AI for Sustainable Development (COMP0173)
Coursework 1

15 February 2022

In this coursework we select a particular sustainability related dataset as well as an AI model that has been applied to the dataset, and critically analyse both in the context of sustainable development goals (SDGs)[1].

1 Description of dataset and baseline

1.1 Chosen dataset and model

The dataset we analyse is the so-called 'brick kiln classification' dataset from [24], and we choose VGG-16 [19] as the baseline machine learning (ML) model from [10].

Lee et al. [10] develop an ML pipeline to detect the presence of brick kilns in satellite images taken above Bangladesh. However, the high-resolution images used in the pipeline were proprietary and therefore could not be released publicly. Instead, SustainBench [24] provides a lower-resolution alternative using Sentinel-2[2] imagery, which is the dataset we will analyse. This dataset comprises of images taken during the same time period as images used in [10]. The main elements are the images, each with size (64, 64, 13) consisting of 13 wavelength channels as defined by Google Earth Engine[3], and ground-truth labels which are 1 if the image contains a kiln (6,329 such samples), 0 otherwise (67,284 samples).

The baseline model for the brick kiln dataset, as implemented by [10], is a convolutional neural network (CNN) with VGG-16 [19] architecture. Specifically, the authors make use of the VGG-16 implementation from the Tensorflow Keras[4] library, and take a transfer learning approach by initialising the network with weights learnt during training on the ImageNet [3] dataset.

1.2 Assess its documentation

The documentation of both the dataset and baseline model are of a good quality. The authors of SustainBench have documented all datasets referenced in the paper on a website[5], including the brick kiln dataset. The documentation for this dataset includes a brief description of its relation to SDG-13 Climate Action, it also includes a description of the dataset itself, data format and code for loading the dataset to aid reproducing the results of the baseline model.

However, we believe the documentation could be improved by including information on how the data was specifically gathered. This additional information could assist contributors in extending the dataset to include other countries as it currently only has data for Bangladesh. Alternatively, it could serve as a baseline method for collecting similar datasets, e.g. satellite images of coal power stations, in a consist manner.

The baseline model methodology has been clearly documented in [10], including the method by which the final model performance was evaluated. This method differs from the standard approach used to evaluate classification models (where predicted labels are compared to true labels in a test dataset) since the model was run on images covering Bangladesh, not all of which were labelled. Since the dataset used in [10] could not be made public, an alternative "de-geo-identified" version stripped of geospatial metadata was released along with Python and R scripts for the entire ML pipeline, these are made available on Harvard Dataverse[6]. These materials make the baseline model replicable, however the exact results cannot be reproduced due to the difference in datasets. Therefore, to allow for fully reproducible results we believe it would be beneficial for the authors to include the model performance achieved on the alternative dataset provided.

2 Connection to SDGs

2.1 Related goals and targets

The brick kiln dataset and baseline VGG-16 model aim to address SDG-13: *"take urgent action to combat climate change and its impacts"*. One of the targets of SDG-13 is 13.2: *"integrate climate change measures into national policies, strategies and planning"*[15]. In order to achieve this target, governments need to be able to monitor compliance of their nations with, nationally and internationally, agreed environmental regulations. This can be difficult especially for low-income governments with limited resources and where the population may rely on informal industries as a main source of income. Brick manufacturing is one such informal industry, particularly in South Asia, which is a major source of pollution and carbon emissions [5].

The dataset and model also aim to address SDG-11: *"make cities and human settlements inclusive, safe, resilient and sustainable"*. In particular, target 11.6: *"by 2030, reduce the adverse per capita environmental impact of cities, including by paying special attention to air quality and municipal and other waste management"*.

2.2 SGD enabler and inhibitor

This ML solution enables progress towards SDG-13. Specifically, by identifying the number of brick kilns across the country, together with an estimate for the carbon emissions produced by one of the informal brick producers, estimates for the total carbon emissions from this informal industry can be obtained. These can then be used to monitor performance against national climate change policies and strategic targets, as well as to help define future policies and strategies.

Moreover, as the solution identifies the specific location of brick kilns, their proximity to regulated entities, such as schools, hospitals, railways, can be calculated. This information can support target 11.6 by measuring the impact of brick kilns on the air quality within cities across Bangladesh. By understanding the contribution of brick kilns to the degradation of air quality, particularly in terms of distance to populous regions, officials could take steps to regulate the brick producers and help reduce adverse impact to the population and environment.

