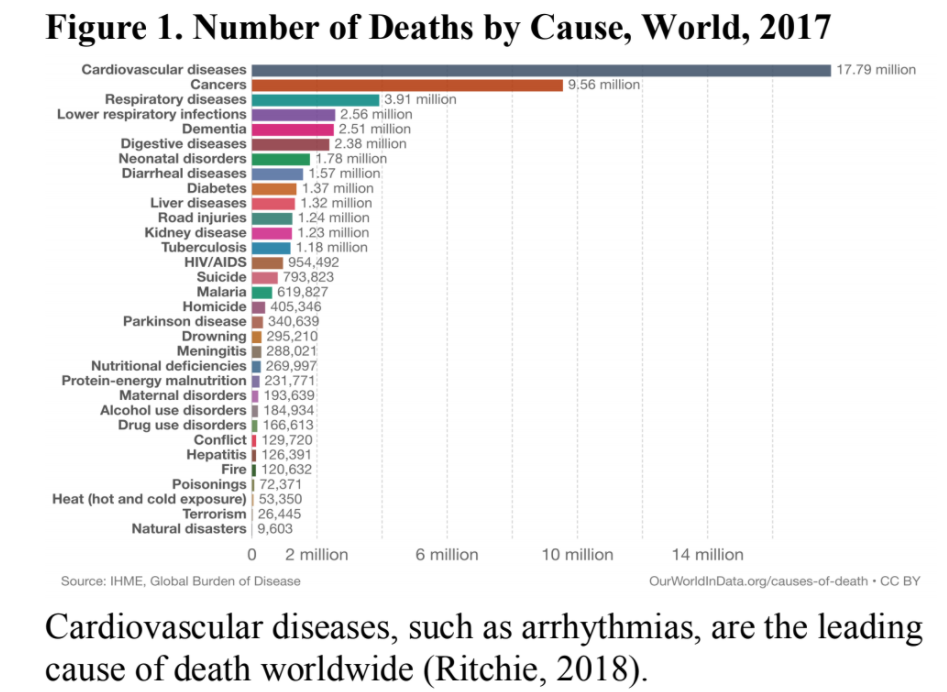
**Developing a Convolutional Neural Network Classifier to Identify Various Arrhythmias from Electrocardiogram Recordings**

Iva and Ella Chee

Yorktown High School

**INTRODUCTION**

**Overview of Research Problem**

Cardiovascular diseases are the main cause of death across the globe, as seen in Figure 1, with more than 17 million people dying each year (Ritchie & Roser, 2018). Furthermore, the prevalence and mortality of cardiovascular diseases continue to grow, underscoring the importance and necessity of regular heart rhythm monitoring to manage and curb cardiovascular diseases (World Health Organization: WHO, 2019). Cardiac arrhythmia is one common cardiovascular disease, characterized by an irregular rhythm of the heartbeat when electrical impulses in the heart fail to work properly. When those electrical impulses fail to function, arrhythmias can result in dangerous complications such as blood clots, stroke, heart failure, cardiac arrest, and sudden death. One major issue with arrhythmia diagnosis is that various arrhythmias can present the same symptoms: palpitations, chest pain, weakness, shortness of breath, fainting (Mayo Clinic, 2017). Fortunately, electrocardiograms or ECGs record the electrical activity of the heart, allowing for the discovery of specific abnormalities in the heart’s rhythm. Furthermore, deep learning with the use of artificial neural networks when used in tandem with ECGs to identify specific cardiac arrhythmias, can provide an even more accurate diagnosis. 

**Review of Literature**

Recently, deep learning techniques have accumulated much attention in biomedical engineering, in specific, the convolutional neural network (CNN), due to its variety of successful applications. Several studies have applied convolutional neural networks to detect, predict, or classify arrhythmia from electrocardiograms (Krittanawong et al., 2019). However, due to an imbalance of patients with various arrhythmias, there is an imbalance of electrocardiogram data presenting the less common arrhythmias. In the United States, though approximately 16.41 million individuals have cardiac arrhythmias, some, like atrial fibrillation (AFIB), are more common (CDC, 2020), whereas others, like supraventricular tachycardia (SVT), are more rare (Rehorn et al., 2021). This class imbalance results in biased deep learning models as they train on datasets containing more common abnormal rhythms. Cardiac arrhythmias present severe health risks, and techniques to improve deep learning models for diagnosis require more research.

