

RECONHECIMENTO FACIAL INFANTIL EM IMAGENS: PARA FUTURAS DETECÇÕES DE CONTEÚDOS ILÍCITOS

Gabrielle Mitie Suzuke Tengan¹, Leila Cristina C. Bergamasco²
^{1,2}Departamento de Ciência da Computação, Centro Universitário FEI
unifgtengan@fei.edu.br leila.cristina@fei.edu.br

Resumo: Com o advento da Internet, e consequentemente do entretenimento online - principalmente infantil - surge a problemática da vulnerabilidade dos menores de idade, que não possuem conhecimento suficiente dos perigos abundantes e por isso, o aumento de assédios destes na Internet têm sido uma preocupação. Órgãos governamentais como o Ministério Público Federal (MPF) tem se utilizado de técnicas de inteligência artificial (IA) com o intuito de contribuir no aperfeiçoamento de sua atuação na defesa da sociedade, especialmente no âmbito da segurança digital, disponibilizando um portal para denúncias anônimas, nas quais os usuários podem mandar evidências em formato de imagens, textos e vídeo, que são analisadas de forma manual por especialistas criminais. Nesse sentido, o presente projeto de pesquisa visa auxiliar a análise de denúncias relacionadas a pornografia infantil, implementando métodos computacionais de processamento de imagens e IA para a detecção e classificação de faces infantis em imagens, que objetiva acelerar o processo, na medida que possibilitará separar imagens com evidências válidas ou não.

1. Introdução

A Internet é um vasto ambiente visto que conta com mais de 64,4% da população [1] mundial ativa, onde grande parte dessa porcentagem são menores de idade. Com isso, é possível perceber que com as diversas opções de entretenimento infantil observa-se a maior vulnerabilidade dos jovens dentro do meio digital, uma vez que, por ser um ambiente amplo, somado à inocência das crianças, sua vulnerabilidade acaba sendo exposta para criminosos da Internet.

Com mais de 23 mil denúncias registradas somente em território brasileiro nos primeiros quatro meses de 2023 [2], é evidente que a pornografia infantil é uma das problemáticas que mais afligem a segurança dos menores em meio digital atualmente. Dito isso, nota-se um aumento considerável das denúncias de pornografia infantil, comparado ao mesmo período no ano de 2022, houve um aumento de 70% [2].

Para a análise dessas denúncias e possível prisão dos criminosos, os especialistas criminais precisam realizar a análise individual desses relatos e categorizá-los de acordo com o crime cometido e se a denúncia possui evidências válidas ou não. O fato de que nem todas as denúncias são válidas é um problema grave, pois acaba gerando a subutilização dos especialistas criminais.

Ao mesmo tempo técnicas de processamento de dados e mais especificamente de imagens, tem evoluído de forma consistente nos últimos anos agregando

técnicas de Inteligência Artificial (IA) para o processamento e classificação de imagens em diferentes classes. Essa metodologia pode ser aplicada em diferentes setores da sociedade, como na saúde, na análise de imagens médicas para detectar doenças e na agricultura para identificar tipos de pragas [3]. Na área de segurança pública, essas técnicas podem identificar padrões de imagens ilícitas e bloquear o conteúdo ou avisar autoridades, empresas como Google e Apple já aplicam isso em suas plataformas [4].

Dessa forma, esse projeto de pesquisa busca aplicar e validar técnicas de processamento de imagem e IA para a classificação de faces infantis. Essa primeira classificação visa facilitar e agilizar o primeiro atendimento de denúncias focadas em pornografia infantil, na medida que poderia filtrar denúncias que efetivamente possuam crianças nas imagens como evidências ou não. A partir dessa filtragem, as imagens com crianças seriam separadas para uma análise futura a ser realizada por um especialista criminal, otimizando o seu tempo para a análise.

Essa pesquisa é parte de um projeto maior que envolve a parceria do Centro Universitário FEI e o Ministério Público Federal de São Paulo (MPF-SP) - divisão de crimes cibernéticos, no qual se planeja implementar algoritmos que auxiliem na classificação de discursos de ódio e pornografia infantil.

2. Metodologia

Dado o objetivo do presente projeto, ele será decomposto em 5 blocos de atividades:

- a) **Pesquisa Bibliográfica:** o aluno pesquisador irá se aprofundar na literatura sobre conceitos, métodos e documentação necessárias para o desenvolvimento do projeto;
- b) **Formação da base de dados:** na primeira fase, o aluno realizará testes em bases de dados públicas como a UTKFace (ZHANG; SONG; QI, 2017). Após a validação inicial da estratégia proposta, será avaliado junto ao MPF a possibilidade da utilização de bases reais.
- c) **Desenvolvimento de técnicas computacionais:** nessa etapa serão implementadas:
 - 1) Técnicas de pré-processamento de imagens para tratamento de ruídos e normalização da imagem;
 - 2) Extração de características considerando aspectos de forma e textura da imagem;
 - 3) Implementação do classificador de Naive-Bayes, é possível que seja considerado a

implementação de outros classificadores, a partir da pesquisa bibliográfica realizada.

