The Thirty-Fifth AAAI Conference on Artificial Intelligence (AAAI-21)

What are GANs?: Introducing Generative Adversarial Networks to Middle School Students

Safinah Ali*, **Daniella DiPaola***, **and Cynthia Breazeal**
MIT Media Lab
Cambridge, Massachusetts, 02139
{safinah, dipaola, cynthiab}@media.mit.edu

Abstract

Applications of Generative Machine Learning techniques such as Generative Adversarial Networks (GANs) are used to generate new instances of images, music, text, and videos. While GANs have now become commonplace on social media, a part of children's lives, and have considerable ethical implications, existing K-12 AI education curricula do not include generative AI. We present a new module, "What are GANs?", that teaches middle school students how GANs work and how they can create media using GANs. We developed an online, team-based game to simulate how GANs work. Students also interacted with up to four web tools that apply GANs to generate media. This module was piloted with 72 middle school students in a series of online workshops. We provide insight into student usage, understanding, and attitudes towards this lesson. Finally, we give suggestions for integrating this lesson into AI education curricula.

Introduction

In the summer of 2020, middle school student Jessica[1] started using the social media applications Instagram, Snapchat and TikTok. TikTok, like several other social media platforms, makes use of user demographics and interests to display advertisements. While browsing through TikTok, Jessica received an advertisement for "Reface: Swap your faces now", which piqued her interest. After downloading the app, Jessica uploaded a selfie and created her first video: Hermione Granger casting a spell, but with Jessica's face. The Reface app makes use of a generative machine learning technique called Generative Adversarial Networks (or GANs) to swap faces on popular media (Lomas 2020). GANs are generative models: they create new data instances of data that resemble your training data (Goodfellow et al. 2014a). GANs can be used to transfer the style of one kind of media (such as a photograph) onto another. The GAN in this application is used to generate a Deepfake, which is synthetic media in which a person in an existing image or video is replaced with someone else's likeness. Like Deepfakes, applications of GANs such as style transfer in face filters, generative music, and generative text have now become both features and content within social media applications.

Even though applications of GANs are already commonplace in children's lives, there aren't enough frameworks to teach children about how GANs work and their use in the real world. This is especially concerning because GANs are accompanied by several ethical concerns such as the generation of fake media through Deepfakes which can be used in malicious ways, the spread of misinformation, data collection and distribution, bias in datasets, and the ownership of machine generated art. Further, GAN-created media is often hyper-realistic and difficult to tell apart from real media. This might lead to children not being able to tell fake media apart from real media.

While there are several efforts to teach middle school students about AI (Ali et al. 2019), not many K-12 AI Education efforts focus on generative machine learning techniques such as GANs. In this work, we describe a lesson to introduce middle school students to GANs. The goal of the lesson was for students to conceptualize how the two parts of a GAN (generator and discriminator) work in opposition with one another to generate new data instances. The first part of the lesson utilizes a web-based interactive tool to teach students how GANs work. Students use a web-tool and take on the generator and discriminator roles to create pixel art and give feedback. The goal of the game is to have students understand the role of the generator and the discriminator, and how they work in opposition in a GAN. The second part of the lesson invites students to explore four existing generative AI web tools. Through the full lesson, students are able to apply their knowledge of how GANs work to existing web tools that utilize this technology.

Background

In this work, we develop a middle school AI Education lesson focusing on the theory, practice and ethical implications of Generative Adversarial Networks.

Generative Adversarial Networks (or GANs)

First introduced by Ian Goodfellow in 2014, Generative Adversarial Networks (GANs) are a new kind of generative machine learning model (Goodfellow et al. 2014b). GANs create new data instances of data that resemble training data.

*Equal contribution.

