AI for Mitigating Effects of Climate and Weather Changes in Agriculture

Narmada Balasooriya *1, C.D. Athuraliya †1,2 and Janak Gunatilleke ‡1

[1]ConscientAI Labs, Colombo, Sri Lanka
[2]LIRNE*asia*, Colombo, Sri Lanka

Problem

In recent years, floods, landslides and droughts have become an annual occurrence in Sri Lanka. Despite the efforts made by the government and other entities, these natural disasters remain challenging mainly to the people who live in high risk areas. It is also crucial to predict such disasters early on to facilitate evacuation of people living in these areas. Furthermore, Sri Lankan economy largely depends on agriculture, yet this sector still remains untouched by recent advancements of AI and other predictive analytics techniques. There is an increased tendency amongst Sri Lankan youth to refrain from agriculture sector due to the lack of technology adoption and risks associated with it [1]. The work by Peiris et.al [2] demonstrates how seasonal climate data is used to predict coconut yield in Sri Lanka. Another Sri Lankan tech company has initiated the project AiGrow [3] to increase the use of state-of-the-art technology in agriculture sector by utilizing AI, smart sensor systems and fertigation systems (automated fertilizer delivery machine). Regionally, Pulse Lab Jakarta has developed a visualization tool to track the impact of forest and peatland fires in Indonesia [4].

Proposal

The solution is an AI based platform that generates insights from emerging data sources. It will be modular, extensible and open source. Data science and machine learning techniques are mainly employed to automate the analysis tasks of heterogeneous and complex data. More specifically deep learning based computer vision is used for tasks such as semantic segmentation and object

*narmadabalasooriya@gmail.com
†cd@conscient.ai
‡janak@conscient.ai

detection to identify land and crop attributes such as land cover, land use, water bodies, water levels, crop stage etc [5]. Other remote sensing attributes such as atmospheric conditions, soil structure are analyzed using machine learning and data science techniques. These low level attributes, extracted from both remote sensing data and aerial imagery will form aggregate features and will be used to build predictive models. Similarly to any other real world AI system, the end solution will consist of multiple data pipelines to extract data, analyse and present results through APIs. Drones can be used to capture higher resolution (both spatial and temporal) aerial imagery in regions of interest, e.g. high disaster risk regions and crop fields. The presentation layer will be a public API that can be consumed through a portal such as Disaster Management Centre of Sri Lanka [6].

Impact

Timely insights on climate and land conditions extracted from remote sensing data and aerial imagery will allow managing disasters better through early warnings, prompt relocations and well-planned post-disaster relief operations. Also for agriculture sector, the solution will allow continuous monitoring of crops using aerial imagery and automated data analysis will allow performing modeling and prediction at scale. These outcomes enable efficient distribution of the crop yield, better economic planning for the agriculture sector, effective price controls and optimized distribution channels. Analysing these two interrelated aspects (disasters and crop) in the long term will give a better understanding about seasonal effects. It will allow authorities to make more timely adjustments and take data-driven decisions in each related sector. All these outcomes will positively affect living conditions of the people since they are directly related to the basic needs such as food and shelter.

Evaluation

Evaluation of the system will be performed at two levels: low-level attributes and high-level, aggregate features. Availability of ground truth for low-level attributes is relatively high since they are based on historical and official data. However acquiring ground truth for certain attributes can be still challenging. Evaluating high-level predictions can be more challenging, mainly due to the unavailability or sparsity of historical data caused by seasonal occurrence. Yet the effectiveness of the proposed method will be quantified by evaluating the platform both before and after, crop seasons and disasters. Initially these evaluations will be performed for smaller regions and consequently extended to larger regions such as districts and provinces. Large-scale predictions may have significant errors due to uncaptured fluctuations in weather such as low pressure in the Bay of Bengal. Incorporating features from more sophisticated climate models is a potential solution for such challenges.

Risks

One of the biggest risks of this solution is its vulnerability to uncaptured fluctuations in weather and climate attributes. In such cases the platform may signal false alarms to the end users. This will cause unnecessary panic and may even cause distrust of both general public and policy makers. The solution is trying to predict critical events that directly affect day-to-day operations of the people. Hence failing to predict a severe threat or a condition will also question the effectiveness of using novel data sources and techniques. Another major risk is relying on external data sources. Unexpected interruptions or terminations of key data streams can make the solution obsolete. These risks can be mitigated to some extent by validating the robustness of predictive models with complementary data sources. Lack of compliance with conventional systems is another potential risk that can be reduced with the involvement of domain experts.

