

Apprentissage par transfert et données mixtes pour évaluer l'importance de crues à partir d'articles d'information

Pierrick Bruneau, Yoann Didry, Thomas Tamisier

LIST, L-4362 Esch-sur-Alzette
pierrick.bruneau@list.lu

Résumé. Dans cet article applicatif, nous décrivons l'utilisation de propriétés textuelles et visuelles par des réseaux de neurones profonds pour évaluer l'importance des crues dans les articles d'information. En particulier, nous estimons la pertinence du transfert de modèles pré-entraînés sur des corpora conceptuellement proches. Nous évaluons également l'apport de modèles à branche double, qui combinent les représentations denses d'un texte et d'une image associée. Nous comparons la performance de ces variantes méthodologiques au moyen des données distribuées dans le cadre de l'atelier MediaEval MultiMedia Satellite (MMSat) 2019. Les résultats présentés ici ont fait l'objet d'une communication à l'atelier : le présent article propose une version significativement étendue des notes techniques accompagnant les prédictions réalisées sur l'ensemble de test de l'atelier MMSat.

1 Introduction

L'identification d'actualités en ligne liées à un événement catastrophique tel qu'une crue, et l'évaluation de l'étendue de cet événement grâce à l'information collectée peuvent être d'une importance critique en vue du secours aux victimes. Les tâches de désambiguïsation de sujet à partir de l'image (*News Image Topic Disambiguation* - NITD) et d'estimation multimodale du niveau de crue (*Multimodal Flood Level Estimation* - MFLE) de l'atelier MediaEval MultiMedia Satellite 2019 (Bischke et al., 2019) encouragent l'application de méthodologies d'apprentissage automatique dans ce contexte. Nous n'avons délibérément pas utilisé de modèle spécialisé à ces tâches, e.g. nous n'avons pas eu recours à des techniques de détection de pose ou d'occlusion (Bulat et Tzimiropoulos, 2016) pour la tâche MFLE.

Dans cette optique, la tâche NITD est un problème de classification d'images, qui peut être traité grâce un réseau de neurones convolutif tel que VGG16 (Simonyan et Zisserman, 2014), InceptionV3 (Szegedy et al., 2016), ou MobileNetV2 (Sandler et al., 2018). Ces modèles définissent de nombreux paramètres (souvent de l'ordre de plusieurs millions) qui sont coûteux à estimer à partir d'une initialisation aléatoire, et *a priori* difficilement sans un échantillon d'apprentissage conséquent. Une solution courante est alors l'apprentissage par transfert, qui consiste à réutiliser des modèles pré-entraînés sur une tâche voisine. De tels modèles pré-entraînés sont même souvent mis à disposition par des bibliothèques d'apprentissage profond telles