2021 PJAS Region 4 Presenter Registration Form

Please type or print very neatly:

Aditya	Kendre	Male	
First Name	Last Name	Gender	
717-620-8494	kendreaditya@gmail.com	12	4
Home Phone	Email	Grade	Region
12 Hamlet Circle	Mechanicsburg	PA	17050
Student Address	City	State	Zip Code

Number of years participated in PJAS (include this year) 1 2 3 4 5 6

Research Title (limit to 60 characters): **Generative Adversarial Networks for PCG Arrhythmia Detection**

School Name: **Cumberland Valley HS** Sponsor: **Michael Floreck**

Research Area: If research could overlap areas, please rank as 1, 2, 3.

____	Biochemistry (BC)	____	Behavioral/Psychology (BEH	2	Biology general (BIO)
____	Botany (BOT)	____	Chemistry (CHM)	1	Computer (CPS)
____	Ecology (EC)	____	Earth and Space (ES)	3	Engineering (ENG)
____	Mathematics (MAT)	____	Microbiology (MIC)	____	Physics (PHY)
____	Zoology (ZOO)				

We certify that this research has been conducted by the student in accordance with the Pennsylvania Junior Academy of Science rules and with advice only from others. PJAS at the regional and the state level utilizes judges from the area surrounding the competition site. The Regional and State competitions are held at a central location on a single day. In that the presenters, judges, and support personnel converge on the specific venues on a specific date, certain restrictions must apply. First, all presentations must take place on that day and on that site. Secondly, the results are deemed final. There cannot be "rejudging" of any presentations. Judges are polled before submitting their final results to ensure that the scores and the awards are acceptable to the entire judging panel. If there is a disagreement, it must be rectified before the judges' final submission. Once the judging panel is released the score and award is final. The large number of participants competing across the state will undoubtedly produce some results that are not expected. PJAS tries to make the evaluations as objective as possible. Using specifically designed rubrics along with orientation of judges, our judging committees do an exceptional job to ensure a fair result. We further agree to accept the Judge's evaluation of this research as final. In addition, we agree to make every attempt to have the student present at the Region 4 Competition and at the State Competition should he/she receive a First Award at the Region 4 Competition.

Signature of Student aditya Date 1/4/2021

Signature of Parent Date 1/4/2021

Signature of Sponsor Date

Checklist for Adult Sponsor (1)

This completed form is required for ALL projects.

To be completed by the Adult Sponsor in collaboration with the student researcher(s):

Student's Name(s): Aditya Kendre

Project Title: Generative Adversarial Networks for PCG Arrhythmia Detection

1. ☑ I have reviewed the ISEF Rules and Guidelines, including the science fair ethics statement.
2. ☑ I have reviewed the student's completed Student Checklist (1A) and Research Plan/Project Summary.
3. ☑ I have worked with the student and we have discussed the possible risks involved in the project.
4. ☐ The project involves one or more of the following and requires prior approval by an SRC, IRB, IACUC or IBC:
   - ☐ Humans
   - ☐ Vertebrate Animals
   - ☐ Potentially Hazardous Biological Agents
     - ☐ Microorganisms ☐ rDNA ☐ Tissues
5. ☑ Items to be completed for **ALL PROJECTS**
   - ☑ Adult Sponsor Checklist (1)
   - ☑ Student Checklist (1A)
     - ☐ Regulated Research Institutional/Industrial Setting Form (1C) (when applicable; after completed experiment)
     - ☐ Continuation/Research Progression Form (7) (when applicable)
   - ☑ Research Plan/Project Summary
   - ☑ Approval Form (1B)

Additional forms required if the project includes the use of one or more of the following (check all that apply):

☐ **Humans,** including student designed inventions/prototypes. (Requires prior approval by an Institutional Review Board (IRB); see full text of the rules.)
- ☐ Human Participants Form (4) or appropriate Institutional IRB documentation
- ☐ Sample of Informed Consent Form (when applicable and/or required by the IRB)
- ☐ Qualified Scientist Form (2) (when applicable and/or required by the IRB)

☐ **Vertebrate Animals** (Requires prior approval, see full text of the rules.)
- ☐ Vertebrate Animal Form (5A) - for projects conducted in a school/home/field research site (SRC prior approval required.)
- ☐ Vertebrate Animal Form (5B) - for projects conducted at a Regulated Research Institution. (Institutional Animal Care and Use Committee (IACUC) approval required prior experimentation.)
- ☐ Qualified Scientist Form (2) (Required for all vertebrate animal projects at a regulated research site or when applicable)

☐ **Potentially Hazardous Biological Agents** (Requires prior approval by SRC, IACUC or IBC, see full text of the rules.)
- ☐ Potentially Hazardous Biological Agents Risk Assessment Form (6A)
- ☐ Human and Vertebrate Animal Tissue Form (6B) - to be completed in addition to Form 6A when project involves the use of fresh or frozen tissue, primary cell cultures, blood, blood products and body fluids.
- ☐ Qualified Scientist Form (2) (when applicable)
- ☐ The following are exempt from prior review but require a Risk Assessment Form 3: projects involving protists, archae and similar microorganisms, for projects using manure for composting, fuel production or other non-culturing experiments, projects using color change coliform water test kits, microbial fuel cells, and projects involving decomposing vertebrate organisms.

