Electronic Stethoscope with Heart Sound Classifier Using Convolutional Neural Network Algorithm and Jetson Nano Deep Learning Accelerator

David Isaac Agsaway, Renz Gregor M. Bartolome, Chademiere Jake C. Castro,
Nathaniel C. Rilloraza, Dominic James S. Intacto
Electronics Engineering Department, College of Engineering Technological University of the Philippines - Manila
Ayala Boulevard, Ermita, Manila, Philippines

Email: agapkiosk.healthcare@gmail.com

***Abstract*— This research addresses the escalating challenge of healthcare service distribution in the face of global population growth, particularly in underserved rural communities. To tackle this issue, the study introduces Electronic Stethoscope with Heart Sound Classifier Using Convolutional Neural Network Algorithm and Jetson Nano Deep Learning Accelerator. This Electronic Stethoscope features a wired interface connecting Jetson Nano and modified Stethoscope. The device is capable of automatically collecting crucial health data, specifically, heart sound. The integration with a mobile application on an Android device enhances user interaction, while a machine learning algorithm on NVIDIA Jetson Nano accurately processes heart sound identification. The study evaluates the Electronic Stethoscope's efficacy, reliability, speed in classifying heart sound, and the user accessibility through a structured User Acceptance Testing framework, aiming to provide a pioneering solution that combines technological advancements and healthcare services for efficient health screening in rural areas.**

***Keywords*— (Electronic Stethoscope, Health assessments, Health Screening Healthcare services, Machine learning algorithm, NVIDIA Jetson Nano, Underserved rural communities, User Acceptance Testing (UAT))**

I. INTRODUCTION

IN In the realm of public health, health workers (HWs) serve as the frontline warriors, bridging the gap between communities and formal healthcare systems. Their dedication and commitment to ensuring the well-being of individuals within their respective populations are invaluable. However, amidst their noble efforts, challenges persist, particularly in the scope of basic health check-ups, specifically in the identification and assessment of heart conditions using stethoscopes.

The identification and assessment of heart conditions using stethoscopes are essential components of basic health check-ups in rural communities. However, the effectiveness of such diagnostic procedures is often hindered by limited access to specialized training and resources among healthcare providers in rural areas.

These given circumstances drive the researchers to conduct a study that aims to investigate innovative solutions that address the importance of resolving the challenges. It seeks to enhance the capability of healthcare providers in rural areas, specifically in the conduct of basic health check-ups. The focus lies on assessing heart sounds using a fabricated electronic stethoscope.

Significant to this exploration is the integration of technology, specifically the utilization of advanced tools such as Jetson Nano as a deep learning accelerator and Convolutional Neural Networks (CNNs) as algorithms. The use of Convolutional Neural Network (CNN) serves as appropriate deep learning to classify heart sound. This deep learning algorithm model uses input image that is suitable in examining visual data and engineered itself to filter the correct optimization.

Meanwhile, NVDIA Jetson Nano is an effective and well-built small computer that allows multiple neural networks that easily works in image classification. This device paired with the Convolutional Neural Network (CNN) will have the finest heart sound classifier and emphasizes the commitment to develop an efficient and newer version of stethoscope.

By leveraging these technological advancements, the study seeks to empower healthcare providers in rural communities with enhanced diagnostic capabilities, enabling more accurate and efficient identification of heart-related conditions during basic health check-ups.

II. BACKGROUND OF THE PROBLEM

Rural communities across the globe face significant healthcare disparities, often characterized by limited access to medical resources and infrastructure. According to recent statistics from the World Health Organization (WHO), rural populations are more likely to experience higher rates of cardiovascular diseases compared to urban counterparts (WHO, 2020). Additionally, a study conducted by the National Rural Health Association revealed that rural residents are 23% more likely to die from heart-related conditions than those living in urban areas (NRHA, 2019). These statistics underscore the critical need for improved healthcare services in rural settings, including enhanced diagnostic capabilities for common health issues such as cardiovascular diseases.

