Détection et classification de visages pour la description de l'égalité homme-femme dans les archives télévisuelles

Stage de fin d'études d'Ingénieur ou de Master 2 – 2018-2019

Mots clés: Deep Learning, Détection de visages, Egalité femme-homme, Machine Learning, Humanités Numériques, Big Data

Contexte

Les missions de l'Institut national de l'audiovisuel (Ina) consistent à archiver et à valoriser la mémoire audiovisuelle française (radio, télévision et web média). A ce jour, plus de 17 millions d'heures de documents issus de la télévision et de la radio sont conservées.

La diversité de ces collections offre un terrain d'expérimentation particulièrement riche pour les humanités numériques. L'utilisation de méthodes d'analyse automatique des archives permet ainsi de décrire les évolutions de la société française en fonction de différents critères d'analyse.

Lors de travaux précédents, les différences de traitement et de représentation existant entre les hommes et les femmes ont été décrites à l'aide de méthodes fondées sur l'estimation du temps de parole et réalisées sur plus de 500 000 heures d'archives [Dou18a, Dou18b, Dou18c].

Le stage proposé s'inscrit dans la continuité de ces travaux. Il vise à concevoir des systèmes de détection et de reconnaissance du genre des visages [Naj17, Ran17]. Ces systèmes seront utilisés pour enrichir la description des différences de représentation existant entre les hommes et les femmes dans les médias à l'aide de métriques à définir dans le cadre du stage : temps d'apparition à l'écran, corrélation entre temps d'apparition et temps de parole, etc.

Objectifs du stage

Le stage vise à mettre au point une chaîne de traitement permettant d'analyser les archives télévisuelles de l'Ina, de détecter les visages, et de déterminer le sexe des visages détectés. Les solutions proposées seront déployées à grande échelle dans un second temps afin de décrire les différences de traitement et de représentation existant entre les hommes et les femmes.

Le stage sera décomposé en plusieurs étapes complémentaires :

- Réaliser un état de l'art et un benchmark des méthodes de détection de visage et de détection du genre (Facenet, OpenFace, ...). Différents critères seront pris en compte lors de cette évaluation, tels que le temps d'exécution nécessaire pour le passage à l'échelle et la robustesse pour l'estimation du pourcentage d'apparition à l'écran.
- Utiliser les propositions retenues pour extraire les visages des personnalités référencées dans le dictionnaire de locuteurs de l'Ina [Sal14]
- Réaliser un recensement des ressources annotées existantes pour la détection et la classification de visages (FDDB, LFW, AFLW, MS_Celeb_1N, ...). Participer à la définition et à l'annotation de nouvelles ressources spécifiques aux usages de l'Ina.

- Déployer les systèmes à grande échelle pour une analyse automatique de l'ensemble des archives de l'Ina
- Définir les indicateurs d'égalité de traitement entre homme et femmes dans les médias

Le langage de programmation utilisé dans le cadre de ce stage sera Python. Le stagiaire aura accès aux ressources de calcul de l'Ina (serveurs GPU).

Valorisation du stage

Différentes stratégies de valorisation des travaux du stagiaires seront envisagées, en fonction du degré de maturité des travaux réalisés :

- Diffusion des outils d'analyse réalisés sous licence open-source via le dépôt GitHub de l'Ina: https://github.com/ina-foss
- Rédaction de publications scientifiques
- Participation à l'organisation de challenges internationaux

Conditions du stage

Le stage se déroulera sur une période de 4 à 6 mois, au sein du service de la Recherche de l'Ina. Il aura lieu sur le site Bry2, situé au 18 Avenue des frères Lumière, 94366 Bry-sur-Marne. Le stagiaire sera encadré par David Doukhan (ddoukhan@ina.fr) et Jean Carrive (jcarrive@ina.fr), et percevra une gratification mensuelle de 527,75 euros/mois.

Profil Recherché

Etudiant en 5ème année de cursus en informatique et intelligence artificielle
Aisance en programmation avec le langage Python
Expérience dans l'utilisation de bibliothèques de machine learning et big data
Capacité à effectuer des recherches bibliographiques
Rigueur, Synthèse, Autonomie, Capacité à travailler en équipe
Un intérêt pour les humanités numériques serait en plus

Bibliographie

[Dou18a] Doukhan, D., & Carrive, J. (2018). Description automatique du taux d'expression des femmes dans les flux télévisuels français. Journée d'études de la parole.

[Dou18b] Doukhan, D., Carrive, J., Vallet, F., Larcher, A., Meignier, S., & Le Mans, F. (2018). An open-source speaker gender detection framework for monitoring gender equality. Acoustics, Speech and Signal Processing (ICASSP).

[Dou18c] Doukhan, D., Poels, G., & Carrive, J. (2018). Describing gender equality in french audiovisual streams with a deep learning approach (accepted). *Journal of European Television History and Culture (VIEW)*.

[Naj17] Najibi, M., Samangouei, P., Chellappa, R., & Davis, L. S. (2017, October). SSH: Single Stage Headless Face Detector. In *ICCV*(pp. 4885-4894).

[Ran17] Ranjan, R., Patel, V. M., & Chellappa, R. (2017). Hyperface: A deep multi-task learning framework for face detection, landmark localization, pose estimation, and gender recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

[Sal14 Salmon, F., & Vallet, F. (2014). An Effortless Way To Create Large-Scale Datasets For Famous Speakers. In *LREC* (pp. 348-352).