Project Report : CS 7643
Detecting Chagas Disease from Electrocardiograms with Transformers

Marc Lafargue
Georgia Institute of Technology
225 North Avenue NW, Atlanta, GA
mlafargue3@gatech.edu

Matheus Rama Amorim
Georgia Institute of Technology
225 North Avenue NW, Atlanta, GA
matamorim@gatech.edu

Trevor Gratz
Georgia Institute of Technology
225 North Avenue NW, Atlanta, GA
tgratz6@gatech.edu

Abstract

The ABSTRACT is to be in fully-justified italicized text, at the top of the left-hand column, below the author and affiliation information. Use the word "Abstract" as the title, in 12-point Times, boldface type, centered relative to the column, initially capitalized. The abstract is to be in 10-point, single-spaced type. Leave two blank lines after the Abstract, then begin the main text. Look at previous CVPR abstracts to get a feel for style and length. The abstract section should contain a brief summary of your work that includes the problem statement, proposed solution and results.

1. Introduction/Background/Motivation

Chagas disease is a parasitic illness transmitted by insects, affecting approximately 8 million people globally, [1] with over 4,750 deaths each year. [2] The vast majority of people living with Chagas disease develop no signs or symptoms upon initial infection. [1] However, chronic infection can cause heart disease, including heart failure. [3]

Because the acute phase of Chagas disease is often asymptomatic, it goes undiagnosed in up to 95% of vector-borne cases i.e. those via insect transmission. [1] If Chagas disease is suspected, then multiple serological tests may be needed to correctly diagnose it. Furthermore, access to serological testing remains limited. [1] There is a critical need for new diagnostic approaches to detect Chagas disease, particularly for the large number of undiagnosed cases.

Chronic cases of Chagas disease can cause changes in the Electrocardiogram (ECG) readings of positive patients, specifically changes associated with a right bundle branch block. [4] Recent work has shown that low-cost widely available diagnostic tools, such as ECGs paired with deep learning algorithms, may be utilized to diagnose Chagas disease. [4] Jidling et al. (2023) developed a convolutional neural network with residual connections achieving a 0.8 Area Under the Receiver Operating Characteristic curve (AUROC) in a binary classification task of Chagas disease. [4] To our knowledge, there are no other studies using deep learning to diagnose Chagas disease from ECG.

The lack of empirical studies using deep learning for Chagas disease detection is surprising given the well-developed field of deep learning based methods for diagnosing other disease using ECGs. For instance, state of the art models for ECG classification have been developed by combining convolution layers with layers that handle the sequential nature of ECGs i.e. RNN or LSTM layers. [5] Other work has explored using convolution layers in combination with transformer blocks for arrhythmia classification. [6]

Our goal is to create a deep learning model that can accurately diagnose Chagas disease from ECG readings, offering a transformative tool that could save millions of lives by enabling timely diagnosis and treatment. To do so we have **ONE SENTENCE OVERVIEW OF APPROACH ONCE DONE**. A performative model could impact the lives of millions of undiagnosed individuals who, with a correct diagnosis, could seek treatment for a debilitating and deadly disease.

This project is part of the 2025 George B. Moody PhysioNet Challenge. As such, we have leveraged the data made available through the challenge: CODE-15 [7] and SaMi-Trop. [8] CODE-15 and SaMi-Trop are comprised of 12-

lead ECG readings with associated meta-data indicating the age and gender of the patient, as well as labels for the positive or negative presence of Chagas. Both data sets were collected from Brazil between 2010 and 2016 (CODE-15) or 2010 through 2011 (SaMi-Trop), have a sampling frequency of 400 Hertz and are between 7.3 and 10.2 seconds long. In the challenge data there are over 343,424 samples in CODE-15 and 815 in SaMi-Trop. Unlike CODE-15, all SaMi-Trop records are positive and have been validated by serological tests.

2. Approach

Two primary challenges when working with the CODE-15 and SaMi-Trop data are computational costs and class imbalance. CODE-15 and SaMi-Trop together contain more than 100 gigabytes of data. 2.1% of the data belong to the positive class while belong to the 97.9% negative class. To address this, we built an analytical sample from all positive class instances and sampled an equal number of negative class instances.

