

Florida International University

Spring 2024 Senior Design Project

Quantifying Copyright Infringement in AI-Generated Images

Student: Uriel Juarez

Mentor: Wenbin Zhang, Product Owner

Instructor/Faculty: Seyedmasoud Sadjadi, Florida International University

PROBLEM

Today's digital landscape, differentiating between original and potentially infringing content is increasingly difficult. This project aims to tackle this issue by using similarity metrics to compare copyrighted images with AI-generated ones.

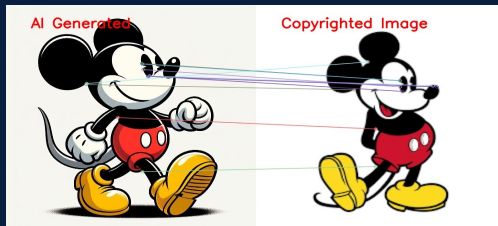
By analyzing metrics like ffsim and ssim, we aim to create a method to gauge the likelihood of copyright infringement based on visual similarity. Through this work, we hope to improve the ability to identify and protect intellectual property rights online

Technologies

Google Colab | Python | OpenCV | SerpAPI | Selenium | Dalle-3



Results



SSIM similarity index: 78.0871%
Potential copyright infringement detected based on SSIM.

FSIM similarity index: 30.7280%
Potential copyright infringement detected based on FSIM.

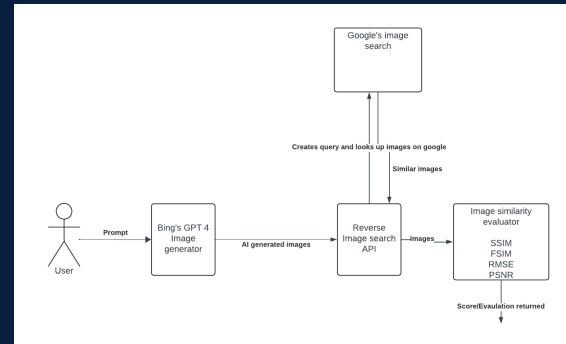
RMSE similarity index: 0.02550827002882957
Potential copyright infringement detected based on RMSE.

PSNR similarity index: 31.5318 dB
Potential copyright infringement detected based on PSNR.

Obtaining copyrighted material

In obtaining copyrighted material for our analysis, a crucial step involved using a reverse image search API. This API helped us find images similar to the provided copyrighted material, much like how Google's image search works. Additionally, the API allowed us to refine our search by specifying certain criteria such as rights or dimensions. This was helpful because it let us filter the results to only include images with specific usage rights or dimensions that we needed. By using this API, we efficiently collected a variety of copyrighted material for our analysis

OBJECT DESIGN



SUMMARY

•**Copyright Infringement Detection:** Developed an application that utilizes advanced similarity metrics and reverse image search APIs to detect potential copyright infringement. The application compares query images with AI-generated images, assessing their similarity to determine the likelihood of infringement.

•**Efficient Data Collection:** Implemented a reverse image search feature using an API, allowing for the efficient retrieval of images closely resembling the copyrighted material. The application also incorporates query options to filter search results based on specific criteria such as rights or dimensions, ensuring the relevance and legality of the collected data.

•**Ease of Use:** Created an application that streamlines the process for users by implementing reverse image search functionality. Users simply input a prompt, and the application performs reverse image searches to measure the level of copyright infringement. This straightforward process eliminates the need for users to upload images manually, enhancing usability and efficiency.

REQUIREMENTS

- Linux/Unix based machine
- AMD64 architecture
- Google colab/Jupyter Notebook
- Git/Github
- Python/PIP



Obtaining AI generated content



•**BingImageCreator Integration:** Integrated BingImageCreator, a publicly available API, into the project's workflow to tap into Bing's free ChatGPT 4 and DALL-E 3 usage capabilities. This integration enabled the project to leverage advanced AI models for generating diverse and high-quality AI-generated images.

•**Cost-Free Image Generation:** By leveraging the user's Bing account and API token, the project accessed Bing's resources for AI image generation at no cost. This cost-effective approach provided the project with a valuable means of testing and refining the program without incurring any expenses associated with image generation.

•**User Authentication and API Token Usage:** Incorporated user authentication mechanisms to facilitate the usage of the user's Bing account via their API token. This authentication process ensured secure access to Bing's resources while enabling seamless integration of BingImageCreator into the application's functionality.

REFERENCES

- Wang, Z., Bovik, A. C., Sheikh, H. R., & Simoncelli, E. P. (2003). "Image Quality Assessment: From Error Visibility to Structural Similarity." Retrieved from <https://www.cns.nyu.edu/pub/eero/wang03-reprint.pdf>
- Sankar, Vinu Sadasivan et al. (2023). "Can AI-Generated Text be Reliably Detected?" arXiv:2303.11156. Retrieved from <https://arxiv.org/pdf/2303.11156.pdf>
- Kariyawasam, Kanchana. (2020). "Artificial intelligence and challenges for copyright law." International Journal of Law and Information Technology, vol. 28, no. 4, pp. 279-296. Retrieved from <https://doi.org/10.1093/ijlit/ieaa023>
- Elkhatat, Ahmed M., Elsaid, Khaled, & Almeer, Saeed. (2023). "Evaluating the efficacy of AI content detection tools in differentiating between human and AI-generated text." International Journal for Educational Integrity. Retrieved from <https://edintegrity.biomedcentral.com/articles/10.1007/s40979-023-00140-5>
- Lancaster, Thomas. (2023). "Artificial intelligence, text generation tools and ChatGPT – does digital watermarking offer a solution?" International Journal for Educational Integrity. Retrieved from <https://edintegrity.biomedcentral.com/articles/10.1007/s40979-023-00131-6>

ACKNOWLEDGEMENT

The material presented in this poster is based upon the work supported by XXXXX (name of the project sponsor: federal, state or local government, or corporate partners, where applicable). We/I thank XXXX for his/her/their assistance and mentorship that I/we received throughout the senior design project.