High-quality annotated data for accurate and reliable plant disease predictions to help farmers make informed decisions and take necessary preventive measures.

AI4CROPHEALTH project team members held a meeting on 5 April 2024 at Morogoro hotel to evaluate the progress on the data collection, data annotation and data labelling. The target was to establish a high-quality annotated dataset to enhance development of an accurate Artificial Intelligence (AI)/Machine Learning (ML) models that accurately detects and classifies plant disease patterns using remote-sensed (Unmanned Aerial Vehicles images) and digital images. Specifically, four crops considered for data collection are Beans, Maize, Cassava and Rice collected in four regions including Pwani, Morogoro, Mbeya and Songwe.

Figure 1 AI4CROPHEALTH project members proceeding with discussion during the Meeting held at Morogoro Hotel.

In his introductory remarks, Dr. Farian S. Ishengoma who is CO-PI, emphasized that Image annotation and labeling are vital processes in the world of computer vision and machine learning**.** With the increasing demand for accurate object detection and image recognition**,** it is crucial to understand the best practices for annotating images to train machine learning algorithms effectively and improve their prediction performance.

Then, Dr. Ishengoma oriented the team members on how to use AI-supported image annotation tool to enhance data annotation process. He explained that image annotation begin by determining what should be labeled in the images and then instructing people involved in annotation the key steps required to perform the annotation tasks using an image annotation tool. After acquiring skills on proper data annotation procedures for the project, the collected images were distributed to each annotator to continue annotating images using an image annotation tool. The AI-supported image annotation tool helps project team to organize and store the structured and unstructured data while providing a model training framework. He further emphasized that the data annotation and labelling should be carried out diligently so that can improve the prediction accuracy of ML algorithms, also the annotated data will be submitted to the experts for validation

Also, Mr Florence Kifyoga, MSc students supported by AI4CROPHEALTH project as part of capacity building component presented progress on image annotation and labelling on behalf of the team. Figure 2, 3 and 4 demonstrate the overview of annotated images for beans, rice and maize extracted from the data annotation tool. During the data annotation process, necessary features within a dataset are accurately labeled, allowing the ML model to identify recurring patterns and develop robust ML algorithms. This labeled data serves as the foundation for training ML models and is instrumental in achieving optimal performance.