In Philippine status, cardiovascular diseases pose a significant health problem, contributing to a notable portion of the overall disease burden. According to recent statistics from the Philippine Statistics Authority (PSA), Ischaemic heart diseases rank among the leading causes of death in Philippines, with a reported prevalence rate of approximately 18.4% (PSA, 2022). These statistics underscore the urgent need for targeted interventions and improved healthcare infrastructure to address the growing burden of cardiovascular diseases in rural communities throughout the Philippines.

Recognizing the significance of addressing these challenges, this study endeavors to explore innovative solutions aimed at enhancing the capacity of healthcare providers in rural communities, particularly in conducting basic health check-ups with a focus on heart sound assessment using stethoscopes.

Electronic Stethoscope with Heart sound Classifier Using Convolutional Neural Network algorithm and Jetson Nano Deep Learning Accelerator perform duties that contribute to the comfortable collection and categorization of heart sound. The Electronic Stethoscope developed in a different method that differs from the classical acquisition of heart sound.

III. OBJECTIVES

A. Main Objective

The core purpose of the study is to integrate Jetson Nano as a deep learning accelerator and Convolutional Neural Network as an algorithm for classifying heart sounds using electronic stethoscope. This objective will play a pivotal role in assisting health workers in conducting basic health check-ups.

B. Specific Objectives

1) To modify a Stethoscope that automatically recognizes and classifies heart sound using Jetson Nano as deep learning accelerator and convolutional neural network as algorithm.

2) To develop a graphical user interface for the heart sound classification of the electronic stethoscope.

IV. RELATED STUDIES

In the study conducted by Nazeer et al., Jetson Nano is used as a computing edge platform that utilizes machine learning and deep learning application to create a real-time object detection and recognition. The device deploys a mechanism that is capable of training models using Jetson Nano developer kit with its deep learning models and libraries. Afterwards, it involves in identifying a given object and determine the appropriate category using machine learning, statistics, et al. [1]

Yamamoto et al. developed automated heart sound classification through the advance signal processing and deep learning technologies. The Convolutional Neural Network (CNN) was applied to the heart sound classification which learn the representation at the time-frequency level that improves the stability of CNN and accelerate the convergence speed. [2]

In a similar study of Wu et al., an electronic stethoscope with a traditional stethoscope with microphone attached to collect cardiopulmonary sounds. The device evaluates the cardiopulmonary sounds as healthy and abnormal using an effective classification algorithm within the electronic stethoscope. [3]

Mandal et al. developed an electronic stethoscope with software development using a CNN-based deep learning model. This device analyzed the heart sound with real-time process. It is also trained to predict normal sounds and abnormality with high accuracy. [4]

V. METHODOLOGY

In this section, the research methods are outlined to provide a proper understanding of the processes behind the system. The overall framework of the system unveils the integration of Jetson Nano as part of machine learning processes, the development of the fabricated electronic stethoscope, and software development encompassing the application for the display output. The aforementioned components collectively facilitated the classification of the patient's heart sounds.

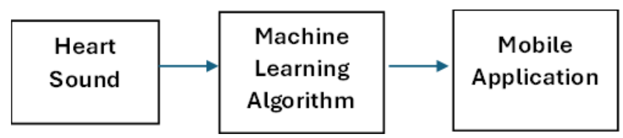

Figure 1. General Block Diagram of Electronic Stethoscope

The general block diagram, depicted in ***Figure 1***, encompasses three (3) key components involved in the heart sound classification process. The process begins with the acquisition of a five (5) second recording of the heart sound. A Machine Learning Algorithm handles the preprocessing and classification of the heart sound. Upon completion of the process, the output label is displayed using a mobile application.

