**Maestría en Ingeniería con Énfasis en Ingeniería de Sistemas y Computación**

**Introducción a la Investigación**

**Universidad del Valle**

**Harold Armando Achicanoy Estrella**

**Taller 1: Definición del Tema de Investigación**

***1. Fuentes de datos exploradas***: SCOPUS, ResearchGate, Google Scholar

***2.a. Búsqueda por keywords***: common bean cropland mapping, landraces remote sensing, image classification, land-cover mapping

***2.b. Búsqueda por autores***: Perushan Rajah, Priscylla Ferraz Câmara Monteiro, Jianfeng Zhou

***3. Búsqueda por revistas y conferencias***: Scientia Agricola, Precision Agriculture, IEEE Journal of selected topics in Applied Earth Observations and Remote Sensing

***4. Búsqueda por centros o grupos de investigación***: School of Remote Sensing and Information Engineering (Wuhan University), Faculty of Environmental Science and Engineering (California Institute of Techonology), Cartography & Remote Sensing (University of Maryland)

***5. Búsqueda por otros artículos, publicaciones en conferencias, libros y documentos relevantes***: Los artículos agregados corresponden a otras aplicaciones donde se han aplicado algoritmos de clasificación sobre otros cultivos.

***6. Referencias bibliográficas en formato bibtex***: La lista completa de referencias se encuentra en el archivo *thesis\_references.bib*

***7. A partir de la información adicional actualizar lista inicial de keywords, autores, instituciones, revistas y centros o grupos de investigación y repetir todos los pasos anteriores***

***8. Resumen de los artículos top 5 más relevantes***: En cuanto al tema de clasificación a trabajar en el proyecto de tesis, el top 5 de los artículos más relevantes ordenados en orden cronológico está dado por:

* Rajah et al., (2017). Determining the optimal phenological stage for predicting common dry bean (*Phaseolus vulgaris*) yield using field spectroscopy
* Rajah et al., (2015). Varietal discrimination of common dry bean (*Phaseolus vulgaris L.*) grown under different watering regimes using multitemporal hyperspectral data
* Monteiro et al., (2013). Vegetation indices simulated from different sensors for the estimation of biophysical variables of common bean
* Monteiro et al., (2012). Assessing biophysical variable parameters of bean crop with hyperspectral measurements
* Epiphanio et al., (1996). Vegetation indices for remote sensing of beans (*Phaseolus vulgaris L.*)

En términos generales, los anteriores artículos están enfocados en la evaluación de los caracteres más importantes a la hora de identificar áreas de cultivo de frijol, como lo son los índices de vegetación y parámetros biofísicos, esto a partir de la utilización de imágenes multi-espectrales e híper-espectrales usando los diferentes niveles espectrales para la composición de dicho índices. Las dos fuentes de datos utilizadas son por lo general, imágenes de sensores remotos satelitales e instrumentos espectroscópicos en campo. Adicional a esto, dos papers se enfocan en tareas de clasificación donde se menciona la clasificación entre variedades de frijol y la determinación de la etapa fenológica óptima para realizar la predicción del rendimiento del cultivo. En ambos casos, los algoritmos de clasificación entran a jugar un rol principal utilizando los caracteres más importantes para realizar tareas de clasificación.

***9. Cuál cree usted que podría ser su aporte en el tema enfocado***: Desde mi punto de vista, creo que el aporte más importante en la detección de áreas de cultivo de variedades criollas de frijol a partir de imágenes satelitales está centrado en la integración de metodologías de clasificación que exploten al máximo las características de los datos de entrada: imágenes provenientes de diferentes sensores, diferentes niveles de medición del espectro electromagnético, diversa resolución espacial e información a través del tiempo.

***10. Descripción de la idea inicial del proyecto de investigación***: La idea inicial del proyecto de investigación consiste en realizar la clasificación e identificación de áreas donde está presentes variedades criollas de frijol a partir del uso de imágenes satelitales.

En términos generales existen dos grandes clases de materiales que se cultivan: por un lado están las variedades de alto rendimiento que se producen a partir de mecanismos de mejoramiento genético formal, y por otro lado, están las variedades criollas que son mejoradas a partir de procedimientos de selección natural e intervención de las comunidades locales.

El interés del estudio surge debido a la necesidad de colectar y conservar las variedades criollas dado que estas ofrecen una rica fuente de diversidad genética que puede ser aprovechada en el mejoramiento genético formal, transmitiendo caracteres de adaptación a ciertas condiciones climáticas y como fuente de nutrientes.

En este sentido, los algoritmos de inteligencia artificial, específicamente los que se enfocan en el reconocimiento de patrones en imágenes, entran a jugar un rol primordial al momento de abordar el problema, dado que al contar con imágenes satelitales tomadas en diferentes momentos del tiempo y a partir del conocimiento de la fenología del cultivo, es decir, las etapas de desarrollo que abarcan el periodo de siembra hasta cosecha, se necesita construir un algoritmo clasificador que permita:

*a)* Diferenciar entre área sembrada de frijol vs área sembrada de otros cultivos

*b)* Discriminar entre área sembrada de variedades de frijol de alto rendimiento vs área sembrada de variedades criollas de frijol, y

*c)* Aprovechar la información espacio-temporal de las imágenes satelitales para obtener una clasificación más precisa.