

AVALIAÇÃO DE MÉTODOS DO ESTADO DA ARTE DE SEGMENTAÇÃO SEMÂNTICA

Gabriel de Almeida Spadafora¹, André Eugênio Lazzaretti², Anderson Brilhador³
, Heitor Silvério Lopes⁴

RESUMO: A segmentação semântica tem se destacado como uma ferramenta essencial no estudo de tatuagens, permitindo a extração de informações biométricas altamente relevantes para a segurança pública. Além disso, esse processo pode contribuir significativamente para o reconhecimento de vítimas e suspeitos, auxiliando investigações forenses e identificações em cenários críticos. No entanto, uma das principais dificuldades deste estudo é a escassez de conjuntos de dados na área, agravada pela dificuldade em equilibrar as classes, o que se deve à popularidade desigual de certos tipos de tatuagens e às limitações de acessibilidade impostas por direitos autorais que protegem muitas dessas imagens. Diante desses desafios, o presente estudo tem como objetivo desenvolver um gerador de imagens semi-sintéticas e analisar seu impacto na segmentação semântica de tatuagens. O gerador automatiza a combinação de uma imagem de pele sem tatuagem com um modelo de tatuagem, gerando variações de imagens. Além disso, ele facilita a etapa de rotulagem, uma vez que uma de suas saídas inclui a máscara rotulada correspondente. Para avaliar o desempenho do gerador semi-sintético, foram criados novos dados a partir dele, formando um conjunto de dados semi-sintético. Em seguida, dois modelos de redes neurais foram treinados: um utilizando esse conjunto semi-sintético e outro utilizando um conjunto de dados reais. Ambos os modelos foram validados e testados exclusivamente em um conjunto de dados reais, permitindo uma comparação direta entre os resultados obtidos com dados reais e semi-sintéticos. Essa análise foi realizada utilizando o IoU por classe e o mIoU como métricas de avaliação. Os resultados obtidos são promissores, demonstrando uma melhora de desempenho do modelo em certas classes do conjunto de dados semi-sintéticos em comparação com o conjunto de dados real.

Palavras-chave: Segmentação Semântica Conjunto Semi-Sintético Tatuagem

1 INTRODUÇÃO

Nos últimos anos, a identificação biométrica através de tatuagens emergiu como um campo promissor tanto na segurança pública quanto na análise forense. As tatuagens, com seus traços quase únicos e sendo formas de expressão humana e arte, carregam informações discriminativas que auxiliam na identificação de indivíduos. De acordo com estudos de Ngan e Grother (2015), as tatuagens têm sido empregadas há muitos anos como biométricas "suaves" em contextos forenses, destacando-se de outras biométricas como gênero, idade e raça por conterem mais informações discriminativas.

Esse aspecto é particularmente relevante em um mundo onde o volume de conteúdo digital cresce exponencialmente, e a necessidade de ferramentas automatizadas para a identificação de suspeitos ou

¹  gspadafora@alunos.utfpr.edu.br;  <https://orcid.org/0009-0002-0982-3997>.

²  andrelazzaretti@gmail.com;  <https://orcid.org/0000-0003-1861-3369>.

³  andersonbrilhador@gmail.com;  <https://orcid.org/0000-0002-1936-140X>.

⁴  heitorslones@gmail.com;  <https://orcid.org/0000-0003-3984-1432>.

vítimas torna-se cada vez mais crítica. A segmentação semântica de tatuagens, que visa identificar e separar cada objeto dentro de uma imagem para análise detalhada, desempenha um papel crucial neste contexto. No entanto, a complexidade das tatuagens, aliada à escassez de conjuntos de dados públicos e abrangentes, tem sido um desafio significativo para o avanço da pesquisa e desenvolvimento de métodos eficientes de segmentação.

