

Génération automatique de comportements sociaux de personnages virtuels à partir de réseaux adverses génératifs (GAN)

Supervision : [Magalie Ochs](#) (TALEP - Pôle Sciences des données – LIS) – Stéphane Ayache (QARMA - Pôle Science des données – LIS)

L'objectif de ce stage de master est de mettre en place une architecture GAN permettant la génération automatique de comportements multimodaux socio-émotionnels pour des personnages virtuels.

Ce stage s'intègre dans plusieurs projets en cours visant la simulation d'attitudes sociales sur des acteurs virtuels (e.g. persuasif, agressif, conciliant) : le projet TRUENESS qui vise à développer une plateforme de réalité virtuelle peuplée de personnages virtuel pour former les individus à lutter contre la discrimination sociale (génée et ethnique) et le projet ANR COPAINS qui vise à développer un personnage virtuel persuasif pour inciter les personnes âgées à faire du sport.

Les réseaux adversaires génératifs (GANs) (Goodfellow et al., 2014) sont reconnus aujourd'hui comme des techniques d'apprentissage particulièrement puissantes pour la génération de nouvelle données, même si l'ensemble de données de départ est relativement restreint (Karras et al., 2020). L'efficacité des architectures GANs pour la génération automatique de comportements multimodaux a été très récemment démontrée (Ahuja et al., 2020 ; Habibie et al., 2021). En s'inspirant de ces travaux, dans ce projet de stage, l'objectif est d'utiliser des modèles GANs pour générer de nouvelles données de comportements socio-émotionnels. Une attention particulière sera donnée aux architectures GANs conditionnels (Conditional GANs – CGANs) (Mirza et al., 2014) pour la génération des comportements avec des étiquettes de différentes attitudes sociales. Différentes architectures de réseaux GANs seront explorées (e.g. GAN autoencoder, Multimodal attention mechanism). Différents corpus de données seront exploités pour la construction du modèle (corpus TRUENESS, Corpus POM, corpus POTUS). Les signaux multimodaux des individus (e.g. prosodie, expressions faciales, mouvements de têtes) dans les corpus ont été annotés automatiquement. Les performances de génération des comportements des différents réseaux GANs construits seront évaluées à travers des tests perceptifs utilisateurs : les animations générées seront présentées à des utilisateurs qui devront juger du naturel et de l'expressivité des comportements multimodaux.

Références

Ahuja, C., Lee, D. W., Ishii, R., & Morency, L. P. (2020, November). No Gestures Left Behind: Learning Relationships between Spoken Language and Freeform Gestures. In Proceedings of the 2020 Conference on Empirical Methods in Natural Language Processing: Findings (pp. 1884-1895).

HABIBIE, Ikhsanul, XU, Weipeng, MEHTA, Dushyant, et al. Learning Speech-driven 3D Conversational Gestures from Video. arXiv preprint arXiv:2102.06837, 2021.

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... & Bengio, Y. (2014). Generative adversarial nets. *Advances in neural information processing systems*, 27.

Karras, T., Aittala, M., Hellsten, J., Laine, S., Lehtinen, J., & Aila, T. (2020). Training generative adversarial networks with limited data. *arXiv preprint arXiv:2006.06676*.

Mirza, M., & Osindero, S. (2014). Conditional generative adversarial nets. arXiv preprint arXiv:1411.1784.
