Mapping for Public Health: Initial Plan for Using Satellite Imagery for Micronutrient Deficiency Prediction

Elizabeth Bondi
ebondi@g.harvard.edu
Harvard University

Andrew Perrault
aperrault@g.harvard.edu
Harvard University

Fei Fang
feifang@cmu.edu
Carnegie Mellon University

Benjamin L. Rice
b.rice@princeton.edu
Princeton University

Christopher D. Golden
golden@hsph.harvard.edu
Harvard University

Milind Tambe
milind_tambe@harvard.edu
Harvard University

ABSTRACT

The lack of micronutrients is a major threat to the health and development of populations, and it is challenging to detect such deficiency at large scale with low cost. In this work, we plan to use data from a study on micronutrient deficiency in Madagascar, which include blood draw results and corresponding questionnaires, along with satellite imagery, to determine whether there are certain cues visible in satellite imagery that could more easily and quickly suggest areas where people may be susceptible to micronutrient deficiency. We propose an approach that will (i) determine important predictors of micronutrient deficiency from blood draws and corresponding questionnaire data, such as type of food consumed, (ii) automatically detect related areas in satellite imagery, such as forest regions where sources of important food may be found, and (iii) use these to predict regions of micronutrient deficiency. As the prediction of micronutrient deficiency in satellite imagery will be done using meaningful predictors, and we anticipate using an inherently interpretable model for prediction based on these objects, such as logistic regression, we aim to create a model that will be intuitive to those in the public health community.[1]

CCS CONCEPTS

• Computing methodologies → Computer vision problems; **Machine learning**; **• Applied computing → Life and medical sciences**.

KEYWORDS

interpretability, remote sensing, micronutrient deficiency

ACM Reference Format:
Elizabeth Bondi, Andrew Perrault, Fei Fang, Benjamin L. Rice, Christopher D. Golden, and Milind Tambe. 2020. Mapping for Public Health: Initial Plan for Using Satellite Imagery for Micronutrient Deficiency Prediction . In KDD 2020 Workshop on Humanitarian Mapping.

1 INTRODUCTION

According to the World Health Organization, the lack of micronutrients "...represents a major threat to the health and development of populations the world over, particularly children and pregnant women in low-income countries." [1]. However, micronutrient deficiency is sometimes called "hidden hunger", as the effects often become visible only when the deficiency is already severe [29]. In other words, the majority of the types of malnutrition do not have physical manifestations. Wasting is only one type of malnutrition that presents itself in physical manifestations. Therefore, regions with micronutrient deficiency are largely unknown to public health organizations until direct measurements are made, such as blood draws. However, these blood draws and questionnaires are costly and time-consuming, and furthermore, quantifying micronutrient levels in a sample requires specialized laboratory equipment that is not widely available. We envision a new approach to detect micronutrient deficiency at large scale with low cost.

We hypothesize that the availability of foods with sufficient micronutrients in a community depends in part on environmental factors, such as the state of forests or agriculture nearby. Madagascar is a unique area for the purposes of studying this hypothesis, as it has many different ecological regions within it [15]. A recent cross-sectional study examined micronutrient deficiency in several of these regions, in addition to anemia, malaria, and detailed questionnaire data [9], which we describe further in Section 3. We propose an approach that will use satellite imagery to identify important environmental factors that may help predict regions of micronutrient deficiency.

With such an approach, interventions such as supplementation, fortification, biofortification, and infectious disease management, could be targeted based on more up-to-date information [2]. In brief, supplementation, e.g. vitamins, is costly and therefore more targeted at an individual level, while fortification, e.g., adding iodine to salt, is inexpensive and could be applied to regions, though it requires population-wide deficiencies and an inability to reach toxic limits with too much of the micronutrient. Biofortification is similar to fortification in that it seeks to add nutrition to foods, but in this case adds nutrients during growth, e.g., adding iron to crops during growth. Finally, infectious disease management is useful to curb micronutrient deficiency, as infectious diseases often deplete micronutrients. Specific interventions would depend on the disease(s). Throughout the rest of this proposal, we will provide a brief literature review for different aspects of this hypothesis, including literature in artificial intelligence and in public health, highlight the proposed methodology, and conclude by summarizing our vision.

