

## CCM-311: Tópicos em Computação Gráfica, Visão Computacional e Processamento de Imagens

**Docente: João Paulo Gois**

**Q1.2025**

**Ementa:**

Revisão do estado da arte em Computação Gráfica, Processamento de Imagens e Visão Computacional, destacando avanços recentes e suas aplicações. Estudo de *Visual Transformers*, com foco em arquiteturas inovadoras e suas utilizações em tarefas de visão computacional. Exploração de *Diffusion Models*, abordando suas arquiteturas e aplicações, incluindo técnicas avançadas para animação 3D de personagens humanos. Flow Matching e suas aplicações.

Análise de redes generativas adversariais (GANs), com ênfase nas *StyleGANs* e seu impacto na geração de imagens realistas. Discussão sobre *Neural Radiance Fields* (NeRF) para síntese de cenas, com aplicações na visualização e representação tridimensional de ambientes.

Atividades práticas individuais ou em grupo no laboratório computacional. Apresentação dos resultados e produção de relatórios técnicos.

### Referências bibliográficas

Ansari, S. (2023). *Building Computer Vision Applications Using Artificial Neural Networks: With Examples in OpenCV and TensorFlow with Python* (Latest ed.). Publisher. <https://link.springer.com/book/10.1007/978-1-4842-9866-4>

Croitoru, F.-A., Hondru, V., Ionescu, R. T., & Shah, M. (2023). Diffusion Models in Vision: A Survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 45(9), 3214-3230. <https://doi.org/10.1109/TPAMI.2023.3050472>

Dosovitskiy, A., Beyer, L., Kolesnikov, A., Weissenborn, D., Zhai, X., Unterthiner, T., Dehghani, M., Minderer, M., Heigold, G., Gelly, S., Uszkoreit, J., & Houlsby, N. (2020). An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale. *CoRR*, *abs/2010.11929*. Retrieved from <https://arxiv.org/abs/2010.11929>

Hu, V. T., Yin, W., Ma, P., Chen, Y., Fernando, B., Asano, Y. M., Gavves, E., Mettes, P., Ommer, B., & Snoek, C. G. M. (2023). Motion flow matching for human motion synthesis and editing. <https://arxiv.org/abs/2312.08895>

Karras, T., Laine, S., & Aila, T. (2018). A Style-Based Generator Architecture for Generative Adversarial Networks. *CoRR*, *abs/1812.04948*. Retrieved from <http://arxiv.org/abs/1812.04948>

Khan, S., Naseer, M., Hayat, M., Zamir, S. W., Khan, F. S., & Shah, M. (2021). Transformers in Vision: A Survey. *ACM Computing Surveys (CSUR)*, 54(10s), Article 200, 1-41. <https://doi.org/10.1145/3505244>

LIPMAN, Yaron; HAVASI, Marton; HOLDERRIETH, Peter; SHAUL, Neta; LE, Matt; KARRER, Brian; CHEN, Ricky T. Q.; LOPEZ-PAZ, David; BEN-HAMU, Heli; GAT, Itai. Flow Matching Guide and Code. 2024. Disponível em: <https://arxiv.org/abs/2412.06264>. Acesso em: 10 dez. 2024.

Mildenhall, B., Srinivasan, P. P., Tancik, M., Barron, J. T., Ramamoorthi, R., & Ng, R. (2020). NeRF: Representing Scenes as Neural Radiance Fields for View Synthesis. *CoRR*, abs/2003.08934. Retrieved from <https://arxiv.org/abs/2003.08934>

Nagasaka, H., & Izuhara, M. (2021). Interactive Visualization of Deep Learning Models in an Immersive Environment. In *Proceedings of the 27th ACM Symposium on Virtual Reality Software and Technology (VRST '21)*, Article 82. Association for Computing Machinery. <https://doi.org/10.1145/3489849.3489956>

Tevet, G., Raab, S., Gordon, B., Shafir, Y., Cohen-Or, D., & Bermano, A. H. (2023). Human motion diffusion model. In *The Eleventh International Conference on Learning Representations*. <https://arxiv.org/abs/2209.14916>

Zhou, W., Dou, Z., Cao, Z., Liao, Z., Wang, J., Wang, W., Liu, Y., Komura, T., Wang, W., & Liu, L. (2024). EMDM: Efficient Motion Diffusion Model for Fast and High-Quality Motion Generation. <https://arxiv.org/abs/2312.02256>