[1]Name changed

For example, GANs can create images that look like photographs of human faces, even though the faces don't belong to any real person. GANs are used to create many different kinds of media, such as visual art, photographs, music, videos, and text. They are able to produce realistic new data by pairing a generator, which learns to produce the target output, with a discriminator, which learns to distinguish true data from the output of the generator. The generator tries to fool the discriminator, and the discriminator tries to keep from being fooled. The generator and the discriminator work against one another to ultimately produce new data that replicates patterns in the training data. GANs have found several applications in art, entertainment, healthcare and robotics. However, GANs also enable the production of Deepfakes, or fake media in which a person in an existing image or video is replaced with someone else's likeness (Kim et al. 2018). Hence, GANs raise some ethical concerns such as the production and circulation of fake media and its implications as well as the ownership and the environmental cost of training big AI models (Strubell, Ganesh, and McCallum 2019).

Teens, Social Media, and GANs

Students are already exposed to the applications of generative machine learning techniques. Children in middle school are active consumers and creators of social media content. Children are familiar with apps such as Snapchat and Instagram that make use of photo filters that often use generative machine learning techniques (Anderson, Jiang et al. 2018). Many students acquire their first personal mobile device during middle school, which is when they start consuming data on social media websites such as Twitter and Facebook, where there can be GAN-generated content (Brandon 2019). Further, children also upload personal data such as images, videos and text on these sites that can then be used as a part of the data sets used to train these models. Given the presence of GAN generated media on social networking applications that children frequent, they are likely to be exposed these generative media.

AI Education

In 2019, AI4K12 released a paper defining the 5 "Big Ideas" of AI education: Perception, Representation & Reasoning, Learning, Natural Interaction, and Societal Impact (Touretzky et al. 2019). Many curricula have been developed to address these big ideas. Payne (2020) surveyed some of the most popular AI education curricula for K-12. Of these, the most common topics covered were neural networks, machine learning, and perception (Payne 2020). Though these topics are integrated into generative AI, no curriculum addressed GANs directly. There exist several collegiate Generative Machine Learning curricula that especially focus on generative art. For instance, ML4A has several courses focusing on the artistic applications of machine learning, such as, The Neural Aesthetic or Machine Learning for Artists[2]. AMI is a program at Google that brings artists and engineers together to realize projects using Machine Intelligence[3]. There are also several graduate level courses that introduce students to training GANs and their applications (Lab 2019; Renaud Danhaive 2019). All of these courses and activities require prior programming experience and a sophisticated computing set-up, often requiring cloud computing or GPU capabilities.

AI literacy focused on generative machine learning techniques such as GANs is imperative for students to be informed about media they encounter online, such as, being able to spot fake data. This curriculum introduces students to GANs, how they work, what kind of data they use, and their different applications. Students also think critically about the benefits and harms of creating and using GANs.

In our work, we designed a web-based interaction that abstracted the ideas of a generator and a discriminator and the roles they play in a GAN to generate synthetic data into a game-based interaction suited for young children. We designed a competitive web game that can be played on any browser and does not require sophisticated computational setup. The game is played by two teams, whose roles are analogous to generators and discriminators, and effectively conveys how GANs work through this role-play interaction.

Target Age Group

Given the importance of teaching generative machine learning to middle school students, this lesson plan is designed for 5th to 8th grade students who are fluent in English with little to no experience in programming. The activity is designed for synchronous remote learning. In order to make it accessible to students, we designed a web-based synchronous interactive tool that is functional on all mainstream browsers. Further, we simplified ideas of probability, neural networks and back propagation, using metaphors where appropriate to make the concepts more understandable.

Our workshop and lessons were designed for middle school students with no previous knowledge of artificial intelligence, machine learning, or neural networks. We took an abstraction approach to teach a complex concept in a simpler understandable manner. Originating in Mathematics Education, the reducing abstraction framework shows how learners attempt to find, either explicitly or implicitly, less abstract forms of their problems to aid their understanding (Hazzan 1999). We reduced the concept of a generator as generating new media to creating images in a small pixel grid, and the concept of discriminators providing feedback in terms of weights or probabilities, as providing positive/negative feedback using green or red markers. We also made use of the experiential learning methodology, or learning through reflection on doing, where students take the role of a generator and discriminator and enact what each neural network does through a game interaction. We structured the interaction as competitive game play between players acting as the generators and the discriminators, which was not only beneficial for engagement, but also accurately depicted the roles of a generator and a discriminator, which work in opposition in a GAN.