***2D-CNN-LSTM Technology for Arrhythmia Diagnosis via ECG (Zheng et al., 2020)***

In a recent study for arrhythmia diagnosis via electrocardiogram, Zheng et al. used a two-dimensional convolutional neural network with long short-term memory (LSTM) which allowed for greater capability of retaining and feeding back information from selectively stored information. In classifying eight ECG beat signals and normal sinus rhythm, they were able to achieve 99.01% accuracy, 99.57% specificity, and 97.67% sensitivity, confirming the efficacy of 2D-CNN-LSTM technology in helping doctors diagnose arrhythmias with ease (Zheng et al.*,* 2020). While this model classified with very high performance, it classified beat type arrhythmias (i.e.: atrial premature beat, premature ventricular contraction, paced beat, ventricular escape beat), rather than rhythm types (i.e.: sinus rhythm, atrial fibrillation, ventricular fibrillation).

***2D-CNN for Arrhythmia Classification via ECG (Hannun et al., 2017)***

In a 2017 study, a dataset was built with over 500 times the number of patients than previous studies and was used to train a 34-layer two-dimensional convolutional neural network that maps ECG samples to a sequence of rhythm categories. The developed algorithm’s performances in both sensitivity and positive predictive value were superior to that of a board of certified cardiologists in the detection of several cardiac arrhythmias from electrocardiograms (Hannun et al., 2017). The findings of this study indicate the success in using a large dataset for an arrhythmia rhythm classifier.

***Fully Automatic ECG ACE-GAN Classification System (Zhai, Zhou & Tin, 2021)***

Zhai, Zhou, and Tin developed a system that utilized a generative adversarial network (GAN) to create more arrhythmic heartbeats for data augmentation and used the discriminator as a classifier (as shown in figure 2). The system outperformed several state-of-the-art automatic systems in detecting supraventricular ectopic beats and ventricular ectopic beats. While their findings demonstrated the effectiveness of implementation of a GAN in ECG classification, they utilized the MIT-BIH Arrhythmia database and classified beat type, whereas we hope a larger, more comprehensive dataset and classification based on rhythm type will produce better results. 

***ECG Arrhythmias Detection Using ACGAN and Residual Network (Wang et al. 2019)***

One 2019 study worked to investigate the use of a GAN to overcome the issue of an imbalanced ECG dataset by constructing a deep learning ECG abnormality detection framework composed of two models: data augmentation model and classification model. Wang et al. created an auxiliary classifier GAN to generate authentic artificial data (as seen in Figure 3) from the MIT-BIH dataset and competition dataset, and used a stacked residual network with LSTM for both single heartbeat detection and consecutive heartbeat detection. Despite successful data augmentation and network classification performance, the accuracy of the consecutive heartbeat detection is not as high as that of the standard benchmark. Thus, Wang et al. established the usefulness of a GAN for data augmentation, while also highlighting the necessity for further research in terms of rhythm (consecutive heartbeat) classification. 

***Our 2020 Study***

Our past study aimed to create and evaluate the effectiveness of a one-dimensional CNN (1D-CNN) that automatically detects and identifies AFIB from short-term ECG recordings out of a dataset including various arrhythmias, AFIB, and normal sinus rhythm from the MIT-BIH Arrhythmia database. The 1D-CNN 8-layer model was developed and evaluated with several performance metrics, resulting in high accuracy, precision, recall, specificity, and F1 score: 98.89%, 99.34%, 97.66%, 99.62%, and 98.50%, respectively, indicating its high efficacy. We combined the MIT-BIH Arrhythmia and Atrial Fibrillation in order to avoid a class imbalance issue, however several other arrhythmias (previously mentioned) exist and must be addressed. Further research in classifying multiple arrhythmias and utilizing additional procedures may provide even more effective methods of diagnosis.

Past research has demonstrated the success in utilizing convolutional neural networks to classify various arrhythmias. However, many of those models are employed to classify arrhythmias by rhythm type (a series of irregular heartbeats) rather than beat type (a single heartbeat). Furthermore, GANs can be implemented to synthesize data to allow for even larger datasets, as well as resolve the imbalance in data—like that of arrhythmias present in ECGs. And similar to the use of CNNs and arrhythmia classification, many of the generative adversarial networks that have been developed were used to augment specific beats, rather than whole rhythm arrhythmias. Combining a CNN and a GAN may result in an even more effective and useful system to advance diagnostic procedures for rhythm arrhythmias.

