AI Detection Software for AI Altering Images and Creating Samples in 2023

Joseph DiGennaro

Old Dominion University, Norfolk Virginia

CYSE 494 Entrepreneurship

Professor Porcher

June 21, 2023

AI Detection Software for AI Altering Images and Creating Samples in 2023

Introduction

Artificial intelligence, also known as AI, has recently been at the forefront of attention in the media. AI has gained popularity for many reasons. Some like to show how it can give different perspectives to problems, and some like the aspect that it imitates human nature. However, some like AI for many malicious reasons. First of all, AI mimics human behavior in the way that it outputs information, so hackers or other people with malicious intentions can use this tool for bad. The main problem of AI stems from trust. Everyone knows that you should not believe everything that is on the internet, but AI is taking that to a whole new level. AI is causing a divide in people knowing what is real, and what is not real. This is not only a problem of trust, but it can also be a problem of fake information. AI is in the beginning stages as of this writing, and it has already been shown to be quite a powerful tool. As mentioned before, AI is a tool, and it all depends on that way in which it is used. A simple Youtube or Google search can show what many people have already done with the AI that is currently available. In some cases, it is nearly impossible to determine what is actually real versus that which is a product of AI. For example, AI has the capability to generate various outputs based upon data sets. Using these sample data sets, AI can output its own information. This is proven to be a problem in AI producing deep fake images or videos of people. These people are usually politicians, celebrities, or people that are well known. However, this is not only limited to these individuals. AI can take and modify data to falsify it, or make it alter it from the original form. When this happens, a major problem is presented in the cybersecurity realm. That problem is integrity. Integrity refers to the

“wholeness”, or “originality” of the data. In order for data to be considered secure, it needs to be unmodified and in the exact condition that it started with. AI can modify this data, therefore corrupting the integrity of the data.

The solution to the problem of AI sampling and altering data would be AI detection software. This software would allow for the image, video, or media in question to be run through the software, which would then detect if AI altered the data. The software will also be able to determine the percentage of data originality, and show a report to the user after analysis. For example, if an image is run through the software, the software should show the percentage of originality of the image, and the percentage of AI within the image. The numbers would likely have a plus or minus tolerance of a few percent to account for any variations. This software would likely start as a small service, operating on an app or website, but hopefully it can grow into web pages, code, and much more through ads. If the software is implemented on a platform, such as social media, users can see in real time if the content that they are viewing is AI generated/altered, or if the content is original. This is sort of how digital signatures operate in the cyber world. However, AI detection software can most definitely be a tool to help detect if data has been altered or tampered with. Although this mitigation method is not bulletproof, it can certainly help users feel more secure as to what content they were viewing in the respect of originality and integrity, which is harder and harder to find in the cyber world.

Scholarly Literature

AI poses a major problem for the average user, and even advanced technology users. AI is very hard to discern from what is real. Someone may not know the difference between fact and fiction, and this can cause problems in the real world. Wagner states, “It is near-impossible for casual consumers of images to authenticate digitally-altered images without a keen understanding of how to "read" the digital image” (Wagner, p. 1) This indicates that it takes skill and precision to decipher AI content. Most users do not have this knowledge, and as AI continues to advance so will the skill of discernment need to increase. This is most prominently known to the public as *deep fakes*, which is when AI takes personal information to use it in a counterfeit manner. This actually happened to famous actor Nicolas Cage, where people swapped his face using AI into situations that he was not present in in reality. “After enough time and training, the computer is then able to create new images that have never existed but can, with surprising validity, appear to accurately depict an entity. The computer is then able to take individual frames from a video and map the expression of a person’s face in the video to the computer-generated replication of the generated model of a person‘s face matching that expression, ostensibly swapping in the new computer-generated image for each frame” (Wagner, p. 36). This highlights the major need of trustworthiness, and reputation protection in the very near future. AI could ruin someone’s reputation if they are falsely presented through AI content. Wagner states, “However, the distribution of videos often go well beyond the original producer” (Wagner, p. 40)Having a way to discern AI created content can better help users to understand what they are viewing. Also, it can help them make better decisions knowing whether what they are seeing is authentic human made content, or AI created content.