[2]See more at https://ml4a.github.io/classes/

[3]See more at https://ami.withgoogle.com/

Concepts Addressed and Expected Learning Outcomes

This lesson provides students with a conceptual framework for how a GAN works. Students learn about the generator and discriminator, the two neural networks that work against one another to generate something new. Students are given examples of deployed GANs and are asked to look at them through the lens of a generator and discriminator. At the end of the activity, students will be able to:

- Describe the relationship between the generator and the discriminator, the two neural networks that make up a GAN.
- For a specific GAN, identify what the generator is trying to create.
- For a specific GAN, identify the dataset that the discriminator is making its decisions off of.

Curriculum Design

The lesson plan constitutes an introduction to GANs, followed by 2 activities: (1) *How do GANs work?* and (2) *Exploring GANs*. This is a 2-hour lesson designed for synchronous online learning. The lesson plan was taught using Google Slides (Introduction), a web-based interactive game (Activity 1), and 4 interactive tool websites (Activity 2).

Introduction to GANs

The lesson begins by introducing GANs as a type of AI technology that is used to generate new data in contrast with AI that classifies data. We explained the terms: Generative as "to create", Adversarial as "to oppose" and Networks as "Neural Networks", a kind of AI algorithm. We then introduce how the two networks in a GAN are called generator and discriminator. We explain the roles of generators and discriminators using the following analogy:

Let's simplify this a bit and pretend that you are a generator and your art teacher is a discriminator.

Your art teacher would tell you that you need to create a Picasso painting. If you do it correctly, you won't have homework for a week, but if you don't, you get double the homework. The only problem is that you have never seen a Picasso painting before, but your teacher has seen many. So for your first try, you must come up with a random painting, and hope your teacher thinks it is a Picasso.

Of course, you don't get it right the first time. Your teacher gives you feedback though, like "more color" and "more edges" and "fewer circles". You take the feedback and keep making changes to your painting. Each time you come back to your teacher, you feel like you are getting closer to creating a Picasso painting. Finally, your teacher believes that you've actually created a real Picasso! Her feedback has helped you become so good at imitating Picasso's style that she can't tell the difference between your painting and one that Picasso made. You get a full week of no homework for being a good generator!

In short, a generator has to come up with something new with limited feedback. Over time, it gets better and better. The discriminator knows what that new thing should be, and won't let the generator "pass" until they come up with something that fits those requirements.

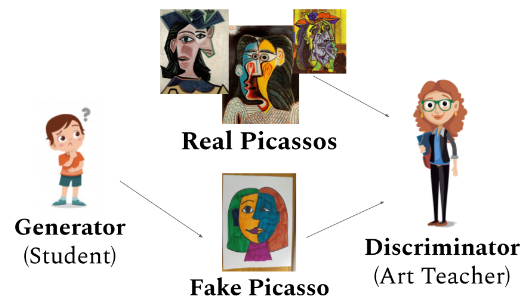

Figure 1: A visual explanation of the art teacher generator/discriminator analogy.

Following the introduction, students played an interactive web-based game that explored the roles of the generator and discriminator.

Activity 1, How do GANs work?

After students understand that GANs are used to generate media, they are introduced to how GANs work. Students learn that a GAN is made up of two neural networks, one called the generator, and one called the discriminator. They are told that the generator and discriminator have two different goals that are in competition with one another:

- The goal of the generator is to create something new that the discriminator will classify as "real".
- The goal of the discriminator is to detect if what the generator creates is "fake".

Then, the students break up into a "generator" group and a "discriminator" group using Zoom breakout rooms. The generators are given a 6x6 grid and told that they need to insert 7-9 squares into the grid to create an arrangement of blocks that passes by the discriminator. The discriminators are given a dataset of images (Figure 2), and told that they must accept images that resemble the images they were given. For this dataset, we chose pixelated faces, because they could be configured in many different ways and were common enough that students would recognize them, allowing students to create their own mental models of the types of images that would "pass" without being too prescriptive.