2 RELATED WORK

First, we reviewed several potential uses of AI in the context of nutrition, including AI and malnutrition and AI and micronutrient deficiency. There is work on predicting anemia and/or hemoglobin levels [6, 8, 24, 32], as well as predicting mutations that lead to

[1] KDD 2020 Workshop on Humanitarian Mapping.

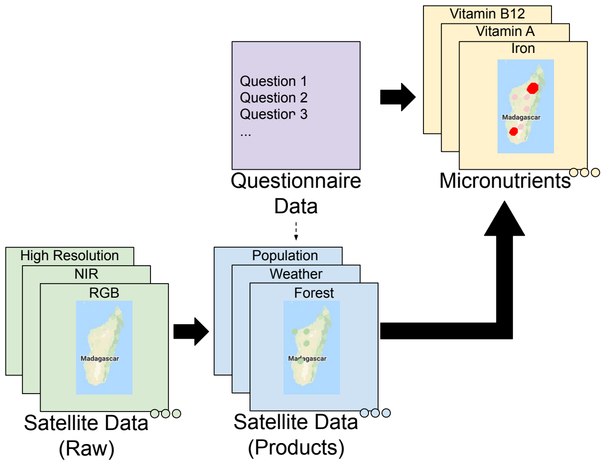

Figure 1: (i) Starting in the first row, we will use questionnaire data to determine what factors can predict micronutrient deficiency. (ii) We will then use these predictors to determine the satellite data products that are necessary (dotted arrow). (iii) We will use these products to predict micronutrient deficiency.

disease related to micronutrients [12]. However, these works do not consider the prediction of micronutrient deficiency at a regional level using satellite imagery. For malnutrition, there is a great deal of work on detecting malnutrition in children via smartphone apps and computer vision, as well as questionnaires (e.g., [11, 25, 26]). Similarly, these do not use satellite data, nor do they focus at a regional level.

There is another body of work on food security in the context of agriculture and food rescues, as well as foodborne illnesses [27]. There is also a great deal of work on AI and agriculture, for example in aiding smallholder farmers [16], finding crop disease [20], and developing systems to improve agricultural output [4, 7]. While these works may apply at a wider level than those mentioned previously, they still do not make use of satellite imagery, nor do they specifically pertain to micronutrient deficiency.

One main focus of AI and agriculture work is on remote sensing with satellite images, such as crop yield prediction [31], estimates of crop area and cropping intensities, agricultural productivity assessments, and water planning and irrigation management [19]. Other areas of work include land cover mapping [13], which has implications for conservation and deforestation monitoring (e.g., [21]), poverty prediction [5], and infrastructure assessment [18] to name a few. The most similar works in terms of topic are probably those concerning agriculture, but they do not consider the human nutrition component to the best of our knowledge.

Finally, there has been a great deal of work towards understanding the relationship between the environment and public health. For example, one focuses on disease prevalence [22], while another focuses on the relationship between conservation practices and human well-being [3]. Our work will provide additional quantitative results to support and build on many of these ideas.

3 PROPOSED METHODOLOGY

We propose to (i) use the questionnaire data, which specifically consist of responses of 6292 individuals from 1125 households within 24 communities in four of the ecological regions of Madagascar, to determine what predicts micronutrient deficiency in the 2881 individuals who also have micronutrient data, then (ii) use these predictors to determine relevant satellite data products to include and/or derive, such as forest cover [28], climate [14], agriculture (e.g., determined via computer vision techniques), and/or roads [17], and finally (iii) use these products to predict micronutrient deficiency. We plan to use logistic regression for (i) and (iii) to foster interpretability, and deep learning to create/interpret relevant satellite data products in (ii). This is illustrated in Fig. 1.

To compute intermediate satellite image products in (ii), we will likely use segmentation techniques, e.g., U-Net [23], to determine exact regions of forest, road, agriculture, etc. within a satellite image. We plan to use information from all respondents (about 6000) to provide labels for these intermediate products, and to augment our training data through the use of additional datasets such as the Demographic and Health Survey/Malaria Indicator Survey in Madagascar [30]. The most recent survey includes anemia results and information collected regarding food and lifestyle, such as roof material, which may be used as additional labels. Training and testing data would therefore be done on regions within Madagascar, with hopes of creating a more widely generalizable process for use in other places. To predict micronutrient deficiency in (iii), we will train using 2000 inputs derived from the intermediate products in (ii) and their corresponding micronutrient result, and test on the remaining 881 micronutrient results.

