[Re] Diffusion-Based Adversarial Sample Generation for Improved Stealthiness and Controllability

William Kang,
syu@student.ubc.ca

Christina Yang
chryang@student.ubc.ca

Abstract

We are doing a reproducibility report based on the paper "Diffusion-Based Adversarial Sample Generation for Improved Stealthiness and Controllability" by Haotian Xue, Alexandre Araujo, Bin Hu, Yongxin Chen. Their Github Repo is here: `https://github.com/xavihart/Diff-PGD/tree/main`.

The paper uses a novel framework to generate adversarial samples. They use a gradient based method guided by a pre-trained diffusion model to try to generate images that appear realistic to the human eye, can fool a wide range of models, and is easy to control how certain regions are modified.

1 Introduction

This is why you should care about our project.

Main paper content can be **up to six pages**, followed by an unlimited number of pages containing references and appendices (which the course staff, like NeurIPS reviewers, won't *necessarily* read).

2 What goes in the report

Part of your grade is based on the contribution of your project: did you do something meaningful in your project, that adds something new to the world? Again, we don't expect you to necessarily make a high-impact totally novel result, and it's fine if you end up with a negative result, but you should have something here. You should make sure that the contribution of your project is clear in the introduction of your paper, and that the rest of the writeup clearly demonstrates that you actually made this contribution.

A large portion of your grade is based on writing a clear and structured paper that addresses the relevant questions that would be asked of a generic scientific/engineering publication. Your writeup should look more or less like a scientific paper as published at, say, NeurIPS, or a NeurIPS workshop. To achieve that, most papers should use something like the following traditional outline:

1. Introduction: Clearly state the problem being addressed. **Explain why it is an important problem to work on**. At a high level, briefly summarize what the limitations of existing approaches that your work will be addressing, and what the contribution of your project is.
2. Related Work: Identify at least three publications on related topics; usually these will be papers that have worked on slightly different problems or papers that have proposed an approach to your problem that is not fully satisfactory. For each paper, briefly say either how the problem addressed is related and/or different (if they address a different problem) or why it doesn't solve the problem you are working on (if it addresses the same problem).
3. Description and justification of what you did, divided up (possibly in multiple sections and/or sub-sections). There's a lot of flexibility here, and it will depend on the type of

Course project for UBC CPSC 440/550, April 2024.

project you are doing. For example, if you're applying standard machine learning methods to a new dataset or doing a Kaggle competition, you could have one (sub-)section describing the dataset and why you think machine learning could help, and one (sub-)section stating the methods you will try and why you think these are appropriate methods (you don't necessarily have to go into detail describing the methods). If you're extending an existing technique, you could have one sub-section describing the existing technique, and one sub-section for each of the extensions you explored. If your project has a theoretical component, you might have one sub-section discussing the assumptions, one sub-section describing the results, and one sub-section describing implications.

4. Experiments and/or analysis (if you have an experimental component): Describe each experiment that you did. Say what each experiment is trying to test. Ideally, each experiment should only try to test one thing and you should control for as many other factors as possible. Subsequently, summarize the result of your experiment, in both the text and in a nice visual form such as a figure; in most cases, **a table with a huge list of numbers is not a nice way to summarize information**.
5. Discussion and future work: State the main conclusions that are obtained from this course project. List at least one strength and one weakness of your contribution. Briefly state what you would do with more time.

An example outline for a perspective paper might be something like

1. Introduction: Clearly state the problem being addressed. Explain why it is an important problem. At a high level, briefly summarize the history of the works that will be discussed in the project.
2. Review: Go through the different works in some logical order, such as chronologically or by going from simple to complex models. Don't just list the methods, but say how they relate to each other (going through the strengths/weaknesses of the different methods, both in comparison to each other and compared to an ideal method that solves the problem).
3. Discussion: Discuss the trends that have occurred over time. Speculate about where the next steps in the trend could lead. Point out issues that are not properly addressed by existing methods. State some interesting directions to explore, or opportunities to use existing tools in new applications.

You don't have to stick exactly to these structures; many of my papers don't. But if you're not using this format, you should make sure what you're doing makes sense, and answers the important questions about your project.

Writing style Mark has some advice for writing here, most of which I agree with: `https://www.cs.ubc.ca/~schmidtm/Courses/Notes/writing.pdf`. (He has a bunch of minor grammatical errors in that document, though. :/)

You're not going to lose points for making some minor grammatical mistakes or anything like that – the published literature is full of them! But you will lose points if it's to the point of making your work hard to understand. Clear, easy-to-understand papers are far more likely to succeed; note that this involves a lot more than just having a good grasp of English.

3 How to format stuff

3.1 Citations

Something a lot of people don't know at first: use `\citet{ref}` (or `\textcite`) if the citation plays a grammatical role in the sentence, e.g. "Vaswani et al. (2017) demonstrated that ...." Use `\citep{ref}` (or `\parencite`) if it doesn't, e.g. "Machine learning is fun (Schmidt and Hüber 1832)."

3.2 Figures

See Figure 1 for how to include a figure.

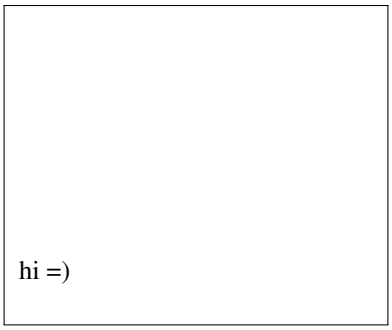

Figure 1: Sample figure caption.

Table 1: Sample table title

	Part	
Name	Description	Size (μm)
Dendrite	Input terminal	$\sim$100
Axon	Output terminal	$\sim$10
Soma	Cell body	up to 10^6

3.3 Tables

Table captions go *above* the table, because that's the usual style, idk. There's an example in Table 1. Avoid vertical rules.

3.4 Math

Note that display math in bare TeX commands will not create correct line numbers for submission. Please use LaTeX (or AMSTeX) commands for unnumbered display math. (You really shouldn't be using $$ anyway; see `https://tex.stackexchange.com/questions/503/why-is-preferable-to` and `https://tex.stackexchange.com/questions/40492/what-are-the-differences-between-align-equation-and-displaymath` for more information.)

3.5 Supplementary Material

You can include extra information in appendices, like Appendix A. You don't have to if you don't want to.

Please *don't* include code in your writeups, unless you did something particularly cool and want to *briefly* describe the way something works as a contribution (this should be relatively unusual). Please instead link to it somewhere, preferably a GitHub repo or similar, especially if the code is a significant contribution of your project.

4 Discussion

In the end, you should give us an A.

Acknowledgments

You can acknowledge useful discussion with other people who aren't coauthors here (or leave the section out). Typically you'd also put funding here, acknowledgements for computing clusters, etc.

References

Schmidt, Yourgan and Mygan Hüber (1832). *Machine Learning is Fun*.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin (2017). "Attention is All you Need." *Advances in Neural Information Processing Systems*.

A Supplementary material

This stuff doesn't count towards your page limit.