MACHINE LEARNING FOR CLASSIFYING ROAD CONDITIONS IN TANZANIA

CASE STUDY

Start Date: January 2017

Completion date: November 2019

FRONTIER TECHNOLOGY LIVESTREAMING

UKaid
from the British people

Can a combination of machine learning and high resolution imagery produce an accurate, automated reflection of Tanzanian road conditions?

THE CONTEXT

The problem:
Road conditions across the developing world offer a very real barrier to development. Roads are often of extremely poor quality, particularly for rural communities, which can hinder economic activity and physically isolate whole communities. Assessing road conditions in order to help decision-makers best allocate resources is therefore crucial.

The challenges of poor road conditions are reflected in Tanzania, where the majority of roads are unpaved. Although having an existing protocol to survey roads, these are costly, time consuming, and struggle to keep up with dynamic change. The island of Zanzibar in particular suffers from demographic pressures on an already poor infrastructure, as well as tropical rainfall, which often causes damage to the road network.

The idea:
With decreasing costs of satellite imagery and Unmanned Aerial Vehicles (UAVs, or drones), and recent advances in Artificial Intelligence (AI) methods, there is strong potential to develop an automated system to catalogue road networks, diagnose their condition, and automatically update entries, which could potentially fit within existing government systems in Zanzibar.

The team:

The Pioneer:
Tim Bushell

The Tech Partners
N/Lab (University of Nottingham)

50km
per day by traditional survey methods

100,00km
of rural roads in Tanzania

10 YEARS
time needed to survey country

THE JOURNEY

The pilot started in January 2017 and ran sets of experiments - called Sprints - which tested key assumptions. For each chapter there is either a pivot point or a significant event which influenced the programme. Here's a storyboard describing the main steps in this pilot's journey:

Conduct a field survey

Preliminary approval of the project was achieved thanks to a pilot field survey, across 50km of rural roads on the island of Unguja (Zanzibar). This was done in collaboration with the Zanzibari Department of Roads (DoR). There was a specified desire from the DoR to integrate data collection with existing practices as much as possible. It was quickly found that the Bump Integrator, an alternative method for tracking road quality in Tanzania to traditional visual inspection, was classifying almost all rural roads into a single 'very poor' category. A new, more actionable, system was required

Pilot field survey conducted ahead of time to avoid bad weather conditions.

Engage with stakeholders

Pilot inception workshop in Zanzibar

Stakeholder engagement was established predominantly through a pilot inception workshop in Zanzibar, with the Zanzibar Road Fund Board, representatives from DROMAS (District Road Management System for Tanzania), and the Dar es Salaam Institute of Technology. The decision to work with Zanzibari authorities was made because of their small size compared with national government authorities, as well as the existence of UAV imagery data for the island.

Open UAV data was provided to the project, given satellite data is not currently commercially available at the required resolution to identify features such as road markings, potholes or rocks. Drone imagery was seen as a promising solution that was less expensive, potentially scalable, and more responsive to technological improvements in the near future.

Wider engagement with actors such as the Tanzania Road Fund Board and the Tanzania National Roads Agency (TANROADS), showed that there was an appetite for knowledge sharing and the potential to scale to contry-wide solutions. This engagement also came from a desire to integrate further along the process with broader road inventory and condition surveys being made with DFID and the World Bank.

Drone image collection delayed

1200km of roads surveyed

Multi-sensor approach adopted

Establish the survey protocol

The development of a survey protocol led to an understanding that a multi-sensor approach was needed, in recognition of the difficulties of determining road quality. Five data streams were used to establish a ground-truth baseline.

Alongside a bump integrator collecting IRI (International Roughness Index) data, the team used smartphone's to collect their own accelerometer data using SensorLab. Data was also recorded via the RoadLabPro software app, human observation and finally video data from a GoPro attached to the vehicles' front bumper.

It was discovered that DoR surveyors had not been using bump integrators on unpaved roads, so this was encouraged, in order to create a universal "gold-standard" survey protocol for the study.

Create a baseline assessment

Manual correction of the network reference dataset

To create a road network reference dataset, Open Street Maps (OSM) was seen as more geographically accurate than the Shapefile-based data provided by the Commission for Lands. Because geographical alignment was crucial for an analysis of the image data, the team chose the OSM dataset.

