

Geração de imagens sintéticas de laboratórios de engenharia elétrica utilizando stable diffusion: uma análise da aplicabilidade de modelos de difusão para visualização educacional

Vitor Amadeu Souza¹; 0009-00-02-1857-6799

1 – UniFOA, Centro Universitário de Volta Redonda, Volta Redonda, RJ.
vitor.amadeu@foa.org.br

Resumo: Este trabalho apresenta uma investigação sobre a aplicação de modelos de difusão, especificamente o Stable Diffusion v1.5, para a geração de imagens sintéticas de laboratórios de engenharia elétrica. O objetivo principal foi avaliar a capacidade do modelo em produzir representações visuais realistas de ambientes laboratoriais típicos desta área de conhecimento, incluindo equipamentos como osciloscópios, multímetros, transformadores e circuitos elétricos. A metodologia consistiu na implementação de um pipeline de geração utilizando a biblioteca Diffusers, com processamento em GPU CUDA para otimização computacional. Os resultados demonstraram que o modelo consegue gerar imagens coerentes e tecnicamente apropriadas de laboratórios de engenharia elétrica, apresentando elementos característicos do ambiente acadêmico e profissional da área. A análise qualitativa das imagens geradas revela potencial significativo para aplicações educacionais, desenvolvimento de materiais didáticos e visualização de conceitos técnicos. As conclusões indicam que modelos de difusão representam uma ferramenta promissora para a criação de conteúdo visual especializado em engenharia, oferecendo alternativas inovadoras para a educação tecnológica.

Palavras-chave: Stable Diffusion. Inteligência Artificial. Engenharia Elétrica. Geração de Imagens. Modelos de Difusão. Educação Tecnológica.

INTRODUÇÃO

A evolução dos modelos de inteligência artificial generativa tem revolucionado diversos campos do conhecimento, proporcionando novas possibilidades para a criação de conteúdo visual automatizado (Rombach *et al.*, 2022). Entre as tecnologias emergentes, os modelos de difusão destacam-se pela capacidade de gerar imagens de alta qualidade a partir de descrições textuais, estabelecendo um novo paradigma na síntese de imagens (Ho; Jain; Abbeel, 2020). O Stable Diffusion, desenvolvido pela Stability AI em colaboração com a RunwayML, representa um marco significativo neste domínio, oferecendo um modelo de código aberto capaz de produzir imagens fotorrealistas com precisão semântica notável (Rombach *et al.*, 2022).

No contexto educacional, particularmente na engenharia elétrica, a visualização de conceitos técnicos e ambientes laboratoriais constitui um elemento fundamental para o processo de ensino-aprendizagem (Felder; Silverman, 1988). A disponibilidade de recursos visuais adequados pode significativamente impactar a compreensão de estudantes sobre equipamentos, procedimentos e configurações laboratoriais típicas da área (Kolb, 1984). Tradicionalmente, a criação de materiais visuais para fins educacionais demanda recursos consideráveis, incluindo acesso a laboratórios físicos, equipamentos fotográficos especializados e tempo substancial para produção (Clark; Mayer, 2016).

O presente trabalho surge da necessidade de investigar sistematicamente as capacidades dos modelos de difusão na geração de imagens especializadas para engenharia elétrica, avaliando tanto a qualidade técnica quanto a adequação pedagógica das representações visuais produzidas. A hipótese central baseia-se na premissa de que modelos como o Stable Diffusion, treinados em vastos conjuntos de dados visuais, possuem conhecimento implícito suficiente sobre equipamentos e ambientes de engenharia elétrica para gerar representações úteis e pedagogicamente relevantes.

Estudos anteriores sobre aplicações educacionais de inteligência artificial generativa têm focado predominantemente em disciplinas como arte, literatura e ciências sociais (Rudolph *et al.*, 2023). A investigação específica sobre aplicações em engenharia, particularmente na

geração de conteúdo visual técnico, permanece relativamente inexplorada, constituindo uma lacuna significativa na literatura científica atual.