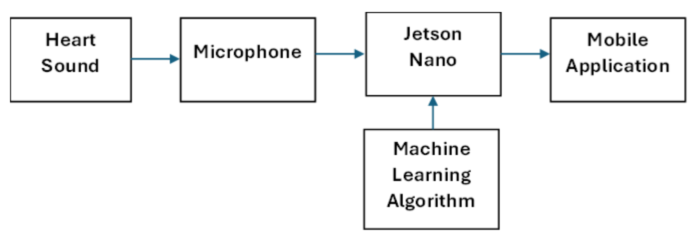

Figure 2. Specific Block Diagram Approach of Electronic Stethoscope

Figure 2 illustrates the procedure for acquiring heart sounds from the patient and categorizing them into three distinct types: normal, murmur, and extrasystole. The initial step outlined in the block diagram involves capturing the heart sounds using a stethoscope. Subsequently, the sound is recorded through an electret microphone, rendering the data suitable for preprocessing and classification procedures. Acting as the accelerator and processor, the Jetson Nano assesses the recorded sound based on the classification model and the integrated algorithm. The resulting classification output of heart sound is then transmitted through a mobile application and presented via a software application or graphical user interface.

A. *System Architecture*

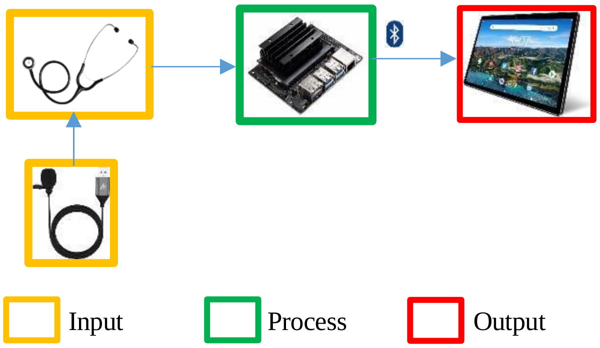

Figure 3. System Architecture of Electronic Stethoscope

As depicted in ***Figure 3***, the input components employ the stethoscope and electret microphone as primary instruments for recording heart sound data. The sound waves are transmitted to the Jetson Nano via a USB-A wired connection. Preprocessing techniques are applied to the data to diminish noise and generate a spectrogram image. This preprocessed spectrogram image undergoes classification using a 2D-CNN model to ultimately categorize the heart sound. Subsequently, the classified output text is transmitted to the mobile tablet via Bluetooth connection. The graphical user interface output on the mobile tablet displays the classification of heart sounds.

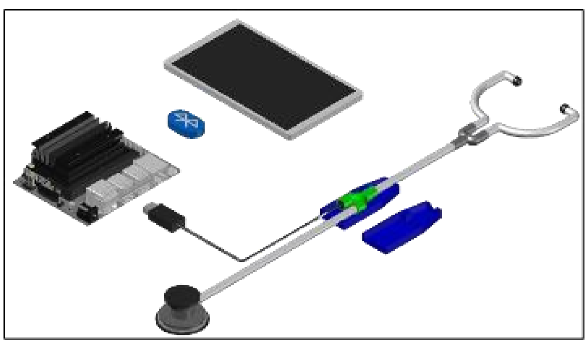

Figure 4. 3D Design of Electronic Stethoscope

Figure 4 illustrates the current configuration of the device, showcasing the connection between the fabricated electronic stethoscope and the Jetson Nano. Precise measurements and a custom-designed 3D-printed casing for the microphone were implemented to enhance both the aesthetic appeal and the organization of the wiring. The device, specifically designed for heart sound auscultation, features an acoustic stethoscope and an electret microphone connected to a 3.5mm audio jack for recording heart sounds directly to a mobile device. Additionally, the device runs an algorithm to determine the status of the patient's heart sound.

VI. RESULTS AND DISCUSSION

Electronic Stethoscope is a heart screening system that uses a machine learning algorithm to automate the classification of heart sounds. The system can distinguish between normal (1) and abnormal (2) heart sounds. Moreover, the system is expected to meet the standards of health associations, hospitals, and clinic tests with its convenience and efficiency.

In addition, MIT App Inventor was used in designing the application and for connecting the Jetson Nano to the android application. Meanwhile, Jetson Nano utilized an open-source operating system, specifically Linux-Ubuntu. This allowed the integration of machine learning to the edge computing device.

The data is initially stored in a temporary format using TinyDB. As the process proceeds, the data is transferred to the main SQLite database for long-term storage and accessibility. The system can export the data as a comma-separated values (CSV) file for further analysis.

