



Fundação Universidade Federal do ABC

Pró reitoria de pesquisa

Av. dos Estados, 5001, Santa Terezinha, Santo André/SP, CEP 09210-580

Bloco L, 3º Andar, Fone (11) 3356-7617

iniciacao@ufabc.edu.br

Projeto de Iniciação Científica submetido para avaliação no Edital: 04/2022

Título do projeto: Uso de computação gráfica na criação de datasets sintéticos para treinamento de modelos de inteligência artificial.

Palavras-chave do projeto:

dataset sintético, computação gráfica, inteligência artificial, aprendizado de máquina

Área do conhecimento do projeto:

Computação Gráfica, Inteligência Artificial, Desenvolvimento de software

Sumário

1	Resumo	3
2	Introdução	3
3	Objetivos	4
4	Metodologia e Viabilidade	5
5	Cronograma de atividades	6

1 Resumo

Este projeto propõe a concepção e desenvolvimento de uma ferramenta para a geração de datasets visuais a partir do uso intensivo de técnicas de computação gráfica. A ferramenta a ser desenvolvida deverá ser capaz de gerar imagens e informações semânticas associadas a estas imagens. Como resultado, este conjunto de dados poderá ser utilizados em projetos correlatos para treinamento de modelos de inteligência artificial, entre eles modelos de redes neurais artificiais, redes de convolução, redes generativas artificiais, entre outras. O projeto deve capacitar o aluno no desenvolvimento de software e introduzir propriedades consideradas essências em datasets sintéticos para inteligência artificial. O software desenvolvido fará uso de recursos disponíveis e de acesso livre, tal como geometrias 3D, texturas, softwares de modelagens 3D e motores gráficos 3D a fim de sintetizar cenas 3D, que serão posteriormente convertidas em bases de imagens e dados descritivos. Ao fim do projeto o software será documentado e disponibilizado para uso em futuras pesquisas.

2 Introdução

Nos últimos anos, houve um grande avanço científico na aplicação de métodos de inteligência artificial, em particular no emprego de técnicas de aprendizado de máquina para a resolução de problemas em múltiplas áreas do conhecimento. As aplicações de métodos de aprendizado de máquina extrapolam as áreas de computação e engenharias e se estendem a áreas como medicina [5], biologia [7], bioquímica [6], física [4], sociologia [1], psicologia [10], entre outras. Esta característica multidisciplinar se deve principalmente a capacidade de generalização proporcionada pelos modelos de aprendizado de máquina.

Muitas destas aplicações utilizam modelos que empregam o chamado aprendizado supervisionado. Nesta metodologia os modelos de aprendizado de máquina são construídos e treinados para uma aplicação específica a partir de exemplos, onde um exemplo é constituído de uma tupla (x, y) onde x é uma amostra e y é um rótulo que caracteriza esta amostra. Um conjunto de exemplos constitui um *dataset* [3].

A partir de exemplos conhecidos presentes no *dataset*, o modelo é ajustado de maneira a aprender a distribuição estatística e as características inerentes aos dados. Um modelo treinado é capaz de generalizar o conhecimento sendo capaz de reconhecer e identificar o padrão de novos dados, que não foram vistos anteriormente no treinamento.

Como os modelos que utilizam aprendizado supervisionado dependem de *datasets* específicos a suas aplicações para extrair o conhecimento durante a fase de treinamento. Estes dados são de extrema importância para o desenvolvimento destas aplicações. Um

desafio recorrente a este tipo de abordagem é a disponibilidade de dados adequados a aplicação. Os dados devem possuir características que descrevam uma boa representação do problema, sendo necessário inevitavelmente uma grande quantidade de dados.

A primeira dificuldade surge em obter dados em determinados domínios, muitas vezes não há dados públicos disponíveis e obter tais dados pode ser um trabalho árduo e até mesmo impraticável. Um exemplo prático pode ser dado para a área de medicina, no reconhecimento de características de uma doença rara, onde há poucas amostras e obter novas amostras pode incorrer em altos custos. Mesmo quando há grande quantidade de amostras para uma determinada aplicação, outro problema recorrente é a atribuição de rótulos a essas amostras.

Valan et al. [13] cita em seu trabalho o desafio de coletar amostras em quantidades suficientes de espécimes de insetos e a identificação de suas características, preservando a integridade das mesmas, processo ao qual demanda muito tempo e recursos.