However, geospatial misalignment of the OSM data with ground truths - often by as much as 20-50 metres - meant that many pieces of the road network were inaccurate. This was due to the precision of road digitisation, whereby a road is only composed of a few points. This means that bends in the road, for example, are not accounted for. Missing roads, as well as more accurate alignment of the network, was added manually.

GPS point and sensor readings were then used to form a baseline by which to compare the effectiveness of the prototype. The pilot team learned that a comprehensive assessment of roads in Zanzibar had not been made for many years. Survey was therefore made over approximately 1200km of paved and unpaved roads in the creation of the baseline assessment.

Manual correction of raw image data

Prepare data for prototyping

Data preparation was surprisingly time-consuming, which resulted from the fact that multiple issues arose surrounding the quality of raw image data. Firstly, the drone imagery hadn't been correctly geo-referenced in flight, meaning that the 70,000 road tiles used in the project could not be accurately extracted. This problem was exacerbated by the fact that the OSM road-network was not sufficiently aligned with the ground-truth.

Furthermore, it also became clear that the UAV imagery in the project also came with its own challenges, in that images were "warped". This is because the UAV was unable to guarantee a completely steady flight path, and meant that manual corrections were again needed at the input stage.

Finally, it was discovered that only DoR road engineers knew which sections of the road network were unpaved. The lack of digital distinction between paved and unpaved roads in the dataset meant that the model was unable to distinguish between the fine graduations of unpaved road conditions targeted in the pilot. Department of Roads engineers had to therefore manually annotate road maps with this information.

Go to prototype

7-point classification system adopted

The pilot team then focused on the rapid prototyping of road condition classification algorithms. They quickly learned that the results were bunching all unpaved roads into the discrete, lowest-quality road category, which meant that results were not granular enough to allow for the appropriate diagnosis of road conditions. The 5-point schema of road classification brought in from developed country contexts was deemed unsuitable, and a new, data-driven 7-point classification scheme - developed using the IRI data from the pilot survey - was made. This was able to widen the definition of road quality for unpaved roads in particular.

Evaluate

Final iteration of model produces 73% accuracy

Using the new system, results were initially, surprisingly low, with an accuracy of 35-40%. This was above the level of chance, but nevertheless disappointing.

An extensive manual inspection of road tiles was undertaken. It found that information other than the road surface was being considered in the model. This included information on foliage or house type, for example, which were affecting the dataset. At the crux of this problem was that of image quality. Many of the input images were blurred, overexposed, or suffered some form of interpolation. Given that the raw drone data were obtained from an external source, options to correct this were limited.

