Section Sciences et Ingénierie de l'environnement
Design Project 2023

Proposition n°23

SmartViti: Detection and prediction of diseases in vineyards

Partenaire externe ou laboratoire IIE
Dr. Saskia Gindraux + Dr. Antonio Abellan
saskia.gindraux@crealp.vs.ch , direction@crealp.vs.ch
Téléphone : 079 791 8625
CREALP
Taille de l'entreprise (nbre de collaborateurs) : 20
Rue de l'industrie 45, 1950 Sion
www.crealp.ch

Encadrant EPFL (proposition facultative qui sera validée par la Section)
Prof. Devis Tuia
devis.tuia@epfl.ch Téléphone: +41 21 693 82 83
(ECEO), ENAC, IIE, EPFL
EPFL Valais Wallis, Rue de l'Industrie 17, Case postale 440, CH-1951 Sion
www.epfl.ch/labs/eceo/

DESCRIPTIF DU PROJET

Context: Viticulture is an important sector of Swiss agriculture in terms of economic income and job creation: vineyards are present in all cantons and managed by a large number of producers. At the same time, the intensification of viticulture has made it increasingly susceptible to diseases and pest outbreaks, causing considerable economic losses to wine growers as they affect grape production as well as fruit quality.

Aim: The aim of this design project is to create a model able to predict the outbreak and evolution of vineyard diseases based on easily acquired input information (leaf images, meteorological data, etc.) and modern Machine Learning (ML) models. Assuming the model would yield accurate results, this design project will propose a proof of concept tool able to inform winegrowers about disease evolution in pseudo real-time. With this information, the producers can apply the necessary treatments at the optimal timing, which will reduce the environmental impact of the activity (i.e., amount of chemical products required to increase the production yield).

Available data: The main input dataset consist on set of 11'000 images of grapevine leaves acquired with during the 2022 summer. Images were collected at three different pilot study areas of the Federal center of agronomic research (Agroscope): Changins (VD), Pully (VD) and

Leytron (VS). The image collection was weekly acquired in order to obtain the complete disease evolution (=epidemiology curve). Images were then labeled by experts with the corresponding disease (e.g. *Oïdium, Mildiou or Flavescence dorée* diseases). Meteorological datasets were also collected at the same timeframe from the meteorological network owed by Agroscope located at the different test sites.

OBJECTIF ET BUTS

(Décrire 1 objectif général et 3-4 buts réalistes)

The design project is divided into three main phases (Figure 1):

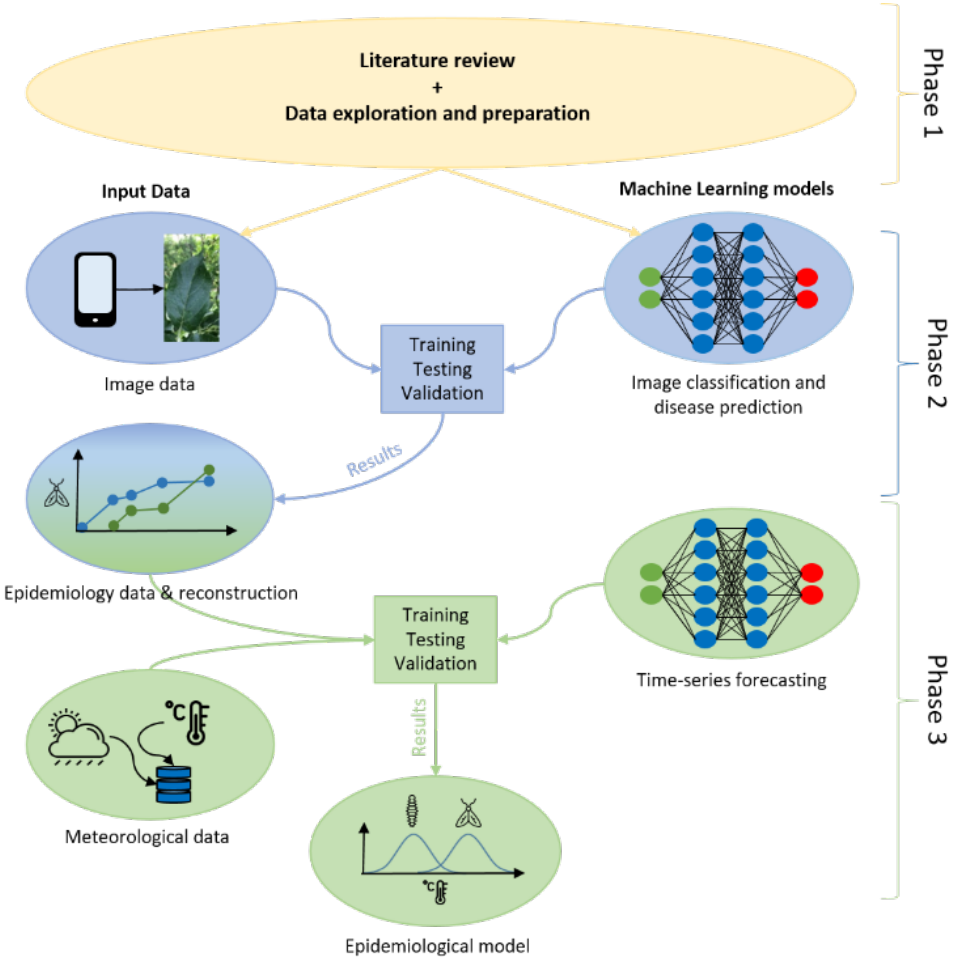

Figure 1: The tree main project phases. On the right are the input data and on the left the models.

