Imperial College London
Department of Earth Science and Engineering
MSc in Environmental Data Science and Machine Learning

Independent Research Project Plan

Managing illegal gold mining in Ghana from space

by
Olga Kostur

Email: ok121@ic.ac.uk
GitHub username: edsml-ok121
Repository: https://github.com/ese-msc-2021/irp-ok121

Supervisors:
Professor Yves Plancherel

Pablo Brito Parada

July 2022

Introduction & Literature Review

Non-fuel mineral resources significantly contribute to the development of national economy in Ghana (Ericsson and Löf, 2019). Interestingly, Ghana was famous for its gold resources since pre-colonial era and was known as the Gold Coast (Rodney, 1969). Gold mining remains crucial for Ghana's economy and contributed 7.8 billion Ghanaian cedis (1.3 billion US dollars) in 2020 to GDP, which was the lowest value added of the metal since available records. In contrast, highest contribution to national GDP by gold was in 2018, accounting for 9.8 Ghanaian cedis. Hence, effective management of Ghana's key resource is vital for prosperity of communities and investment in local infrastructure.

Existing records of Ghana's gold mining are very incomplete and lack essential data about illegal activities - galamsey, a combination of Ghanaian words "gather" and "sell." (Balaniuk, 2020). Center for Remote Sensing and Geographic Information Services (CERSGIS) estimates that approximately 35% more gold is extracted through small-scale mines informally and without a licence, supporting 4.5 million people. The situation is problematic not only for government and environment, but for people themselves – most of them do not have knowledge about mercury and heavy metals. This can lead to severe kidney and neurological diseases, while contaminating drinking water for entire local communities. Overall, 1.8 million ill-health years or early deaths can be attributed to mercury poisoning in small-scale gold mining globally (www.theworldcounts.com).

Furthermore, the environmental degradation due to illegal gold mining is severe, as galamsey operate at a broader spatial scale. Galamsey is 7 times greater (Barenblitt, 2021), than large-scale mining involving heavy machinery. Water is essential for the mining process, as tropical forests and riverbanks are being excavated due to presence of natural water that quickly fills up ponds. This results in conversion of primary forests to mining ponds and bare land, significantly changing the biodiversity, ecosystems, and global climate (Gerson, 2021). Moreover, miners use acids such as pyrite to weather sulphide minerals, which raises acidity of water and contaminates environment with metals (Viljo et al. 2003). If unmonitored and unregulated this activity may lead to knock-on effect and unexpected environmental feedbacks. Hence, it is vital to detect illegal mining and monitor its effects on surrounding environment in real time in order to prevent environmental degradation as well as design effective environmental restoration practices to reduce vulnerabilities to acceptable levels.

Many studies attempted to detect these mining activities or analyse their impact on environment using satellite imagery. Csillik (2019) used satellite remote sensing coupled with airborne LiDAR to estimate aboveground carbon stocks from gold mining and deep learning models in Peruvian Amazon. Adamek (2019) used indices derived from spectral bands to highlight presence of environmentally dangerous metals, concluding that it is effective in revealing the basic environmental condition of gold mining sites. Labbe (2021) presented two alternative methods in detecting mining ponds - Object Detection (RetinaNet algorithm creates bounding box around each pond) and Sematic Segmentation (U-Net algorithm allocates each pixel to class). He established that U-Net is more appropriate in detecting ponds, however hypothesising that using a combination of two (first RetinaNet, then U-Net) would lead to best results. Camalan's (2022) work is a more complex object segmentation algorithm Re-CNN (RNN & CNN combined) that is based on manually labelled Sentine-2A images of mining ponds where ponds were given a category – active, transition or inactive. When ponds get abandoned algae starts to colonise the micro-environment, changing the pond reflectance colour from brown to green and dark green, depending on how long the pond was abandoned for. The algorithm, however detects temporal changes between 2 timestamps outputting images of binary and multi-class changes.