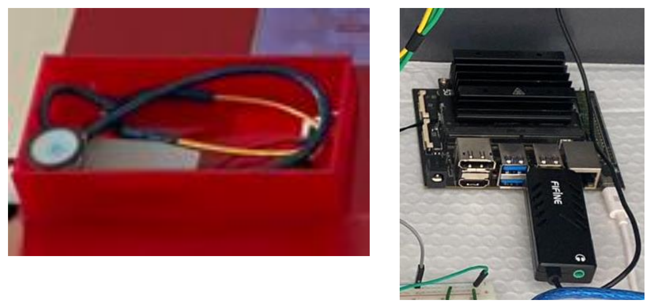

Figure 5. Modified Electronic Stethoscope with Jetson Nano

The development of the Electronic Stethoscope as seen in ***Figure 5,*** prioritized optimal functionality and the overall shape of the traditional stethoscope. To achieve direct implementation, the researchers divided the stethoscope into two sections. The first section, containing the chest piece, was connected to a 3D-printed Y-shaped tubing. One branch of the Y-tubing connected to a lapel microphone, enabling the device to record incoming audio signals from the stethoscope. The other branch remained open, allowing for manual auscultation. All components of the Y-tubing were then housed within a larger 3D-printed enclosure for improved protection and isolation of heart sounds. As seen in ***Figure 5***, sound waves produced by vibration in stethoscope were subjected to NVIDIA Jetson Nano, specifically, Okdo Nano C100 version. The Jetson Nano, as an edge computing device, allowed the incoming wav form to be processed in such a way that the heart sound is classified according to its spectrogram characteristics.

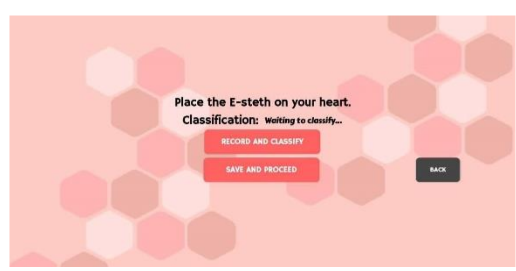

Figure 6. Mobile Application for Electronic Stethoscope

Figure 6 is the actual interface for classification of heart sound using electronic Stethoscope and Jetson Nano. The mobile application was created with the use of MIT App inventor, which enabled the system to create a two-way communication from Jetson Nano and graphical user interface (GUI).

TABLE 1. Summary of Parametric Values for Tested Algorithms

Machine Learning Models	Parameters			
	ACCURACY	PRECISION	RECALL	F - SCORE
KNN with 96 means	0.86	0.76	0.42	0.54
KNN with 40 means	0.86	0.76	0.48	0.59
KNN with spectral amplitude	0.8	0.52	0.39	0.45
KNN with wavelet features	0.71	0.16	0.11	0.13
KNN with both 40 mean and wavelet	0.85	0.68	0.46	0.55
Long Short-Term Memory	0.83	0.74	0.58	0.59
1D-Convolutional Neural Network	0.81	0.71	0.51	0.48
2D-Convolutional Neural Network	**0.90**	**0.81**	**0.93**	**0.87**

As seen in **Table 1**, the least appropriate model to use for classifying heart sound is the KNN utilizing wavelet features. Values resulted from testing the model were classified as the least in all parameters among all the tested models. The resulted value for using wavelet features include 0.71 accuracy value, 0.16 precision value, 0.11 recall value, and 0.13 F-Score value. Low scores in parameters might be a cause of wavelet being not able to effectively extract the features since heart sounds contain complex patterns. More than that, information might be lost during the feature extraction using wavelet leading to misrepresentation of heart sound signal and poor performance in classifying.

Among all the models trained, 2D-Convolutional Neural Network garnered an outstanding numerical value; accuracy valuated at 0.89, precision with 0.81, recall value of 0.93, and lastly, f-score of 0.87. In connection, 2D CNN gathered highest value due to the architecture of CNNs involves building a hierarchy of features. This process starts with identifying low-level features such as edges and textures in the data. Subsequently, the network progresses to capturing higher-level features, such as specific patterns or anomalies within heart sounds. This hierarchical approach empowers 2D CNNs to effectively distinguish between various heart sound categories.