☐ **Hazardous Chemicals, Activities and Devices** (No SRC prior approval required, see full text of the rules.)
- ☐ Risk Assessment Form (3)
- ☐ Qualified Scientist Form (2) (required for projects involving DEA-controlled substances or when applicable)

☐ **Other**
- ☐ Risk Assessment Form (3)

☑ **I attest to the information checked above and that I have read and agree to abide by the science fair ethics statement.**

Mike Floreck		10/30/20
Adult Sponsor's Printed Name	Signature	Date of Review (mm/dd/yy)
(717) 506-3413	mfloreck@cvschools.org	
Phone	Email	

Student Checklist (1A)

This form is required for ALL projects.

1. a. Student/Team Leader: Aditya Kendre Grade: 12

   Email: kendreaditya@gmail.com Phone: (717) 622-1281

   b. Team Member: ______ c. Team Member: ______

2. Title of Project:

   Generative Adversarial Networks for PCG Arrhythmia Detection

3. School: Cumberland Valley High School School Phone: (717) 506-3413

   School Address: 6746 Carlisle Pike

   Mechanicsburg, PA 17050

4. Adult Sponsor: Mike Floreck Phone/Email: mfloreck@cvschools.org

5. Does this project need SRC/IRB/IACUC or other pre-approval? ☐ Yes ☒ No Tentative start date: ______

6. Is this a continuation/progression from a previous year? ☒ Yes ☐ No

   If Yes:

   a. Attach the previous year's ☑ Abstract **and** ☑ Research Plan/Project Summary

   b. Explain how this project is new and different from previous years on

   ☑ Continuation/Research Progression Form (7)

7. This year's laboratory experiment/data collection:

   10/30/20 — Actual Start Date: (mm/dd/yy)

   03/01/21 — End Date: (mm/dd/yy)

8. Source of Data:

   ☐ Collected self/mentor ☑ Other Describe/url: Physionet Database

9. List name and address of all non-home and non-school work site(s):

Name: ______ ______

Address: ______ ______

______ ______

Phone/email ______ ______

10. **Complete a Research Plan/Project Summary following the Research Plan/Project Summary instructions and attach to this form.**

11. **An abstract is required for all projects after experimentation.**

Research Plan/Project Summary Instructions

A complete Research Plan/Project Summary is required for ALL projects and must accompany Student Checklist (1A).

- All projects must have a Research Plan/Project Summary
  a. Written prior to experimentation following the instructions below to detail the rationale, research question(s), methodology, and risk assessment of the proposed research.
  b. If changes are made during the research, such changes can be added to the original research plan as an addendum, recognizing that some changes may require returning to the IRB or SRC for appropriate review and approvals. If no additional approvals are required, this addendum serves as a project summary to explain research that was conducted.
  c. If no changes are made from the original research plan, no project summary is required.

- Some studies, such as an engineering design or mathematics projects, will be less detailed in the initial project plan and will change through the course of research. If such changes occur, a project summary that explains what was done is required and can be appended to the original research plan.

- The Research Plan/Project Summary should include the following:
  a. **RATIONALE:** Include a brief synopsis of the background that supports your research problem and explain why this research is important and if applicable, explain any societal impact of your research.
  b. **RESEARCH QUESTION(S), HYPOTHESIS(ES), ENGINEERING GOAL(S), EXPECTED OUTCOMES:** How is this based on the rationale described above?
  c. Describe the following in detail:
- **Procedures:** Detail all procedures and experimental design including methods for data collection, and when applicable, the source of data used. Describe only your project. Do not include work done by mentor or others.
- **Risk and Safety:** Identify any potential risks and safety precautions needed.
- **Data Analysis:** Describe the procedures you will use to analyze the data/results.
  d. **BIBLIOGRAPHY:** List major references (e.g. science journal articles, books, internet sites) from your literature review. If you plan to use vertebrate animals, one of these references must be an animal care reference.

Items 1–4 below are subject-specific guidelines for additional items to be included in your research plan/project summary as applicable.