We hypothesize that several environmental factors may be particularly important, including the presence and health of forests and agriculture. We also believe that micronutrient deficiency and its potential environmental influences could be confounded by poverty, so *plan to include poverty measures (e.g., from [10]) to control for these effects.* Furthermore, we hypothesize that easy access to larger food markets reflected by proximity to roads would lead to less micronutrient deficiency. If this proved to be true in our analysis, for example, it may indicate that there is increased participation in markets or healthcare along roads, and interventions could focus on widening access to these necessities nearby, for example via supplementation. As drawing this conclusion will rely heavily on model interpretability, we draw inspiration from [5] in our goal to create an interpretable pipeline for policy makers, but are not able to use the same methodology, as we must examine larger regions to determine environmental factors linked to micronutrient deficiency. We may be able to similarly consider more fine-grained areas such as households depending on the factors that prove to be important predictors of micronutrient deficiency.

4 OVERALL VISION

To summarize, our vision is to provide predictions for micronutrient deficiency, a global public health concern, based on cues in satellite imagery, i.e., environmental factors. We propose an interpretable model in order to do this, both to determine which factors will be important to study, and how they effect micronutrient deficiency prediction. We hope that such a system could be used around the world to aid public health understanding and interventions, more frequently and more easily than traditional questionnaire-based methods.

5 ACKNOWLEDGMENTS

This work was partially supported by NSF CCF-1522054 and IIS-1850477. We are also grateful for the support from the United States Agency for International Development (grant AID-FFP-A-14-00008) implemented by Catholic Relief Services (CRS) in consortium with four local implementing partners in Madagascar. The views and opinions expressed in this paper are those of the authors and not necessarily the views and opinions of the United States Agency for International Development.

We would like to thank Hervet J. Randriamady for his role in developing and leading field data collection efforts. We would like to thank the many people across Madagascar who welcomed us to conduct research within their communities. We would like to thank some key individuals of Ministère de de la Santé Publique (MSANP) who deeply supported this project: Dr. Herlyne Ramihantaniarivo (former Directeur Général de la MSANP), Pr. Jean de Dieu Marie Rakotomanga (Directeur Général de l'INSPC), Dr. Marie Rolland Ratsimbazafy (Responsable de l'INSPC), Celestin Tara (Assistant Technique de l'INSPC), Dr. Harinelina Randriamasiarijaona (Chef Service de la Nutrition), Dr. Orieux Tovotsimihefo (Médecin Inspecteur Morombe), Dr. Raphael Hotahiene (Directeur de Lutte contre les Maladies Transmissibles), Dr. Maurice Randriarsion (Médecin Inspecteur Mananjary), Dr. Georges Rakotondranaivo (Médecin Inspecteur Ambatofinandrahana), Dr. Noromalala Sylvie Tidahy (Directeur de la Santé Familiale), Hery Rakotoniary (Médecin Inspecteur Fandriana). We express our thanks to the mayors and clinic chiefs of the rural communes of Mananjary (Tsaravary, Mahatsara Atsimo, Ankatafana and Antsenavolo), Toliara II (Tsianisiha, Maromiandra, and Belalanda), Morombe (Morombe and Basibasy), and Amoron'i Mania (Alakamisy Ambohimahazo, Ambondromisotra, and Soavina). We would like to thank Bureau de Développement ECAR de Mananjary (BDEM), Caritas Morombe (CMB), and Conseil Diocésain de Développement de Toliara (CDD), and the ADRA ASOTRY team in Amoron'i Mania that have facilitated our work with the local communities.