MÉTODOS

A implementação do sistema de geração de imagens baseou-se na utilização do modelo Stable Diffusion v1.5, disponibilizado pela RunwayML através da plataforma Hugging Face (Wolf *et al.*, 2020). A escolha deste modelo específico fundamentou-se em sua ampla adoção pela comunidade científica e na disponibilidade de documentação técnica abrangente (Rombach *et al.*, 2022). O pipeline de geração foi implementado utilizando a biblioteca Diffusers, desenvolvida pela Hugging Face, que oferece uma interface padronizada para diversos modelos de difusão (Von Platen *et al.*, 2022). Esta biblioteca proporciona abstração adequada dos detalhes implementacionais complexos dos modelos de difusão, permitindo foco nos aspectos aplicativos da pesquisa.

O desenvolvimento do prompt textual constituiu elemento vital da metodologia, dado que a qualidade e especificidade das descrições textuais impactam diretamente a adequação das imagens geradas (Liu *et al.*, 2023). O prompt utilizado "An electrical engineering laboratory, with circuits, oscilloscopes, multimeters and transformers", foi elaborado considerando terminologia técnica específica da engenharia elétrica, incluindo equipamentos fundamentais comumente encontrados em laboratórios acadêmicos e industriais. A seleção dos equipamentos mencionados no prompt baseou-se em literatura especializada sobre laboratórios de engenharia elétrica e nas diretrizes curriculares nacionais para cursos de graduação em engenharia (Brasil, 2019).

O processo de geração seguiu o pipeline padrão do Stable Diffusion, iniciando com a codificação do prompt textual através do modelo CLIP (Contrastive Language-Image Pre-training), responsável pela conversão da descrição textual em representação vetorial adequada para orientar o processo de difusão (Radford *et al.*, 2021). Subsequentemente, o modelo de difusão propriamente dito realizou o processo iterativo de remoção de ruído, refinando progressivamente uma representação inicial aleatória até convergir para uma imagem coerente com a descrição fornecida.

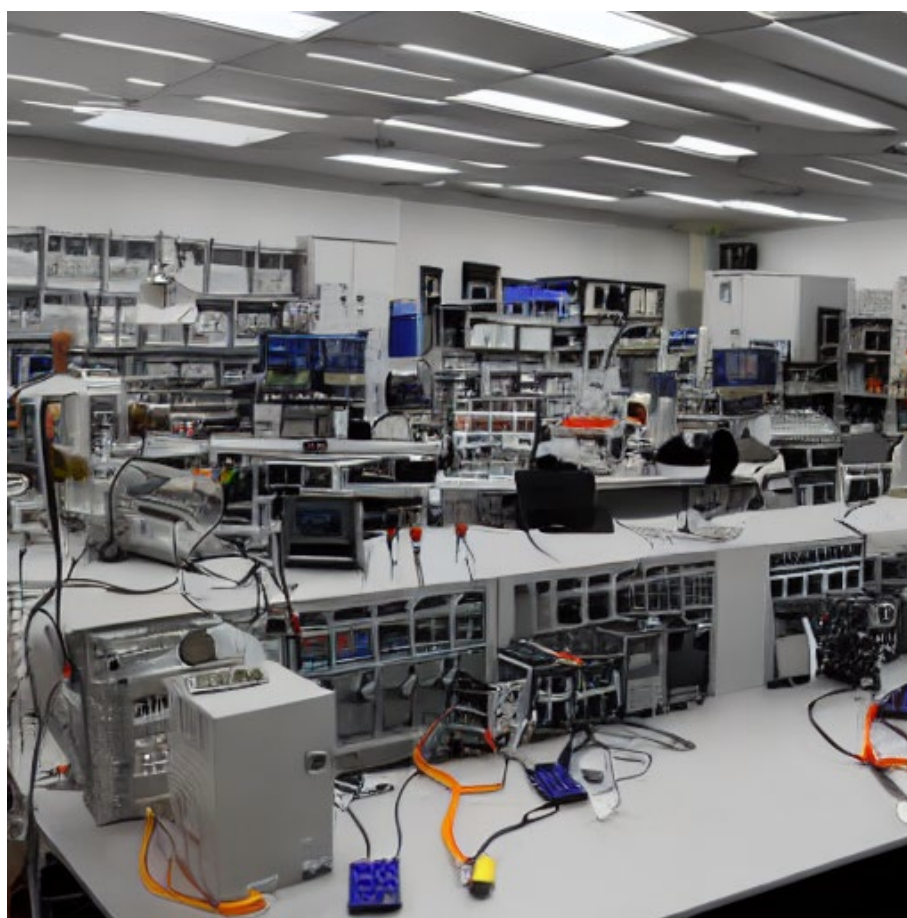


O código-fonte está disponível para download através do link: <https://github.com/vitor-souza-ime/sd>.