1. Human participants research:
  a. Participants: Describe age range, gender, racial/ethnic composition of participants. Identify vulnerable populations (minors, pregnant women, prisoners, mentally disabled or economically disadvantaged).
  b. Recruitment: Where will you find your participants? How will they be invited to participate?
  c. Methods: What will participants be asked to do? Will you use any surveys, questionnaires or tests? If yes and not your own, how did you obtain? Did it require permissions? If so, explain. What is the frequency and length of time involved for each subject?
  d. Risk Assessment: What are the risks or potential discomforts (physical, psychological, time involved, social, legal, etc.) to participants? How will you minimize risks? List any benefits to society or participants.
  e. Protection of Privacy: Will identifiable information (e.g., names, telephone numbers, birth dates, email addresses) be collected? Will data be confidential/anonymous? If anonymous, describe how the data will be collected. If not anonymous, what procedures are in place for safeguarding confidentiality? Where will data be stored? Who will have access to the data? What will you do with the data after the study?
  f. Informed Consent Process: Describe how you will inform participants about the purpose of the study, what they will be asked to do, that their participation is voluntary and they have the right to stop at any time.

2. Vertebrate animal research:
  a. Discuss potential ALTERNATIVES to vertebrate animal use and present justification for use of vertebrates.
  b. Explain potential impact or contribution of this research.
  c. Detail all procedures to be used, including methods used to minimize potential discomfort, distress, pain and injury to the animals and detailed chemical concentrations and drug dosages.
  d. Detail animal numbers, species, strain, sex, age, source, etc., include justification of the numbers planned.
  e. Describe housing and oversight of daily care.
  f. Discuss disposition of the animals at the end of the study.

- **Potentially hazardous biological agents research:**
  a. Give source of the organism and describe BSL assessment process and BSL determination.
  b. Detail safety precautions and discuss methods of disposal.

4. Hazardous chemicals, activities & devices:
  - Describe Risk Assessment process, supervision, safety precautions and methods of disposal.
  - Material Safety Data Sheets are not necessary to submit with paperwork.

ISEF Sample Abstract & Certification

Generative Adversarial Networks for PCG Arrhythmia Detection

Aditya Kendre
Cumberland Valley HS, Mechanicsburg, PA, USA

With the rapid growth of computational power and complex algorithms, we propose a novel approach to detect arrhythmias in Phonocardiograms (PCGs). Typically, Electrocardiograms are used to diagnose arrhythmias, requiring medical-grade equipment to accurately recognize cardiac illnesses. PCGs provide ease of access to everyone who has a device capable of recording audio, allowing medical professionals to treat arrhythmias in the developmental stages. The new design comprises two subsystems; one is based on the relationship between Electrocardiograms (ECGs) and PCGs, and the other between PCGs and arrhythmias. The association between ECGs and PCGs is amended to translate from one space to another, where ECGs become dimensionally reduced, then reconstructed into a PCG signal. The second subsystem uses a Generative Adversarial Networks (GAN), in which both arbitrary PCG signals are generated, and preexisting ECG datasets are recreated into PCG signals (using subsystem one). These signals are fed into a classifier that detects if an arrhythmia is present. This proposed system's advantage is that PCG data is more readily available than ECG data; hence, more heart diagnostics can be made.

Category
Pick one only--
mark an "X"
in box at right

- [ ] Animal Sciences
- [ ] Behavioral and Social Sciences
- [ ] Biochemistry
- [ ] Biomedical and Health Sciences
- [ ] Biomedical Engineering
- [ ] Cellular & Molecular Biology
- [ ] Chemistry
- [x] Computational Biology and Bioinformatics
- [ ] Earth & Environmental Sciences
- [ ] Embedded Systems
- [ ] Energy: Sustainable Materials and Design
- [ ] Engineering Mechanics
- [ ] Environmental Engineering
- [ ] Materials Science
- [ ] Mathematics
- [ ] Microbiology
- [ ] Physics and Astronomy
- [ ] Plant Sciences
- [ ] Robotics & Intelligent Machines
- [ ] Systems Software
- [ ] Translational Medical Science

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check all that apply):

   - [ ] human participants
   - [ ] potentially hazardous biological agents
   - [ ] vertebrate animals
   - [ ] microorganisms
   - [ ] rDNA
   - [ ] tissue

2. This abstract describes only procedures performed by me/us, reflects my/our own independent research, and represents one year's work only.
   [x] yes [ ] no

3. I/We worked or used equipment in a regulated research institution or industrial setting.
   [ ] yes [x] no

4. This project is a continuation of previous research.
   [x] yes [x] no

5. My display board includes non-published photographs/visual depictions of humans (other than myself):
   [x] yes [ ] no

6. I/We hereby certify that the abstract and responses to the above statements are correct and properly reflect my/our own work.
   [x] yes [ ] no