Another key topic of AI is the application into the music industry. AI can be used in a multitude of ways in the studio. For example, it can generate lyrics, sounds, or synthesized results. It can even replicate the voice of an individual. This is where the fine line between using AI as a tool, and using AI illegally comes into play. Stealing someone's song, voice, or lyrics is likely a copyright infringement on the artist. This is why AI should be evaluated carefully in music sampling. It is one thing to use it as a tool, but when it is used to take the work of others it becomes a problem. Xu Tan states, "The typical tasks in AI music composition include melody generation, song writing, accompaniment generation, arrangement, performance generation, timbre rendering, sound generation, and singing voice synthesis, which cover different modalities (e.g., symbolic music score, sound) and well match to the theme of ACM Multimedia " (Tan, 2021). These many attributes of music production can be influenced by AI in order to create music, specifically samples.

At this current time, AI still has some limitations, but as time progresses AI will get more intricate and involved in the actions it can accomplish. Voice recognition and synthesis are the new hot topics with AI, as they can be used to make voices and speech. This can cause a problem when someone's voice is impersonated without their permission. As mentioned previously, AI is in the beginning stages of growth. As it advances, the bugs will eventually be worked out of the AI system. AI can even write its own code, so it will most definitely change. Music composition and AI are a combination that is likely to stay. However, our product of AI detection will help mitigate some of the negative effects of this technology. For example, our detection tool will help protect artists and their work that they make in the studio. Our AI tool will help to find copyright infringements, and can even notify artists when there is unauthorized use of their content.

Intellectual property rights are critical to understand when applying AI into the mix. It needs to be understood where the line is crossed that using someone else's content becomes an intellectual property rights violation. DeCosta states, "Attorneys are more frequently being engaged by clients confronting the impact of artificial intelligence technology on businesses". This helps to explain that the increase in AI has already led to an increase in legal inquiries about property rights. As AI advances, many legality means and measures are being set up to help safely regulate AI. "Copyrights can be used as another form of protecting AI, because AI software can be copyrightable". (DeCosta, 2017) The new legal battle for today is determining whether AI can produce a copyrightable item, and if it can contain a copyright. AI continues, and will continue to produce multitudes of outputs. Some of these outputs may be used by individuals or companies seeking to make a profit, so understanding the legal implications helps.

Our AI tool will seek to protect previously copyrighted information by checking current AI content with existing records. These record databases can be filled with trademarks, patents, copyrights, trade secrets, and much more. DeCosta states, "Trade secrets do not require governmental approval, and there is no application or examination process—and consequently no prosecution costs or application fees" (DeCosta, 2017), so it is highly likely that AI technology will likely be used by many under the veil of *trade secrets*. This is simply due to the lack of bureaucracy in legal methods and minimal financial cost that it requires to operate under a trade secret status. Our AI detection system will likely follow under the trade secrets category, as our AI detection algorithm is proprietary. This can help avoid unnecessary legal trouble that a small business like ours simply cannot take at this time.

One critical issue that is still being debated today is determining the author of AI produced content. Should credit be given to generators such as ChatGPT, or does the authorship transfer to the individual who requested information from AI? Nowak-Gruca states, “Advanced technologies, such as Artificial Intelligence (AI) systems, have been pushing nowadays societies toward new ethical and legal challenges, including copyright law dilemmas” (Nowak-Gruca, p. 1, 2021) This introduces the topic referred to as *ghostwriting*. Although this may sound like a spooky term, it simply refers to when something is written, but the writer does not take credit for the work. This is why this term is commonly used when talking about AI, as AI authors much of the work that many users take credit for. So, this brings the question to light as to who actually gets credit for AI generated information. Can authors copyright something that they got from using AI? Will AI generators need to be given credit for outputted data? All of these questions are still heavily debated, and will likely continue to be for the foreseeable future.

In addition, “ The works in question, produced by AI, are eligible for copyright, although nonhuman copyright is treated with suspicion by many countries’ laws. Continental law stipulates that copyrighted works must be human-made” (Nowak-Gruca, p. 30, 2017). This statement further ties into copyright laws and regulations. AI is obviously not human, so theoretically it can produce no works that could be copyrighted. However, if users take possession of AI outputs after the fact, then the situation could change as they could file a copyright claim on their behalf. AI presents this unique problem of work not having an author attached to it. With these modern problems, modern solutions will be required to deal with such issues. It is likely that copyright law will change to accommodate AI media, but until that time comes there will be plenty of uncertainty as AI evolves faster than the laws and regulations.

