

Analyses statistiques multivariées et Machine learning. Application à la caractérisation de la croissance de populations de plantes.

Contexte : La maîtrise des rendements des cultures, le développement de nouvelles maladies des plantes, le dérèglement climatique, amènent les acteurs de l'agriculture en charge de la sélection variétale, à identifier des semences performantes, résistantes aux maladies, à des périodes de sécheresse ou de brusques variations environnementales durant leur développement. Ces travaux peuvent bénéficier d'avancées technologiques récentes en matière d'instrumentation et d'imagerie automatisées et des traitements de données associés applicables pour le phénotypage de larges populations de plantes. Une échelle particulièrement importante est celle de la graine, dont la qualité germinative conditionne la suite du développement de la plante. Dans ce stage, nous nous intéressons à des graines de betterave sucrière pour laquelle la France est l'un des premiers producteurs mondiaux. Le phénotypage automatisé de la germination des graines et de l'élongation des plantules résultantes est de nature à produire de grandes quantités de données qu'il importe d'analyser de façon quantitative.

Problématique : Le travail vise l'analyse statistique multivariée de la croissance de plantules comme illustré sur la Fig. 1. Il s'agit en particulier d'analyser si des lots de plantules au patrimoine génétique déterminé ou avec des caractéristiques germinatives mesurées par ailleurs se distinguent dans cette phase de développement. On procèdera à une analyse multivariée en prenant notamment en compte l'évolution temporelle de différents organes au sein de ces plantules. On pourra également ajouter dans l'analyse, des paramètres issus d'imagerie par tomographie X. Des méthodes d'apprentissage automatique de type non supervisé sous forme de clustering (K-nearest neighbors, mean shift, dynamic time warping, réseaux de neurones de type autoencoders) seront mises en œuvre.

Stage de niveau master 2, de 5 à 6 mois, de février à septembre 2018, rémunéré au niveau de la gratification réglementaire de l'ordre de 500 à 600 € mensuels.

Compétences attendues : Connaissances de base en data sciences, expérience dans un langage de programmation adapté (au choix Python, R, Matlab ...).

Compétences développées : Méthodes de traitement du signal et des images avancées (Analyses statistiques multivariées, Machine learning), expérience pluridisciplinaire (STIC – sciences du vivant) impliquant des partenaires académiques et privés.

Environnement de travail : Le stage se déroulera au laboratoire LARIS de l'Université d'Angers, en collaboration avec la plateforme d'instrumentation et d'imagerie du végétal PHENOTIC de l'INRA d'Angers. Le stage est financé dans le cadre du programme ANR Investissements d'Avenir AKER en partenariat avec la Station Nationale d'Essais de Semences (SNES-GEVES) et le groupe industriel français Florimond Desprez.

Encadrement : David ROUSSEAU, Etienne BELIN, François CHAPEAU-BLONDEAU.

Contact : david.rousseau@univ-angers.fr

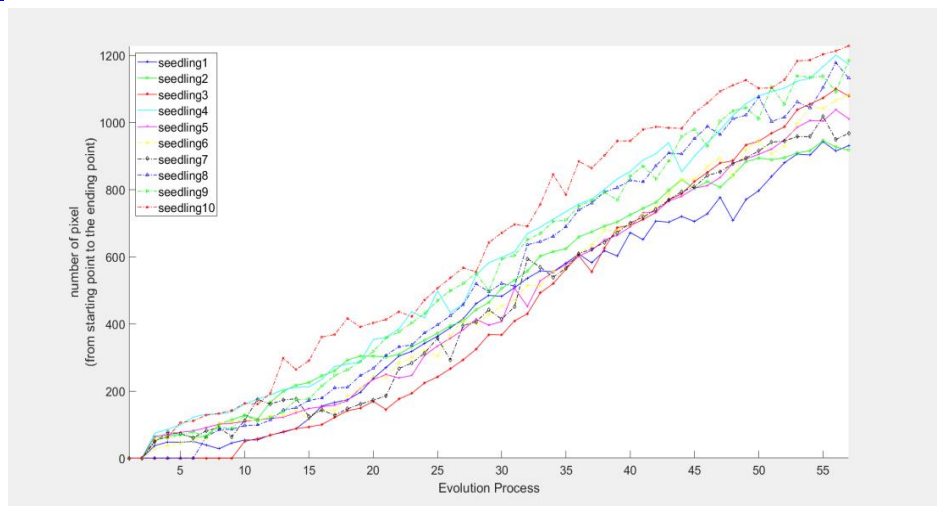
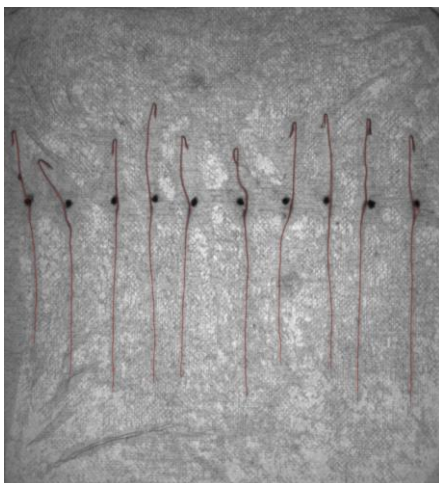


Figure 1: Croissance de plantules de betterave sucrière. À gauche 10 plantules à un stade de développement donné et que l'on a segmentées individuellement dans l'image. À droite l'évolution au cours du temps de la longueur des 10 plantules.

[1] L. Benoit, D. Rousseau, E. Belin, D. Demilly, F. Chapeau-Blondeau; "Simulation of image acquisition in machine vision dedicated to seedling elongation to validate image processing root segmentation algorithms."; *Computers and Electronics in Agriculture*, vol. 104, pp. 84-92, 2014.