FOR ISEF OFFICIAL USE ONLY

Generative Adversarial Networks for PCG Arrhythmia Detection

Aditya Kendre
Cumberland Valley High School
Mechanicsburg, PA 17050

January 3, 2021

Abstract

With the rapid growth of computational power and complex algorithms, we propose a novel approach to detect arrhythmias in Phonocardiograms (PCGs). Typically, Electrocardiograms are used to diagnose arrhythmias; requiring medical grade equipment to accurately recognize cardiac illnesses (Rajpurkar et al., 2017). PCGs, however, provide ease of access to everyone who has a device capable of recording audio, allowing medical professionals to treat arrhythmias in the developmental stages. The new design comprises two subsystems; one is based on the relationship between Electrocardiograms (ECGs) and PCGs, and the other between PCGs and arrhythmias. The association between ECGs and PCGs is amended to translate from one space to another, where ECGs become dimensionally reduced, then reconstructed into a PCG signal. The second subsystem uses a Generative Adversarial Networks (GAN), in which both arbitrary PCG signals are generated, and preexisting ECG datasets are recreated into PCG signals (using subsystem one). These signals are fed into a classifier that detects if an arrhythmia is present. This proposed system's advantage is that PCG data is more readily available than ECG data; hence, more heart diagnostics can be made.

1 Introduction

Problem Statement. Every physical examination done with a stethoscope should aim to diagnose any arrhythmias present within a patient.

Question. Is it possible to create a model capable of surpassing the accuracy of Cardiologists in identifying heart arrhythmias in Phonocardiograms?

Hypothesis. It is possible to exceed the accuracy of Cardiologists when compared to that of a Generative Adversarial Network‘s, to identify heart arrhythmias in Phonocardiograms.

Materials List. Computer.

Electrocardiograms have created a profound impact in the field of cardiology, specifically in recognizing heart arrhythmias, a problem with the rhythm of one's heartbeat. Noninvasive arrhythmia analysis is based on multiple electrodes that reflect the electrical activity on ECGs. However, with the recent surge of heart-related medical cases, it is getting difficult to diagnose heart conditions at an early stage. As most treatments rely on detecting the disease in it's infancy stages. Traditionally, arrhythmias are diagnosed by cardiologists by analyzing ECG recordings (Jordaens, 2018). Some clinics have adopted a new technique in which ECG and PCG signals are simultaneously recorded and then computationally analyzed. This, however, still requires an instrument capable of recording ECG data. Such instruments are only available during scheduled appointments, often which are recommended by physicians. If a physician fails to detect symptoms of arrhythmia, a patient may never receive a diagnosis. One study found 44% of cardiologists were not able to detect common cardiac events with stethoscopes (Mangione et al., 1993); in another study, delays in cardiac-related illness diagnosis and treatment impacted procedural success rates by as much as 24% (Bunch et al., 2013). We propose a method where it is now possible to accurately detect arrhythmias with only PCG recordings. This provides an opportunity for physicians to check for potential developments of cardiac arrhythmias at every physical exam accurately.

Current PCG arrhythmia diagnosis methods only recognize between Normal and Abnormal (binary classification), providing minimal information about what is present within the PCG signal (Aziz et al., 2020). This is because no PCG datasets exist that include more than 3 classes of arrhythmia. Therefore, it is necessary to transform pre-existing ECG datasets with multiple classes to PCG signals. This enables models to detect a larger range of arrhythmia without explicitly collecting new PCG recordings. Currently, no technology attempts to construct PCG signals from existing ECG data.