Moreover, two-dimensional Convolutional Neural Networks (2D CNNs) excel in automatically extracting relevant features directly from raw input data. This capability allows them to identify intricate patterns within heart sound spectrograms or other two-dimensional representations. By recognizing these patterns, 2D CNNs achieve superior performance in classifying different heart sound types.

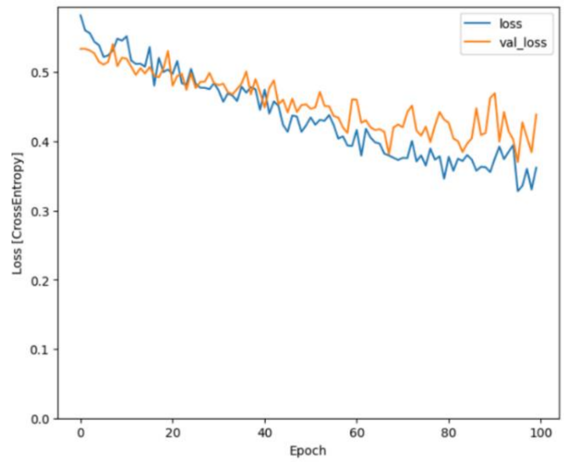

Figure 7. Line Graph for Loss in Train and Validation Sets Using 2D – Convolutional Neural Network at 100 Epochs

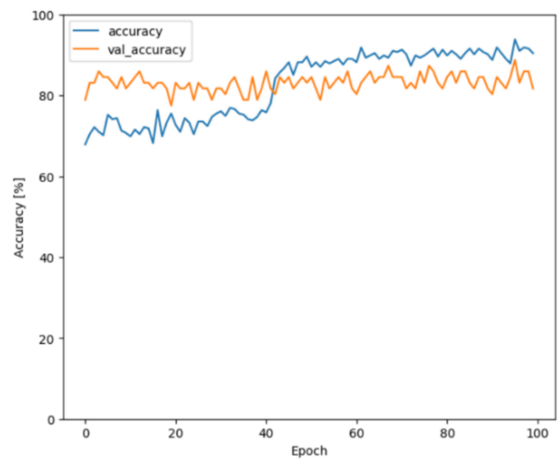

Figure 8. Line Graph for Accuracy in Train and Validation Sets Using 2D – Convolutional Neural Network at 100 Epochs

The line graph served as a visual representation of the model's learning process throughout a series of training epochs. The model underwent training for 100 epochs, utilizing a batch size of 16 data points per training iteration. The Adam optimizer was employed to guide the model's parameter updates during training, aiming to minimize the loss function and improve performance. Additionally, the model architecture incorporated four (4) convolutional layers, responsible for extracting spatial features from the input data. Evaluation of the training process revealed a promising outcome. The model achieved an average loss of 0.329, indicating a successful reduction in the discrepancy between the model's predictions and the ground truth labels. Furthermore, the model attained a high accuracy of 0.899, signifying its ability to correctly classify a significant portion of the testing data.

Analysis of the line graph revealed a crucial aspect of the training process. Notably, the testing and validation sets exhibited minimal discrepancies in values throughout the 100 epochs. This observation suggests that the model had not yet encountered gradient explosion, a phenomenon that can hinder training progress. The absence of gradient explosion implies that the model's learning rate remained appropriate for the current stage of training, allowing the model to effectively optimize its parameters.

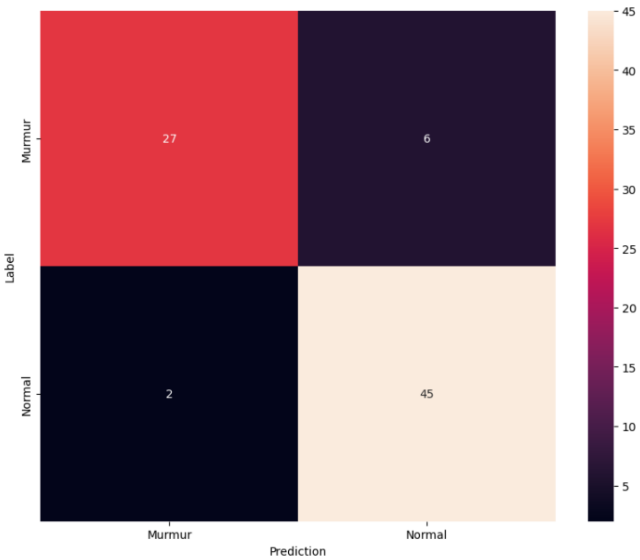

Figure 9. Confusion Matrix for Most Accurate Algorithm (2D – CNN)

