**Draft for KAZA – LCLU Classification Project**

**Objectives:**

* Conduct pixel based classification of Sichifulo area for 2023
* Conduct OBIA classification using colab (and GEE/geemap) for 2023
* Comparing accuracy and finding best approach

**Work Packages**

|  |  |  |  |
| --- | --- | --- | --- |
| Pixel based | | Object based | |
| Classic LC classification | Agricutlrual fields detection | Object based | Comparison between pixel and Obia |
| Checking training data for errors | | | |
| Collect training data for the other 2 areas (only agriculture) | | | |
| Aquire and preprocess high resolution satelite data 🡪 s2 | | | |
|  |  | Conduct the segmentation using the SAM algorythm in colab | Compare different accuracies/confusion matrix, etc. |
|  |  | Vailidfy the segmentation | create bar plot of accuracy per class or both pixel- and objectz based |
| Calculate suiting Indices | Calculate NDVI metrices | Extract feautres such as texture, shape, area etc. | Bar plot of accuracy metrices of both classes (overall, producers, users, kappa coefficiant) |
| Conduct classification - including Accuracy assessment | Conduct classification - including Accuracy assessment | Conduct classification - including Accuracy assessment  (both for classic classification and also for agricultural field detection?!) | McNemar’s Test?! |
| (Map classification) | (Map classification) | (Map classification) | Visual inspection and selection of screenshots (overly both classifications/ identify patterns via gis tools) |
| * 16th june |  | 16th june – 14th july | 14th july – 21th july |

**Detailed Methodology**

**Preparation**

* Checking training data for errors
  + Metadata inspection
    - Check for doubled IDs
    - Check for empty properties e.g. landcover property
* Aquire and preprocess hgih resolution satelite data 🡪 s2
  + S2
  + Clip to aoi
  + Cloud mask
  + Filter for years 2020 and 2023
  + Filter needed bands
    - B2-4, B8, B11-12, red edge bands helpful?
  + Identify and calculate helpful Indices/paramters for classification
    - NDVI, NDWI, NDBI, EVI, radar backscatter?! Others?!
    - Include paramters like slope, aspect? 🡪 SRTM
  + Rescale bands Red, Green, Blue to represent true reflection values
  + Create median composite
    - Over whole year or specified season for example?
* Collect training data for the other 2 areas
  + Just by visual inspection of the median composites??
  + 🡪 ask nuno how to collect the training data
  + But before, do own research
  + Ask him to scip the new area and collecting the samples
  + How many per class?

**Pixel based classification**

* Conduct and iterate classification (including validation) to find most suitable band combinations
  + Overlay training collection of respective year with median composite for respective year
  + Split in training and validation data
  + Create the classification function which includes:
    - Create and train a classifier 🡪 probably random forest is best
    - Apply the classifier
    - Majority filter to exclude single class pixels
    - Classify validation data and create error matrix and accuracy metrices
  + Apply the classification function
  + Iterate the classification with different band/indices combinations
    - Iterate each band/indices combination multiple times to get a stable accuracy assessment?
    - Chose best fit based on literature review, feature importance scores, maybe automated band selection ?! (probably to much)
* Each step has to be done for both target years seperatly
* Create an appealing map of the most promising classification output

**Object based classification**

* Conduct the segmentation using the SAM algorythm in colab
* validify the segmentation (no standardized way exists)
  + visual inspection
  + Conduct evaluation via indicator assessment ??
    - Common indicators: Precision, recall, entropy, Moran’s1,
    - Common methods to evaluate the indicators
      * OA (Li, Xao 2004)
      * entropy (Mantilla and Yari 2017)
      * Recall (Jordan and Angelopoulou 2012)
      * Inter, Intra, Intra-inter, Borsotti, Rosenberger (Chabrier et al. 2004)

🡪 probably to time consuming!!

* + Automated hyperparameter tuning ?! 🡪 libaries exist
    - E.g. Optuna, Hyperopt
  + Maybe ChatGPT suggestions
* Import the segmented scene into GEE
* Extract feautres such as texture, shape, area etc.
* Conduct and iterate classification (including validation) to find most suitable band combinations
  + Same methodology as in pixel based
* Create an appealing map of the most promising classification output

Segmentation validation – ChatGPT suggestions

* Visual inspection
* Intra segment homogeneity
  + Calculate the variance or standard deviation of pixels values within each segment. Low intra-segment variance indicates homogenoues segments
* Inter-segment discrepancy
  + Evaluate the differences between adjacent segments. Ideally adjacent segments should have distinct properties
* Multi-scale segmentation consistency
  + Perform segmentation at multiple scales and check for consistency. Robust segments should be relatively stable across different scales

**Agricultural fields mapping**

NDVI metrices

* Max per year
* Min per year
* Stdv per year
* Timeframe of growing season using threshold values
  + Start of season, end of season
  + What thresholds are suiting a savannah region in sambia?
* Sudden drop in ndvi 🡪 extract the biggest difference/drop between 2 successive dates in NDVI values
* NDVI range
* Cumlative ndvi