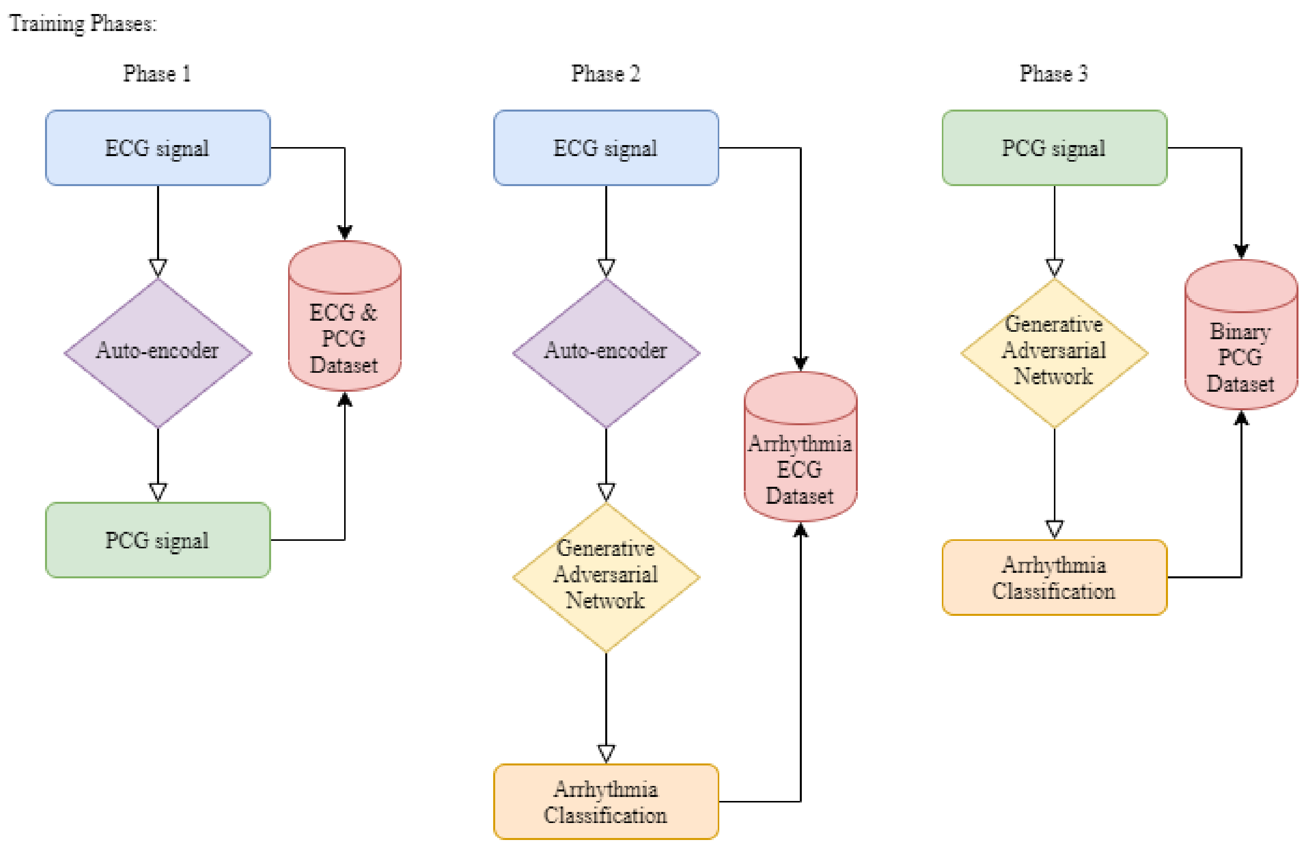
2 Methodology

2.1 Approach

The model contains two sub-models, an Autoencoder (AE), and a Generative Adversarial Network (GAN). The AE is responsible for extracting relevant features from an ECG signal and constructing a PCG signal from the latent features. The GAN is responsible for extracting relevant features and classifying the PCG signals.

The training phase involves 3 stages: AE training, AE+GAN training, and GAN fine-tuning. During training phases, all datasets will follow the following split: 70% - training, 15% - validation, 15% - testing; this cross-validation step validates that both models are not overfitting during the training phase. The first stage involves training the AE with a supervised dataset of ECG and PCG signals (Liu et al, 2016). The second stage involves training both the AE and the

GAN with a supervised dataset of arrhythmias within ECGs (Goldberger et al., 2017). During the training process, the AE model will be frozen (the weights and biases of the AE model won't be trained) as this process is already done in the preceding stage. The last stage is fine-tuning the GAN on a binary supervised dataset of PCG signals (Normal vs Abnormal). This validates the model's metrics in the previous step.



2.2 Data Analysis

While testing and training, the model will be validated against with metrics such as recall, precision, accuracy, loss, FBeta, F1 score, and ROC/AUC score. These tests will ensure that the model is accurately predicting the classes, and identifying important features within the datasets. Each step in the training phase will represent a milestone and an accuracy of 97% will mark the completion criteria.

2.3 Potential Problems

Overfitting: One of the largest problems in Deep Learning overall, which possesses a threat to our model is overfitting. Overfitting typically happens when the model metrics of the training and validation set diverge. This suggests that the model is not generalizing, but rather memorizing the training dataset. To combat overfitting, researchers typically implement data argumentation techniques to reinforce important features in a dataset.

Domain Shift: A domain shift occurs when a source dataset performs well but on a different dataset distribution, the performance drastically decreases. Typically, domain adaptation is often used to improve performance on target datasets. This is done by training the model itself on multiple datasets to improve the model's capacity to generalize.

Traning Time: With large multi-model architectures, it becomes tough to train models on a single GPU. This can happen for a number of reasons, but the main reason is because the model takes up too much memory of the GPU. Generally, parallel processing is used to split tasks and assign them to different GPUs. For instance, the AE model will run on a single GPU, while the GAN will run on another GPU.