Problem Description and Objectives

There is a vast literature on managing, detecting, and monitoring environmental effects of illegal gold mining, ranging from traditional machine learning such as random forest and support vector machine to unsupervised convolutional auto encoders for change detection in illegal gold mining (Scammacca, 2021; Ibrahim, 2020). However, Moomen (2022) published an overview of applications of satellite imagery in monitoring artisanal gold mining, concluding that despite vast availability of Earth observation data, little progress has been achieved in creation of consistent and effective data. Nevertheless, it would be possible to derive robust internationally acclaimed methods of detecting these activities in real-time to prevent environmental degradation.

This study aims to create a monitoring framework to detect gold mining ponds and classify them into active, inactive and transition state by using dataset created by Camalan et.al (2022) of labelled images of gold mining ponds in Peru, Indonesia, Venezuela, and Myanmar at 2 timesteps - August 2021 and 2019. The study will add 6 manually classified images of Ghana's gold mining, expanding open-source dataset for future use, and ensuring that resulting tool is generic and applicable worldwide. Data augmentation techniques will be crucial for this project, as manually labelling Sentinel-2 images is very time consuming.

Some common data augmentation techniques include horizontal and vertical flipping, rotation, zooming, rotation, and translation (Abdelhack, 2020).

While existing training data will be used for this project, after an extensive literature review, no studies were found that detect and classify ponds with satellite images, using free cloud platforms – Google Earth Engine (GEE) and Google Colaboratory (GC). While Akpah (2021) developed a successful CAE network (type of CNN) to classify ponds as legal or illegal, the drone DJI Phantom 4 UAV, capturing images of 4000 × 2250 pixel resolution is expensive and difficult to reproduce. Conversely, Ibrahim (2020) used Sentinel-2 images from GEE platform chose broad classes - “mining areas”, other “non-vegetated areas”, “vegetated areas”, and “water”. Such method produced an effective tool for preliminary localisation of mining activities despite not distinguish between pond types.

Thus, the software developed will be accessible, reproducible, and a free tool that is able to detect and differentiate ponds depending on their use status, which is essential for effective environmental management. In addition, presence of 3 heavy metals, commonly present in gold mining areas (Lupa, 2015) - Ferrous iron mineral group, Minor ferric iron mineral group, Clay-sulphate-mica-marble mineral group will be assessed using spectral indices to analyse influence of pond category on metal abundance. This approach will help to interpret the states of environmental chemistry and estimate the scale of damage as well as reveal how it recovers naturally post site abandonment. The project will start by identifying areas of interest in Ghana and creating data pre-processing pipelines, followed by the generation of pond labels in LabelBox tool – example of RGB input and output is visible in Figure 1.