The game begins when the generators send over a configuration of 7-9 blocks to the discriminator. The discriminator team then has to decide whether or not the configuration should "pass". If it does not pass, then they must give the generator team feedback by giving them two blocks that are correct (marked in green) and two blocks that are incorrect (marked in red). The generator then gets another chance to produce a configuration of squares that will pass through the discriminator. The process continues until the generator produces a configuration that the discriminator determines fits with the rest of the data. An example of the back and forth from this game can be found in Figure 3.

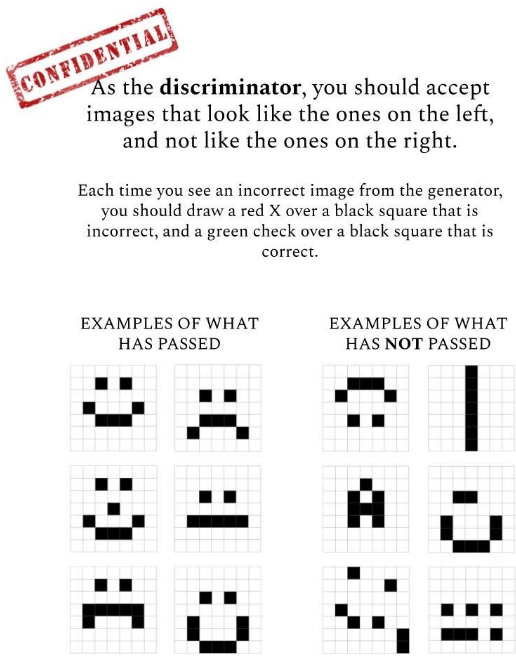

Figure 2: The data set given to the discriminator team

.

After the activity, students are told that a GAN goes through this process, but many more times to create much more detailed pieces of media. Since we taught this workshop remotely, the game was designed on a web-interface, and the facilitators from each group screen-share the game interface, while the entire team works together to make decisions. We made use of Javascript and Socket.io, which is a web-socket library, to develop this tool. The generator and discriminator screens are two clients of the application. The tool updates live as generators mark their input boxes. After they are done, they click on "Submit" and the discriminators mark two green and two red squares to provide positive and negative feedback respectively. The update live on the discriminator screen. When they are done, they click on "Submit". The feedback from the discriminator fades into light-green and light-red boxes, and the generators can now use this feedback to attempt their next round. The teams play the game until the generators get it correct. The discriminators then click on "You got it", and the game ends. On the back-end, we log in a 6x6 matrix the squares each team marks in every round.

Activity #2, Exploring GANs

After students understand how GANs work, they are asked to explore some interactive web-based tools that use GANs to create media. There were four GANs to explore:

- **AI Duet:** Built by Yotam Mann and Google, this web tool utilizes generative piano music to let users play a duet with the computer (Mann 2016). Users press keys to play a music note, and AI Duet adds some notes to form a duet. The tool utilizes Tensorflow and Tone.js and has been trained on many MIDI examples and it learns about musical concepts, building a map of notes and timings.

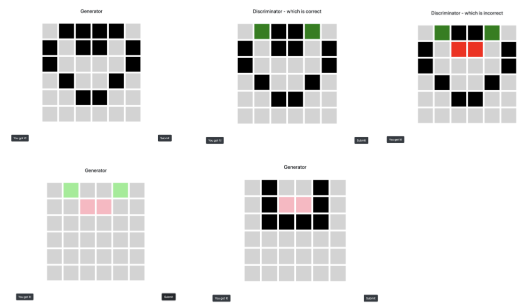

Figure 3: Example back and forth game play between the generator and the discriminator teams

.

- **Sketch RNN**: Built by the Google Creative Lab, Sketch RNN is an interactive web experiment that lets you draw together with a recurrent neural network model (Ha and Eck 2017). The neural net has learned to draw by training it on millions of doodles collected from the Quick, Draw! Dataset [4]. Once the user starts drawing an object, Sketch-RNN will come up with many possible ways to continue drawing this object based on where they left off.
- **AI News Anchor**: Developed by Xinhua and the Chinese search engine, these AI-powered news anchors were developed through machine learning to simulate the voice, facial movements, and gestures of real-life broadcasters, to present "a lifelike image" of a human news anchor (Kuo 2018).
- **This Person Does Not Exist**: This tool utilizes StyleGan2 (Karras et al. 2019) and has been trained on human faces to generate fake human faces using GANs [5].