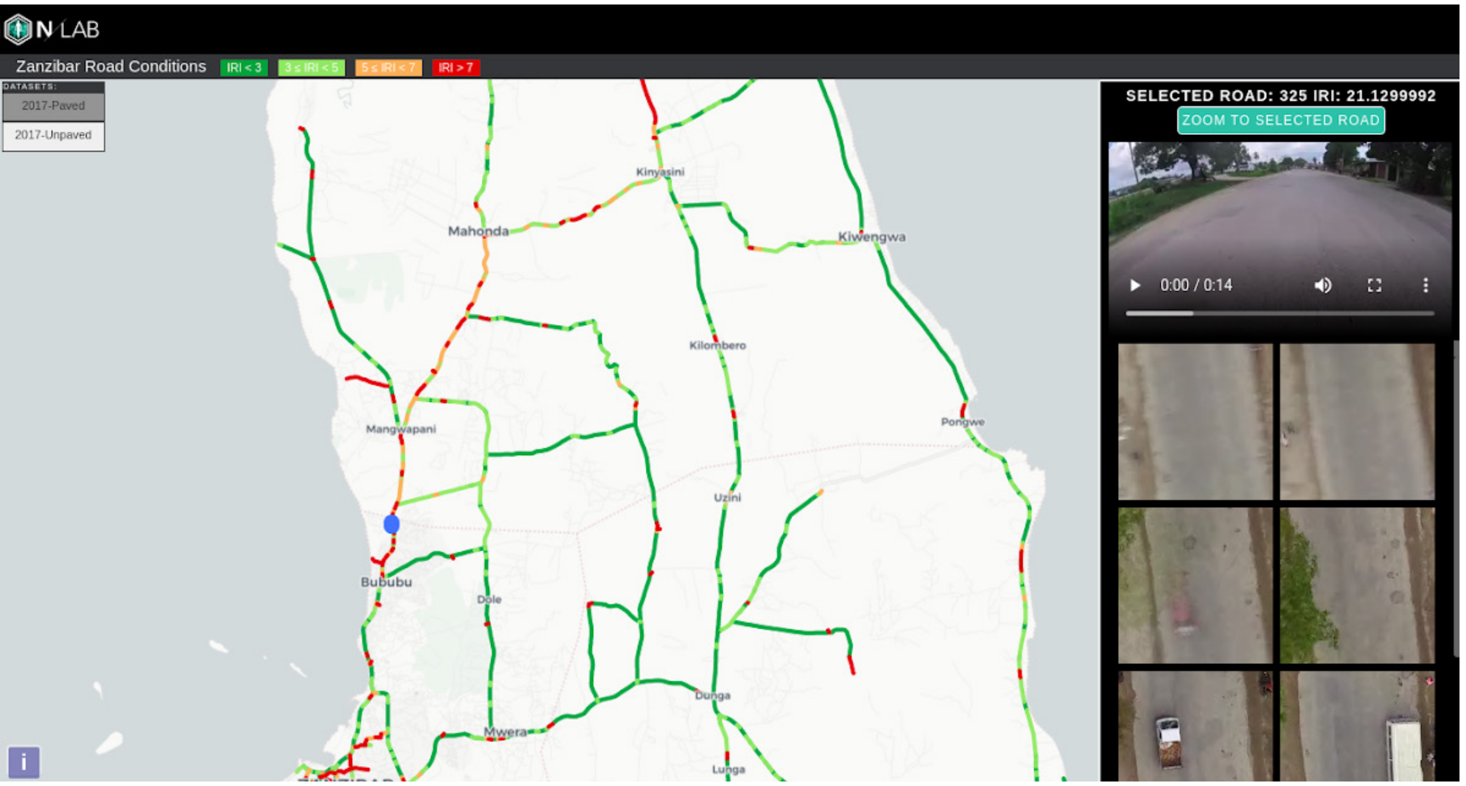
Without the ability to rectify image quality, the project conducted a process of quality control, with only 5000 of 70,000 images usable in the next round of model iterations. Results were transformed, with new models 73% accurate in differentiating between "Good" and "Poor" unpaved road conditions (with another 10% flagged for review).

WHY ARE ROADS A PRIORITY FOR DFID?

Cities across East Africa are growing at a challenging rate, as much as 10% for some of the region's largest urban centres. Dar-es-Salaam, for example, was recorded as having less than 3.5 million in 2010, yet has reached 6.5 million by 2020. What this means for transport is congestion, pollution, and roads that are put under increasing stress.

The importance of road quality to development outcomes is clear. Good road links reduce trade barriers and their associated costs, which are disproportionately high across the developing world. On the other hand, helping to connect rural and urban areas is seen as equitable, by providing economic opportunities as well as improving overall quality of life and access to services for rural communities.

DFID helps countries build the long-term development of their infrastructure. Its strengths lie in producing high quality research, informing partner action, and fostering public and private investment. Between 2015 and 2016, DFID bilateral programming for transport and urban infrastructure totalled £3.9 billion, across 57 programmes. In East Africa, DFID is investing in infrastructure, to strengthen towns and cities as hubs of growth, facilitate international trade, and achieve greater regional integration. The key rationale for this project was helping to identify whether machine learning can be included in such DFID programming, specifically for road inventorying in Tanzania.



"Normally when people apply AI it is to finely crafted data sets. But in the real world of developmental projects, where data has been either opportunely or coarsely collected, lack of data quality becomes crucial.

Dealing with this, and designing for biased, sparse, noisy and incomplete data, simply has to become a core part of the workflow in developing contexts."