Phase 1: The objective of the first phase (Figure 1, in yellow) is to gather the required knowledge of vineyard diseases, as well as the ML models that could be applied within this project (e,g, Fuentes and Tongson, 2018). Students will familiarize with the available datasets and prepare them for processing. One input dataset consists of 11'000 optical images of grapevines, which were tagged with disease names. Preliminary image treatment (i.e. cropping, resizing, etc.) and data augmentation should be applied to the images in order to refine the dataset that will be used for the development of the ML model in Step 2. The second dataset are time-series of the evolution of the different diseases (epidemiology curves). These time-series are derived from field observations (tags of the image collection). They are the validation data from the outputs of the ML model of Phase 2 and input data for the ML model of Phase 3. The third input dataset consists of meteorological time-series that cover the same timeframe than the image data collection

(June-August 2022). The times-series will need to be filtered and aggregated before being used as an input for the ML model of Step 3.

Phase 2: The second phase (Figure 1, in blue) involves using datasets of wine leaves to train a ML model that will classify the images per disease. The student can either build the ML model from scratch, or modify an existing model developed for a similar study but on apple tree leaves (Grünig et al. 2021). The ML model will be trained and validated with the input image dataset. The objective is to have a model that classifies/predicts which disease is present on each leaf. This will enable the reconstruction of the epidemiology curves (i.e. the evolution of the disease over time) for each disease and each site. These curves will be validated with the ones obtained from the field data.

Phase 3: The third phase (Figure 1, in green) aims to predict the above-mentioned epidemiology curves for different diseases. Collecting thousands of images every year to obtain this information is too time consuming, so the idea is to predict the curve based on meteorological data only. Therefore, a ML model which integrates the meteorological time-series from the Agrometeo stations as well as the epidemiology curves will be developed. The objective is to detect and select which meteorological variables explain the outburst and the evolution of the disease best, in order to obtain a forecasting model for disease evolution.

DESCRIPTIF TÂCHES

(Décrire 3 à 4 étapes de la démarche de projet en spécifiant s'il y a une partie expérimentale - terrain, mesures, prototypage)

The first month of the Design Project will be dedicated to gain understanding of the problematic and rationale behind the vineyard diseases and disease evolution, as well as making data exploration and preparation (See Table 1, activity 1). Once the input data will be ready, it will feed the two different ML models that could be developed by both students simultaneously (Table 1, activities 2. and 3.). The time dedicated for that activity is estimated in around 6 to 7 weeks. Finally, the results and analysis should be synthetized in a final report, and prepare the final codes for further work (i.e. annotated scripts and user manual) (Table 1, activity 4.).

ACTIVITY	February		March				April				May				
	20 - 24	27 - 03	06 - 10	13 - 17	20 - 24	27 - 31	03 - 07	10 - 14	17 - 21	24 - 28	01 - 05	08 - 12	15 - 19	22 - 26	29 - 02
	week 1	week 2	week 3	week 4	week 5	week 6	week 7	week 8	week 9	week 10	week 11	week 12	week 13	week 14	week 15
1. LITTERATURE REVIEW AND DATA PREPARATION															
1.1 Wineyard diseases															
1.2 ML models for image classification															
1.3 ML models for time-series forecasting															
1.4 Datasets' exploratory analysis															
1.5 Data preparation for ML integration															
2. MACHINE LEARNING: IMAGE CLASSIFICATION AND DISEASE PREDICTION															
2.1 Model choice															
2.2 Model development and testing															
2.3 Model validation															
2.4 Results: Derivation of the epidemiological curve															
3. MACHINE LEARNING: TIME SERIES FORECASTING															
3.1 Model choice															
3.2 Model development and testing															
3.3 Model validation															
3.4 Results: Find the predictive variables for the epidemiological curve															
4. DELIVERABLES															
4.1 Reports (intermediate / final)															
4.2 Code documentation															
4.2 User Manual															

Table 1: Calendar of the Design Project

BIBLIOGRAPHY

Grünig, M., Razavi, E., Calanca, P., Mazzi, D., Wegner, J.D., Pellissier, L., 2021. Applying deep neural networks to predict incidence and phenology of plant pests and diseases. Ecosphere 12, e03791. https://doi.org/10.1002/ecs2.3791

Fuentes, S., & Tongson, E. (2018). Advances and requirements for machine learning and artificial intelligence applications in viticulture. Wine & Viticulture Journal, 33(3), 47–52. https://search.informit.org/doi/10.3316/informit.274506637392389