Data

We have extracted and utilized data from various remote sensing and aerial imagery data sources such as Google Maps, USGS Earth Explorer [7], DigitalGlobe Open Data [8] and static datasets such as SpaceNet [9] and SAT-4/SAT-6 airborne datasets [10]. This involved multiple extraction methods including direct downloads, API access and using data providers' user interfaces to access data. We evaluated potential ground truth data sources such as OpenStreetMap (OSM) to validate the outputs of predictive models especially in regions where up to date, official data is unavailable. Most of these sources are publicly available and can be accessed freely. Social media sources that contain point of interest information such as Foursquare also have been used to obtain ground truth in our ongoing work. Social media data needs to be accessed through their respective APIs and some platforms have released their own datasets specifically for humanitarian efforts [11].

Labels

Our solution requires additional data annotation. As a result of supervised learning algorithms employed to build predictive models, ground truth labels are essential to make this solution effective. Some of the aforementioned datasets for low-level attributes inherently contain labels (e.g. high frequency, temporal data) but ground truth data is largely unavailable for seasonal and static features. In such cases we have used proxy and complementary data sources to achieve ground truth. Reaching out to relevant government organizations also has been an option when official data is available. Nevertheless there are cases where ground truth data collection and manual annotation is unavoidable. In such cases our approach is to initially focus on smaller regions to reduce the required manual labour for collection and annotation tasks. Crowdsourcing is another viable option that can be scaled when initial groundwork such as platforms and training are in place.

Social System

Our team consists of experts that utilize AI technologies such as machine learning and computer vision in both commercial and non-commercial applications across domains. We have conducted research in related fields and published at international conferences such as European Conference on Computer Vision (ECCV). We are competent in applying these technologies in solutions that address real world problems. Compared to traditional software engineering which is mostly deterministic, developing systems with (probabilistic) AI elements is challenging. We have an adequate understanding about this new engineering challenge from our previous and current work. Our background in software engineering has made us knowledgeable in other essential engineering aspects such as cloud computing, version control, integration and testing. Additionally, development of this platform will be benefited from the expertise of climatologists, disaster scientists as well as the stakeholders in policymaking process. Our partners are working closely with government organizations of Sri Lanka.

Technical System

To date we have conducted relevant research and implemented below given core modules of this solution in collaboration with our partners.

1. Data extraction for remote sensing data, aerial imagery and point of interest data from multiple sources
2. Identifying built-up areas using deep semantic segmentation
3. Crop extent estimation and crop yield prediction using remote sensing data

We have published the results of above work as research publications and technical reports including quantitative results. These results have demonstrated the potential of using big data sources with AI technologies. We have built predictive models using remote sensing data such as vegetation index to estimate crop extent and crop levels. We explored the feasibility of using deep semantic segmentation to estimate land cover, more specifically built-up areas from satellite imagery and tried to infer land use by combining estimated land cover with point of interest data extracted from sources such as Foursquare.

References

[1] Madura Thivanka. Why is Sri Lanka's youth stepping down from agriculture? http://www.ft.lk/article/627465/Why-is-Sri-Lanka-s-youth-stepping-down-from-agriculture. [Online; accessed 22-March-2019].

[2] T.S.G. Peiris, J.W. Hansen, and L. Zubair. Use of seasonal climate information to predict coconut production in sri lanka. *International Journal of Climatology*, 28:103–110, 2008.

[3] AiGrow. https://aigrow.lk. [Online; accessed 22-March-2019].

[4] Haze Gazer: A Crisis Analysis Tool. https://www.unglobalpulse.org/projects/haze-gazer-a-crisis-analysis-tool, 2016. [Online; accessed 22-March-2019].

[5] C.D. Athuraliya, Sameera Ramasinghe, and Sriganesh Lokanathan. Deep semantic segmentation for built-up area extraction and mapping from satellite imagery. Technical report, IDRC Digital Library, 2018.

[6] Disaster Management Centre. http://www.dmc.gov.lk. [Online; accessed 22-March-2019].

[7] USGS EarthExplorer. https://earthexplorer.usgs.gov. [Online; accessed 22-March-2019].

[8] Open Data Program - DigitalGlobe. https://www.digitalglobe.com/ecosystem/open-data. [Online; accessed 22-March-2019].

[9] SpaceNet on AWS. https://spacenetchallenge.github.io. [Online; accessed 22-March-2019].

[10] Saikat Basu, Sangram Ganguly, Supratik Mukhopadhyay, Robert DiBiano, Manohar Karki, and Ramakrishna Nemani. DeepSat: A Learning Framework for Satellite Imagery. In *Proceedings of the 23rd SIGSPATIAL International Conference on Advances in Geographic Information Systems*, SIGSPATIAL '15, pages 37:1–37:10, New York, NY, USA, 2015. ACM.

[11] Derrick Bonafilia, James Gill, Danil Kirsanov, and Jason Sundram. Mapping for humanitarian aid and development with weakly- and semi-supervised learning. https://ai.facebook.com/blog/mapping-the-world-to-help-aid-workers-with-weakly-semi-supervised-learning, 2019. [Online; accessed 22-March-2019].