



English below

Détection et classification supervisée depuis des données 3D

Laboratoire : Centre de Recherche Astrophysique de Lyon – UMR 5574 (<https://cral.univ-lyon1.fr/>)

Directeur : Matthew Lehnert

Direction de thèse (HDR ou équivalent) : Nicolas Bouché, équipe /GALPAC/

Co-direction :

Adresse électronique et téléphone : nicolas.bouche@univ-lyon1.fr

Contexte et Description du projet :

Contexte: L'instrument [MUSE](#) du VLT est un spectrographe à intégral (Integral Field Spectrograph ; IFS) de champ qui fournit des données hyperspectrales, c'est-à-dire fournit un spectre pour chacun des 90,000 pixels dans le champ de vue d' $1' \times 1'$. Ce champ de vue qui fait qu'une observation typique de MUSE contient environ 100 galaxies distribuées sur une large gamme de redshift, entre 0 to 6.

Détecter chaque source dans le champ en aveugle ainsi que l'**identification** de la solution du décalage vers le rouge (redshift) sont deux tâches très difficiles. En effet, concernant la détection, certaines galaxies ont principalement des raies d'émissions, alors que d'autres ont principalement du continuum stellaire. Si des algorithmes de détection existent soit pour les raies d'émission, soit pour le continuum, aucun algorithme n'existe pour faire ces deux tâches simultanément.

En outre, trouver la bonne solution du redshift pour chacun des objets dans le champ est une tâche difficile, car cela requiert souvent des inspections visuelles des spectres, pour chacune des solutions possibles, tâche qui est chronophage et difficilement `scalable`.

Ce projet vise à potentiellement transformer la faible efficacité pour extraire les sources et pour identifier les redshifts de toutes les galaxies présentes dans des données 3D telle que MUSE.

Avec la grande popularité de l'instrument MUSE, avec les IFS à grand champ tel que BlueMUSE, et le [WST](#), il sera important de développer de nouvelles façons d'exploiter de telles données qui fourniront des milliards de spectres du mode IFUs et des centaines de millions de spectres du mode multi-objets.

Objectifs: La première partie du projet consiste à développer des outils pour faciliter la détermination de la solution en redshift grâce aux techniques de Machine Learning (ML), tel que la classification supervisée utilisant des Convolutional Neural Networks (CNN) qui seront entraînés sur des échantillons de galaxies extraites des différents programmes GTO MUSE par notre équipe. Ces outils seront développés sur les spectres 1D extraits des données MUSE, et seront ensuite adaptés sur les données 3D (mini-cubes) de chaque objet. La seconde partie de la thèse consiste à développer soit des méthodes avancées ML hybride soit des méthodes non-supervisées pour détecter et classifier les sources depuis les cubes de données 3D directement.

CRAL est l'institut pilote de plusieurs Integral Field Units de prochaine génération telle que [BlueMUSE/VLT](#), [Harmoni/ELT](#), [WST](#), et a construit les spectrographes de 4MOST qui fournira des millions de spectres.

Date de début de thèse : **Sep 2024**

Autres sources de financement envisagées que l'ED52 :

au lycée du par.



Français ci-dessus

Detection and supervised classification from 3D data

Institute: Centre de Recherche Astrophysique de Lyon – UMR 5574 (<https://cral.univ-lyon1.fr/>)

Director: Matthew Lehnert

PhD supervisor (HDR or equivalent): Nicolas Bouché, GALPAC team

Co-supervisor:

Email address and phone number: nicolas.bouche@univ-lyon1.fr

Context and Project description:

Context: The [MUSE](#) instrument on the European Very Large Telescopes is a powerful integral field spectrograph (IFS) yielding hyperspectral data, with a spectra for each of the 90,000 pixels in the field-of-view. A typical MUSE observations contains about 100 galaxies over a wide range in redshifts from 0 to 6.

Detecting each source blindly and **finding out** the proper redshift solution are both challenging tasks. Regarding the detection, some galaxies have emission lines, while others have mainly broad stellar continuum. Detection algorithms exist to detect either the emission lines or on the broad continuum light, but none exist to perform both tasks simultaneously.

In addition, finding the redshift solution for each galaxy is a challenging task, as this usually requires manual inspection of the possible solutions for each source, a time consuming task.

This project will thus potentially transform the way one can extract and identify the redshifts of all galaxies in deep wide-field IFU observations.

Given the large demand and usage of MUSE, with the upcoming wide-field IFS such as blue-MUSE, and [WST](#), it is of paramount importance to develop new ways to exploit such data as it will lead to billions of spectra from the IFUs and hundreds of millions of spectra from the multi-object spectrograph.

Objectives: The first part the proposed thesis project will consists in developing tools for the redshift identification with Machine Learning (ML) techniques, such as super-vised classification using deep Convolutional Neural Networks (CNN) trained on galaxy spectra with redshifts extracted from various MUSE large programs from our team. Such tools will be developed on the 1D extracted spectra, and adapted to 3D data mini-cubes. The second part of the thesis will consist in developing either hybrid techniques or unsupervised ML methods in order to perform the source detection and classification tools from the 3D data-cubes directly.

CRAL is leading or co-leading several next generation of Integral Field Units such as [BlueMUSE/VLT](#), [Harmoni/ELT](#), [WST](#), and the 4MOST spectrograph which will lead to millions of spectra.

Starting PhD date: **Sep 2024**

Other foreseen funding than ED52: N/A