Em resposta a esses desafios, este trabalho explora as capacidades da aprendizagem de modelos para a segmentação semântica de tatuagens, propondo uma solução para a escassez de dados. O estudo desenvolve um gerador não supervisionado capaz de criar imagens semi-sintéticas de indivíduos tatuados, juntamente com suas respectivas máscaras rotuladas de forma automatizada. O objetivo principal é possibilitar a criação de grandes conjuntos de dados, mesclando imagens e modelos de tatuagens por meio de técnicas de processamento de imagem, gerando assim uma variedade de imagens e suas respectivas máscaras.

2 METODOLOGIA

A metodologia começa com a preparação dos dados necessários para o gerador semi-sintético, que exige dados rotulados com máscaras de pele. Como o conjunto de dados disponível é limitado em amostras e regiões de pele, será realizada uma etapa de data augmentation para ampliá-lo e assegurar áreas de pele adequadas em cada amostra. As subseções seguintes detalharão o tratamento dos dados, o funcionamento do gerador semi-sintético e as métricas de avaliação utilizadas.

2.1 Conjunto de Dados Disponível

No campo da segmentação semântica, diversos estudos complementares têm sido desenvolvidos nos últimos anos. Para este trabalho, o conjunto inicial de imagens base foi fornecido pelos estudos de [Gonzalez-Soler, Rathgeb e Fischer \(2023\)](#). Esse conjunto inicial de 140 imagens foi expandido para 461 imagens e máscaras por meio de técnicas de data augmentation.

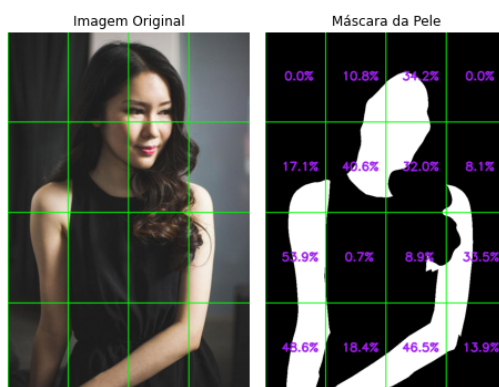


Figura 1 – Divisão da imagem e máscara em subpartes. Fonte:
Elaborado pelos autores.

O processo de expansão envolveu a divisão das imagens originais em subpartes utilizando uma grade 4x4, conforme ilustrado na Figura 1. Apenas as subimagens contendo pelo menos 50% de pele foram selecionadas, visando garantir uma área suficiente para a aplicação de tatuagens no gerador. Cada quadrante selecionado foi redimensionado para 224x224 pixels, dimensão esta escolhida devido à arquitetura da rede neural utilizada neste estudo, que demanda essa proporção em razão do elevado custo computacional associado ao processamento de imagens de maior tamanho.

Quando a subimagem selecionada era menor que 224x224, utilizamos o modelo EDSR (Enhanced Deep Super-Resolution) para aumentar a resolução da imagem, garantindo que todas as amostras mantivessem o tamanho adequado para o treinamento do modelo. Este processo de augmentation possibilitou a criação de um conjunto de dados mais robusto e equilibrado, adequado para a segmentação semântica.

2.2 Gerador de Imagem Semi-Sintética

Neste trabalho, propomos um gerador que adiciona automaticamente tatuagens sintéticas a imagens reais. A Figura 2 apresenta uma visão geral da arquitetura do gerador, destacando a etapa de tratamento do conjunto inicial de imagens base, a mesclagem da tatuagem na imagem original e o resultado final da imagem com sua respectiva máscara rotulada. Um pseudocódigo que resume o processo de geração é detalhado no Algoritmo 1.