On the other hand the solution may inhibit SDG-11, target 11.1: *"by 2030, ensure access for all to adequate, safe and affordable housing and basic services and upgrade slums"*. If the solution is successful in providing The Government of Bangladesh with enough information to take action in support of targets 11.6 and 13.2, this may result in the closure of some brick kilns or their relocation to less populous areas of the country. This could in turn increase the cost of developing housing, as bricks would need to be transported into the populous regions at an additional financial cost, and environmental cost of carbon emissions, from the transportation process. Thereby hindering the development of affordable housing and upgrading slums, where 47% of the urban population in Bangladesh lived in 2018 according to figures from The World Bank[7].

As a subsequent consequence, the solution may also inhibit SDG-13, target 13.1: *"strengthen resilience and adaptive capacity to climate-related hazards and natural disasters in all countries"*. Without more affordable housing people may have no option but to live in temporary housing structures such as tents within slums. These structures are unlikely to provide sufficient protection against climate-related hazards such as floods, or natural disasters such as earthquakes. Hence, this will not be strengthening resilience, as measured by indicator 13.1.1 in [15], against such events in Bangladesh.

3 Impact on sustainability

3.1 Social sustainability

Does it contribute to people's quality of life now and in the future? The dataset and solution could help reduce the amount of particulate matter (PM) and black carbon (BC) emitted by local unregulated and informal brick kilns. In Dhaka, Bangladesh, PM is the air pollutant most harmful to public health and the environment when compared to other measured criteria pollutants [2]. Moreover, 84% of $PM_{2.5}$ (PM with a diameter less than 2.5μm) in Dhaka is emitted by the brick sector [5]. $PM_{2.5}$ can be inhaled deep into the lungs and long-term exposure can cause increased morbidity and mortality [23]. By locating brick kilns in Bangladesh, the dataset and solution could help the government identify non-compliant producers and require adherence to certain emission standards or shut them down where required. In turn this would improve people's quality of life both now and for the future by way of improved air quality.

What are the risks if the data/models are compromised? The original proprietary dataset contained high-resolution satellite images, which may have associated privacy risks if compromised. As the solution aims to support the government in identifying brick kilns, they could have access to the dataset and may use it for purposes other than kiln identification, e.g. surveillance on members of the public, and therefore violate certain privacy rights. Alternatively, hostile actors could gain access to the images through a cyber attack and use it for nefarious activities. Similar issues have been raised in the domain of conservation monitoring, and principles for socially responsible use of technology and data have been suggested in the area [17].

3.2 Environmental sustainability

Is the application resource intensive? What about the data collection? For this application the SustainBench images are collected by Sentinel-2 satellites which were launched using Vega[8] rockets [7]. The initial launch is resource intensive, however once in orbit the satellites are solar powered [6] taking up little resource. Moreover, the satellites provide data for many applications, including agriculture, land ecosystems monitoring, disasters mapping, and so the resource cost is spread across many use cases rather than just the brick kiln application.

The model used for the brick kiln solution is VGG-16 which requires about 0.1 petaflop/s-day for training [16]. From [16] we observe the computational resource required to train VGG-16 is comparable to ResNets [9] and Seq2Seq [21], while requiring significantly less than larger models such as AlphaZero and AlphaGoZero [18].

Taking into account dataset reusability and comparable computational requirements for other popular neural network models, we believe the solution is not very resource intensive.

Does the application have a positive or negative impact on our carbon footprint? The application has the potential to decrease our carbon footprint, as it provides a mechanism for government to monitor brick kilns across Bangladesh, and possibly other countries in South Asia. This can improve compliance with environmental regulations and prevent the release of additional carbon emissions from illegal small-scale brick kilns, which are usually highly inefficient using coal as fuel [11].

3.3 Economic sustainability

Could it increase wealth/power concentration? The solution highlights small-scale informal brick kilns, which may lead the government to close down non-compliant kilns. However, since bricks are central to construction in Bangladesh [11], production levels will need to be maintained and this may lead to more efficient kilns being commissioned in the country. Such plants would likely be larger in scale, requiring large upfront capital expenditure and therefore run by larger corporations rather than individuals. This would take income away from local residents and transfer the wealth, generated by brick production, to corporations.

Do some people benefit economically at the expense of others? The dataset and solution might not be equitable. The solution is a classification based neural network, which can rely on certain regions within an image to make a classification decision, techniques such as *Class Activation Mapping* can be used to highlight these regions [4]. Indeed, the authors of the solution [10] highlight rows of drying bricks as one of the distinct features of a brick kiln visible in satellite images. Depending on how strong this feature is for classification, producers who are able to cover the drying bricks may go undetected by the solution and therefore by authorities. This could result in certain producers being able to evade detection whilst others are scrutinised on their production and made to comply with certain regulations, adding to production costs, or to perhaps even close their kilns. In both cases the latter set of producers would be negatively impacted financially whilst the former set may benefit by capturing additional business.