References

[1] Jordaens, L. (2018). Aclinical approach to arrhythmias revisited in 2018. Netherlands Heart Journal. doi:10.1007/s12471-018-1089-1

[2] Mangione, S. (1993). The Teaching and Practice of Cardiac Auscultation during Internal Medicine and Cardiology Training: A Nationwide Survey. Annals of Internal Medicine, 119(1), 47. doi:10.7326/0003-4819-119-1-199307010-00009

[3] Bunch, T. J., May, H. T., Bair, T. L., Johnson, D. L., Weiss, J. P., Crandall, B. G., … Day, J. D. (2013). Increasing time between first diagnosis of atrial fibrillation and catheter ablation adversely affects long-term outcomes. Heart Rhythm, 10(9), 1257–1262. doi:10.1016/j.hrthm.2013.05.013

[4] Aziz, S., Khan, M. U., Alhaisoni, M., Akram, T., Altaf, M. (2020). Phonocardiogram Signal Processing for Automatic Diagnosis of Congenital Heart Disorders through Fusion of Temporal and Cepstral Features. Sensors, 20(13), 3790. doi:10.3390/s20133790

[5] Felipe Alonso "Detection of life threatening arrhythmias using feature selection andsupport vector machines", IEEE Transactions on Biomedical Engineering,Vol 61No.3,pp.832-840,March 2014.

Approval Form (1B)

A completed form is required for each student, including all team members.

1. To Be Completed by Student and Parent

a. **Student Acknowledgment:**

- I understand the risks and possible dangers to me of the proposed research plan.
- I have read the ISEF Rules and Guidelines and will adhere to all International Rules when conducting this research.
- I have read and will abide by the science fair ethics statement.

Student researchers are expected to maintain the highest standards of honesty and integrity. Scientific fraud and misconduct are not condoned at any level of research or competition. Such practices include but are not limited to plagiarism, forgery, use or presentation of other researcher's work as one's own, and fabrication of data. Fraudulent projects will fail to qualify for competition in affiliated fairs and ISEF.

Aditya Kendre — aditya — 10/29/20

Student's Printed Name | Signature | Date Acknowledged (mm/dd/yy) (Must be prior to experimentation.)

b. **Parent/Guardian Approval:** I have read and understand the risks and possible dangers involved in the **Research Plan/Project Summary**. I consent to my child participating in this research.

Nivrutti Kendre — 10/29/20

Parent/Guardian's Printed Name | Signature | Date Acknowledged (mm/dd/yy) (Must be prior to experimentation.)

2. To be completed by the local or affiliated Fair SRC

(Required for projects requiring prior SRC/IRB APPROVAL. Sign 2a or 2b as appropriate.)

a. **Required for projects that need prior SRC/IRB approval BEFORE experimentation** (humans, vertebrates or potentially hazardous biological agents).

The SRC/IRB has carefully studied this project's **Research Plan/Project Summary** and all the required forms are included. My signature indicates approval of the **Research Plan/Project Summary** before the student begins experimentation.

SRC/IRB Chair's Printed Name

Signature | Date of Approval (mm/dd/yy) (Must be prior to experimentation.)

OR

b. **Required for research conducted at all Regulated Research Institutions with no prior fair SRC/IRB approval.**

This project was conducted at a regulated research institution (**not home or high school, etc.**), was reviewed and approved by the proper institutional board before experimentation and complies with the ISEF Rules. **Attach (1C) and any required institutional approvals (e.g. IACUC, IRB).**

SRC Chair's Printed Name

Signature | Date of Signature (mm/dd/yy) (May be after experimentation)

3. Final ISEF Affiliated Fair SRC Approval (Required for ALL Projects)

SRC Approval After Experimentation and Before Competition at Regional/State/National Fair

I certify that this project adheres to the approved **Research Plan/Project Summary** and complies with all ISEF Rules.

Regional SRC Chair's Printed Name | Signature | Date of Approval (mm/dd/yy)

State/National SRC Chair's Printed Name *(where applicable)* | Signature | Date of Approval (mm/dd/yy)

Qualified Scientist Form (2)

May be required for research involving human participants, vertebrate animals, potentially hazardous biological agents, and hazardous substances and devices. Must be completed and signed before the start of student experimentation.

Student's Name(s) Aditya Kendre

Title of Project Generative Adversarial Networks for PCG Arrhythmia Detection

---

To be completed by the Qualified Scientist:

Scientist Name: Lifang He

Educational Background: Machine Learning/Deep Learning/Biomedical Informatics Degree(s): B.S., Computational Mathematics; Ph.D., Computer Science

Experience/Training as relates to the student's area of research

Assistant Professor	Lehigh University
Position:	Institution:
BC 327, 113 Research Drive, Bethlehem, PA 18015	lih319@lehigh.edu
Address:	Email/Phone:

1. Have you reviewed the ISEF rules relevant to this project and the science fair ethics statement relevant to this project? ☒ Yes ☐ No

2. Will any of the following be used?
   a. Human participants ☐ Yes ☒ No
   b. Vertebrate animals ☐ Yes ☒ No
   c. Potentially hazardous biological agents (microorganisms, rDNA and tissues, including blood and blood products) ☐ Yes ☒ No
   d. Hazardous substances and devices ☐ Yes ☒ No