Para problemas onde as amostras são imagens, a atribuição de rótulos pode abranger desde identificação de objetos ilustrados na imagem em uma lista pré-determinada de objetos possíveis, até casos mais complexos de atribuir características que descrevam cada pixel da imagem. Nestes últimos casos, a atribuição de rótulos a uma única imagem pode demorar um tempo considerável, se tornando um problema para *datasets* contendo milhares de exemplos.

A síntese de amostras e anotação utilizando técnicas de computação gráfica é uma forma de contornar a aquisição manual de amostras e rotulação das mesmas. A partir de um ambiente virtual 3D é possível criar imagens e vídeos sintéticos, onde há grande controle sobre o ambiente. Esta abordagem vem sendo utilizada por pesquisadores a fim de gerar grande quantidade de dados rotulados e se mostrou eficaz na generalização do aprendizado quando aplicados em métodos como redes neurais artificiais, redes de convolução, redes adversariais, entre outras. [8, 9, 12, 14]

Em um ambiente 3D podemos controlar aspectos como posição da câmera e transformações de posição, escala e rotação de modelos 3D, além de rótulos por pixel de informações espaciais e semânticas, tais como profundidade, material, classe de objetos, entre outros. Tais características influenciam diretamente a qualidade e usabilidade dos dados no treinamento de redes neurais artificiais e outras técnicas de aprendizado de máquina.

3 Objetivos

O aluno terá contato com a área de síntese de imagens, que é parte integral da computação gráfica. Além de adquirir experiência no desenvolvimento de software, o aluno

deve se familiarizar com aspectos importantes dos *datasets* utilizados em aprendizado de máquina e desenvolver um senso crítico em relação a características importantes destes *datasets*.

Os objetivos deste projeto encontram-se elencados nos seguintes itens:

- O desenvolvimento de uma ferramenta de computação gráfica para a criação de *datasets* sintéticos, onde o usuário possui um alto controle sobre o ambiente e objetos estudados.
- Busca, pesquisa e obtenção de recursos, tais como modelos 3D e imagens de texturas, que serão utilizados na ferramenta desenvolvida.
- Criação de *datasets* sintéticos de vídeos e imagens que serão aplicados em outros projetos de pesquisa.
- Familiarização e capacitação do aluno com o desenvolvimento de software multidisciplinar.

4 Metodologia e Viabilidade

O aluno devera se familiarizar com os conceitos de computação gráfica para a síntese de imagens, entre elas a manipulação de objetos e malhas 3D, modelos de preenchimento e iluminação, e sistemas de animação. Para isso, será realizado um estudo teórico sobre o tema. Não obstante, o aluno deverá realizar uma pesquisa bibliográfica dos *datasets* sintéticos existentes, e a metodologia utilizada para gerar tais *datasets*.

Além do estudo teórico, será realizado um trabalho de busca e pesquisa de recursos digitais para auxiliar no desenvolvimento e aplicação da ferramenta desenvolvida. Estes recursos incluem modelos 3D, imagens utilizadas como texturas, imagens em HDR, dados de animações, entre outros.

O software principal será desenvolvido utilizando o motor gráfico do software Unity [11]. O Unity é um software popular de desenvolvimento de jogos disponível gratuitamente. O aluno deverá explorar as capacidades gráficas e de manipulação geométrica deste software, tais como criação e edição de malhas 3D, transformações geométricas, edição de materiais e criação de animações em 3D.

Dentre as funcionalidades planejadas para a ferramenta, estão o desenvolvimento dos seguintes itens:

- manipular e criar variações de materiais;

- manipular e criar variações de ambiente através de mapas de ambientes;
- manipular e criar variações de geometria em um modelo 3D;
- manipular e criar variações de animação em um modelo 3D;
- manipular e gerar imagens que representam profundidade, mapas de normais, vetores de movimento;
- gerar imagens com informações semânticas de objetos;
- manipular características e posição da câmera;
- manipular e criar efeitos que simulem desenho ou destaquem características importantes dos modelos;
- manipular e criar efeitos de *shader* [2] que simulem características de imagens capturadas em ambientes reais, tais como desfoque, brilho, ruído.

Será utilizado a linguagem de programação C# para o desenvolvimento de componentes no software Unity. Eventuais scripts poderão ser desenvolvidos em outras linguagens de programação, tais como Python e C++, a fim de auxiliar no pré-processamento de dados. Durante o processo de desenvolvimento será utilizado softwares de controle de versão, tal como Git, e realizado o processo de documentação do software de maneira adequada.

Por fim, o aluno terá contato com aplicações multidisciplinares, em vista que a ferramenta desenvolvida deverá ser utilizada em projetos correlatos que envolvem o treinamento de redes neurais artificiais para tratar de aplicações onde não há datasets disponíveis publicamente.