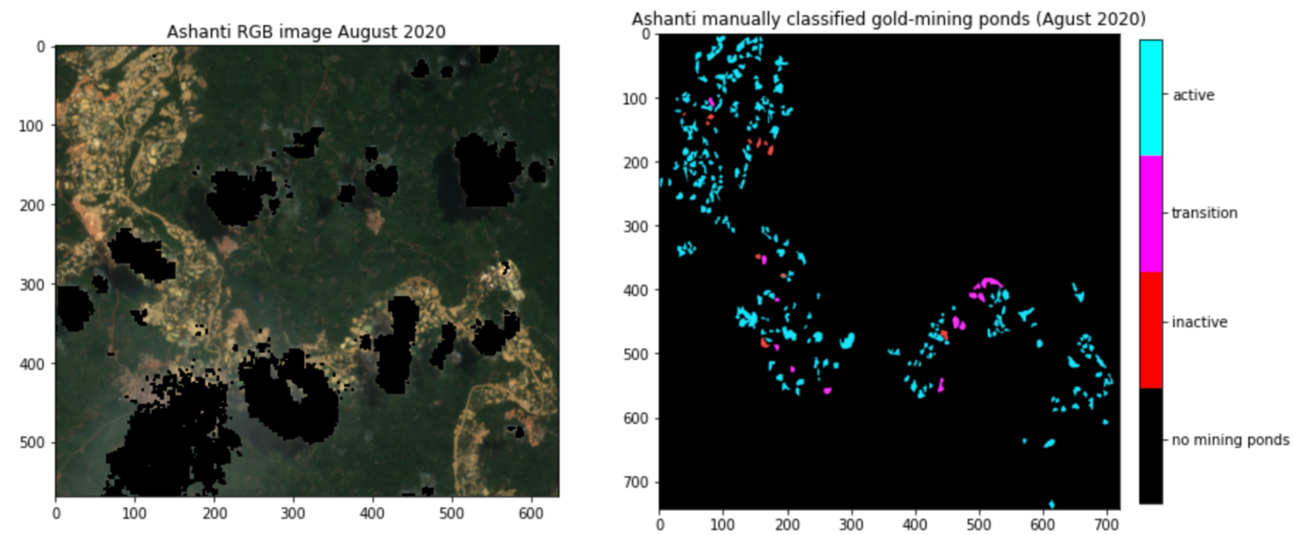


Figure 1: Ashanti Region. A - RGB image with min-max enhancement. B - manually classified mining ponds

Then, several CNNs and traditional ML classifiers such as support vector machine (SVM) or random forest (RD), will be trained to select most robust model. Finally, by overlaying data of heavy metals with classes' labels the chemical profile of ponds will be built, revealing how concentrations of metals changes as pond transitions from active to inactive state (Fig 2).