3. Will this study be a sub-set of a larger study? ☐ Yes ☒ No

4. Will you directly supervise the student? ☒ Yes ☐ No
   a. If no, who will directly supervise and serve as the Designated Supervisor?
   b. Experience/Training of the Designated Supervisor:

To be completed by the Qualified Scientist:

I certify that I have reviewed and approved the Research Plan/Project Summary prior to the start of the experimentation. If the student or Designated Supervisor is not trained in the necessary procedures, I will ensure her/his training. I will provide advice and supervision during the research. I have a working knowledge of the techniques to be used by the student in the Research Plan/Project Summary. I understand that a Designated Supervisor is required when the student is not conducting experimentation under my direct supervision.

Lifang He
Qualified Scientist's Printed Name

Lifang He
Signature

01/04/21
Date of Approval (mm/dd/yy)

To be completed by the Designated Supervisor when the Qualified Scientist cannot directly supervise.

I certify that I have reviewed the Research Plan/Project Summary and have been trained in the techniques to be used by this student, and I will provide direct supervision.

Designated Supervisor's Printed Name

Signature

Date of Approval (mm/dd/yy)

Phone

Email

Risk Assessment Form (3)

Must be completed before experimentation.

Student's Name(s) Aditya Kendre

Title of Project Generative Adversarial Networks for PCG Arrhythmia Detection

To be completed by the Student Researcher(s) in collaboration with Designated Supervisor/Qualified Scientist: (All questions must be answered; additional page(s) may be attached.)

1. List all hazardous chemicals, activities, or devices that will be used; identify microorganisms exempt from pre-approval (see Potentially Hazardous Biological Agent rules).

   The only device used in this research project is a laptop.

2. Identify and assess the risks and hazards involved in this project.

   N/A

3. Describe the safety precautions and procedures that will be used to reduce the risks.

   N/A

4. Describe the disposal procedures that will be used (when applicable).

   N/A

5. List the source(s) of safety information.

   N/A

To be completed and signed by the Designated Supervisor (or Qualified Scientist, when applicable):
I agree with the risk assessment and safety precautions and procedures described above. I certify that I have reviewed the Research Plan/Project Summary and the International Rules, including the science fair ethics statement and will provide direct supervision.

Designated Supervisor's Printed Name

Signature

Date of Review (mm/dd/yy)

Position & Institution

Phone or email contact information

Experience/Training as relates to the student's area of research

Continuation/Research Progression Projects Form (7)

Required for projects that are a continuation/progression in the same field of study as a previous project. This form must be accompanied by the previous year's abstract and Research Plan/Project Summary.

Student's Name(s) Aditya Kendre

To be completed by Student Researcher: List all components of the current project that make it new and different from previous research. The information must be on the form; use an additional form for previous year and earlier projects.

Components	Current Research Project	Previous Research Project: Year: 19-20
1. Title	Generative Adversarial Networks for PCG Arrhythmia Detection	ECG-Based Abnormal Heartbeat Classification: A Deep Learning Approach for Arrhythmia Detection
2. Change in goal/ purpose/objective	To create a lightweight, precise, and accurate model for predicting heart arrhythmias in Phonocardiograms using a Generative Adversarial Network capable of exceeding Cardiologists' accuracy.	To create a model capable of surpassing the accuracy of Cardiologists in identifying heart arrhythmias in Electrocardiograms.
3. Changes in methodology	A Generative Adversarial Networks comprises of two models: a generator model and a classifier model (which contains a Convolutional Neural Network). The generator creates artificial PCG data to deceive the classifier into predicting the data is a real PCG signal while simultaneously being fed true PCG data from a dataset.	A Convolutional Neural Network extracts latent features from an electrocardiogram database following a fully-connected Linear layer that predicts whether an arrhythmia is present within the electrocardiogram, based upon the features extracted by the CNN.
4. Variable studied	Manipulated variables include: Learning Rate, Batch size, Number of Epochs, Hidden Layers, Hidden Units, Activations Functions, and level of Data Augmentation. Responding variables include: Loss, Accuracy, Recall, Precision, F-Beta Score, F1 Score, and ROC and AUC.	Manipulated variables include: Number of layers, Hidden Units, and the level of Data Augmentation. Responding variables include: Loss and Accuracy.
5. Additional changes	Conversion between ECG and PCG signals using an Autoencoder.	ECG signal with a one-dimensional CNN.

Attached are:

☑ Abstract and Research Plan/Project Summary, Year 19-20

I hereby certify that the above information is correct and that the current year Abstract & Certification and project display board properly reflect work done only in the current year.