RESULTADOS E DISCUSSÃO

A execução do código implementado resultou na geração de uma imagem sintética de laboratório de engenharia elétrica que demonstra capacidades do modelo Stable Diffusion v1.5 para aplicações técnicas especializadas. A imagem produzida (Figura 1) apresenta características visuais consistentes com ambientes laboratoriais reais, incluindo elementos arquitetônicos típicos como iluminação fluorescente, mobiliário técnico e organização espacial apropriada.

Figura 1 - Imagem gerada por stable diffusion



Fonte: O autor.

A análise qualitativa da imagem revela presença clara de equipamentos eletrônicos diversos, dispostos sobre bancadas de trabalho em configuração típica de laboratórios educacionais. Particularmente notável é a capacidade do modelo em gerar uma representação espacial coerente do ambiente laboratorial, incluindo elementos como armários de armazenamento, estações de trabalho individuais e infraestrutura de suporte técnico. Esta organização espacial reflete padrões estabelecidos em normas de segurança e ergonomia para laboratórios de engenharia elétrica (IEC, 2019).

A qualidade fotorrealística da imagem gerada demonstra evolução significativa das capacidades dos modelos generativos em comparação com tecnologias anteriores baseadas em redes adversariais generativas (GANs) (Goodfellow et al., 2014). A ausência de artefatos visuais óbvios e a manutenção de consistência global na composição da cena indicam maturidade técnica adequada para aplicações práticas.

Do ponto de vista pedagógico, a imagem apresenta potencial considerável para aplicações educacionais. Estudantes iniciantes em engenharia elétrica podem beneficiar-se da exposição prévia a representações visuais de ambientes laboratoriais, facilitando a posterior adaptação aos espaços físicos reais (Bloom, 1956).

Contudo, a análise crítica também revela limitações importantes. Embora a imagem apresente equipamentos com características visuais apropriadas, detalhes técnicos específicos como ausência de alguns elementos fornecidos através do prompt não são suficientemente detalhados para substituir completamente recursos educacionais convencionais. Esta limitação é consistente com as características dos modelos de difusão, que priorizam coerência visual global em detrimento de precisão técnica detalhada (Dhariwal; Nichol, 2021).

A comparação com recursos visuais tradicionalmente utilizados em educação em engenharia revela vantagens e desvantagens distintas. Fotografias reais de laboratórios oferecem precisão técnica superior e fidelidade completa aos equipamentos específicos, mas apresentam limitações de disponibilidade, custo de produção e flexibilidade de customização (Mayer, 2001). Por outro lado, as imagens sintéticas proporcionam

flexibilidade significativa para adaptação a contextos específicos, geração sob demanda e personalização de acordo com objetivos pedagógicos particulares.

Estudos comparativos recentes sobre aplicações educacionais de inteligência artificial generativa sugerem que a eficácia pedagógica destes recursos depende fundamentalmente da integração adequada com metodologias de ensino estabelecidas (Holmes *et al.*, 2023). A utilização isolada de imagens sintéticas não substitui experiências práticas essenciais, mas pode complementar significativamente recursos educacionais tradicionais quando aplicada estrategicamente.

As implicações para desenvolvimento de materiais didáticos são substanciais. Editoras educacionais e instituições de ensino podem beneficiar-se da capacidade de gerar ilustrações técnicas customizadas com custos reduzidos e flexibilidade temporal significativa. A possibilidade de adaptar representações visuais para contextos culturais específicos ou configurações institucionais particulares representa vantagem competitiva importante.