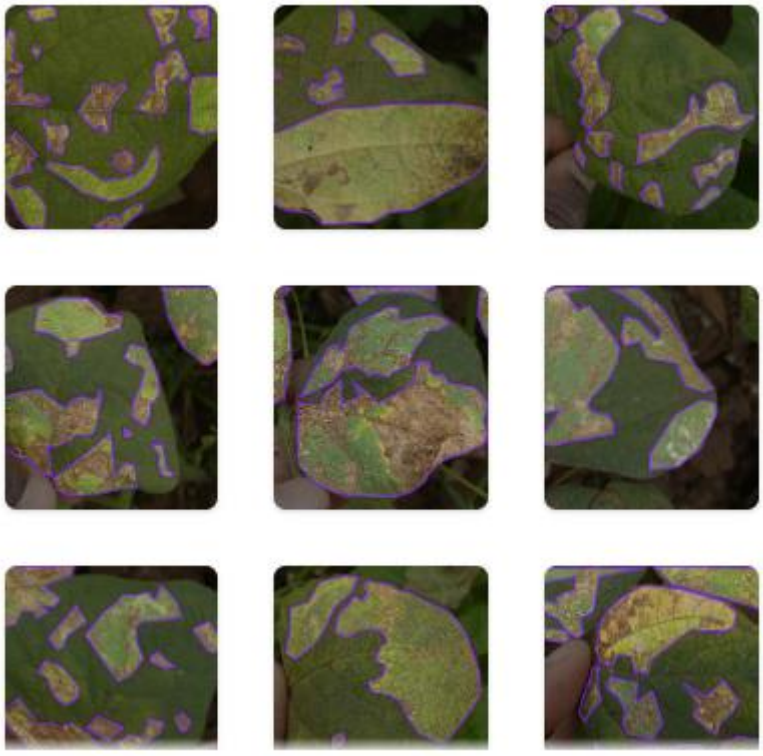

Figure 2 Sample annotated bean leaves images indicating the affected areas

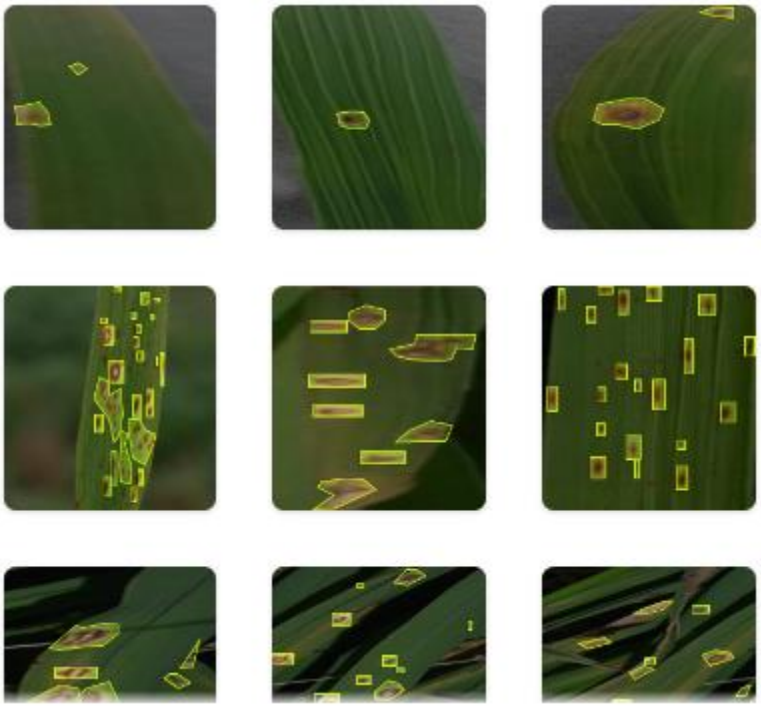

Figure 3 Sample annotated rice leaves images indicated the affected areas

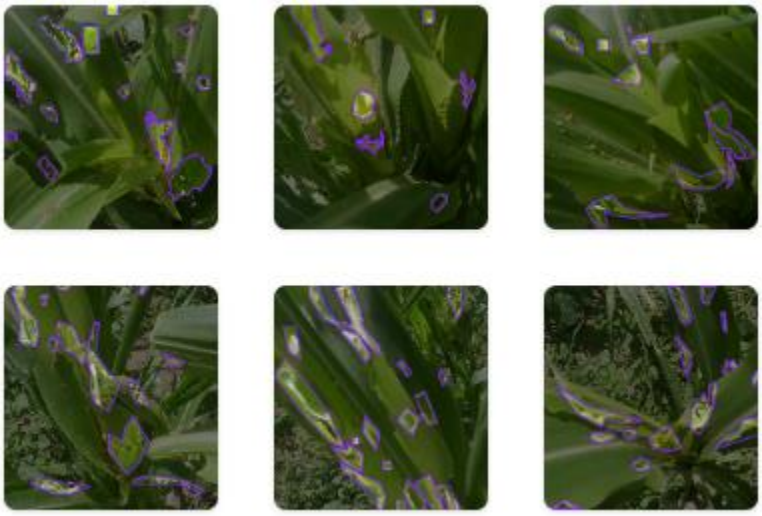

Figure 4 Sample annotated maize leaves images indicating the affected areas

Dr Michael P.J. Mahenge (PI), further explained that data annotation process is crucial to the project because we target to enhance accuracy and reliability of plant disease predictions to help farmers make informed decisions and take appropriate preventive measures timely. Undertaking this process efficiently guarantees well trained AI/ML algorithms that will efficiently predict disease occurrence and recommend preventive measures, thereby contributing to the 2nd Sustainable Development Goal (SDG) goals of ending hunger, achieving food security and

improved nutrition, and promoting sustainable agriculture by 2030. Also, will contribute to the 4th national goal aiming to create a well-educated and learning society, and our project will use innovative AI/ML-empowered tools to spread knowledge on plant disease and their good management practices. Moreover, the project will contribute to the 13th SDG target of strengthening resilience and adaptive capacity to climate-related disasters and the 5th African Development Agenda 2063 target of modern agriculture for increased productivity and production.

This project is funded by the United Republic of Tanzania through Commission for Science and Technology (COSTECH), the project team greatly appreciate the supports from the funder.