



## AVIS DE SOUTENANCE DE THESE

*Le Doyen de la Faculté des Sciences Dhar El Mahraz –Fès – annonce que*

Mme **Soukaina Chraa Mesbahi.**

Soutiendra : **le Samedi 24/06/2023 à 9H00**

Lieu : **FSDM – Salle de réunion département de Mathématique.**

*Une thèse intitulée :*

**Reconnaissance et classification des gestes humaines 2D en temps réel.**

*En vue d'obtenir le Doctorat*

*FD : Sciences et Technologies de l'Information et de la Communication*

*Spécialité : Informatique*

*Devant le jury composé comme suit :*

Nom et prénom	Etablissement	Grade	Qualité
Pr YAHYAOUY Ali	Faculté des Sciences Dhar El Mahraz Fès	PES	Président
Pr BELLACH Benaïssa	Ecole Nationale des Sciences Appliquées ENSAO Oujda	PES	Rapporteur & examineur
Pr AGHOUTANE Badraddine	Faculté des Sciences Meknès	PH	Rapporteur & examineur
Pr El FAZAZI Khalid	Faculté des Sciences Dhar El Mahraz Fès	PH	Rapporteur & examineur
Pr ENNAHNAHI Nouredine	Faculté des Sciences Dhar El Mahraz Fès	PH	Examineur
Pr SABRI My Abdelouahed	Faculté des Sciences Dhar El Mahraz Fès	PES	Examineur
Pr RIFFI Jamal	Faculté des Sciences Dhar El Mahraz Fès	PH	Examineur
Pr MAHRAZ Mohamed Adnane	Faculté des Sciences Dhar El Mahraz Fès	PH	Co-Directeur de thèse
Pr TAIRI Hamid	Faculté des Sciences Dhar El Mahraz Fès	PES	Directeur de thèse



## Résumé :

De nos jours, de nombreux systèmes ont été développés pour faciliter l'interaction entre l'homme et la machine. Parmi ces systèmes, on trouve des dispositifs couramment utilisés tels que la souris, le clavier, la reconnaissance vocale, faciale, le suivi de mouvement et les interfaces gestuelles. La main joue un rôle essentiel dans le langage corporel et constitue un moyen indispensable pour interagir avec les systèmes informatiques. Les gestes sont une forme de communication non verbale où les mouvements de la main, du visage ou d'autres parties du corps transmettent un message spécifique. Ils fournissent des informations cruciales sur l'état de l'interaction en fonction du geste de la main et de la position des doigts. Cette capacité est également essentielle pour interpréter la langue des signes utilisée par les personnes sourdes ou malentendantes pour une communication efficace à travers le monde. Ainsi, la traduction automatique de la langue des signes revêt une grande importance. Cependant, la détection et la reconnaissance des gestes de la main dans des scènes complexes représentent un défi majeur pour de nombreuses applications de vision par ordinateur, telles que l'analyse des actions de la main, l'interaction homme-machine, la reconnaissance de la langue des signes, la réalité virtuelle et la surveillance des mouvements des mains.

Dans ce travail, nous développons trois nouvelles approches de détection du mouvement, la première est basée sur la soustraction de l'arrière-plan, ce qui permet de détecter les pixels et les régions contenant l'objet en mouvement à partir des caméras statiques. Dans cette méthode nous combinons le modèle de mélange gaussien GMM et l'algorithme du flot optique.

Notre seconde approche, vu le progrès significatif que connaît la vision par ordinateur ces dernières années, s'appuie sur l'apprentissage en profondeur. Nous proposons un enchaînement de plusieurs étapes qui va nous permettre d'extraire onze caractéristiques que nous avons mises en place grâce à des calculs mathématiques précis.

Enfin, pour la bonne raison d'améliorer la communication des personnes avec limitations physiques, nous présentons une approche qui repose sur les algorithmes YOLO (You Only Look Once) version : V3, V4, V4-tiny et V5. À l'aide de ces algorithmes et suivant un schéma que nous avons conçu, nous sommes parvenus à un système qui permet de détecter, classifier et prédire la classe des gestes de manière plus précise. Cette méthode vise à faciliter la communication via des messages ou des appels vidéo. L'objectif ultime est de rendre la communication accessible à tous, en offrant des solutions technologiques innovantes qui permettent aux personnes handicapées de s'exprimer et de communiquer de manière plus libre et efficace.

**Mots clés :** *GMM, Soustraction d'arrière-plan, Flot optique, Extraction de caractéristiques, LSTM, CNN, YOLO, Détection d'objets, Geste de la main.*



## Abstract:

Nowadays, numerous systems have been developed to facilitate human-machine interaction. Among these systems, commonly used devices include the mouse, keyboard, voice recognition, facial recognition, motion tracking, and gesture interfaces. The hand plays an essential role in body language and is an indispensable means of interacting with computer systems. Gestures are a form of nonverbal communication where hand movements, facial expressions, or other body movements convey a specific message. They provide crucial information about the state of interaction based on hand gestures and finger positions. This ability is also vital for interpreting sign language used by deaf or hearing-impaired individuals for effective communication worldwide. Thus, automatic sign language translation holds great importance. However, detecting and recognizing hand gestures in complex scenes pose a major challenge for many computer vision applications, such as hand action analysis, human-computer interaction, sign language recognition, virtual reality, and hand motion monitoring. In this work, we develop three novel motion detection approaches. The first is based on background subtraction, which enables the detection of pixels and regions containing moving objects from static cameras. In this method, we combine the Gaussian Mixture Model (GMM) and the optical flow algorithm.

Our second approach leverages the significant progress in computer vision in recent years and is based on deep learning. We propose a series of steps to extract eleven features that we have established through precise mathematical calculations.

Finally, with the aim of improving communication for individuals with physical limitations, we present an approach based on the YOLO (You Only Look Once) algorithms, specifically versions V3, V4, V4-tiny, and V5. Using these algorithms and following a designed framework, we have achieved a system that can detect, classify, and predict gesture classes more accurately. This method aims to facilitate communication through messaging or video calls. The ultimate goal is to make communication accessible to everyone by providing innovative technological solutions that enable disabled individuals to express themselves and communicate more freely and effectively.

**Key words:** *GMM, Background subtraction, Optical flow, Feature extraction, LSTM, CNN, YOLO, Object detection, Hand gesture.*