- d) **Avaliação e Análise:** o método desenvolvido será avaliado considerando as métricas descritas na seção 2.2.1. Serão feitas avaliações com bases de dados genéricas e posteriormente com a base de dados real do MPF.
- e) **Documentação e Divulgação:** a partir dos resultados parciais e totais do presente projeto de pesquisa, relatórios e artigos científicos serão elaborados para divulgação em congressos científicos. Será utilizada linguagem Python e bibliotecas de processamento de imagem como Scikit-Image e Opencv.

Em relação a avaliação serão consideradas métricas de acurácia, precisão e revocação para avaliar o método desenvolvido.

3. Resultados Parciais

Até o momento foram realizadas diversas comparações entre os histogramas das imagens com o auxílio da biblioteca numpy e matplotlib, ambas pertencentes a linguagem Python. As comparações foram feitas usando o conceito da distância Euclidiana.

Primeiramente foi realizada a transformação da imagem para a escala de cinza e a delimitação da região dos olhos. Posteriormente, foi criado o histograma da imagem da área delimitada que foi normalizado utilizando a técnica min-max, que consiste em redimensionar cada dado no intervalo de [0,1] ou [-1,1]. Os histogramas do olho esquerdo de adultos e crianças foram separados e analisados por gênero, e depois foi calculado a distância Euclidiana entre eles que define que, quanto mais próximo de zero o resultado for, mais parecidos os histogramas são. A Figura 1 apresenta a matriz de distâncias geradas.

Nessa matriz foi possível identificar que entre os adultos houve uma melhor separação do que entre crianças, por exemplo: homens são mais parecidos entre si, assim como as mulheres. Entretanto, esse comportamento não é observado para crianças.

| Olho Esquerdo | Menina 1 | Menina 2 | Menino 1 | Menino 2 | Mulher 1 | Mulher 2 | Homem 1 | Homem 2 |
|---------------|----------|----------|----------|----------|----------|----------|---------|---------|
| Menina 1 | 0 | 3.72 | 4.13 | 4.31 | 4.38 | 3.7 | 3.48 | 3.73 |
| Menina 2 | 3.72 | 0 | 3.37 | 5.69 | 5.16 | 5.15 | 2.4 | 2.46 |
| Menino 1 | 4.13 | 3.37 | 0 | 6.33 | 4.13 | 5.68 | 2.98 | 2.73 |
| Menino 2 | 4.31 | 5.69 | 6.33 | 0 | 4.25 | 2.76 | 5.84 | 5.43 |
| Mulher 1 | 4.38 | 5.16 | 4.13 | 4.25 | 0 | 3.93 | 4.84 | 3.97 |
| Mulher 2 | 3.7 | 5.15 | 5.68 | 2.76 | 3.93 | 0 | 5.27 | 4.89 |
| Homem 1 | 3.48 | 2.4 | 2.98 | 5.84 | 4.84 | 5.27 | 0 | 2.25 |
| Homem 2 | 3.73 | 2.46 | 2.73 | 5.43 | 3.97 | 4.89 | 2.25 | 0 |

Figura 1 - Matriz de distâncias

4. Etapas Futuras

As próximas etapas a serem percorridas no projeto serão focadas na implementação da ferramenta DeepFace, afim de tentar realizar a identificação da idade de cada rosto encontrado na imagem e assim prosseguir para a sua classificação.

5. Referências

- [1] The Changing World of Digital In 2023 - We Are Social UK --- wearesocial.com. <https://wearesocial.com/uk/blog/2023/01/the-changing-world-of-digital-in-2023/>, [Accessed 02-10-2023]
- [2] Denúncias de imagens de abuso e exploração sexual infantil online compartilhadas pela SaferNet com as autoridades têm aumento de 70% em 2023 --- new.safernet.org.br. <https://new.safernet.org.br/content/denuncias-de-imagens-de-abuso-e-exploracao-sexual-infantil-online-compartilhadas-pela>, [Accessed 02-10-2023].
- [3] MARQUES, Bruno Henrique Pereira. Avaliação de algoritmos baseados em deep learning para localizar placas veiculares brasileiras em ambientes complexos. In.
- [4] CARVALHO, Itamar Almeida de. Classificação de imagens de pornografia e pornografia infantil utilizando recuperação de imagens baseada em conteúdo. [S.l.: s.n.], 2012.

Agradecimentos

À instituição FEI, seu corpo docente, direção e administração que oportunizaram a construção deste projeto.

¹ Aluno de IC do Centro Universitário FEI. Projeto com vigência de 04/2023 a 05/2024.