We are aware of one known paper using deep neural networks to diagnose Chagas disease. Jidling et al. (2023) constructed a convolution neural network with residual connections an performed binary classification. However, convolutional neural networks on their own ignore the sequential nature of ECG data. Building off of this literature, we built three primary models to explore the data:

1. A convolutional neural network with residual connections

2. A bi-directional encoder transformer

3. A bi-directional encoder transformer stacked on top of a convolutional neural network with residual connections

For the first model we recreated the network used by Jidling et al (2023). [4] Recreating this network was necessary because we performed data curation to address the class imbalance and computational challenges presented with the SaMi-Trop and CODE-15 data. It would be challenging to ascertain whether differences in performance in our transformer and/or transformer plus convolution models were attributable to differences in architecture or differences in the underlying data.**MATT PLEASE BRIEFLY DESCRIBE - IF WE NEED SPACE WE CAN ALSO POINT THE READER TO THE ORIGINAL PAPER**.

Our second model uses a bidirectional transformer to diagnose Chagas disease and intentionally. **MARC PLEASE DESCRIBE, please add something on the classification token and how the model incorporates positional information**

Lastly, we combine these two models into one, so as to incorporate the advantages of each into a single model e.g. the feature extractive capabilities of convolution and the sequential advantages of transformers. The first section of the model consists of three parallel tracks of one-dimensional convolution blocks. The three tracks vary the kernel size to capture features constructed from different widths. [9] A single convolution block is a convolution layer with 32 filter, batch normalization, a Rectified Linear Unit, drop out, and a residual connection. Within each track the convolution blocks are stacked in three sequential layers. The output of these layers is projected into a dimension compatible with our transformer, down sampled by a factor of 2 to reduce computational complexity, and then fed into the bidirectional transformer encoder. As discussed above, we extract the classification token. Next we created age and sex embeddings and concatenated them with our classification token. This concatenated vector was fed through a fully connected linear layer with layer normalization, a rectified linear unit layer, and dropout. This structure allows ECG data to interact in non-linear ways with demographic data on patients. This is important as there are known differences in ECG data by age and sex. [10] The intent of this architecture was to use a 1-dimensional convolutional neural network to perform feature extraction and then to use a transformer to model the sequential nature of ECG data.

The full architecture can be seen in Figure 1. The loss function used was cross-entropy, data were zero-padded to sequence length of 4,096 (10.2 seconds at a sampling rate of 400 per second), and fed through a bandpass filter (0.5 and 100 for high and low pass filters respectively). Dropout was applied with a hyper parameter value of 0.1.

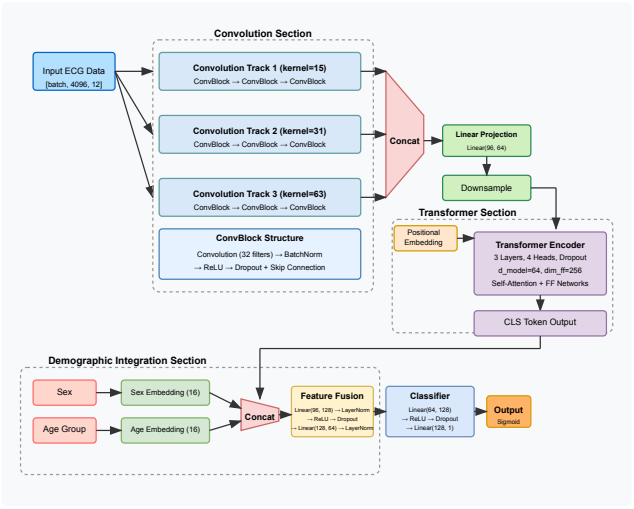

Figure 1. Convolution + Transformer Model: Architecture

Given the noted challenges with data set size and the

sparsity of positive labeled data, we attempted to augment the positive labeled data. Specifically, a random amount of time between 0 and 1 second from the front of the positive sequence was cropped, provided that the crop did not reduce the total length of the sequence to less than 7.3 seconds. Similarly, we cropped the positive-labeled sequences at the end of the sequence. This left three versions of a single positive sequence. We then sampled new negative labeled data to create a balanced data set. Moreover, we performed pre-training on negative labeled data from the PTB-XL dataset and performed full fine-tuning. For pre-training, we masked a continuous random half-second of the data and made the model predict the mask. For reasons discussed in the next section, we prefer the model trained on the unaugemented dataset with no pre-training.