**Problem Statements**

* Class imbalance issue requires a solution—currently proposed techniques to resolve this issue (random resampling or cost-sensitive learning) could overfit the training data and reduce generalization when actually applied (Zhou et al., 2021).
* Few CNNs classify arrhythmias by rhythm type (i.e.: atrial fibrillation and supraventricular tachycardia); past studies focus on beat type (i.e.: atrial premature beat, premature ventricular contraction) and those that did use rhythm type were not as accurate or used less realistic data.

**Objective**

* Create and evaluate the effectiveness of a deep convolutional neural network with a generative adversarial network that automatically detects and identifies arrhythmia rhythms from short-term ECG recordings, in order to advance diagnostic procedures.
* Classify arrhythmias from electrocardiograms by rhythm type: sinus bradycardia (SB), atrial fibrillation (AFIB), sinus tachycardia (ST), atrial flutter (AF), sinus irregularity (SI), supraventricular tachycardia (SVT), atrial tachycardia (AT), atrioventricular node reentrant tachycardia (AVNRT), atrioventricular reentrant tachycardia (AVRT).

**Hypothesis**

* A generative adversarial network can be used to augment existing short-term 10-second 12-lead ECGs of several types of arrhythmias and resolve class imbalance issue and enlarge the dataset.
  + Wang et al. developed a GAN to create a larger dataset of ECG signals, and used the generated data to train/test a LSTM for arrhythmia classification; their model achieved high accuracy, sensitivity, and specificity (Wang et al., 2019).
* A convolutional neural network can be trained to accurately and efficiently identify several types of arrhythmias present in a short-term 10-second 12-lead ECG from a dataset of patients with various arrhythmias and healthy sinus rhythms.
  + Summer 2020, we conducted a project developing a 1D-CNN that classified AF from 30-second 12-lead ECGs with 98.89% accuracy.
  + Hannun et al. developed a 1D-CNN that classified 12 classes of arrhythmias, sinus rhythm and noise from 30-second ECGs; it had a higher F1 score, precision and recall than certified cardiologists (Hannun et al., 2017).

**CITATIONS (for introduction)**

CDC. (2020, September 8). *Atrial Fibrillation*. Centers for Disease Control and Prevention. https://www.cdc.gov/heartdisease/atrial\_fibrillation.htm

Hannun, A. Y., Rajpurkar, P., Haghpanahi, M., Tison, G. H., Bourn, C., Turakhia, M. P., & Ng, A. Y. (2017). Cardiologist-level arrhythmia detection with convolutional neural networks. *CoRR*, (),. https://doi.org/

Krittanawong, C., Johnson, K. W., Rosenson, R. S., Wang, Z., Aydar, M., Baber, U., Min, J. K., Tang, W. H. W., Halperin, J. L., & Narayan, S. M. (2019). Deep learning for cardiovascular medicine: a practical primer. *European Heart Journal*, *40*(25), 2058–2073. https://doi.org/10.1093/eurheartj/ehz056

Mayo Clinic. (2017). *Heart arrhythmia - Symptoms and causes*. Mayo Clinic; https://www.mayoclinic.org/diseases-conditions/heart-arrhythmia/symptoms-causes/syc-20350668

Rehorn, M., Sacks, N. C., Emden, M. R., Healey, B., Preib, M. T., Cyr, P. L., & Pokorney, S. D. (2021). Prevalence and incidence of patients with paroxysmal supraventricular tachycardia in the United States. *Journal of Cardiovascular Electrophysiology*. https://doi.org/10.1111/jce.15109

Ritchie, H., & Roser, M. (2018). *Causes of death*. Our World in Data. https://ourworldindata.org/causes-of-death

Wang, P., Hou, B., Shao, S., & Yan, R. (2019). ECG Arrhythmias Detection Using Auxiliary Classifier Generative Adversarial Network and Residual Network. *IEEE Access*, *7*, 100910–100922. https://doi.org/10.1109/access.2019.2930882

World Health Organization: WHO. (2019, June 11). *Cardiovascular diseases*. Who.int; World Health Organization: WHO. https://www.who.int/health-topics/cardiovascular-diseases/#tab=tab\_1

Zheng, Z., Chen, Z., Hu, F., Zhu, J., Tang, Q., & Liang, Y. (2020). An Automatic Diagnosis of Arrhythmias Using a Combination of CNN and LSTM Technology. *Electronics*, *9*(1), 121. https://doi.org/10.3390/electronics9010121

Zhou, Z., Zhai, X., & Tin, C. (2021). Fully automatic electrocardiogram classification system based on generative adversarial network with auxiliary classifier. *Expert Systems with Applications*, *174*, 114809. https://doi.org/10.1016/j.eswa.2021.114809