CONCLUSÕES

Este estudo demonstrou a viabilidade técnica e o potencial pedagógico da aplicação de modelos de difusão, especificamente o Stable Diffusion v1.5, para geração de imagens sintéticas de laboratórios de engenharia elétrica. Os resultados obtidos confirmam que modelos generativos contemporâneos possuem capacidades para produzir representações visuais tecnicamente adequadas e pedagogicamente relevantes de ambientes especializados em engenharia.

A qualidade fotorrealística das imagens geradas, combinada com a presença de equipamentos tecnicamente apropriados e organização espacial coerente, estabelece fundamentos para aplicações educacionais práticas. A capacidade do modelo em interpretar terminologia técnica específica e traduzi-la em representações visuais demonstra evolução das tecnologias de inteligência artificial generativa.

Estudos futuros podem investigar aplicações específicas em diferentes subdisciplinas da engenharia elétrica, avaliação quantitativa da eficácia pedagógica através de estudos

controlados com estudantes, e desenvolvimento de metodologias de integração com recursos educacionais existentes. A exploração de técnicas de fine-tuning para especialização em domínios técnicos específicos também representa direção promissora para pesquisas subsequentes.

A contribuição principal deste trabalho reside na demonstração prática da aplicabilidade de modelos de difusão para geração de conteúdo educacional especializado em engenharia, estabelecendo precedente para investigações mais amplas sobre aplicações de inteligência artificial generativa na educação tecnológica.

REFERÊNCIAS

BLOOM, B. S. Taxonomy of educational objectives: The classification of educational goals. New York: Longmans, Green, 1956.

BRASIL. Ministério da Educação. Diretrizes Curriculares Nacionais do Curso de Graduação em Engenharia. Resolução CNE/CES nº 2, de 24 de abril de 2019.

CLARK, R. C.; MAYER, R. E. E-learning and the science of instruction: Proven guidelines for consumers and designers of multimedia learning. 4. ed. Hoboken: John Wiley & Sons, 2016.

DHARIWAL, P.; NICHOL, A. Diffusion models beat GANs on image synthesis. Advances in Neural Information Processing Systems, v. 34, p. 8780-8794, 2021.

FELDER, R. M.; SILVERMAN, L. K. Learning and teaching styles in engineering education. Engineering Education, v. 78, n. 7, p. 674-681, 1988.

GOODFELLOW, I. et al. Generative adversarial nets. Advances in Neural Information Processing Systems, v. 27, p. 2672-2680, 2014.

HO, J.; JAIN, A.; ABBEEL, P. Denoising diffusion probabilistic models. Advances in Neural Information Processing Systems, v. 33, p. 6840-6851, 2020.

HOLMES, W. et al. Artificial intelligence in education: promises and implications for teaching and learning. Educational Technology Research and Development, v. 71, n. 2, p. 1-20, 2023.

IEC - INTERNATIONAL ELECTROTECHNICAL COMMISSION. IEC 61010-1: Safety requirements for electrical equipment for measurement, control, and laboratory use. Geneva: IEC, 2019.

KOLB, D. A. Experiential learning: Experience as the source of learning and development. Englewood Cliffs: Prentice-Hall, 1984.

LIU, V. et al. Design guidelines for prompt engineering text-to-image generative models. Proceedings of the CHI Conference on Human Factors in Computing Systems, p. 1-23, 2023.

MAYER, R. E. Multimedia learning. Cambridge: Cambridge University Press, 2001.

RADFORD, A. et al. Learning transferable visual models from natural language supervision. International Conference on Machine Learning, p. 8748-8763, 2021.

ROMBACH, R. et al. High-resolution image synthesis with latent diffusion models. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, p. 10674-10685, 2022.

RUDOLPH, J. et al. War of the chatbots: Bard, Bing Chat, ChatGPT, Ernie and beyond. The new AI gold rush and its impact on higher education. Journal of Applied Learning and Teaching, v. 6, n. 1, p. 364-389, 2023.

VON PLATEN, P. et al. Diffusers: State-of-the-art diffusion models. GitHub Repository, 2022. Disponível em: <https://github.com/huggingface/diffusers>. Acesso em: 15 jan. 2025.

WOLF, T. et al. Transformers: State-of-the-art natural language processing. Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: System Demonstrations, p. 38-45, 2020.