To facilitate the Moody PhysioNet Challenge, an example repository with several helper functions was made available. From this repository, code to extract the Code-15 and Sami-Top data, as well as, some feature information contained within the header was used. We created the preprocessing code, the transformer code, the convolution plus transformer code, and the code to evaluate the performance of these models.

(10 points) What did you do exactly? How did you solve the problem? Why did you think it would be successful? Is anything new in your approach?

(5 points) What problems did you anticipate? What problems did you encounter? Did the very first thing you tried work?

MATT, see the note below, did we use code from here https://github.com/antonior92/automatic-ecg-diagnosis

Important: Mention any code repositories (with citations) or other sources that you used, and specifically what changes you made to them for your project.

3. Experiments and Results

MATT

MARC

Prior to settling on the model described in Figure 1 for the convolution plus transformer model, a number of models designed to reduce the computational complexity of the model were created. Our first model replaced the full convolution layers with depthwise separable convolutions, each convolution track was only two deep, and we omitted the integration of the age and sex characteristics. For all models discussed below, the performance metrics were calculated on a validation set representing 20% of the instances.

The first model achieved an AUROC of 0.8334. While adding the fully-connected layers necessary to incorporate age and sex data is computationally intensive, adding these data may significantly improve performance. Our second model iterated on the first by included the demographic integration section depicted in Figure 1. With the inclusion of these variables the AUROC validation performance jumped to 0.8449.

Additional models varied this second architecture slightly. For instance, we increased the depth of the convolution tracks to three and we increased the number of convolution track to four, with the fourth convolution track using a kernel size of 127. The architecture with a wider convolution showed signs of overfitting in the later epochs. Thus, we conducted an additional experiment where we increased the dropout rate from 0.1 to 0.2 and increased the number of training epochs to 20. However, this variation failed to produce continued improvements.

We focused on variations in the convolution architecture because the existing literature achieved relatively good results utilizing convolution only. For instance, Jidling et al., (2023) was able to achieve a 0.8 AUROC using only convolutions. However, all variations of architecture, with the exception of the model excluding age and sex, achieved AUROCs between 0.84 and 0.85.

In the end, our preferred model presented in Figure 1, achieved a validation AUROC of 0.8437, an accuracy of 75.5%, precision of 72.6%, and a recall of 83.1%. Table XX below presents the confusion matrix followed by Figure XX, the loss curves.

Actual \ Predicted	**Positive**	**Negative**
Positive	1240	252
Negative	469	976

Table 1. Convolution + Transformer: Validation Confusion Matrix

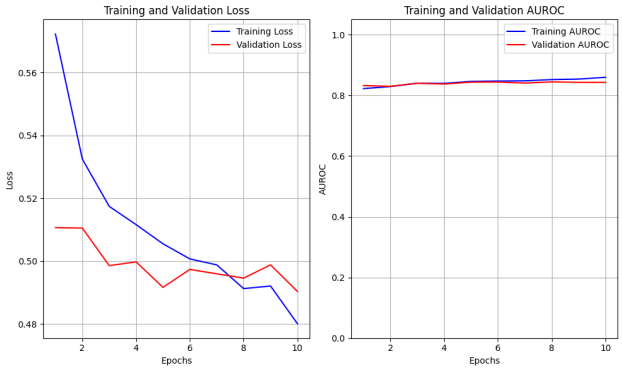

Figure 2. Convolution + Transformer Model: Training and Validation Metrics

Our experiments with different architectures revealed that near the 6th epoch the validation loss and the validation AUROC ceased to improve with more epochs. Combined with our findings on increasing the drop out rate and training epochs, this consistency across minor variations in

architectures indicated that further variations of our architecture were unlikely to yield significant gains. As such, we attempted to augment our model in two ways: data augmentation and pre-training.