5 Cronograma de atividades

As atividades deste projeto estão descritas nos itens abaixo:

1. Estudo preliminar bibliográfico e obtenção dos dados
 - (a) Estudo bibliográfico de datasets sintéticos
 - (b) Estudo do software Unity
 - (c) Obtenção de modelos 3D, texturas e outros recursos computacionais
2. Desenvolvimento da ferramenta

- (a) Desenvolvimento de variações geométricas, transformações de câmera, mapas de ambiente.
- (b) Desenvolvimento de variações relativas a materiais, efeitos de *shader* e iluminação.
- (c) Desenvolvimento de interface gráfica com o usuário.

3. Aplicação da ferramenta

- (a) Criação de datasets de exemplos
- (b) Aplicação em um problema de aprendizado de máquina

4. Escrita

- (a) Relatórios e eventuais artigos científicos
- (b) Documentação do software
- (c) Distribuição do software em um repositório

Etapa	Mês											
	01	02	03	04	05	06	07	08	09	10	11	12
1 (a)	x	x	x	x								
1 (b)		x	x	x	x							
1 (c)				x	x							
2 (a)		x	x	x	x	x						
2 (b)				x	x	x	x	x	x			
2 (c)						x	x	x	x	x		
3 (a)										x	x	
3 (b)											x	x
4 (a)										x	x	x
4 (b)								x	x	x	x	x
4 (c)												x

Referências

- [1] Giovanni Di Franco and Michele Santurro. Machine learning, artificial neural networks and social research. *Quality & quantity*, 55(3):1007–1025, 2021.
- [2] John P Doran. *Unity 2021 Shaders and Effects Cookbook*. Packt Publishing, Birmingham, England, 4 edition, October 2021.
- [3] Ian Goodfellow, Yoshua Bengio, and Aaron Courville. *Deep Learning*. MIT Press, 2016. <http://www.deeplearningbook.org>.
- [4] George Em Karniadakis, Ioannis G Kevrekidis, Lu Lu, Paris Perdikaris, Sifan Wang, and Liu Yang. Physics-informed machine learning. *Nature Reviews Physics*, 3(6):422–440, 2021.
- [5] Mike May. Eight ways machine learning is assisting medicine. *Nat Med*, 27(1):2–3, 2021.
- [6] Max Mowbray, Thomas Savage, Chufan Wu, Ziqi Song, Bovinille Anye Cho, Ehecatl A Del Rio-Chanona, and Dongda Zhang. Machine learning for biochemical engineering: A review. *Biochemical Engineering Journal*, 172:108054, 2021.
- [7] Alexander Popkov, Fedor Konstantinov, Vladimir Neimorovets, and Alexey Solodovnikov. Machine learning for expert-level image-based identification of very similar species in the hyperdiverse plant bug family miridae (hemiptera: Heteroptera). *Systematic Entomology*, 47(3):487–503, 2022.
- [8] Mike Roberts, Jason Ramapuram, Anurag Ranjan, Atulit Kumar, Miguel Angel Bautista, Nathan Paczan, Russ Webb, and Joshua M. Susskind. Hypersim: A photorealistic synthetic dataset for holistic indoor scene understanding. In *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pages 10912–10922, October 2021.
- [9] German Ros, Laura Sellart, Joanna Materzynska, David Vazquez, and Antonio M Lopez. The synthia dataset: A large collection of synthetic images for semantic segmentation of urban scenes. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 3234–3243, 2016.
- [10] Hannes Rosenbusch, Felix Soldner, Anthony M Evans, and Marcel Zeelenberg. Supervised machine learning methods in psychology: A practical introduction with annotated r code. *Social and Personality Psychology Compass*, 15(2):e12579, 2021.

- [11] Unity Technologies. *Unity Real-time Development Engine*. Unity, 2022. <https://unity.com/>.
- [12] Jonathan Tremblay, Thang To, and Stan Birchfield. Falling things: A synthetic dataset for 3d object detection and pose estimation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 2038–2041, 2018.
- [13] Miroslav Valan, Dominik Vondráček, and Fredrik Ronquist. Awakening a taxonomist’s third eye: exploring the utility of computer vision and deep learning in insect systematics. *Systematic Entomology*, 46(4):757–766, 2021.
- [14] Gul Varol, Javier Romero, Xavier Martin, Naureen Mahmood, Michael J. Black, Ivan Laptev, and Cordelia Schmid. Learning from synthetic humans. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, July 2017.