

Lieu du stage

ISIR — Institut des
Systèmes Intelligents
et de Robotique

4, place Jussieu
75005 Paris

À contacter

Kévin Bailly
kevin.bailly@isir.upmc.fr

Arnaud Dapogny
ad@datakalab.com

Génération et manipulation de visages pour la spécialisation de modèles d'analyse faciale avec peu de données

Mots clés: Analyse faciale, peu de données, spécification, reciblage, synthèse de visage.

Contexte

Les progrès récents en apprentissage statistique et, en particulier, en apprentissage profond ont permis des avancées significatives des méthodes d'analyse faciales. Si certaines tâches telles que la localisation de points caractéristiques [1] ou la reconnaissance de visage [2] surpassent les compétences humaines, il subsiste des domaines de l'analyse faciales sensibles aux fortes variations d'apparence entre individus [3], [4]. Ainsi, pour reconnaître une expression faciale, un observateur humain doit s'adapter rapidement à l'identité de la personne et à la particularité de ses traits morphologiques.

Afin de reproduire un tel comportement dans un modèle d'apprentissage profond, on propose de tirer parti du domaine de la génération et de la manipulation de visages synthétiques [5]–[8] afin de générer des exemples à partir d'une image de la personne à analyser (ou d'une courte séquence vidéo du visage de cette personne). Pour y arriver, les méthodes de l'état de l'art prennent généralement en entrée une image source et un ensemble de points caractéristiques et génèrent en sortie une image de ce visage qui conserve l'identité de la personne tout en appliquant la pose et l'expression décrite par les points caractéristiques sources.

De cette manière, nous souhaitons générer un ensemble de données synthétiques spécifique à une personne, avec des configurations de pose et d'expression variées en nous appuyant sur des méthodes de manipulation de visage avec différentes configurations de points caractéristiques. Nous souhaitons dans un second temps utiliser cette base de données pour ajuster un modèle de reconnaissance d'expressions faciales et d'Action Units afin d'obtenir un détecteur spécifique à une identité capable de généraliser pour un grand nombre de configurations.

Objectifs du stage

L'objectif principal de ce stage est d'étudier les méthodes de génération et de manipulation de visage et en identifier les limites avant de proposer des solutions originales pour lever ces verrous scientifiques, et ainsi améliorer la qualité des images générées (qui peut être évaluée par des métriques telles que le score inception ou la distance de Fréchet). En particulier, quelques Actions Units telles que l'AU6 (remontée des joues) ou 17 (remontée du menton) ne sont pas bien décrites par le déplacements de quelques points caractéristiques sur le visage. Une attention particulière sera apportée à l'intégration de cette information sémantique dans la génération de ces exemples d'apprentissage.

En fonction de l'avancée du stage, ces méthodes de manipulation de visage pourront être utilisées pour générer des exemples équilibrés en terme de pose et d'expression (ou d'activation d'AU) à partir d'une simple photo d'une personne (ou d'une courte séquence vidéo de cette personne). Ces exemples synthétiques pourront alors être utilisés pour ré-entraîner ou ajuster un système d'analyse faciale (par exemple de reconnaissance d'expression ou d'action units). Un intérêt particulier sera porté aux méthodes issues des communautés de l'apprentissage sans ou avec peu d'exemples et l'apprentissage en continue (cela permettra de trouver un compromis entre rigidité et plasticité du système proposé).

Profil et compétences recherchées

Edudiant de Master ou Grande École. Compétences requises :

- Machine Learning / Deep Learning
- Vision par ordinateur
- Programmation Python et librairie deep learning (tensorflow ou pytorch)

- Excellentes capacités relationnelles et rédactionnelles (français et anglais)

Modalités de candidature

Pour postuler à ce stage, le candidat est invité à communiquer par mail à ad@datakalab.com et kb@datakalab.com :

- Son CV
- Ses résultats académiques des deux dernières années universitaires
- (optionnel) Un lien vers un des ces projets en machine learning (lien GitHub / GitLab ou Colab)

Références

- [1] A. Dapogny, K. Bailly and M. Cord, "Decafa: Deep convolutional cascade for face alignment in the wild", in ICCV, 2019.
- [2] J. Deng, J. Guo, N. Xue and S. Zafeiriou, "Arcface: Additive angular margin loss for deep face recognition", in CVPR, 2019.
- [3] A. Dapogny, K. Bailly and S. Dubuisson, "Confidence-weighted local expression predictions for occlusion handling in expression recognition and action unit detection", International Journal of Computer Vision, 2018.
- [4] G. Tallec, A. Dapogny and K. Bailly, "Multi-order networks for action unit detection", arXiv preprint, 2022.
- [5] J. Thies, M. Zollhofer, M. Stamminger, C. Theobalt and M. Nießner, "Face2face: Real-time face capture and reenactment of rgb videos", in CVPR, 2016.
- [6] E. Zakharov, A. Shysheya, E. Burkov and V. Lempitsky, "Few-shot adversarial learning of realistic neural talking head models", in ICCV, 2019.
- [7] J. Zhang, X. Zeng, M. Wang et al., "Freenet: Multi-identity face reenactment", in CVPR, 2020.
- [8] R. Tolosana, R. Vera-Rodriguez, J. Fierrez, A. Morales and J. Ortega-Garcia, "Deepfakes and beyond: A survey of face manipulation and fake detection", Information Fusion, 2020.