As AI picks up the pace, it will likely influence its way into almost all aspects of life. For example, the political world is now being used to showcase AI technology, specifically through the *deep fakes* mentioned previously in the paper. These deep fakes may be viewed as a harmless joke at the surface level, but they actually have power to persuade and change individual's opinions. The political world is all about persuasion and bias, and it seems like AI is being used to make that happen. This is likely to occur through disinformation tactics, which are intended to mislead people. Dobber states, "Deepfakes consist of largely real images, and producers only manipulate relatively small elements of the video (e.g., facial expressions, voice), which contributes to the realism of the deep fake. In this sense, a deepfake is qualitatively different from a photoshopped image: a deep fake deceives not just the eyes, but the ears as well" (Dobber, p. 70, 2021). AI technology is very powerful, and when used for malicious purposes, it can bring real life consequences to pass.

AI is likely to bring much confusion to the political realm, specifically targeting high ranking political officials. Dobber states, "If a political actor has enough training data, the actor can make many different, realistic deepfakes of the same person in a short period of time" (Dobber, p. 70, 2021). Political figures typically have ample airtime media and coverage to the public, so AI has a lot of source media to choose from. This can be very dangerous, as the more input data AI has, the more advanced its outputs will be. It is also likely to suspect that times of election and the ending of terms will show a trend of increased AI deep fakes. These will likely be used to slander political individuals, and have a goal in mind to manipulate the view into believing a lie. This further highlights the need for AI transparency, especially when it comes to political discussions that seek to change an individual's mind and thinking.

Moving onward from the political aspects, not even scientific publications are safe. Similar to the deep fakes mentioned previously, AI can be used to alter images, which can cause trouble for the scientific and academic community. Tampered or altered images of scientific evidence can make such institutions lose credibility. AI can change an image in many ways. For example, if scientists have determined a picture of a cell model, AI can be used to change the quantity of items within the image, thus voiding the authenticity of the work. Gu states, "Sampling using trained generative models can produce fake images that follow patterns similar to the real images. Images generated by these models are visually realistic and even scientifically self consistent" (Gu, 2022). This can be used for someone to falsify study, or make the results be more in their favor. AI can be used to achieve these goals, which is why an AI detection software is needed. It needs to be known if an image was artificially made, especially if it is masquerading as a genuine image. Furthermore, peer-reviewers are not experts on the AI subject, so they may not be able to identify AI content on their own.

With AI detection software, the image in question could be run through the software after watching a quick advertisement to support the company. Once the advertisement is viewed in its entirety, the image/media would then be processed through the program. Once analyzed and cross referenced with multitudes of information from around the world, the program would then output a result as to whether the image was genuine or not. Also, if the image was made by AI, then a percentage would be stated on how much the image differs from the original image. As more and more people use the AI detection program, it would work more efficiently, as it has more source files to operate from. Not to mention, the detection software would run off of AI to keep up with the AI it is trying to analyze and interpret.

As mentioned previously, AI operates and continues to advance based upon a great multitude of input data. Oftentimes, these exist in the forms of images, text, audio, or video files. However, there have been many instances where AI has taken images that it was not supposed to. Appel states, “Similar cases filed in 2023 bring claims that companies trained AI tools using data lakes with thousands — or even many millions — of unlicensed works. Getty, an image licensing service, filed a lawsuit against the creators of Stable Diffusion alleging the improper use of its photos, both violating copyright and trademark rights it has in its watermarked photograph collection” (Appel, 2023). This raises the question of whether AI has been corrupted due to using this information. For example, if future AI is based upon using unauthorized images and media in the past, will it incorporate the unauthorized materials in future outputs?

This very question is the reason for the AI detection software that we have developed in our group. AI detection software will help determine if content contains AI material, and if so, the percentage ratio of AI to the original content. Not to mention, this type of detection software can help protect content creators in their endeavors, knowing that their work will be protected. Appel states, “Content creators actively should monitor digital and social channels for the appearance of works that may be derived from their own” (Appel, 2023). However, most content creators simply do not have time to deal with such matters, which is why our AI detection tool will help with this issue. Also, as AI advances it will become more difficult to discern AI media from original content. Our AI detection will help content creators have peace of mind that their content remains their content. AI detection will also help users understand the content that they are viewing. The tool can show if content is partially AI, fully AI, or verified original content by the original source.

Expanding Beyond The Major of Cybersecurity

Currently, I'm working towards an education in the cybersecurity world. Even though AI has a great deal of potential in cybersecurity, it expands way outside of the boundaries of cybersecurity. AI seems to be working its way into all parts of life, and it shows no partiality to majors, minors, or education level. AI is used by many people, from beginners all the way to senior program developers. It is capturing the interest of many worldwide! AI is a tool that is used by many to help them with their tasks. For example, people who work with stock portfolios may consult with AI to help them determine market cycles, or to see if a stock is worth buying. Another example would be a writer who wants help brainstorming new ideas for a project. Someone who is in the music industry may want to ask AI for a certain melodic chord setup that is catchy for the listener. AI can even be used to write song lyrics based upon certain input values. In the near future, it is likely that AI will be able to write music, which may allow for more entry level music producers to get started in the industry. The uses for AI seem to be almost limitless, and it has not been available to the public for that long in 2023.

