

### Fiche de stage

**Sujet du stage : TargetFinder: Developing an External Access System for Widget Information in Graphical User Interfaces**

Encadrant·e·s : Julien Gori

Date de début du stage : Janvier - Mars 2025

Durée du stage : 5-6 mois

Niveau d'études souhaité : Master 2

Laboratoire d'accueil : ISIR (*Institut des Systèmes Intelligents et de Robotique*), Campus Pierre et Marie Curie, 4 place Jussieu, 75005 Paris.

### Personne à contacter

Prénom Nom : Julien Gori

Email : gori@isir.upmc.fr

Envoyer votre candidature par mail, avec [sujet du stage] en objet, un CV et une lettre de motivation.

### Description du stage (en français)

#### Résumé :

Dans la plupart des applications graphiques, les utilisateurs interagissent avec divers éléments - comme des boutons, des menus et des barres de titre - appelés widgets. Les informations sur ces widgets (telles que leur taille, leur emplacement et leurs étiquettes) ne sont généralement accessibles qu'au sein de l'application elle-même. L'objectif de ce stage est de développer des méthodes permettant d'accéder à ces informations depuis l'extérieur de l'application, pour permettre des nouvelles utilisations, comme l'analyse de l'interface utilisateur et l'amélioration de l'accessibilité.

#### Objectifs du stage :

Ce projet s'inscrit dans le cadre d'une collaboration entre Sorbonne Université et l'Université de Lille. Au cours de ce stage, l'étudiant travaillera à la conception d'un système capable de recueillir des informations sur les widgets d'interfaces utilisateurs graphiques (GUI). L'étudiant s'appuiera sur un démonstrateur existant qui permet d'identifier l'emplacement et la taille d'une cible unique. Le démonstrateur s'appuie sur un réseau de détection d'objets pré-entraîné YOLOv8. Le premier objectif de ce projet est d'améliorer le démonstrateur pour qu'il puisse identifier plusieurs emplacements et tailles de cibles en même temps, ainsi que de déterminer leurs label. Pour ce faire, le stagiaire appliquera des techniques existantes de traitement d'images pour pré-étiqueter un ensemble de données existant qui, après vérification, sera utilisé pour ré-entraîner le réseau existant.

Le deuxième objectif de ce projet est de mettre en œuvre une technique d'interaction existante qui exploite les informations sur la cible, par exemple le pointage sémantique ou le bubble cursor, afin de montrer les possibilités offertes par ce projet.

Sous la co-tutelle de :

### Profil recherché :

Nous recherchons une personne ayant une formation en informatique, avec des compétences en apprentissage automatique / vision par ordinateur / traitement d'images.

La connaissance de Python est requise, la connaissance de C++ n'est pas nécessaire mais est appréciée.

### Description du stage (en anglais)

#### Subject:

In most graphical applications, users interact with various elements – like buttons, menus, and title bars – called widgets. Information about these widgets (such as their size, location, and labels) is typically accessible only within the application itself. The objective of this internship is to develop methods to access this information from outside the application, enabling a broader range of uses, such as cross-application automation, user interface analysis, and accessibility enhancements.

#### Internship Objectives:

This project is part of an ongoing collaboration between Sorbonne Université and Université de Lille. During this internship, the student will work on designing a system that can gather widget information from graphical user interfaces (GUIs) externally. The student will build on an existing demonstrator that can identify a single target's location and size. The demonstrator builds on a YOLOv8 pre-trained object detection network. The first goal of this project is to improve the demonstrator to identify multiple target locations and sizes at once, as well as determine their labels. To do so, the intern will apply existing image processing techniques to pre-label an existing dataset, which will, after verification, be used to retrain the existing network. The second goal of this project is to implement an existing interaction technique that leverages target information, for example semantic pointing, to showcase the possibilities offered by this project.

#### Required Profile:

We are looking for a person with a background in computer science, with skills in machine learning / computer vision / image processing. Knowledge of Python is required, knowledge of C++ is not needed but is appreciated.

Sous la co-tutelle de :