Aditya Kendre — Student's Printed Name(s)

aditya — Signature

10/30/20 — Date of Signature (mm/dd/yy)

OFFICIAL ABSTRACT and CERTIFICATION

A Deep Learning Approach for Arrhythmia Detection

Aditya Kendre
Cumberland Valley High School, Mechanicsburg PA, Adams County

Early detection of cardiac arrhythmia has the potential to prevent the millions of moralities that the disease causes globally. However, there are few automated systems to identify arrhythmia. A significant impediment in achieving successful methods include the lack of a large training dataset. Despite this difficulty, processes like data augmentation allow for anincreased amount and diversity of data. Here, the electrocardiogram (ECG) datasets were obtained from the PhysioNet database. The dataset was used to train a Convolutional Neural Network (CNN) on classifying cardiac arrhythmia. Experimental results illustrate advantages such as better responsiveness and higher accuracy of deep learning-based models when compared to the traditional analysis on ECGs.

Category
Pick one only — mark an "X" in box at right

- Animal Sciences
- Behavioral & Social Sciences
- Biochemistry
- Biomedical & Health Sciences
- Biomedical Engineering
- Cellular & Molecular Biology
- Chemistry
- Computational Biology & Bioinformatics ■
- Earth & Environmental Sciences
- Embedded Systems
- Energy: Sustainable Materials and Design
- Engineering Mechanics
- Environmental Engineering
- Materials Science
- Mathematics
- Microbiology
- Physics & Astronomy
- Plant Sciences
- Robotics & Intelligent Machines
- Systems Software
- Translational Medical Sciences

1. As a part of this research project, the student directly handled, manipulated, or interacted with (check ALL that apply):

   ☐ human participants ☐ potentially hazardous biological agents
   ☐ vertebrate animals ☐ microorganisms ☐ rDNA ☐ tissue

2. I/we worked or used equipment in a regulated research institution or industrial setting: ☐ Yes ■ No

3. This project is a continuation of previous research. ☐ Yes ■ No

4. My display board includes non-published photographs/visual depictions of humans (other than myself): ☐ Yes ■ No

5. This abstract describes only procedures performed by me/us, reflects my/our own independent research, and represents one year's work only ■ Yes ☐ No

6. I/we hereby certify that the abstract and responses to the above statements are correct and properly reflect my/our own work. ■ Yes ☐ No

This stamp or embossed seal attests that this project is in compliance with all federal and state laws and regulations and that all appropriate reviews and approvals have been obtained including the final clearance by the Scientific Review Committee.

ECG-Based Abnormal Heartbeat Classification: A Deep Learning Approach for Arrhythmia Detection

Aditya Kendre

Cumberland Valley High School

Rationale

Electrocardiograms (ECG) have created a profound impact in the field of cardiology, specify in recognizing of heart arrhythmias. Non-invasive arrhythmia analysis is based on 10 electrodes that reflect the electrical activity on ECGs. An estimated three million cases of arrhythmia occur in the United States yearly (Mayo Clinic). Diagnosing this disease early is the key to one's wellness, yet 18% of cardiologists misinterpreted ECGs containing atrial fibrillation (Anh et al, 2006). With the recent advancements in technology, Machine Learning algorithms such as Deep Neural Networks (DNNs), allow a computer to learn features and identify patterns within a given dataset. On the basic level, DNNs receive input data, and through a series of weights and biases, outputs a confidence value in all possible labels of the dataset, similar to a human's neural network. Furtherance in the accuracy of abnormal heartbeat classification will allow cardiologists to accurately, and efficiently recognizing arrhythmia before becoming prevalent in one's wellbeing.

Research

Research Question: This research project will examine whether a classifier will be able to accurately identify abnormal heartbeat in ECGs.

Hypothesis: If an image classifier received a supervised dataset of heart arrhythmia of ECGs, then the image classifier will allow an accurate identification of arrhythmia.

Expectation: The image classifier should reach an accuracy of above 82%.

Procedure:

1. Gather a dataset of annotated ECGs
2. Determine type of classifier used to learn dataset features
3. Analyze results using Gradient Decent and Mean Loss function

Risks and Safety:

This research project involves no risks or safety concerns.

References

Alfaras, Miquel, Soriano, & Silvia. (2019, July 3). A Fast Machine Learning Model for ECG-Based Heartbeat Classification and Arrhythmia Detection. Retrieved October 30, 2019, from https://www.frontiersin.org/articles/10.3389/fphy.2019.00103/full.

Mayo Clinic. (2019, April 2). Heart arrhythmia. Retrieved October 30, 2019, from https://www.mayoclinic.org/diseases-conditions/heart-arrhythmia/symptoms-causes/syc-20350668?utm_source=Google&utm_medium=abstract&utm_content=Cardiac-arrhythmia&utm_campaign=Knowledge-panel.

Srinivasan, N. T., & Schilling, R. J. (2018, June). Sudden Cardiac Death and Arrhythmias. Retrieved October 30, 2019, from https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6020177/.