4 Other sustainability factors

4.1 Other critical questions we should be asking

- (Social) How does the application impact people's livelihoods? I.e. does it threaten to displace people's source of income, or provide new opportunities?
- (Environmental) Does the application offset the carbon emissions it produces, both during development and in its ongoing usage?
- (Economic) Does the application promote or stunt economic development? I.e what is its short and long term impact on economic development?

5 Interactions between SDGs

We believe the brick kiln detection application can have a positive impact on SDG-3: *"ensure healthy lives and promote well-being for all at all ages"*, since it helps the government detect brick kilns illegally built too close to schools and hospitals. By identifying such kilns, the solution can aid in these being relocated or shut down thereby preventing children and hospital patients inhaling pollutants such as PM, which have been linked to respiratory diseases [1, 12]. This would promote well-being and contribute towards target 3.9 as measured by indicator 3.9.1 in [15].

Along a similar line of reasoning the application could also have a positive impact on SDG-4: *"ensure inclusive and equitable quality education and promote lifelong learning opportunities for all"*. Studies have shown air pollution can have an adverse impact on children's neurobehavioral and cognitive functions [25, 14, 22]. In particular, one such study [13] found results in Chile that suggest higher annual particulate matter (PM_{10}) and ozone (O_3) levels are clearly associated with a reduction in test scores. Moreover, pollution can also indirectly affect a child's education by resulting in illnesses which prevents them from being able to regularly attend school [8]. By helping to locate and prevent kilns from operating nearby to schools, the application could help children learn and develop in a safer environment.

We note that the application may have a negative impact on SDG-8: *"promote sustained, inclusive and sustainable economic growth, full and productive employment and decent work for all"*. In the short term the solution may result in some brick kilns being shut down which could lead to increased unemployment rates [20] hindering SDG-8 as measured by indicator 8.5.2 in [15]. Moreover, with an increase in unemployment amongst adults, families may require children to work in order to gather enough income to survive. This would not only negatively impact target 8.7, which includes *"ending child labour in all its forms by 2025"*, but also cascades to SDG-4 as the children who have to work to support their families would not be able to attend school and thus miss out on a quality education.

6 Speculative solutions

6.1 Changes to dataset

One of the potential negatives we identified in Section 5 was an increase in unemployment rates. To prevent a sudden increase in redundancies of brick kiln workers across the country, we believe the dataset on which the model identifies brick kilns, could be restricted to densely populated regions such as Dhaka where reduction in pollution could improve the air quality for a significant portion of the population. By restricting initial identifications to such regions it would provide the government time to understand the social and economical impact of closing kilns, and potential job alternatives for a displaced workforce to transition into, across the rest of the country.

6.2 Changes to ML model

Another risk identified in Section 3.1 was to invasion of people's privacy using the high-resolution version of the dataset. To mitigate the risk, it would be useful to adapt the model to perform the same classification using low-resolution images. Some features of a brick kiln are very distinctive [10], e.g. rows of drying brick, chimney stack, and due to their size would remain so in low-resolution satellite images. As such, it may be worth exploring whether an ML model trained on low-resolution images, and focussed on identifying such distinct features of brick kilns, could achieve a similar level of performance to the current proposed solution.

6.3 Technology governance

Moreover, the privacy risk could be further mitigated by introducing governance processes to restrict the set of people who can manually analyse the dataset, and when required to do so must provide justification for viewing the data, thereby creating an audit trail. Additionally, policies could be created for interacting with the data itself. For example, a person only being able to view a sample of the data so as not to have exhaustive access to all of the images.

6.4 The ideal scenario

In the ideal scenario the dataset would consist of satellite images obtained in a environmentally sustainable way, and/or such that the images could be used for multiple use cases which contribute towards achieving SDGs, particularly SDG-13 to combat climate change, helping to offset the emissions produced in obtaining the images.

Moreover, it would be ideal if the ML model could use minimal computational resources to be trained and still achieve state-of-the-art classification results using low-resolution images, to protect population privacy. Ideally, the same model could be generalised to detect other high polluting, non-compliant facilities, such as tanneries and small-scale mines, using the same satellite dataset. This would increase the application's contribution towards SDG-13, whilst maintaining a similar absolute environmental impact.

References

[1] D. V. Bates. The effects of air pollution on children. *Environmental health perspectives*, 1995.

[2] B. A. Begum, P. K. Hopke, and A. Markwitz. Air pollution by fine particulate matter in bangladesh. *Atmospheric Pollution Research*, 4(1):75–86, 2013.

[3] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, and L. Fei-Fei. Imagenet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 248–255, 2009.

[4] S. Desai and H. G. Ramaswamy. Ablation-cam: Visual explanations for deep convolutional network via gradient-free localization. In *2020 IEEE Winter Conference on Applications of Computer Vision (WACV)*, pages 972–980, 2020.