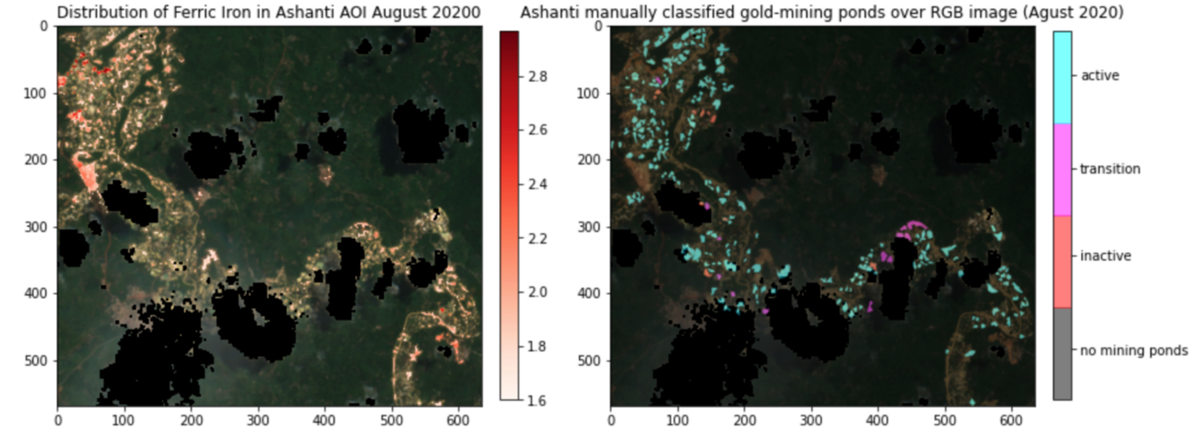
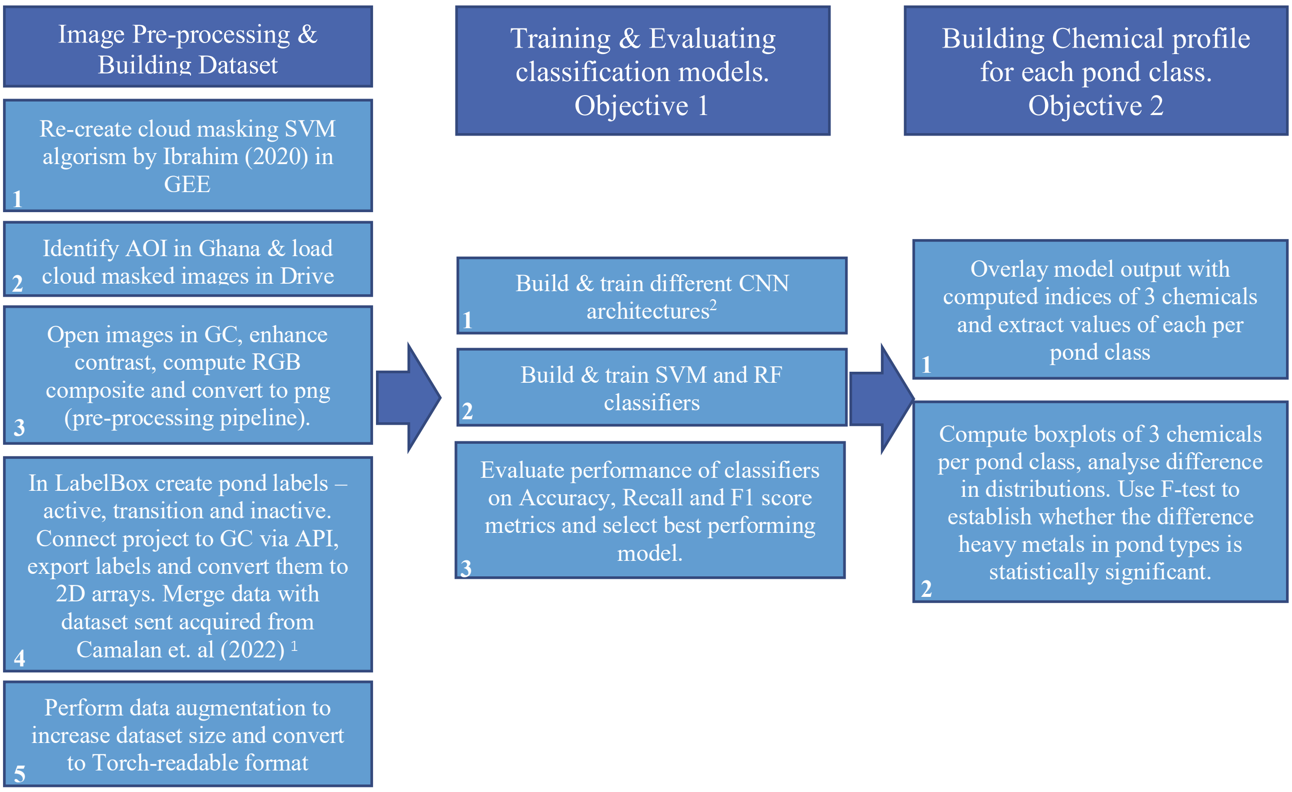


Figure 2: Minor Ferric Acid distribution in Ashanti, August 2020

The workflow of this project is documented in Fig 3.



[1] Data was requested from Seda Camalan, the main author of the paper, but not received yet. There is a risk that data will come late, disrupting project's workflow. A risk mitigation will be creating more data in LabelBox. Data augmentation will be used to create sufficient size dataset.

[2]CNN architectures that will be tried are still to be finalised, but AlexNet, ResNet and GoogleNet will be tried

Figure 3: Project workflow

Progress to Date and Future Plan

To date the focus was data pre-processing and exploration of LabelBox tool to expand dataset created by Camalan (2022). South region of Ghana was searched visually using GEE and 3 AOI were selected that have high concentration of mining ponds – Fig 4. Cloud masking algorithm by Ibrahim (2020) was recreated in GEE and applied to 6 images (3 August 2021 and 3 August 2019). In GC functions to load and to pre-process images (e.i. enhance image contrast, perform histogram matching for images from same AOI and convert to png) were created and tested. After experimenting with RBG and water index images for manual labelling in LabelBox, the most effective approach was discovered – improving image resolution with unblurring tool – Remini (app.remini.ai), then resizing to original resolution with cv2 library in python. This way, edges of ponds become more distinct, increasing precision and accuracy of their manual segmentation. Finally, the LabelBox API was explored and code for exporting images and converting to numpy arrays was written.