- Dr James Goulding, N/LAB

THE RESULTS

All of the critical assumptions behind this idea were tested and proved ✅ or disproved. ❎
We gained insight on all the assumptions, but some had questions remaining. 🤔

VALUE	DID POTENTIAL USERS ENGAGE WITH THE TECH?	
Government engagement was an overall success. The Zanzibar Department of Roads was involved fully in the process, including helping with material needs and the formulation of a survey protocol, as well as conducting the surveys themselves. DoR's strong engagement led to development of a new software suite, allowing the road agency to both upload and visualize their survey datasets, and shifting their practices.		✅
TECH	**DID THE TECH DEPLOYMENT WORK?**	
Although the project's AI models produced a very promising level of accuracy, major issues plagued development. The most central of these was nothing to do with machine learning, but quality of data available. The challenges of regular collection of high-fidelity drone data, that is neither warped, nor over-exposed, is an existential threat towards the application of machine learning in this context. These were issues the pilot team could not affect, but acted as strong barriers to tech deployment.		🤔
GROWTH	**WHAT IS THE LIKELIHOOD FOR SCALE UP?**	
Aside from the issues concerning tech deployment, opportunities for scale will rely on both the availability and cost of high-resolution satellite or UAV data. UAV regulation is a particular inhibitor to projects at scale, because many countries are becoming increasingly heavy-handed on drone regulation. For example, If Beyond Line of Sight visual capability is banned within a country, this will be a serious barrier to the ease of taking aerial photographs over large distances. Drone usage is currently banned in Tanzania. However, if 15cm resolution satellite imagery can be obtained at reasonable cost, or provided through private-sector partnerships, many of the pilots image processing challenges would be bypassed, and the potential for fully automated road-condition systems would become highly viable at scale.		🤔
IMPACT	**WHAT LEVEL OF POSITIVE SOCIAL IMPACT OR INFLUENCE HAS BEEN ACHIEVED?**	
In Zanzibar, the completion of surveys as part of the pilot was useful, because for many years this had not been achieved. The project therefore contributed to a baseline assessment of roads on the island of Unguja. Additionally, the new survey protocol established during the project may also be adopted by local actors. However, the key thesis of the project - that cost and time expenditure for road survey in developing country contexts could be reduced - was not proven, because the technical expertise required to clean and reference image data is, for the moment, constrictive.		🤔
OPPORTUNITIES FOR SCALE		
Has it attracted any co-funding or follow on investment?	No, but funding for two follow-on projects are being explored: (1) exploration with Earth Observation data providers to co-develop AI that uses super high-resolution satellite imagery; and (2) a continuing collaboration with DoR to use AI models, driven by go-pro video capture mounted on motorbikes.	

BY THE NUMBERS

1183km of road surveyed

c.70k image tiles pre-processing

7% of images passed quality control

73% accuracy of final model

REFLECTIONS FROM THE HUB

Insights on Machine Learning

Stakeholder buy-in is crucial

What this pilot shows is that building partnerships with government agencies can allow for crucial flexibility and co-design. The pilot survey was able to be accelerated to avoid seasonal floods thanks to the buy in from DoR Engineers. And transfer of knowledge between stakeholders led to expansion beyond the initial FTL pilot, expanding from drone to Go-Pro video imagery - and leading to development of new software and processes that have changed the way DoR manage their survey data.

Classifications do not "fit-all"

The 5-point road condition classification system adapted from more developed countries was unsuitable in creating a baseline assessment of formal and informal roads in Tanzania. The team had to create a new classification system to better differentiate between informal roads, which were bunching towards the lowest figure for road quality in the 5-point system.

Your output is only as good as your input

The pilot project was framed as a testing of machine learning capabilities, but it could well have been an exercise in producing a dataset. Before testing its core hypothesis, the pilot had to first prove the assumption that there was a high-quality suite of aerial photography, with accurate geolocation. Targeting critical assumptions surrounding data inputs should form part of earlier Sprints if possible.

Be open in your approach to data

This pilot would not have been possible without open data - both in terms of drone imagery and the OSM road network used to isolate roads in that imagery. Openness was a core priority in outputs too, providing a route to engage with actors working on roads across Tanzania with an eye to scaling up. In particular, an extra sprint was implemented due to a need to improve on existing OSM networks, and make its data AI ready. This has led to a new, vastly improved road network dataset that can be fed directly back to the community.

"Overall, the project didn't quite get the level of breakthrough we hoped for, but it did prove the technique so I think it's done fantastically well... The team did amazingly well, a tremendous level of perseverance and determination throughout"

- Tim Bushell, DFID Pioneer

iMC worldwide

BRINK

ideas to impact.

UKaid
from the British people