Students spent approximately 10 minutes on each interactive tool. During this activity, students also make slides of each activity which includes a title, a screenshot of the artwork they created, and what they thought about it.

After the activity, students reflected on 1-4 of the tools by answering the following questions:

- What do you think the generator in this GAN is trying to generate? What dataset is the discriminator basing its decisions on?
- How could this technology do the most good?
- How could this technology do the most harm?

After students complete the activity, we come together as a class and discuss our experiences, specifically what was interesting, novel, or surprising to them.

[4] See more at https://quickdraw.withgoogle.com/

[5] See more at https://www.thispersondoesnotexist.com/

Workshop	1	2	3	4
n	22	12	16	22
Gender	F=13, M=9	F=5, M=7	F=5, M=11	F=13, M=9
Grade	5^{th}(1), 6^{th}(5), 7^{th}(5), 8^{th}(8), 9^{th}(3)	5^{th}(5), 6^{th}(7)	7^{th}(10), 8^{th}(3), 9^{th}(1), 10^{th}(1), 11^{th}(1)	6^{th}(5), 7^{th}(2), 8^{th}(12), 9^{th}(3),

Table 1: 72 participants participated in the activity across four online workshops

User Study

Methods

This activity was piloted in four synchronous online summer workshops. All workshops were held virtually over Zoom, and the activities were made available to students on Google classroom. All courses were taught by a team of researchers and educators. The first two workshops were a part of a larger Introduction to AI curriculum, and the remaining two workshops were a part of a workshop just focused on GANs. Timing varied depending on the workshop, but most workshops met daily for 2-3 hours for anywhere between one and three weeks, for a total of 30 hours of curriculum per workshop.

Participants

72 participants (grade 5 – 11) participated in the activity. The majority of participants were in grades five through nine, with one student in tenth grade and one in eleventh. The participants were spread across four different online summer programs, with distinct students in each program. The workshops were led by two teachers, and were assisted by 2-3 teaching staff. All participants,their parents, and teachers signed the assent and consent forms respectively to participate in these programs.

Assessment

During the ”How do GANs Work?” activity, data was collected during game play. Each round was documented and conversations were transcribed. During the ”Exploring GANs” activity, students recorded their responses in a Google form.

Students were asked to reflect on their learning experience at the end of each day. Multiple lessons may have happened in that day. In the reflection, students answered the following questions:

- What did you learn in today’s session?
- Which activities did you like most today? Why?
- Which activity/activities was/were hard for you? Why was it hard?
- What would you suggest to improve these activities?

Students from the first workshop were interviewed at the end of the entire program to share their experiences with individual activities and their overall perceptions of AI. They were asked the following questions about the GANs lesson:

- Is this activity engaging? Why or why not?
- Did you have any difficulties learning the activity? If yes, could you tell me more about it? If no, why not?
- What suggestions do you have for us to make this activity better?

Results

The game was played as a part of four distinct Middle School AI workshops with 22, 12, 16 and 22 students in each. Students were randomly assigned into the generator and discriminator groups, along with 2-3 instructors in each group. The two groups played the game for approximately 15 minutes, for an average of 5.25 rounds per game. Figure 4 shows three instances of the final images that passed by the discriminator.

During game play, students on the generator side started by making random guesses about what meaningful shape it could be. For instance, during the activity, one student said,

> It could be a heart.

After a few rounds, students started to pick on patterns from the discriminators’ feedback. For instance, in one workshop, a student on the generator side said,

> It is something that is symmetric.

Another student said,

> The two dots on the top could be eyes.