With all of this advanced technology, our AI detection software seeks to mitigate some of the uncertainty that AI brings to the table. In a digital world, it is important to understand what is real, and what is not. Everyone knows the commonly used phrase that you should not trust everything on the internet, which is most definitely a valid statement. However, with AI rapidly coming online, this statement will only become more true, as AI is very hard to discern from real human content. All the while, many users typically do not understand how to see if something was AI generated, so having an AI detection tool can be very helpful. In the early stages, our AI detection software will be a proprietary service that will be available for free use, provided that

users first watch an advertisement to pay for the company and employee costs. If users would like to upgrade their membership to ad-free, then they can pay a small monthly fee. Our AI detection tool can be used by everyday users, academic institutions, scientific laboratories, or anywhere else that wants to ensure that content is original and authentic. In the future, we would like to implement this software to various social media platforms, provided they pay us a commission for the service. This would be a win-win situation, as we would make money, and social media would know in real time whether or not the media that they were consuming was made partially, or in full by AI.

This tool of AI detection is most definitely needed from a worldwide perspective. It is getting more difficult to trust sources and content with rapidly advancing AI. This is especially true because artificial intelligence can deceive by using both visual and audio means. For example, if a video came to surface of another nation's president saying that they want to go to war with the US, the first thing to do would be to run that video through our software to make sure it was real. Chances are, a video like that would result in our software confirming that it was generated by AI, therefore calming the situation. Without our AI tool, people would likely think the video was real and start to panic over something that is completely false! Although this may seem like an unlikely example, it is not impossible. We need to be prepared for all types of AI threats and deep fakes, as they can cause trouble worldwide. This is not just a cybersecurity problem, but rather a worldwide problem! As mentioned above, AI is advancing quickly and even writing its own code, so our AI detection tool needs to be able to stay many steps ahead to help keep users safe. Not to mention, knowledge is power. If users can discern where there information is coming from, they can make better decisions.

The Telling Signs of Success

It will be quite easy to tell if the AI detection software is working based on two factors. The first factor is the analytical data from the AI detection software. This can include the number of AI detection, percentage ratios, altered media, and so on. These results can help make statistical assumptions based on the current situation, and experimental data. Secondly, society is the teller of success. If the AI detection software works according to plan, society will be more knowledgeable of AI, and if the content that they view has been altered or generated by AI. Another way we can judge the success of the AI detection tool is advertising revenue. If advertisement revenue is high, then that explains how many users are using the AI detection software. Also, we can judge the success based on how many paid subscriptions are issued to users of our AI detection software. Users that have a paid subscription are premium customers, and likely to conduct business with us again.

Society needs to understand how AI works, and how it affects them. Training users to be aware is essential in a technological world. Cyber hygiene is becoming more and more important with the development of AI. Though these barriers may seem impossible to overcome, there is hope for a successful society and future! Our AI detection tool will most definitely be able to fight against misinformation, and help users understand the content that they view on the internet. Our AI detection tool essentially acts as a shield to help protect users, similar to how antivirus and anit-malware software works to help protect computer users. Our tool will most definitely be effective because it helps to empower users with knowledge. Knowledge is power, and giving users more knowledge will allow for a better understanding of AI, and how to identify it in all types of situations.

From Innovation to Reality

In order to make this into a reality, the project must first get funding. Thankfully, limited funding is needed because we can use AI to help make our tool, which is currently free to use. It is likely that we would finance the project among the group, and reap the benefits later. It will require an initial investment of time, energy, and many late nights but it is certainly worth it. To get this project started requires a plan. Having a plan allows for the group to have a mission and vision statement to help us stay on track. Once the plan is established, the work can begin. Our AI detection tool will likely use AI to help discern AI content and media. Since AI can write its own code, it can change and evolve depending on the world of AI. The most expensive part of computer software is maintenance. So, using AI as our maintenance tool saves a lot of time and money, which is essential since this is an entrepreneurial effort.

