HAL open science

Étude comparative de méthodes de reconnaissance des émotions à partir d'expressions faciales

Rim El Cheikh, Hélène Tran, Issam Falih, Engelbert Mephu Nguifo

To cite this version:

Rim El Cheikh, Hélène Tran, Issam Falih, Engelbert Mephu Nguifo. Étude comparative de méthodes de reconnaissance des émotions à partir d'expressions faciales. 2023. hal-04086437

HAL Id: hal-04086437

https://hal.science/hal-04086437v1

Preprint submitted on 2 May 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Étude comparative de méthodes de reconnaissance des émotions à partir d'expressions faciales

Rim El Cheikh*, Hélène Tran*,** Issam Falih*, Engelbert Mephu Nguifo*

* Université Clermont-Auvergne, CNRS, Mines de Saint-Étienne,
Clermont Auvergne INP, LIMOS, 63000 Clermont-Ferrand, France
{rim.el_cheikh, helene.tran}@doctorant.uca.fr
{issam.falih, engelbert.mephu_nguifo}@uca.fr
** Jeolis Solutions, 63000 Clermont-Ferrand, France

1 Introduction

La recherche en reconnaissance automatique des émotions est active depuis de nombreuses décennies et ses applications sont diverses, telles que la santé et le divertissement. Notre étude porte sur les méthodes qui prédisent les émotions à partir d'images d'expressions faciales (FER). Trois modèles d'état de l'art pour les tâches FER ont été sélectionnés pour être expérimentés. Ils divergent dans leurs architectures et la méthode utilisée pour améliorer la qualité de l'inférence des émotions. Nos expériences fournissent une comparaison équitable de leurs performances sur trois ensembles de données qui diffèrent en termes de taille, de méthode de collecte d'images et de distribution des classes.

2 Expérimentations et discussion

Les réseaux de neurones étudiés sont ESR (Siqueira et al., 2020), SCN (Wang et al., 2020) et DACL (Farzaneh et Qi, 2021), qui utilisent respectivement l'apprentissage par ensembles, la relabellisation des données incertaines, et l'attention. Ils sont évalués par validation croisée 5-fold sur les trois jeux de données FER+ [1], AffectNet [2] et CK+ [3]. Une description détaillée des réseaux, des données et du protocole expérimental se trouve dans notre article (El Cheikh et al., 2022).

Le tableau 1 reporte les mesures de performance des modèles sur chaque jeu de données. DACL fournit les meilleurs scores sur FER+ et AffectNet, deux jeux de données contenant des expressions faciales spontanées. Cela suggère que son module d'attention aide à se focaliser sur les régions pertinentes pour déduire l'émotion. ESR donne les meilleures performances sur CK+ qui contient des émotions actées. Le réseau utilise des représentations partagées pour l'entrainement de chaque branche (Siqueira et al., 2020), ce qui s'avère efficace dans un contexte où les émotions sont posées et intentionnelles.

1. https://github.com/microsoft/FERPlus/
2. http://mohammadmahoor.com/affectnet/
3. https://www.jeffcohn.net/resources/

		TJ	TJP	PR	RP	F1	AUC ROC
FER+	ESR	0.857	0.617	0.855	0.857	0.856	0.937
	SCN	0.810	0.520	0.808	0.810	0.809	0.956
	DACL	**0.867**	**0.647**	**0.863**	**0.867**	**0.865**	**0.973**
AffectNet	ESR	0.648	**0.439**	0.626	0.648	0.637	0.821
	SCN	0.651	0.390	0.622	0.651	0.636	0.894
	DACL	**0.664**	0.429	**0.633**	**0.664**	**0.648**	**0.901**
CK+	ESR	**0.915**	**0.888**	**0.922**	**0.915**	**0.918**	0.945
	SCN	0.820	0.703	0.798	0.820	0.808	**0.962**
	DACL	0.846	0.790	0.843	0.846	0.844	0.951

TAB. 1 – Moyennes des métriques de performance (taux de justesse (TJ), TJ pondéré (TJP), précision (PR), rappel (RP), score F1, score AUC ROC) sur les données de test.

Les matrices de confusion présentées dans notre article (El Cheikh et al., 2022) mettent en évidence la difficulté des modèles à différencier certaines émotions. Par exemple, de nombreux échantillons sont classés "neutre" alors qu'ils contiennent une émotion, notamment lorsque celles-ci sont naturelles (FER+ et AffectNet). En particulier, l'entraînement sur FER+ induit beaucoup d'erreurs en prédisant "neutre" pour les données étiquetées "triste", mais pas pour ceux étiquetés "surprise", bien que "triste" et "surprise" soient en proportions très proches. Cela suggère que l'expression neutre et la tristesse ont des caractéristiques physiques semblables.

3 Conclusion

Cette étude comparative montre l'impact de l'architecture neuronale sur la classification des émotions, actées ou spontanées. Des défis ont été identifiés, tels que la sous-représentation de certaines classes d'émotion et l'identification ambiguë de l'émotion dans un contexte non contrôlé. Etendre les expériences à plus de modèles et d'ensembles de données fournirait une référence fiable pour choisir un modèle FER adapté en fonction de l'application souhaitée.

Références

El Cheikh, R., H. Tran, I. Falih, et E. M. Nguifo (2022). A comparative study of emotion recognition methods using facial expressions. https ://arxiv.org/abs/2212.03102.

Farzaneh, A. H. et X. Qi (2021). Facial expression recognition in the wild via deep attentive center loss. In *Proceedings of the IEEE/CVF winter conference on applications of computer vision*, pp. 2402–2411.

Siqueira, H., S. Magg, et S. Wermter (2020). Efficient facial feature learning with wide ensemble-based convolutional neural networks. In *Proceedings of the AAAI conference on artificial intelligence*, Volume 34, pp. 5800–5809.

Wang, K., X. Peng, J. Yang, S. Lu, et Y. Qiao (2020). Suppressing uncertainties for large-scale facial expression recognition. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*, pp. 6897–6906.