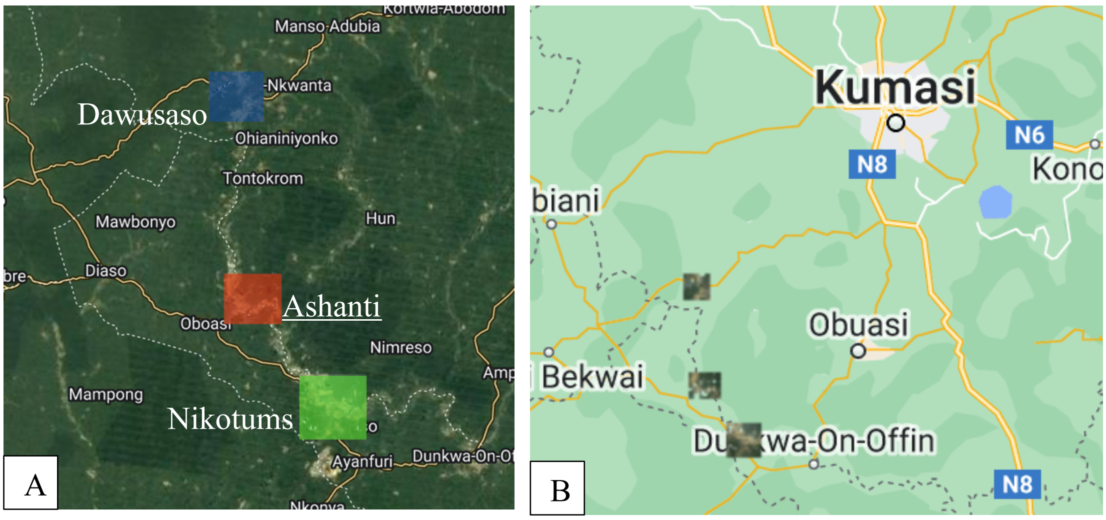


Figure 4: Location of 3 study Regions. A - zoomed in on satellite map 3 AOI in Ashanti region. B - zoomed out base map showing 3 locations are south-west off Kumasi

The next steps of the project will be taken according to provisional timetable, documented Table 1.

Table 1: Key dates of the project and planned processes.

Dates	Process
6 June – 20 June	1. Literature Review, formulating research objective and aims. 2. Identifying AOIs and recreating cloud masking algorism by Ibrahim (2020)
21 June – 1 July	1. Implementing pre-processing functions 2. Labelling data in Labelbox 3. Implementing data export code via API and converting labels to 2D arrays. 4. Project Plan write up
2 July – 25 July	1. Finalising dataset (potentially merging with own data with Calaman's if provided on time), performing data augmentation, and converting it to PyTorch dataset. 2. Building CNN, SVM and RF classifiers & training them. 3. Evaluating results of different classifiers and choosing the most optimal one.
26 July – 7 August	1. Overlay model results with chemical indices, extract data per class. Analyse distributions and test for statistical differences. 2. Revise code, finalise custom python packages and pipelines and write tests. 3. Start writing thesis (make figures, plan sections, write introduction)
8 August – 2 September	1. Thesis write up 2. Finalising code, tests, and GitHub workflows. 3. Revision of GitHub repository and final submission.

Bibliography

1. Abdelhack, M., 2020. A comparison of data augmentation techniques in training deep neural networks for satellite image classification.