Data augmentation and pre-training were described in the approach section; we augmented the positive labeled classes and sampled an equal number of negative instance and pre-trained on a continuous half second mask of the PTB-XL dataset. Note the PTB-XL data was not used in training other than pre-training. During fine-tuning with the augmented data we created IDs for each unaugmented original file. We sampled these IDs and all versions of a file were placed in either the training or validation sets, but not both. This sampling procedure ensure that there was no data leakage between the training and validations sets.

With a larger dataset we expected the model to learn more and present less chance of overfitting. We increased the complexity of the model by increasing the number of hidden dimensions in the transformer block to 128 and the number of encoder layers to 4. We utilized a fourth parallel convolution track with the additional track having a kernel size of 127.

At first glance, the full-fine tuned pre-trained model trained on augmented data performed significantly better than our preferred model achieving a 0.8736 AUROC on the validation set. However, the model exhibited undesirable behavior. Specifically, the recall drops precipitously to 66.1

Actual \ Predicted	**Positive**	**Negative**
Positive	2888	1481
Negative	7	4394

Table 2. Augmented and pre-trained Convolution + Transformer: Validation Confusion Matrix

The high rate of false negatives is particularly concerning for a medical diagnostic model. Utilization of a model such as this would mean many people with Chagas disease are being told they do not have the disease. A potential remedy would be to lower the classification threshold of 0.5. However, this would likely result in many more false positives. This could induce unnecessary, costly, and invasive serological testing. For these reasons, our preferred model is that depicted in Figure 1.

The pre-trained model with augmented data was intended to address the paucity of positive labeled instances. To that end, pre-training ECG models has received notable recent attention.[11][12] Future work on diagnosing Chagas disease with deep neural networks should invest in leveraging these advancements in pre-training to address the scarcity of labeled data.

.

(10 points) How did you measure success? What experiments were used? What were the results, both quantitative and qualitative? Did you succeed? Did you fail? Why? Justify your reasons with arguments supported by evidence and data.

Important: This section should be rigorous and thorough. Present detailed information about decision you made, why you made them, and any evidence/experimentation to back them up. This is especially true if you leveraged existing architectures, pre-trained models, and code (i.e. do not just show results of fine-tuning a pre-trained model without any analysis, claims/evidence, and conclusions, as that tends to not make a strong project).

4. Other Sections

You are welcome to introduce additional sections or subsections, if required, to address the following questions in detail.

(5 points) Appropriate use of figures / tables / visualizations. Are the ideas presented with appropriate illustration? Are the results presented clearly; are the important differences illustrated?

(5 points) Overall clarity. Is the manuscript self-contained? Can a peer who has also taken Deep Learning understand all of the points addressed above? Is sufficient detail provided?

(5 points) Finally, points will be distributed based on your understanding of how your project relates to Deep Learning. Here are some questions to think about:

What was the structure of your problem? How did the structure of your model reflect the structure of your problem?

What parts of your model had learned parameters (e.g., convolution layers) and what parts did not (e.g., post-processing classifier probabilities into decisions)?

What representations of input and output did the neural network expect? How was the data pre/post-processed? What was the loss function?

Did the model overfit? How well did the approach generalize?

What hyperparameters did the model have? How were they chosen? How did they affect performance? What optimizer was used?

What Deep Learning framework did you use?

What existing code or models did you start with and what did those starting points provide?

Briefly discuss potential future work that the research community could focus on to make improvements in the direction of your project's topic.

Student Name	Contributed Aspects	Details
Marc Lafargue 1	Data Creation and Implementation	Scraped the dataset for this project and trained the CNN of the encoder. Implemented attention mechanism to improve results.
Matheaus Rama Amorim	Implementation and Analysis	Trained the LSTM of the encoder and analyzed the results. Analyzed effect of number of nodes in hidden state. Implemented Convolutional LSTM.
Trevor Gratz	Convolution plus Transformer Model	Trained the convolution plus transformer model, conducted the experiments on these models, and wrote up the results

Table 3. Contributions of team members.