[5] A. Eil, J. Li, P. Baral, and E. Saikawa. Dirty stacks, high stakes: An overview of brick sector in south asia. *World Bank*, 2020.

[6] ESA. The Operational Copernicus Optical High Resolution Land Mission, 2013. https://esamultimedia.esa.int/docs/S2-Data_Sheet.pdf.

[7] ESA. About Copernicus Sentinel-2, 2017. https://sentinel.esa.int/documents/247904/4180891/Sentinel-2-infographic.pdf.

[8] F. D. Gilliland, K. Berhane, E. B. Rappaport, D. C. Thomas, E. Avol, W. J. Gauderman, S. J. London, H. G. Margolis, R. McConnell, K. T. Islam, and J. M. Peters. The effects of ambient air pollution on school absenteeism due to respiratory illnesses. *Epidemiology*, 2001.

[9] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition, 2015.

[10] J. Lee, N. R. Brooks, F. Tajwar, M. Burke, S. Ermon, D. B. Lobell, D. Biswas, and S. P. Luby. Scalable deep learning to identify brick kilns and aid regulatory capacity. *Proceedings of the National Academy of Sciences*, 118(17), 2021.

[11] S. P. Luby, D. Biswas, E. S. Gurley, and I. Hossain. Why highly polluting methods are used to manufacture bricks in Bangladesh. *Energy for Sustainable Development*, 28:68–74, 2015.

[12] I. Manisalidis, E. Stavropoulou, A. Stavropoulos, and E. Bezirtzoglou. Environmental and health impacts of air pollution: A review. *Frontiers in Public Health*, 8, 2020.

[13] S. J. Miller and M. A. Vela. The Effects of Air Pollution on Educational Outcomes: Evidence from Chile. *Inter-American Development Bank*, 2013.

[14] P. Mohai, B.-S. Kweon, S. Lee, and K. Ard. Air pollution around schools is linked to poorer student health and academic performance. *Health Aff (Millwood).*, 2011.

[15] U. Nations. SDG indicators, 2020. https://unstats.un.org/sdgs/indicators/indicators-list/.

[16] Open AI. AI and Compute, 2018. https://openai.com/blog/ai-and-compute.

[17] C. Sandbrook, D. Clark, T. Toivonen, T. Simlai, S. O'Donnell, J. Cobbe, and W. Adams. Principles for the socially responsible use of conservation monitoring technology and data. *Conservation Science and Practice*, 2021.

[18] D. Silver, T. Hubert, J. Schrittwieser, I. Antonoglou, M. Lai, A. Guez, M. Lanctot, L. Sifre, D. Kumaran, T. Graepel, T. Lillicrap, K. Simonyan, and D. Hassabis. Mastering chess and shogi by self-play with a general reinforcement learning algorithm, 2017.

[19] K. Simonyan and A. Zisserman. Very deep convolutional networks for large-scale image recognition, 2015.

[20] S. Sumon. Bangladeshi purge on brick kilns could create a serious job crisis, 2019. https://www.arabnews.com/node/1591501/world.

[21] I. Sutskever, O. Vinyals, and Q. V. Le. Sequence to sequence learning with neural networks, 2014.

[22] S. Wang, J. Zhang, X. Zeng, Y. Zeng, S. Wang, and S. Chen. Association of traffic-related air pollution with children's neurobehavioral functions in Quanzhou, China. *Environmental health perspectives*, 2009.

[23] WHO. Ambient (outdoor) air pollution, 2021. https://www.who.int/news-room/fact-sheets/detail/ambient-(outdoor)-air-quality-and-health.

[24] C. Yeh, C. Meng, S. Wang, A. Driscoll, E. Rozi, P. Liu, J. Lee, M. Burke, D. B. Lobell, and S. Ermon. Sustainbench: Benchmarks for monitoring the sustainable development goals with machine learning, 2021.

[25] C. H. Zhang, L. Sears, J. V. Myers, G. N. Brock, C. G. Sears, and K. M. Zierold. Proximity to coal-fired power plants and neurobehavioral symptoms in children. *J Expo Sci Environ Epidemiol.*, 2021.

Notes

1. https://sdgs.un.org/goals
2. https://sentinel.esa.int/web/sentinel/missions/sentinel-2
3. https://developers.google.com/earth-engine/datasets/catalog/COPERNICUS_S2_SR#bands
4. https://www.tensorflow.org/api_docs/python/tf/keras/applications/vgg16/VGG16
5. https://sustainlab-group.github.io/sustainbench/
6. https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/HVGW8L
7. https://data.worldbank.org/indicator/EN.POP.SLUM.UR.ZS?locations=BD
8. https://en.wikipedia.org/wiki/Vega_(rocket)