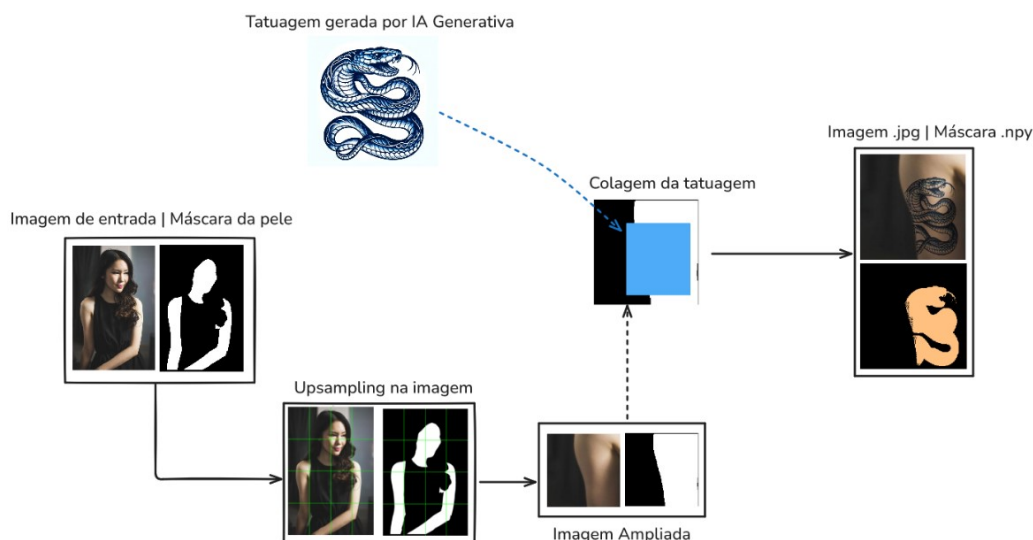


Figura 2 – Arquitetura do gerador de imagem semi-sintética. Fonte: Elaborado pelos autores.

Para posicionar corretamente a tatuagem na imagem, começamos escolhendo um local adequado. Identificamos as regiões de pele R no mapa de segmentação M (as áreas brancas na máscara de pele). Selecionamos aleatoriamente uma dessas regiões $r \in R$ e, dentro dela, escolhemos uma coordenada $x = (x,y)$ de pele, garantindo que o retângulo a partir de x , limitado pelas bordas da pele, tenha pelo

Algorithm 1 Algoritmo para o gerador de imagem semi-sintéticas. Fonte: Elaborado pelos autores.

```

1: procedure GERADOR( $I, M, T, C$ )
2:    $R \leftarrow \text{find\_contours}(M)$ 
3:    $r \leftarrow \text{random\_region}(R), M_T \leftarrow \text{get\_tattoo\_map}(T)$ 
4:    $x \leftarrow \text{random\_position}(r), M' \leftarrow \text{place\_tattoo\_on\_mask}(M, T, x)$ 
5:   while  $|M \cap M'| < \frac{M_T}{2}$  do
6:      $x \leftarrow \text{random\_position}(r), M' \leftarrow \text{place\_tattoo\_mask}(M, T, x)$ 
7:   end while
8:    $I' \leftarrow \text{blend\_tattoo\_on\_image}(I, T, x), S \leftarrow M$ 
9:    $S[M \cap M'] \leftarrow C, S[1] \leftarrow 0$ 
10:  return  $I', S$ 
11: end procedure

```

menos metade do tamanho da tatuagem. Isso assegura que a maior parte da tatuagem seja visível na imagem final.

Posicionamos a tatuagem em x e a sobrepomos na imagem original, multiplicando a camada da tatuagem com a imagem de entrada. As partes da tatuagem que ultrapassam a área de pele são recortadas. Para adicionar realismo, ajustamos as cores, aplicamos desfoque gaussiano e reduzimos a opacidade da tatuagem. Além disso, a imagem gerada I' vem com uma máscara previamente rotulada, onde a região tatuada é marcada com o valor da classe C correspondente. Por fim, removemos a porção de pele da máscara, resultando em uma máscara que mostra apenas o fundo e a rotulação. Implementamos o método em Python, utilizando amplamente a biblioteca OpenCV.

A obtenção do modelo de tatuagem utilizou inteligências artificiais generativas, como DALL-E (OPENAI, 2021) e Adobe Firefly (ADOBE, 2023). A partir delas, foi criado um conjunto de dados de modelos de tatuagens sintéticas, gerando em média de 3 a 6 modelos para 33 classes diferentes. Dessa forma, obtivemos um conjunto total de 180 imagens de tatuagens.