5. Work Division

Please add a section on the delegation of work among team members at the end of the report, in the form of a table and paragraph description. This and references do **NOT** count towards your page limit. An example has been provided in Table 3.

6. Miscellaneous Information

The rest of the information in this format template has been adapted from CVPR 2020 and provides guidelines on the lower-level specifications regarding the paper's format.

6.1. Language

All manuscripts must be in English.

6.2. Paper length

Papers, excluding the references section, must be no longer than six pages in length. The references section will not be included in the page count, and there is no limit on the length of the references section. For example, a paper of six pages with two pages of references would have a total length of 8 pages.

6.3. The ruler

The LaTeX style defines a printed ruler which should be present in the version submitted for review. The ruler is provided in order that reviewers may comment on particular lines in the paper without circumlocution. If you are preparing a document using a non-LaTeX document preparation system, please arrange for an equivalent ruler to appear on the final output pages. The presence or absence of the ruler should not change the appearance of any other content on the page. The camera ready copy should not contain a ruler. (LaTeX users may uncomment the `\cvprfinalcopy` command in the document preamble.) Reviewers: note that the ruler measurements do not align well with lines in the paper — this turns out to be very difficult to do well when the paper contains many figures and equations, and, when done, looks ugly. Just use fractional references (e.g. this line is 095.5), although in most cases one would expect that the approximate location will be adequate.

6.4. Mathematics

Please number all of your sections and displayed equations. It is important for readers to be able to refer to any

particular equation. Just because you didn't refer to it in the text doesn't mean some future reader might not need to refer to it. It is cumbersome to have to use circumlocutions like "the equation second from the top of page 3 column 1". (Note that the ruler will not be present in the final copy, so is not an alternative to equation numbers). All authors will benefit from reading Mermin's description of how to write mathematics: http://www.pamitc.org/documents/mermin.pdf.

Finally, you may feel you need to tell the reader that more details can be found elsewhere, and refer them to a technical report. For conference submissions, the paper must stand on its own, and not *require* the reviewer to go to a techreport for further details. Thus, you may say in the body of the paper "further details may be found in [13]". Then submit the techreport as additional material. Again, you may not assume the reviewers will read this material.

Sometimes your paper is about a problem which you tested using a tool which is widely known to be restricted to a single institution. For example, let's say it's 1969, you have solved a key problem on the Apollo lander, and you believe that the CVPR70 audience would like to hear about your solution. The work is a development of your celebrated 1968 paper entitled "Zero-g frobnication: How being the only people in the world with access to the Apollo lander source code makes us a wow at parties", by Zeus *et al*.

You can handle this paper like any other. Don't write "We show how to improve our previous work [Anonymous, 1968]. This time we tested the algorithm on a lunar lander [name of lander removed for blind review]". That would be silly, and would immediately identify the authors. Instead write the following:

> We describe a system for zero-g frobnication. This system is new because it handles the following cases: A, B. Previous systems [Zeus et al. 1968] didn't handle case B properly. Ours handles it by including a foo term in the bar integral.
>
> ...
>
> The proposed system was integrated with the Apollo lunar lander, and went all the way to the moon, don't you know. It displayed the following behaviours which show how well we solved cases A and B: ...

As you can see, the above text follows standard scientific convention, reads better than the first version, and does not explicitly name you as the authors. A reviewer might think it likely that the new paper was written by Zeus *et al*., but cannot make any decision based on that guess. He or she would have to be sure that no other authors could have been contracted to solve problem B.

FAQ

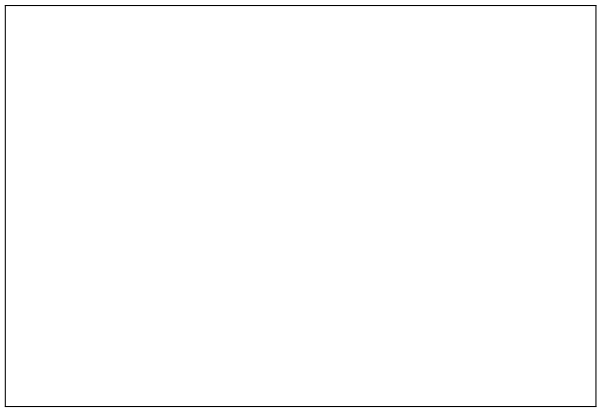

Figure 3. Example of caption. It is set in Roman so that mathematics (always set in Roman: $B \sin A = A \sin B$) may be included without an ugly clash.