REFERENCES

[1] 2020. Micronutrients. https://www.who.int/nutrition/topics/micronutrients/en/
[2] 2020. Nutrition interventions. https://www.who.int/elena/intervention/en//
[3] Matthew Agarwala, Giles Atkinson, Benjamin Palmer Fry, Katherine Homewood, Susana Mourato, J Marcus Rowcliffe, Graham Wallace, and EJ Milner-Gulland. 2014. Assessing the relationship between human well-being and ecosystem services: a review of frameworks. *Conservation and Society* 12, 4 (2014), 437–449.
[4] Wadhwani AI. 2019. We are a Google AI Impact grantee. https://www.wadhwaniai.org/2019/05/07/we-are-a-google-ai-impact-grantee/
[5] Kumar Ayush, Burak Uzkent, Marshall Burke, David Lobell, and Stefano Ermon. 2020. Generating Interpretable Poverty Maps using Object Detection in Satellite Images. *IJCAI* (2020).
[6] Iman Azarkhish, Mohammad Reza Raoufy, and Shahriar Gharibzadeh. 2012. Artificial intelligence models for predicting iron deficiency anemia and iron serum level based on accessible laboratory data. *Journal of medical systems* 36, 3 (2012), 2057–2061.
[7] Elizabeth Bondi, Carl Salvaggio, Matthew Montanaro, and Aaron D Gerace. 2016. Calibration of UAS imagery inside and outside of shadows for improved vegetation index computation. In *Autonomous Air and Ground Sensing Systems for Agricultural Optimization and Phenotyping*, Vol. 9866. International Society for Optics and Photonics, 98660J.
[8] Antonin Dauvin, Carolina Donado, Patrik Bachtiger, Ke-Chun Huang, Christopher Martin Sauer, Daniele Ramazzotti, Matteo Bonvini, Leo Anthony Celi, and Molly J Douglas. 2019. Machine learning can accurately predict pre-admission baseline hemoglobin and creatinine in intensive care patients. *NPJ digital medicine* 2, 1 (2019), 1–10.
[9] Christopher D Golden, Benjamin L Rice, Hervet J Randriamady, Arisoa Miadana Vonona, Jean Frederick Randrianasolo, Ambinintsoa Nirina Tafangy, Mamy Yves Andrianantenaina, Nicholas J Arisco, Gauthier N Emile, Faustin Lainandrasana, Robuste Fenoarison Faraniaina Mahonjolaza, Hermann Paratoaly Raelson, Vololoniaina Ravo Rakotoarilalao, Anjaharinony Andry Ny Aina Rakotomalala, Alex Dominique Rasamison, Rebaliha Mahery, M Luciano Tantely, Romain Girod, Akshaya Annapragada, Amy Wesolowski, Amy Winter, Daniel L Hartl, James Hazen, and C Jessica E Metcalf. In Submission.. A cross-sectional study of the social, demographic, and ecological variation of nutrition and disease across Madagascar. *Frontiers in Public Health* (In Submission.).
[10] Neal Jean, Marshall Burke, Michael Xie, W Matthew Davis, David B Lobell, and Stefano Ermon. 2016. Combining satellite imagery and machine learning to predict poverty. *Science* 353, 6301 (2016), 790–794.
[11] Sangita Khare, S Kavyashree, Deepa Gupta, and Amalendu Jyotishi. 2017. Investigation of nutritional status of children based on machine learning techniques using Indian demographic and health survey data. *Procedia computer science* 115 (2017), 338–349.
[12] Mohamad Koohi-Moghadam, Haibo Wang, Yuchuan Wang, Xinming Yang, Hongyan Li, Junwen Wang, and Hongzhe Sun. 2019. Predicting disease-associated mutation of metal-binding sites in proteins using a deep learning approach. *Nature Machine Intelligence* 1, 12 (2019), 561–567.
[13] Nataliia Kussul, Mykola Lavreniuk, Sergii Skakun, and Andrii Shelestov. 2017. Deep learning classification of land cover and crop types using remote sensing data. *IEEE Geoscience and Remote Sensing Letters* 14, 5 (2017), 778–782.
[14] Amy McNally, Kristi Arsenault, Sujay Kumar, Shraddhanand Shukla, Pete Peterson, Shugong Wang, Chris Funk, Christa D Peters-Lidard, and James P Verdin. 2017. A land data assimilation system for sub-Saharan Africa food and water security applications. *Scientific data* 4, 1 (2017), 1–19.
[15] Justin Moat, Paul Philip Smith, et al. 2007. *Atlas of the vegetation of Madagascar*. Royal Botanic Gardens, Kew.
