

AVIS DE SOUTENANCE DE THÈSE

DOCTORAT (Arrêté du 26 août 2022 modifiant l'arrêté du 25 mai 2016)

Monsieur Félix MERCIER

candidat au diplôme de Doctorat de l'Université d'Angers, est autorisé à soutenir publiquement sa thèse

le 29/09/2026 à 09h00

IRHS

Bâtiment A

Salle du rez-de-chaussée

42, rue Georges Morel

49070 BEAUCOUZÉ

sur le sujet suivant :

Approche d'IA centrée sur la donnée pour le phénotypage à haut débit des plantes par imagerie

Directeur de thèse : **Monsieur David ROUSSEAU**

Composition du jury :

Monsieur Nizar BOUHLEL, Maître de Conférences Institut Agro Rennes-Angers, Co-encadrant

Monsieur Jean-Pierre DA COSTA, Professeur des Universités Bordeaux Sciences Agro, Rapporteur

Madame Angelina EL GHAZIRI, Maîtresse de Conférences Institut Agro Rennes-Angers, Co-encadrante

Monsieur David ROUSSEAU, Professeur des Universités Université d'Angers, Directeur de thèse

Madame Sana SELLAMI, Maîtresse de Conférences HDR Aix-Marseille Université, Rapporteur

Monsieur Jean-Yves TOURNERET, Professeur des Universités Toulouse INP-ENSEEIH, Examineur

Madame Cindy TORRES, PhD Vilmorin-Mikado, La Ménitrie, Membre Invité

Madame Julia BUITINK, PhD INRAe IRHS, Membre Invité



Résumé de la thèse

Le phénotypage des plantes par imagerie joue un rôle clé dans l'analyse des premiers stades du développement végétal, en particulier pour l'étude de la germination, de l'émergence et de l'établissement des plantules. Ces stades précoces, qui concernent directement la vigueur des semences et la réussite de l'implantation des cultures, sont caractérisés par une forte dynamique temporelle, une variabilité biologique marquée et des contraintes expérimentales importantes. Bien que les approches d'apprentissage profond aient démontré leur potentiel dans ce domaine, leur déploiement reste limité par le coût de l'annotation, la diversité des dispositifs expérimentaux et la difficulté à maintenir des modèles fiables face à des données évolutives. Cette thèse s'inscrit dans le cadre du "Data-Centric Artificial Intelligence", une approche qui place la donnée — son acquisition, sa sélection, sa structuration et son évolution — au cœur du processus d'apprentissage. Contrairement aux approches model centric, centrées sur l'optimisation des architectures de réseaux de neurones, le paradigme data-centric vise à améliorer la performance, la robustesse et la généralité des systèmes d'IA en agissant prioritairement sur la qualité et l'usage des données. Les travaux de la thèse s'articulent autour de plusieurs dispositifs expérimentaux complémentaires, appliqués aux stades précoces du développement des plantes, depuis la semence jusqu'à la jeune plantule. Une première contribution porte sur des stratégies d'échantillonnage adaptatif via des méthodes bayésiennes avancées, étudiées sur un dispositif expérimental dédié au suivi en imagerie RGB de la cinétique de germination de semences sur buvard. Dans un second temps, la thèse s'intéresse à la détection d'anomalies sur un autre dispositif expérimental, basé sur l'imagerie RGB-profondeur, dédié au suivi de l'émergence des graines hors du sol et au développement initial des plantules. Avec ce dispositif, la thèse aborde la problématique de la mise à jour des modèles dans des contextes de changement de domaine, notamment lors du passage d'une espèce végétale à une autre, à travers des approches d'apprentissage par transfert actif ou détection conforme. Une dernière contribution repose sur l'introduction de modèles probabilistes non paramétriques, en particulier les mélanges de processus de Dirichlet, pour la modélisation des stades de développement précoces. L'ensemble de ces travaux quantifie les gains (en temps de calcul, en volume de stockage, en volume d'annotation, ...) que l'on peut espérer par une approche "Data-Centric Artificial Intelligence" en phénotypage des plantes sur des cas d'usage d'intérêts agronomiques réels. Les méthodes présentées ont une valeur générique. Toutefois, leur mise en œuvre nécessite l'intégration de connaissances a priori, la calibration de modèles ou encore le design d'espace latent qui demeurent spécifiques à chaque cas d'usage.