Q: Are acknowledgements OK?
A: No. Leave them for the final copy.

Q: How do I cite my results reported in open challenges?
A: To conform with the double blind review policy, you can report results of other challenge participants together with your results in your paper. For your results, however, you should not identify yourself and should not mention your participation in the challenge. Instead present your results referring to the method proposed in your paper and draw conclusions based on the experimental comparison to other results.

6.5. Miscellaneous

Compare the following:

`$conf_a$`	$conf_a$
`$\mathit{conf}_a$`	conf_a

See The TEXbook, p165.

The space after *e.g.*, meaning "for example", should not be a sentence-ending space. So *e.g.* is correct, *e.g.* is not. The provided `\eg` macro takes care of this.

When citing a multi-author paper, you may save space by using "et alia", shortened to "*et al*." (not "*et. al.*" as "*et*" is a complete word.) However, use it only when there are three or more authors. Thus, the following is correct: " Frobnication has been trendy lately. It was introduced by Alpher [14], and subsequently developed by Alpher and Fotheringham-Smythe [15], and Alpher *et al*. [16]."

This is incorrect: "... subsequently developed by Alpher *et al*. [15] ..." because reference [15] has just two authors. If you use the `\etal` macro provided, then you need not worry about double periods when used at the end of a sentence as in Alpher *et al*.

For this citation style, keep multiple citations in numerical (not chronological) order, so prefer [15, 14, 17] to

[14, 15, 17].

6.6. Formatting your paper

All text must be in a two-column format. The total allowable width of the text area is $6\frac{7}{8}$ inches (17.5 cm) wide by $8\frac{7}{8}$ inches (22.54 cm) high. Columns are to be $3\frac{1}{4}$ inches (8.25 cm) wide, with a $\frac{5}{16}$ inch (0.8 cm) space between them. The main title (on the first page) should begin 1.0 inch (2.54 cm) from the top edge of the page. The second and following pages should begin 1.0 inch (2.54 cm) from the top edge. On all pages, the bottom margin should be 1-1/8 inches (2.86 cm) from the bottom edge of the page for 8.5×11-inch paper; for A4 paper, approximately 1-5/8 inches (4.13 cm) from the bottom edge of the page.

6.7. Margins and page numbering

All printed material, including text, illustrations, and charts, must be kept within a print area 6-7/8 inches (17.5 cm) wide by 8-7/8 inches (22.54 cm) high.

6.8. Type-style and fonts

Wherever Times is specified, Times Roman may also be used. If neither is available on your word processor, please use the font closest in appearance to Times to which you have access.

MAIN TITLE. Center the title 1-3/8 inches (3.49 cm) from the top edge of the first page. The title should be in Times 14-point, boldface type. Capitalize the first letter of nouns, pronouns, verbs, adjectives, and adverbs; do not capitalize articles, coordinate conjunctions, or prepositions (unless the title begins with such a word). Leave two blank lines after the title.

AUTHOR NAME(s) and AFFILIATION(s) are to be centered beneath the title and printed in Times 12-point, non-boldface type. This information is to be followed by two blank lines.

The ABSTRACT and MAIN TEXT are to be in a two-column format.

MAIN TEXT. Type main text in 10-point Times, single-spaced. Do NOT use double-spacing. All paragraphs should be indented 1 pica (approx. 1/6 inch or 0.422 cm). Make sure your text is fully justified—that is, flush left and flush right. Please do not place any additional blank lines between paragraphs.

Figure and table captions should be 9-point Roman type as in Figures 3 and 4. Short captions should be centred.

Callouts should be 9-point Helvetica, non-boldface type. Initially capitalize only the first word of section titles and first-, second-, and third-order headings.