[16] Neil Newman, Lauren Falcao Bergquist, Nicole Immorlica, Kevin Leyton-Brown, Brendan Lucier, Craig McIntosh, John Quinn, and Richard Ssekibuule. 2018. Designing and evolving an electronic agricultural marketplace in Uganda. In *Proceedings of the 1st ACM SIGCAS Conference on Computing and Sustainable Societies*. 1–11.
[17] OpenStreetMap contributors. 2017. Planet dump retrieved from https://planet.osm.org . https://www.openstreetmap.org.
[18] Barak Oshri, Annie Hu, Peter Adelson, Xiao Chen, Pascaline Dupas, Jeremy Weinstein, Marshall Burke, David Lobell, and Stefano Ermon. 2018. Infrastructure quality assessment in africa using satellite imagery and deep learning. In *Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*. 616–625.
[19] CD Peters-Lidard, Danielle Wood, Margaret M Hurwitz, and Bradley Doorn. 2018. Advancing the Use of Earth Observations to Benefit Global Food Security and Agriculture. In *98th American Meteorological Society Annual Meeting*. AMS.
[20] John Alexander Quinn, Kevin Leyton-Brown, and Ernest Mwebaze. 2011. Modeling and monitoring crop disease in developing countries. In *Twenty-Fifth AAAI Conference on Artificial Intelligence*.
[21] Caleb Robinson, Le Hou, Kolya Malkin, Rachel Soobitsky, Jacob Czawlytko, Bistra Dilkina, and Nebojsa Jojic. 2019. Large scale high-resolution land cover mapping with multi-resolution data. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 12726–12735.
[22] Benjamin Roche, Andrew P Dobson, Jean-François Guégan, and Pejman Rohani. 2012. Linking community and disease ecology: the impact of biodiversity on pathogen transmission. *Philosophical Transactions of the Royal Society B: Biological Sciences* 367, 1604 (2012), 2807–2813.
[23] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-net: Convolutional networks for biomedical image segmentation. In *International Conference on Medical image computing and computer-assisted intervention*. Springer, 234–241.
[24] Bryan Saldivar-Espinoza, Dennis Núñez-Fernández, Franklin Porras-Barrientos, Alicia Alva-Mantari, Lisa Suzanne Leslie, and Mirko Zimic. 2019. Portable system for the prediction of anemia based on the ocular conjunctiva using Artificial Intelligence. *arXiv preprint arXiv:1910.12399* (2019).
[25] Siddhi Shah, Shefali Naik, and Vinay Vachharajani. 2016. Child Growth Mentor—A Proposed Model for Effective Use of Mobile Application for Better Growth of Child. In *Proceedings of International Conference on ICT for Sustainable Development*. Springer, 153–159.
[26] Mehrab Shahriar, Mirza Shaheen Iqubal, Samrat Mitra, and Amit Kumar Das. 2019. A Deep Learning Approach to Predict Malnutrition Status of 0-59 Month's Older Children in Bangladesh. In *2019 IEEE International Conference on Industry 4.0, Artificial Intelligence, and Communications Technology (IAICT)*. IEEE, 145–149.
[27] Zheyuan Ryan Shi, Claire Wang, and Fei Fang. 2020. Artificial Intelligence for Social Good: A Survey. *arXiv preprint arXiv:2001.01818* (2020).

[28] Masanobu Shimada, Takuya Itoh, Takeshi Motooka, Manabu Watanabe, Tomohiro Shiraishi, Rajesh Thapa, and Richard Lucas. 2014. New global forest/non-forest maps from ALOS PALSAR data (2007–2010). *Remote Sensing of environment* 155 (2014), 13–31.

[29] Klaus von Grebmer, Amy Saltzman, Ekin Birol, Doris Wiesman, Nilam Prasai, Sandra Yin, Yisehac Yohannes, Purnima Menon, Jennifer Thompson, Andrea Sonntag, et al. 2014. 2014 Global Hunger Index: The challenge of hidden hunger. *IFPRI books* (2014).

[30] ICF. The DHS Program website. [n.d.]. Madagascar: MIS, 2016. http://www.dhsprogram.com

[31] Jiaxuan You, Xiaocheng Li, Melvin Low, David Lobell, and Stefano Ermon. 2017. Deep gaussian process for crop yield prediction based on remote sensing data. In *Thirty-First AAAI Conference on Artificial Intelligence*.

[32] Jiali Zhang and Weiming Tang. 2020. Building a prediction model for iron deficiency anemia among infants in Shanghai, China. *Food Science & Nutrition* 8, 1 (2020), 265–272.