Figure 9 defined the characteristic and actual application of the device. A total of eighty (80) samples were tested as part of evaluating the device. As seen in confusion matrix, a total of 27 True Negative (TN) predictions were produced leading to the idea of successfully predicting 27 abnormal heart sounds. On the other hand, 45 of given data were classified as True Positive (TP), having a successful 45 predictions for normal classification of heart sound. Type I error (False Positive) resulted in 6 failed classifications of normal heart sound. Meanwhile, Type II error (False Negative) was not able to correctly classify two (2) of the abnormal input heart sound.

The evaluation process involved testing the device on 80 heart sound samples. Analysis of the results using a confusion matrix revealed successful classification of both normal and abnormal heart sounds, with some minor errors. Among the models tested, 2D-Convolutional Neural Networks (2D CNNs) achieved the best performance due to their ability to automatically extract relevant features and build a hierarchy of features from the raw data, allowing them to effectively distinguish between normal and abnormal heart sounds. Conversely, KNN with wavelet features performed the worst, likely because wavelets are not suitable for capturing the complex patterns present in heart sounds.

VII. CONCLUSION

The integration of an electronic stethoscope into the kiosk proved that it is an essential tool for heart sound detection. The Jetson Nano is efficient for heart sound analysis and with the 2D convolution neural network integrated into it is the most suitable for the purpose of heart sound analysis, as it has the highest accuracy among the machine learning algorithms that are used. With this part of the kiosk, it is significant for the community of Barangay Medicion II-B Imus Cavite to include this feature in the kiosk as it is beneficial to the community.

The continuous testing of the developed application is imperative to ensure its reliability and functionality. Moreover, refining the coding is essential to enhance the application's compatibility for widespread adoption across various health centers and clinics.

VIII. RECOMMENDATIONS

Electronic Stethoscope is considered useful and efficient in data gathering and heart assessment. However, there is more to add and improve the project. The proponents would like to recommend increasing the sample population. This will increase the accuracy and capability of the device.

It is also recommended to further isolate the sound by incorporating various physical and programming technicalities of the device. This will help to minimize the natural noise, white noise, and artifact sounds.

An improvement for graphical user interface is also a viable condition for accessibility of each patient that will be using the modified electronic stethoscope.

Lastly, exploring various machine learning algorithms might help to make the system more optimized, effective, and reliable for different cases of heart sounds.

REFERENCES

[1] Nazeer, Dr. M., Qayyum, M., & Ahad, A. (2022, December 22). Real time object detection and recognition in machine learning using Jetson Nano. SSRN.https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4286087

[2] Ren, Z., Qian, K., Dong, F., Dai, Z., Nejdl, W., Yamamoto, Y., & Schuller, B. W. (2022, May 14). Deep attention-based neural networks for explainable heart sound classification. Machine Learning with Applications. https://www.sciencedirect.com/science/article/pii/S266682702200038 X

[3] Wu, Y.-C., Han, C.-C., Chang, C.-S., Chang, F.-L., Chen, S.-F., Shieh, T.-Y., Chen, H.-M., & Lin, J.-Y. (2022, June 3). Development of an electronic stethoscope and a classification algorithm for Cardiopulmonary Sounds. MDPI. https://www.mdpi.com/1424 8220/22/11/4263

[4] Roy, T. S., Roy, J. K., & mandal, N. (2023, June 8). Design and development of electronic stethoscope for early screening of valvular heart disease prediction. Biomedical Signal Processing and Control. https://www.sciencedirect.com/science/article/abs/pii /S17468094230 05190