

February 05, 2019 | By Kimberly Mann Bruch

SDSC Researchers Team with UNICEF for Liberian Schools Project

Deep Learning Satellite Images Provide Insight to Rural Education Areas

More than 38,000 children were forced to participate in Liberia's second civil war, which lasted 14 years and claimed some 250,000 lives before ending in 2003. The West African country slowly continues to rebuild basic infrastructures, including education.

Toward that end, the United Nations Children's Fund (UNICEF) has been working with researchers at the San Diego Supercomputer Center (SDSC) as well as other parts of UC San Diego to determine the location of existing Liberian schools so they can provide them with resources and work with policy makers to plan for future schools in the country.



Validating the model's performance is accomplished by comparing its prediction of school locations with UNICEF's data. The orange squares represent tiles identified by the model as likely school locations, and the pink circles are actual schools from UNICEF's data. Courtesy of Mai Nguyen, SDSC

The SDSC team recently presented [their findings](#) at the 11th IEEE/ACM International Conference on Utility and Cloud Computing in December 2018. The presentation focused on their use of deep learning, which is a subfield of artificial intelligence and makes use of computational models that can automatically learn patterns from data instead of being explicitly programmed. The 'deep' in deep learning refers to the many layers of interconnected processing units in the model that allows it to learn representations of the data at multiple and increasingly complex levels of abstraction.

How exactly did deep learning in this particular project work, and how did it determine the locations of the schools?

“Using a particular type of deep learning model called a convolutional neural network (CNN), we extracted salient features from the satellite image tiles,” explained Mai Nguyen, a senior SDSC data scientist. “We then applied cluster analysis to organize the image tiles into clusters where tiles with similar features were grouped together. Clusters with the majority of image tiles containing schools were then identified using ground truth data (on-site information provided by UNICEF) if available or by visual inspection of the cluster contents. These school clusters were then used to identify likely locations of schools in a region.”

Deep learning, combined with cluster analysis, is necessary for a scenario like this where actual locations of schools could be inaccurate or outdated. This is the case for many applications, where ground truth data is often insufficient, unreliable, or missing altogether.

The approach used by the SDSC researchers provided a way to focus on the regions of interest, where schools were likely to be located, while discarding a large majority of “noise”, or inaccuracies, in the data. In the test region, Nguyen said this approach was able to locate 80 percent of the schools by considering only less than 2 percent of the total image tiles in the region – a significant reduction in the search space.

“This work is a follow-on to a series of studies we have done using high-resolution satellite data and deep learning with the SDSC Data Science Hub together with the Big Pixel Initiative at UC San Diego’s Qualcomm Institute,” said Ilkay Altintas, SDSC’s Chief Data Science Officer and lead for the Data Science Hub.

Altintas is also a co-PI for the NSF-funded Cognitive Hardware and Software Ecosystem Community Infrastructure (CHASE-CI), a network of fast Graphics Processing Units (GPU) appliances for machine learning and storage managed through Kubernetes on the high-speed Pacific Research Platform (PRP).

“Having the CHASE-CI platform at the Qualcomm Institute at our fingertips with (GPUs), scalable machine learning tools, and high-bandwidth data transfer is ideal to accelerate this kind of research,” added Altintas.

The next step was to apply this approach to larger and geographically different regions to test its robustness and transferability to diverse environments. The SDSC researchers also plan to investigate the use of imagery data that includes additional spectral bands, as well as higher temporal resolution to improve their method’s precision and robustness in locating schools.

Finding and providing assistance to Liberian schools has been a challenge for UNICEF and other organizations, such as a small grassroots group, which is based in San Diego and led by Liberian-American Malia Harris. Upon learning about the SDSC work related to using satellite imagery to locate schools, Harris was ecstatic.

“The areas in this study are incredibly rural and undeveloped because during the war, almost everyone left for safer ground,” said Harris. “The only people left in those counties are people who could not leave for one reason or another. But they also have a right to education.”

Harris, who lived in Liberia for most of her life, led efforts to build a primary and secondary school in Monrovia in 1988. In 1990, the school building was significantly damaged by rebels and three more times during the many years of the war. During the years the school was being rebuilt, the Harris home would serve as a school for the nearby children.

In 2000, Harris attended a World Organization for Early Childhood Education conference in London and from there left for the U.S., asking for asylum. She worked diligently from San Diego to help her Liberian community members rebuild the school, which was completed in 2007. Harris continues to help with the school and is now establishing a non-profit, called the Liberian Perseverance Association, as operational costs continue to be challenging. However, she is grateful that at least there is a functioning school in Monrovia.

“Once the information from these SDSC researchers can be correlated with on-ground data, groups such as the Liberian Perseverance Association can more effectively help the school teachers and send supplies to these more remote areas,” she said.

“We are happy to hear about this project to better help the development and sustainability of schools in Liberia,” said Melvin Gartei, a science teacher at Saint Paul Presbyterian School in Monrovia. “While our biggest challenges are fundamental issues such as students who are hungry, we would greatly welcome learning tools from UNICEF and other groups to help these children someday rise above their extreme poverty. Hopefully, projects like this one at SDSC that are able to map out the working school sites can then work with organizations to help us build things like science laboratories, as they are rarely available in our school systems.”

“Hearing that our work will be helpful from the education community in Liberia is exactly what we want to hear,” said Jessica Block, coordinator of the big Pixel Initiative. “Our goal is to advance and apply these techniques to discover insight that the world needs. We want these tools to be available to everyone.”

This research was partially supported by SDSC's Data Science Hub and NSF-1331615 under CI, Information Technology Research, and SEES Hazards programs and NSF-1730158 for CHASE-CI. UNICEF supported the project and provided school data. *Digitalglobe* sponsored the Big Pixel Initiative, which provides access to the [Global Basemap](#) and the opportunity to use the [Geospatial Big Data Platform](#) (GBDX) during the Sustainability Challenge of 2018. Preliminary methods developed in this work were supported by [PricewaterhouseCoopers \(PwC\)](#).

MEDIA CONTACT

Jan Zverina, 858-534-5111, jzverina@sdsc.edu

UC San Diego's [Studio Ten 300](#) offers radio and television connections for media interviews with our faculty, which can be coordinated via studio@ucsd.edu. To connect with a UC San Diego faculty expert on relevant issues and trending news stories, visit <https://ucsdnews.ucsd.edu/media-resources/faculty-experts>.