2.3 Métricas de Avaliação

Para avaliar o desempenho do gerador, foi criado um conjunto de 34.000 imagens com suas máscaras, geradas a partir de 461 imagens e máscaras tratadas, complementadas por 180 modelos de tatuagens sintéticas. No total, os dados abrangem 33 classes de tatuagens, com uma média de 1.100 imagens por classe. A Figura 3 mostra algumas imagens semi-sintéticas geradas, juntamente com as tatuagens sintéticas usadas na criação dessas imagens.

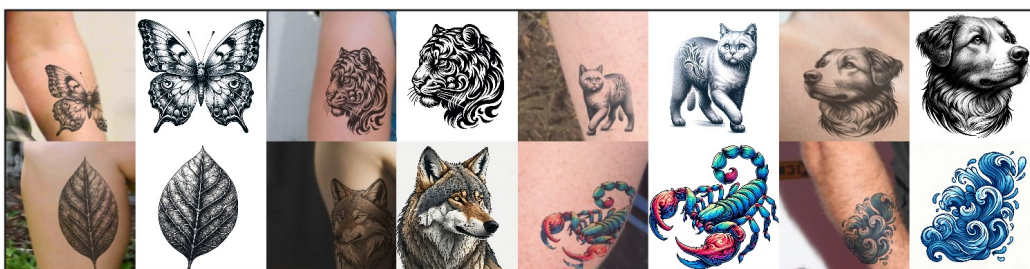


Figura 3 – Imagens semi-sintéticas geradas e suas tatuagens sintéticas. Fonte: Elaborado pelos autores.

Em seguida, dois modelos de redes neurais foram treinados para a avaliação: um utilizando o conjunto semi-sintético e outro utilizando o conjunto de dados reais. Ambos os modelos foram validados e testados em um conjunto de dados reais. O conjunto de dados reais utilizado para o treinamento, validação e teste foi fornecido pelos estudos de [Brilhador et al. \(2024\)](#).

Como métricas de comparação, utilizou-se o Índice de Jaccard (IoU) por classe e o mIoU (IoU médio) para comparar as performances dos modelos. O IoU compara a área de sobreposição entre as previsões do modelo e as anotações reais, refletindo a precisão em uma escala de 0 a 1 para cada classe. O mIoU, ou média do IoU, avalia a consistência da precisão do modelo em todas as classes envolvidas.

O modelo escolhido para essas avaliações foi o SegFormer, uma arquitetura avançada de rede neural que une um Encoder baseado em transformadores com um Decoder MLP (Multilayer Perceptron). Essa combinação permite que o SegFormer processe informações em várias escalas, melhorando a captura de detalhes tanto em grandes áreas quanto em pequenos detalhes, conforme descrito na pesquisa de [Xie et al. \(2021\)](#). Nos experimentos, a rede foi inicializada com os pesos pré-treinados do SegFormer B1.

3 RESULTADOS E DISCUSSÃO

Os resultados do experimento indicaram que a maioria das classes treinadas com o conjunto de dados reais apresentou valores de IoU superiores, como pode ser observado na Figura 4. As classes *branch* e *wolf* foram exceções, tendo melhor desempenho com os dados semi-sintéticos, enquanto classes como *knife*, *crown*, *lion* e *eagle* mostraram valores de IoU semelhantes entre ambos os conjuntos. Por outro lado, classes como *anchor*, *fox*, *gun*, *ribbon*, *rope*, *shark*, *shield*, *skull*, *spider* e *star* obtiveram valores de IoU iguais a zero, sugerindo dificuldades na segmentação dessas classes.