While the generators were drawing, students on the discriminator side were both trying to guess what the generators are making while they are making it, for instance, one student said,

> That looks like a heart.

and were also starting to make decisions about which ones they would mark correct (green) and incorrect (red). They used Zoom’s annotation marker for this, and the instructor facilitated a conversation where all students decided which positive and negative feedback they wanted to go ahead with. While incorrect blocks were easy to recognize, the correct feedback sometimes entailed discussion. For instance, the corner markers in the first image in Figure 5 (4x1 position) could be envisioned as a smile, depending on where

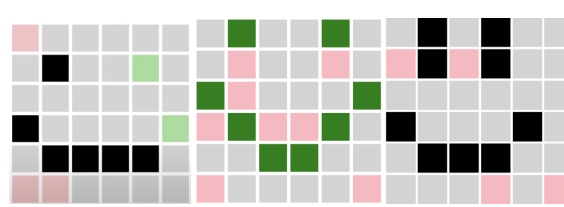

Figure 4: Gameplay screenshots from three different workshops. Three out of four games reached the correct answer, while one game was terminated early.

the eyes are. In cases of dispute, discriminators went with the highest votes.

After the activity, we convened together with all the students and discussed what students on both sides were thinking, and how they made their decisions. Students on the generator side started with random guesses, and continued to alter their inputs depending on the discriminators' correct and incorrect feedback. One student on the generator side said,

> Um, we will look at the hints that you gave us to read in the green block. And based on that, we will try to figure out different things we could make things that would cover up the green blocks, but not cover up the red blocks.

Students on the discriminator side looked for features of a face within the input that they received from the generators. They also referred to the training data set to make their decision. One student on the discriminator side said,

> We were looking for features that resembled a face or stuff that would belong in a face.

Another student said,

> What we used is, we went through six different examples. We had to try to find one, in between, which one looks like something. Oh, you're trying to pick a smiley face. So when we were trying to look through that we're trying to see. Okay, so this is what we want. And this is what we don't want. Then we take. So then we tell you guys, hey, this is what we don't want so take it out.

After playing the generator discriminator game, students explored up to four different web tools that use GANs. After they explored the tools, they were asked to identify 1) what the generator in the GAN was trying to generate, and 2) what dataset the discriminator in the GAN was basing its decisions on. Students were allowed to answer questions for one to four tools. We received 99 completed responses from 72 students. Of these 99 responses, there were 39 for SketchRNN, 28 for AI Duet, 23 for This Person Does Not Exist, and 9 for AI News Anchor. Overall, 88% of student responses were able to correctly identify what the generator was trying to generate, and 60% of student responses were able to correctly identify the dataset that the discriminator used. Results were broken down further by media assessed, and can be found in Table 2.

At the end of the lesson, students filled out a Questionnaire to reflect on their experiences. We received a total of 44 responses to this form. When we asked students to write *"one thing that they learned"*, 27 students said they learned how GANs work, with 15 of those explicitly calling out the *generator* and *discriminator*. When we asked students what their favorite activity of the day was, 18 mentioned the *How do GANs Work?* game,18 mentioned the *Exploring GANs* activity, and 8 mentioned another activity from the day. When we asked students what was challenging or difficult, 26 of them said "*None*", 3 said the *How Do GANs Work?* game, 5 said the *Exploring GANs* activity (mainly because of technical difficulty with the tools), and 10 mentioned things that did not relate to this lesson.

GAN Web Tool	Responses	% Correct Generator	% Correct Discriminator
This Person Does Not Exist	23	100	78.3
AI News Anchor	9	100	55.6
Sketch RNN	39	87.2	64.1
AI Duet	28	78.6	42.9

Table 2: Breakdown of student responses for the *Exploring GANs* activity.

At the end of the first workshop, we interviewed students and asked them to reflect on the full GANs lesson. One student on the generator team shared some frustrations with the game, specifically questioning why the discriminator didn't give direct "right or wrong" feedback:

> What was confusing about the GANs was like how it works. So it generated the discriminator would like, have the right answer. So we were just generating pics to see if it was right. But I'm like, how are you supposed to know if you're right if the discriminators really weren't [giving you] the right answer. I was totally confused. If the discriminator told you if you were right or wrong, I'm like, "oh, so that makes more sense to me if I'm right or wrong".

Another student voiced frustrations at the beginning of the lesson, but mentioned that the scaffolding during class made the concepts more understandable:

> Because I don't know what generator and a classifier is, at first. No, it was very difficult. But then, once the teacher explained, it was a bit more easy.

