

Nom et prénom	Etablissement	Grade	Qualité
SATORI Khalid	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Président
OUANAN Mohammed	Faculté des Sciences, Meknès	PES	Rapporteur
EL AMRANI Aumeur	Ecole Supérieure de Technologie, Meknès	PES	Rapporteur
BOUMHIDI Jaouad	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Rapporteur
EL AKKAD Nabil	Ecole Nationale des Sciences Appliquées, Fès	MCH	Examineur
TAZI El Bachir	Faculté Polydisciplinaire, Taza	PES	Examineur
SATORI Hassan	Faculté des Sciences Dhar EL Mahraz, Fès	PES	Directeur de thèse



## Résumé :

Cette thèse intègre la Reconnaissance Automatique de la Parole (RAP) aux systèmes embarqués afin de concevoir un drone commandé par la voix en temps réel pour la langue arabe. Il répond à un double défi : la rareté des corpus spécialisés pour les variantes arabes peu dotées et les contraintes computationnelles des plateformes embarquées. Pour combler ces lacunes, une plateforme de reconnaissance vocale open source, hautement configurable et adaptée à la fois aux spécificités linguistiques et à l'efficacité matérielle, a été développée. Un corpus interne et spécifique au domaine des commandes de drones a été constitué, intégrant les principales caractéristiques phonétiques, les structures syllabiques représentatives et un ensemble de bruits réels liés à l'environnement des UAV. Deux systèmes complémentaires de reconnaissance vocale ont ensuite été développés. La première repose sur un modèle stochastique de Markov caché (HMM) modulaire et efficace, optimisé pour les plateformes embarquées à ressources limitées. Le second est un modèle de bout en bout, prêt pour informatique en périphérie de réseau, intégrant l'arabe standard moderne (MSA) et l'arabe marocain (Darija), enrichi par une attention légère au niveau des canaux pour renforcer la robustesse face aux interférences acoustiques des drones. La conception du système s'appuie sur une analyse phonologique et syllabique approfondie, mettant en avant les patrons accentués et les structures syllabiques équilibrées, offrant ainsi un cadre rigoureux pour la modélisation de l'arabe dans des conditions à faibles ressources. Un protocole d'évaluation unifié, exécuté sur CPU, GPU et plateformes embarquées, mesure la précision de reconnaissance, le temps d'inférence et la consommation mémoire. Les résultats montrent que le modèle stochastique atteint une précision satisfaisante avec un coût computationnel minimal, tandis que le modèle de bout en bout offre une précision supérieure au prix d'une légère augmentation de la latence et de la mémoire. Les deux systèmes conservent des performances en temps réel et une compacité élevée, démontrant la faisabilité d'une reconnaissance vocale arabe robuste dans des environnements embarqués contraints. Cette thèse propose ainsi une solution pratique et déployable, reliant la conception de corpus, la modélisation statistique et l'apprentissage profond compact pour faire progresser la reconnaissance vocale de l'arabe dans des contextes à faibles ressources et à contraintes matérielles.

## Mots clés :

**Reconnaissance Automatique de la Parole (RAP) ; Control de vol des drones ; Langues Arabes à Faibles Ressources ; Modèles Stochastiques et d'Apprentissage Profond ; Mécanismes d'Attention ; Robustesse Linguistique et au Bruit ; Intelligence Artificielle en Périphérie; Systèmes Embarqués en Temps Réel ; Plates-formes à Ressources Limitées.**

**Automatic Speech Recognition (ASR); Drone Flight Control; Low-Resource Arabic Languages; Stochastic and Deep Learning Models; Attention Mechanisms; Linguistic and Noise-Aware Robustness; Edge AI; Embedded Real-Time Systems; Resource-Constrained Platforms.**