
Master thesis : Drone control through a vocal interface

Auteur : Pirlet, Matthias

Promoteur(s) : Louppe, Gilles

Faculté : Faculté des Sciences appliquées

Diplôme : Master : ingénieur civil en science des données, à finalité spécialisée

Année académique : 2021-2022

URI/URL : <http://hdl.handle.net/2268.2/15850>

Avertissement à l'attention des usagers :

Tous les documents placés en accès ouvert sur le site le site MatheO sont protégés par le droit d'auteur. Conformément aux principes énoncés par la "Budapest Open Access Initiative"(BOAI, 2002), l'utilisateur du site peut lire, télécharger, copier, transmettre, imprimer, chercher ou faire un lien vers le texte intégral de ces documents, les disséquer pour les indexer, s'en servir de données pour un logiciel, ou s'en servir à toute autre fin légale (ou prévue par la réglementation relative au droit d'auteur). Toute utilisation du document à des fins commerciales est strictement interdite.

Par ailleurs, l'utilisateur s'engage à respecter les droits moraux de l'auteur, principalement le droit à l'intégrité de l'oeuvre et le droit de paternité et ce dans toute utilisation que l'utilisateur entreprend. Ainsi, à titre d'exemple, lorsqu'il reproduira un document par extrait ou dans son intégralité, l'utilisateur citera de manière complète les sources telles que mentionnées ci-dessus. Toute utilisation non explicitement autorisée ci-avant (telle que par exemple, la modification du document ou son résumé) nécessite l'autorisation préalable et expresse des auteurs ou de leurs ayants droit.

UNIVERSITY OF LIÈGE
MASTER THESIS

Drone control through a vocal interface

Author: PIRLET Matthias
Academic Supervisor: Prof. LOUPPE Gilles (ULiège)
Industrial Supervisor: GREFFE Christophe (GeneriX)

Master thesis carried out in order to obtain the degree of
Master of Science in Data Science & Engineering
Academic year 2021-2022

Abstract

Today the use of drones is widely spread for many tasks, but for some of these, such as firefighting, it is vital that the operator's hands are kept free to do their job properly. Hopefully speech recognition is an exploding discipline in deep learning. This work therefore focuses on finding a deep learning model that recognises spoken commands to control a drone from a pre-defined set.

The first part of the work was to build a training dataset of commands with the combination of complete commands generated through the use of text-to-speech APIs and hand-crafted commands. These were created thanks to the concatenation of an open source spoken words dataset and another set of spoken words acquired through a web platform created in order to complete the missing words of the vocabulary command set. The testing set was acquired by asking people to record complete commands under real conditions.

The second part of the work focuses more on the different models that could be developed and all the techniques that can be used. These are presented as an ablation study, in order to improve the results on a test set in real conditions. Several methods were applied in order to achieve the final goal: the first is the use of computer vision models where the input of these models is a simple spectrogram of the different commands. The results using these types of models were not as good as those of the new models which take directly the raw waveform as input and combines vision, attention and self-supervised learning. The best version of this model obtains a F1-Score of 0.9973 on a real conditions dataset.