Another student mentioned that the team-based, simulation aspect of the game was unique compared to other lessons in the curriculum:

> Yeah [the activity is engaging], because we were all working together to find out what it could've been and we kept guessing and it was more teamwork in that activity than the ones from before. So it made it more interesting and fun.

Another student reflected on their new understanding of the real-world applications of GANs, referencing the tools that we used in class. This new understanding was different than what they initially thought AI could do:

> *"How GANs Work"* was an interesting experience. Because when I went in, I originally thought GANs would maybe be like, just maybe like a remix on a specific song. Like they wouldn't fake that, but it was paintings and they could even like fake people's pictures. Like they could take a bunch of data from what people's faces look like, or like a cat's face looks like, and then they could actually create a whole live thing and that person doesn't even exist.

Discussion

We designed a two-part lesson to first teach students about how GANs work, and encourage them to apply their thinking to existing GAN tools. The first part, an introductory GANs lesson, which includes interactive web-based game, was designed to help students understand the goals of the generator and discriminator. The web tool is designed to be used for synchronous online learning for virtual classrooms. Students were divided into two goal-driven teams—the generators and the discriminators. We tested the game with 72 middle school students distributed across four summer AI Education camps.

Out of the four instances of this game, we found that three generator teams were able to pass by the discriminator, often after a few rounds of sending incorrect configurations of squares to the discriminator. For the team that did not pass by the discriminator, we shared the data set with the generator team. All final generated images were different for each instance of game play, and none of them matched the data set exactly. This validates our idea that faces are a good example for the game, since they are common enough for students to recognize without focusing too much on the exact positions of the pixels. In future iterations of the game, data sets should be common images that students recognize and that can be represented in multiple ways (i.e. shapes or letters).

We see room for iteration and improvement in the lesson. The student's comment about not understanding the role of the discriminator may be explained by the fact that they were only exposed to the generator perspective. This may explain the student's confusion as to why the discriminator could not "tell you if you were right or wrong". In the Exploring GANs activity, students had a harder time determining what the discriminator was basing its decisions off of compared to what the generator was trying to generate. This may also be related to the fact that both the generator and discriminator group had access to what the generator was trying to generate during the game, but the discriminator group was the only group that had access to the dataset. In the future, two rounds of this game would be played, allowing all students to understand the GAN from both the generator and discriminator perspective. It would also be interesting to have students on the discriminator side develop their own datasets, instead of being given a dataset.

By encouraging children to explore both the technical aspects and the real-world applications of GANs in our lesson, students were able to see the full picture of how the technology was made but also how it is used in real life. Through daily reflections, we found that students found both activities equally engaging. We hope that this framing allows students, like Jessica, to make more informed choices about how to build and use GANs.

Limitations

This lesson was designed to teach students about how GANs work, what different neural networks of a GAN are, and allowed them to explore tools that utilize GANs. While students gained a good understanding of how a generator and a discriminator work in opposition with each other to generate data, through the abstraction of GANs for the game, we were not able to get all technical concepts of how GANs work across. For example, students did not learn about neural networks. Secondly, the use of a binary feedback (red and green) is not representative of how discriminators actually work. In future work, we plan to allow students to proceed to more complex versions of the game after playing the basic version, where they could respond with the closeness to the dataset on a scale of 0 - 1 representing probability. Further, this will also help tackle the current concern that the generators assume the green pixels are completely correct, hence reducing the randomness and novelty in the generated graphics. The lesson could also be accompanied by students viewing a real-life GAN in action, such as a hand-written letter generator GAN trained on the EMNIST dataset (Cohen et al. 2017), to understand how the network learns over time. Finally, our lesson was also conducted remotely, so while we have students' responses, we are not aware of their engagement level during the game and activity. The online implementation of the game excludes students who do not have access to remote learning resources.

Conclusion

In this paper, we present an introductory activity to teach middle school students about GANs. We tested our activity with 72 middle school students and found that students developed a foundational understanding of how GANs work, what generators and discriminators are, and what some applications of GANs are. In the future, we hope to expand this activity to allow students to play the role of both generator and discriminator and to address different types of media generation. We hope that this work inspires more K-12 educational resources for generative AI.