2. Akpah, S., Ziggah, Y.Y. and Mireku-Gyimah, D., 2021. Assessing the suitability of convolutional auto encoder as an unsupervised tool for image classification in artisanal small-scale gold mining environment.

3. Barenblitt, A., Payton, A., Lagomasino, D., Fatoyinbo, L., Asare, K., Aidoo, K., Pigott, H., Som, C.K., Smeets, L., Seidu, O. and Wood, D., 2021. The large footprint of small-scale artisanal gold mining in Ghana. Science of the Total Environment, 781, p.146644.

4. Camalan, S., Cui, K., Pauca, V.P., Alqahtani, S., Silman, M., Chan, R., Plemmons, R.J., Dethier, E.N., Fernandez, L.E. and Lutz, D.A., 2022. Change detection of amazonian alluvial gold mining using deep learning and sentinel-2 imagery. Remote Sensing, 14(7), p.1746.

5. Center for Remote Sensing and Geographic Information Services (CERSGIS) Small scale mining portal. Accessed June 28, 2022

6. Csillik, O. and Asner, G.P., 2020. Aboveground carbon emissions from gold mining in the Peruvian Amazon. Environmental Research Letters, 15(1), p.014006.

7. Ericsson, M., and Löf, O (2019). Mining's contribution to national economies between 1996 and 2016. Mineral Econ. 32, 223–250. doi: 10.1007/s13563-019-00191-6

8. Gerson, J.R.; Topp, S.N.; Vega, C.M.; Gardner, J.R.; Yang, X.; Fernandez, L.E.; Bernhardt, E.S.; Pavelsky, T.M. Artificial Lake Expansion Amplifies Mercury Pollution from Gold Mining. Sci. Adv. 2020, 6, eabd4953.

9. Health Effects of Gold mining, Available at: https://www.theworldcounts.com/challenges/planet-earth/mining/health-effects-of-mining Accessed (June 28, 2022)

10. Ibrahim, E., Lema, L., Barnabé, P., Lacroix, P. and Pirard, E., 2020. Small-scale surface mining of gold placers: Detection, mapping, and temporal analysis through the use of free satellite imagery. International Journal of Applied Earth Observation and Geoinformation, 93, p.102194.

11. Labbe, N., 2021. Detecting illegal gold mining sites in the Amazon forest: Using Deep Learning to Classify Satellites Images.

12. Lupa, M & Lesniak, A 2015, New prospects for automated geological mapping with Sentinel-2 multispectral instrument in 16th Czech - Polish workshop ON RECENT GEODYNAMICS OF THE SUDETY MTS. AND ADJACENT AREAS.

13. Moomen, A.W., Lacroix, P., Benvenuti, A., Planque, M., Piller, T., Davis, K., Miranda, M., Ibrahim, E. and Giuliani, G., 2022. Assessing the Applications of Earth Observation Data for Monitoring Artisanal and Small-Scale Gold Mining (ASGM) in Developing Countries. Remote Sensing, 14(13), p.2971.

14. Naadi T., (11 August 2021) The illegal gold mines killing rivers and livelihoods in Ghana

15. Rodney, W., 1969. Gold and slaves on the Gold Coast. Transactions of the Historical Society of Ghana, 10, pp.13-28.

16. Scammacca, O., Gunzburger, Y. and Mehdizadeh, R., 2021. Gold mining in French Guiana: A multi-criteria classification of mining projects for risk assessment at the territorial scale. The Extractive Industries and Society, 8(1), pp.32-43.

17. The illegal gold mines killing rivers and livelihoods in Ghana, Available at: https://www.bbc.com/news/av/world-africa-58119653, (Accessed 28, Jun 2022)

18. Viljo, K, Hilkka, A, Francis, C, Eija, H, Jukka, L & Maarit, M 2003, Guidelines for Image Processing procedures and algorithms for mining related contamination and impact detection from airborne imaging spectroscopy. Available from: <http://tupa.gtk.fi/>. [10 December 2019].