FIRST-ORDER HEADINGS. (For example, **1. Introduction**) should be Times 12-point boldface, initially capitalized, flush left, with one blank line before, and one blank line after.

Method	Frobnability
Theirs	Frumpy
Yours	Frobbly
Ours	Makes one's heart Frob

Table 4. Results. Ours is better.

SECOND-ORDER HEADINGS. (For example, **1.1. Database elements**) should be Times 11-point boldface, initially capitalized, flush left, with one blank line before, and one after. If you require a third-order heading (we discourage it), use 10-point Times, boldface, initially capitalized, flush left, preceded by one blank line, followed by a period and your text on the same line.

6.9. Footnotes

Please use footnotes[1] sparingly. Indeed, try to avoid footnotes altogether and include necessary peripheral observations in the text (within parentheses, if you prefer, as in this sentence). If you wish to use a footnote, place it at the bottom of the column on the page on which it is referenced. Use Times 8-point type, single-spaced.

6.10. References

List and number all bibliographical references in 9-point Times, single-spaced, at the end of your paper. When referenced in the text, enclose the citation number in square brackets, for example [17]. Where appropriate, include the name(s) of editors of referenced books.

6.11. Illustrations, graphs, and photographs

All graphics should be centered. Please ensure that any point you wish to make is resolvable in a printed copy of the paper. Resize fonts in figures to match the font in the body text, and choose line widths which render effectively in print. Many readers (and reviewers), even of an electronic copy, will choose to print your paper in order to read it. You cannot insist that they do otherwise, and therefore must not assume that they can zoom in to see tiny details on a graphic.

When placing figures in LaTeX, it's almost always best to use `\includegraphics`, and to specify the figure width as a multiple of the line width as in the example below

```
\usepackage[dvips]{graphicx} ...
\includegraphics[width=0.8\linewidth]
                {myfile.eps}
```

6.12. Color

Please refer to the author guidelines on the CVPR 2020 web page for a discussion of the use of color in your document.

[1] This is what a footnote looks like. It often distracts the reader from the main flow of the argument.

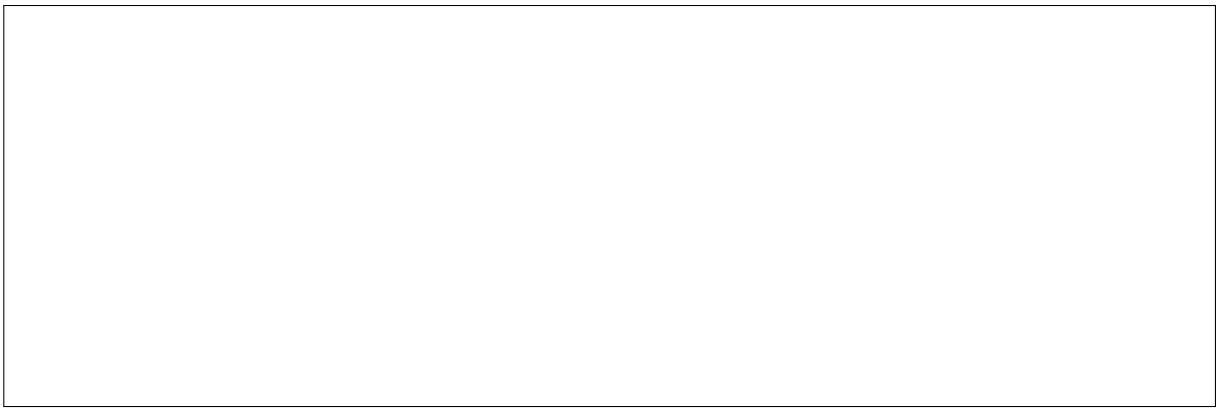

Figure 4. Example of a short caption, which should be centered.

