, 4(4), 243-246. https://doi.org/10.1016/j.icte.2018.10.005

[18] Tegegne, B. S., Man, T., Van Roon, A. M., Riese, H., & Snieder, H. (2018). Determinants of heart rate variability in the general population: The lifelines cohort study. HeartRhythm, 15(10), 1552-1558. https://doi.org/10.1016/j.hrthm.2018.05.006

[19] Trivedi, G. Y., Saboo, B., Singh, R. B., Maheshwari, A., Sharma, K., & Verma, N. (2019). Can decreased heart rate variability be a marker of autonomic dysfunction, metabolic syndrome and diabetes? Journal of Diabetology, 10(2), 48. https://doi.org/10.4103/jod.jod_17_18

[20] What is diabetes? (2020, June 11). CDC. Retrieved August 28, 2021, from https://www.cdc.gov/diabetes/basics/diabetes.html

[21] World Health Organization. (2021, April 13). Diabetes [Fact sheet]. World Health Organization. Retrieved October 12, 2021, from https://www.who.int/news-room/fact-sheets/detail/diabetes

Adeethyia Shankar is a 12th grade student at Brookfield High school in Brookfield, Connecticut, United States of America. He has published "Predicting the Utilization of Mental Health Treatment with Various Machine Learning Algorithms" and "Wavelet based Machine Learning Approaches towards Precision Medicine in Diabetes Mellitus". His research interests include machine learning and biomedicine.

Mr. Shankar has presented a paper at the international Experimental Biology convention and was nominated for the Presidential scholar award at the Junior Science and Humanities Symposium.

How to Cite this Article:

Shankar, A. (2022). Diagnosing Diabetes Using Wavelet-Based Machine Learning and Heart Rate Variability. International Journal of Science and Engineering Investigations (IJSEI), 11(125), 1-7. http://www.ijsei.com/papers/ijsei-1112522-01.pdf