Video walk-through: **https://bit.ly/3bGWoNM**

Acknowledgments

The authors would like to thank Irene Lee, Helen Zhang, Jenna Hong, Grace Kim, Victor Sindato, Carolyn Song, Nisha Devasia, Ryan Blumofe and Yihong Cheng for their teaching assistance and feedback on the activity. Thank you to our partners and teachers at STEAM Ahead for all of your help in organizing the workshop and giving us feedback on the activity. Lastly, thank you to Amazon Future Engineer and NSF EAGER Grant #2022502 for funding this work.

References

Ali, S.; Payne, B. H.; Williams, R.; Park, H. W.; and Breazeal, C. 2019. Constructionism, Ethics, and Creativity: Developing Primary and Middle School Artificial Intelligence Education. In *International Workshop on Education in Artificial Intelligence K-12 (EDUAI'19)*.

Anderson, M.; Jiang, J.; et al. 2018. Teens, social media & technology 2018. *Pew Research Center* 31: 2018.

Brandon, J. 2019. There Are Now 15,000 Deepfake Videos on Social Media. Yes, You Should Worry. URL https://www.forbes.com/sites/johnbbrandon/2019/10/08/there-are-now-15000-deepfake-videos-on-social-media-yes-you-should-worry/#3b2cb2b43750.

Cohen, G.; Afshar, S.; Tapson, J.; and Van Schaik, A. 2017. EMNIST: Extending MNIST to handwritten letters. In *2017 International Joint Conference on Neural Networks (IJCNN)*, 2921–2926. IEEE.

Goodfellow, I.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014a. Generative adversarial nets. In *Advances in neural information processing systems*, 2672–2680.

Goodfellow, I. J.; Pouget-Abadie, J.; Mirza, M.; Xu, B.; Warde-Farley, D.; Ozair, S.; Courville, A.; and Bengio, Y. 2014b. Generative Adversarial Networks.

Ha, D.; and Eck, D. 2017. A neural representation of sketch drawings. *arXiv preprint arXiv:1704.03477* .

Hazzan, O. 1999. Reducing abstraction level when learning abstract algebra concepts. *Educational Studies in Mathematics* 40(1): 71–90.

Karras, T.; Laine, S.; Aittala, M.; Hellsten, J.; Lehtinen, J.; and Aila, T. 2019. Analyzing and Improving the Image Quality of StyleGAN.

Kim, H.; Garrido, P.; Tewari, A.; Xu, W.; Thies, J.; Niessner, M.; Pérez, P.; Richardt, C.; Zollhöfer, M.; and Theobalt, C. 2018. Deep video portraits. *ACM Transactions on Graphics (TOG)* 37(4): 1–14.

Kuo, L. 2018. World's first AI news anchor unveiled in China. *The Guardian* 8.

Lab, M. M. 2019. Computer Visions: Generative Machine Learning Tools for Creative Applications. http://visions.media.mit.edu/.

Lomas, N. 2020. Deepfake video app Reface is just getting started on shapeshifting selfie culture. URL https://techcrunch.com/2020/08/17/deepfake-video-app-reface-is-just-getting-started-on-shapeshifting-selfie-culture/.

Mann, Y. 2016. AI duet. *Experiments with Google. See, https://experiments. withgoogle. com/ai/ai-due t* .

Payne, B. 2020. *Can my algorithm be my opinion?: An AI + Ethics Curriculum for Middle School Students*. Master's thesis, Massachusetts Institute of Technology, Media Lab, Cambridge, MA, USA.

Renaud Danhaive, M. A. 2019. Creative Machine Learning for Design. https://architecture.mit.edu/subject/spring-2020-4453.

Strubell, E.; Ganesh, A.; and McCallum, A. 2019. Energy and policy considerations for deep learning in NLP. *arXiv preprint arXiv:1906.02243* .

Touretzky, D.; Gardner-McCune, C.; Martin, F.; and Seehorn, D. 2019. Envisioning AI for K-12: What should every child know about AI? In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, 9795–9799.