References

[1] Zulma M Cucunubá, Sebastián A Gutiérrez-Romero, Juan-David Ramírez, Natalia Velásquez-Ortiz, Soledad Ceccarelli, Gabriel Parra-Henao, Andrés F Henao-Martínez, Jorge Rabinovich, María-Gloria Basáñez, Pierre Nouvellet, et al. The epidemiology of chagas disease in the americas. *The Lancet Regional Health–Americas*, 37, 2024. 1

[2] Francisco Rogerlândio Martins-Melo, Marcia C Castro, and Guilherme Loureiro Werneck. Levels and trends in chagas disease-related mortality in brazil, 2000–2019. *Acta Tropica*, 220:105948, 2021. 1

[3] Maria Carmo P Nunes, Caryn Bern, Eva H Clark, Antonio L Teixeira, and Israel Molina. Clinical features of chagas disease progression and severity. *The Lancet Regional Health–Americas*, 37, 2024. 1

[4] Carl Jidling, Daniel Gedon, Thomas B Schön, Claudia Di Lorenzo Oliveira, Clareci Silva Cardoso, Ariela Mota Ferreira, Luana Giatti, Sandhi Maria Barreto, Ester C Sabino, Antonio LP Ribeiro, et al. Screening for chagas disease from the electrocardiogram using a deep neural network. *PLoS Neglected Tropical Diseases*, 17(7):e0011118, 2023. 1, 2

[5] Xinwen Liu, Huan Wang, Zongjin Li, and Lang Qin. Deep learning in ecg diagnosis: A review. *Knowledge-Based Systems*, 227:107187, 2021. 1

[6] Chao Che, Peiliang Zhang, Min Zhu, Yue Qu, and Bo Jin. Constrained transformer network for ecg signal processing and arrhythmia classification. *BMC Medical Informatics and Decision Making*, 21(1):184, 2021. 1

[7] Antônio H Ribeiro, Manoel Horta Ribeiro, Gabriela MM Paixão, Derick M Oliveira, Paulo R Gomes, Jéssica A Canazart, Milton PS Ferreira, Carl R Andersson, Peter W Macfarlane, Wagner Meira Jr, et al. Automatic diagnosis of the 12-lead ecg using a deep neural network. *Nature communications*, 11(1):1760, 2020. 1

[8] Clareci Silva Cardoso, Ester Cerdeira Sabino, Claudia Di Lorenzo Oliveira, Lea Campos de Oliveira, Ariela Mota Ferreira, Edécio Cunha-Neto, Ana Luiza Bierrenbach, João Eduardo Ferreira, Desirée Sant'Ana Haikal, Arthur L Reingold, et al. Longitudinal study of patients with chronic chagas cardiomyopathy in brazil (sami-trop project): a cohort profile. *BMJ open*, 6(5):e011181, 2016. 1

[9] Christian Szegedy, Wei Liu, Yangqing Jia, Pierre Sermanet, Scott Reed, Dragomir Anguelov, Dumitru Erhan, Vincent Vanhoucke, and Andrew Rabinovich. Going deeper with convolutions. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 1–9, 2015. 2

[10] Peter W Macfarlane. The influence of age and sex on the electrocardiogram. *Sex-Specific Analysis of Cardiovascular Function*, pages 93–106, 2018. 2

[11] Jessica Y Bo, Hen-Wei Huang, Alvin Chan, and Giovanni Traverso. Pretraining ecg data with adversarial masking improves model generalizability for data-scarce tasks. *arXiv preprint arXiv:2211.07889*, 2022. 4

[12] Temesgen Mehari and Nils Strodthoff. Self-supervised representation learning from 12-lead ecg data. *Computers in biology and medicine*, 141:105114, 2022. 4

[13] Authors. Frobnication tutorial, 2014. Supplied as additional material `tr.pdf`. 6

[14] FirstName Alpher. Frobnication. *Journal of Foo*, 12(1):234–778, 2002. 6, 7

[15] FirstName Alpher and FirstName Fotheringham-Smythe. Frobnication revisited. *Journal of Foo*, 13(1):234–778, 2003. 6, 7

[16] FirstName Alpher, FirstName Fotheringham-Smythe, and FirstName Gamow. Can a machine frobnicate? *Journal of Foo*, 14(1):234–778, 2004. 6

[17] Authors. The frobnicatable foo filter, 2014. Face and Gesture submission ID 324. Supplied as additional material `fg324.pdf`. 6, 7