Also, we can use many platforms to run our service on. For example, we can get a website domain very inexpensively and host our service there. In addition, we could code the AI detection tool to work on Android and Apple iOS devices. This is likely necessary as many users operate primarily on mobile devices today. Thankfully, there is a great multitude of knowledge on youtube and various internet articles that could help us learn to code and create our own apps. Once that has been completed, we can spread the word through social media platforms such as, Instagram, Snapchat, Facebook, TikTok, and much more! Creating an account with these platforms is completely free, so it would be of no expense on our part. The more downloads we get will provide more advertisement revenue that we can use to further invest into the software to make it better, and hopefully scale it larger to make more profits. Gaining contracts with universities, laboratories, and social media corporations would be our bread and butter.

It is important to note that AI is biased in the way that it operates. It can be biased in the respects of societal and data biases. Societal biases are when AI takes assumptions that exist in society, and incorporate it into its algorithm. Data biases are when data is invalid, inaccurate, or not a complete data set. This can make the AI make strange decisions and produce biased outcomes, which are unethical for society. Also, AI produces output very close to actual human behavior, so discernment of human versus AI content will be a harder battle everyday. Not to mention, education is one of the hardest barriers to overcome. Society needs to understand how AI works, and how it affects them. Training users to be aware is essential in a technological world. Cyber hygiene is becoming more and more important with the development of AI. Though these barriers may seem impossible to overcome, there is hope for a successful society and future!

Next Steps

This is just the start of AI detection software, and much is likely to change in the very near future. We most definitely know that there is a need for software like this, it is just the fact of designing the software and moving to distribution, which is easier said than done. However, it is not impossible, and it can most certainly be brought to fruition. Ideally, we would like to secure some type of partnership with an AI intuition, which would help us get front row access to AI algorithms and protocols. Not to mention, collaboration with other technology companies can help make our product better, and it is likely that they will have ample resources to get the ball rolling faster than we can with our limited resources. Design thinking is important to use in this entrepreneurial effort, as we can work the bugs out early and fast. This helps save time, money, and frustration when getting a product ready for debut.

I have learned a lot from this entrepreneurship, especially when it comes to understanding the need for solutions to problems that exist in society. The best products are those that solve problems, and AI detection software is something that will definitely be needed to understand the information that will be surfacing both now and in the future. The lessons learned throughout the modules helped me better understand what it takes to run a business, and the best ways to do it. Design thinking was a critical takeaway that I learned from the course. I like that it is a human based approach that helps people's needs. I think it is important to help others, especially when your product has a positive benefit in someone's life. My thoughts of entrepreneurship have definitely changed since taking this course. Initially, I thought running a business was fairly straightforward, but I now realize that it takes skill, motivation, and persistence to achieve your goals. Entrepreneurship requires the entrepreneur to wear many hats, and try when it seems impossible. Another thing I learned was the importance of networking. Having a good network of people makes business much easier and enjoyable.

Nothing is perfect in life, so there are most definitely a couple things I would have changed in this effort. Firstly, we should have decided on the AI detection tool more early on in the project, as that would have given more time to think about it. Secondly, we all have jobs so it was difficult to have all of our schedules align, which was tough at times. However, we all managed to hustle hard and put in the work. I would say time management is key in such projects. Also, being organized can help make things less stressful when working as a group. All in all, I enjoyed working on this entrepreneurship project, and I think it will surely help me in my future career and endeavors.

References

Aman. (2021). Microsoft word - JIPR-1069 corrected proof. *Journal of Intellectual Property Rights*, *27*.

Appel, G. (2023, April 7). *Generative AI has an intellectual property problem*. Harvard Business Review. https://hbr.org/2023/04/generative-ai-has-an-intellectual-property-problem

DeCosta, F. (2017, August 30). *Intellectual property protection for artificial intelligence*. Finnegan, Henderson, Farabow, Garrett & Dunner, LLP. https://www.finnegan.com/en/insights/articles/intellectual-property-protection-for-artificial-intelligence.html

Dobber, T. (2021). *Do (Microtargeted) Deepfakes Have Real Effects on Political Attitudes?* https://journals.sagepub.com/doi/pdf/10.1177/1940161220944364

Gu, J. (2022). AI-enabled image fraud in scientific publications. *Patterns*, *3*(7). https://doi.org/10.1016/j.patter.2022.100511

Tan, X. (2021, October). *A Tutorial on AI Music Composition*. ACM Digital Library. https://dl.acm.org/doi/abs/10.1145/3474085.3478875

Wagner, T. L., & Blewer, A. (2019). "The word real is no longer real": Deepfakes, gender, and the challenges of ai-altered video. *Open Information Science*, *3*(1), 32–46. https://doi.org/10.1515/opis-2019-0003