Essas limitações podem ser atribuídas a vários fatores. Primeiro, como as tatuagens são geradas por inteligências artificiais generativas, os prompts nem sempre resultam em imagens ideais, e o processo pode ser demorado. Segundo, devido a restrições de hardware, foi necessário utilizar imagens de 224x224 pixels; essa baixa resolução impede o modelo de capturar detalhes importantes, dificultando o treinamento eficaz. Terceiro, imagens com segmentação precisa da pele são difíceis de obter, o que pode levar a erros na aplicação das tatuagens sintéticas e comprometer o desempenho do modelo. Portanto, os resultados indicam a necessidade de ajustes no gerador de imagens semi-sintéticas e no modelo de rede neural para melhorar a segmentação.

4 CONSIDERAÇÕES FINAIS

Embora algumas classes dos dados semi-sintéticos tenham superado os resultados dos Dados Reais, a maioria ainda requer refinamentos. Melhorias no modelo de rede neural são essenciais para alcançar maiores valores de IoU e mIoU. Adicionalmente, será necessária uma revisão nos modelos de tatuagem das classes com IoU nulo ou muito baixo, assim como aprimoramentos no gerador semi-sintético.

Com essas modificações, o trabalho tem potencial para contribuir significativamente para a área de Segmentação Semântica, utilizando grandes conjuntos de dados semi-sintéticos como uma alternativa para lidar com a escassez de dados reais e melhorar a performance em cenários práticos.

Classe	IoU_DadosReais	IoU_DadosSemiSintéticos	Classe	IoU_DadosReais	IoU_DadosSemiSintéticos
background	0.908	0.872	lion	0.0922	0.0373
anchor	0.1911	0.0	mermaid	0.0732	0.0185
bird	0.135	0.0432	octopus	0.2981	0.065
branch	0.0003	0.0281	owl	0.4071	0.0914
cat	0.1671	0.0412	ribbon	0.1625	0.0
crown	0.1748	0.0747	rope	0.0068	0.0
dog	0.2209	0.0146	scorpion	0.2886	0.0767
eagle	0.1329	0.0601	shark	0.0654	0.0
fire	0.0774	0.0071	shield	0.1731	0.0
fish	0.2425	0.0137	skull	0.1512	0.0
flower	0.2349	0.1059	snake	0.1132	0.0103
fox	0.085	0.0	spider	0.0965	0.0
gun	0.0684	0.0	star	0.0759	0.0
heart	0.1249	0.0189	tiger	0.5356	0.0567
key	0.2739	0.2631	water	0.087	0.0068
knife	0.0242	0.0162	wolf	0.0144	0.0488
leaf	0.1941	0.0057	mIoU	0.1787	0.0593

Figura 4 – Comparação entre Conjunto de Dados Reais e Conjunto de Dados Semi-Sintético. Fonte: Elaborado pelos autores.

AGRADECIMENTOS

Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Laboratório de Inteligência Computacional (LABIC), Universidade Tecnológica Federal do Paraná.

REFERÊNCIAS

- ADOBE. **Adobe Firefly: Overview and Capabilities**. [S.l.: s.n.], 2023. <https://www.adobe.com/products/firefly.html>.
- BRILHADOR, Anderson et al. Open-Set Tattoo Semantic Segmentation. **IEEE Access**, 2024.
- GONZALEZ-SOLER, Lazaro J.; RATHGEB, Christian; FISCHER, Daniel. Semi-synthetic Data Generation for Tattoo Segmentation. **2023 11th International Workshop on Biometrics and Forensics (IWBF)**, IEEE, p. 1–6, 2023.
- NGAN, Mei; GROTH, Patrick. Tattoo Recognition Technology-Challenge (Tatt-C): An Open Tattoo Database for Developing Tattoo Recognition Research. In: PROC. Intl. Conf. on Identity, Security and Behavior Analysis (ISBA 2015). [S.l.: s.n.], 2015. P. 1–6.
- OPENAI. **DALL-E: Creating Images from Text**. [S.l.: s.n.], 2021. <https://openai.com/dall-e>.
- XIE, Enze et al. SegFormer: Simple and Efficient Design for Semantic Segmentation with Transformers. **arXiv preprint arXiv:2105.15203**, 2021. Accepted by NeurIPS 2021.