

Full Length Research Paper

High-accuracy detection of malaria mosquito habitats using drone-based multispectral imagery and Artificial Intelligence (AI) algorithms in an agro-village peri-urban pastureland intervention site (Akonyibedo) in Unyama Sub-County, Gulu District, Northern Uganda

Mona Minakshi¹, Tanvir Bhuiyan¹, Sherzod Kariev¹, Martha Kaddumukasa², Denis Loum³, Nathanael B. Stanley⁴, Sriram Chellappan¹, Peace Habomugisha⁵, David W. Oguttu⁶ and Benjamin G. Jacob^{3*}

¹Department of Computer Science and Engineering, University of South Florida, Tampa, FL, USA.

²Department of Parasitology and Entomology, Makerere University P. O. Box 7062, Kampala, Uganda.

³Nwoya District, Local Government, Nwoya, Uganda.

⁴College of Public Health, University of South Florida, Tampa, FL, USA.

⁵The Carter Center, Plot 15 Bombo Road, Kampala, Uganda.

⁶Vector Control Division, Plot 6, Loudrel Road, Kampala Ministry of Health, Uganda.

Received 12 February, 2020; Accepted 27 May, 2020

The purpose of this research is to design, deploy and validate artificial intelligence (AI) algorithms operating on drone videos to enable a real time methodology for optimizing predictive mapping unknown, geographic locations (henceforth, geolocation) of potential, seasonal, *Anopheles (gambiae, and funestus)* larval habitats in an agro-village, epi-entomological, intervention site (Akonyibedo village) in Gulu District, Northern Uganda. Formulae are developed for classifying the drone swath, capture point, land cover in Akonyibedo agro-village. An AI algorithm is designed for constructing a smartphone application (app) in order to enable automatic detection of potential larval habitats from drone videos. The aim of this work is to enable scaling up to larger intervention sites (e.g., district level, sub-county) and then throughout entire Uganda. We demonstrate how capture point, stratifiable, drone swath coverage in Akonyibedo village can be accomplished employing temporal series of re-centered, real time, imaged, Anopheline, larval habitat, seasonal, map projections. We also define a remote methodology for detecting unknown, georeferenceable, capture point geolocations of potential, seasonal, breeding sites employing multispectral, wavelength, signature, reflux emissivities in a drone spectral library. Our results show that high-resolution drone imagery when processed employing state of the art AI algorithms can discriminate a profile of water bodies where *Anopheles* mosquitoes are most likely to breed (overall ground truth accuracy of 100%). Live, high definition, Anopheline larval habitat signature maps can be generated in real-time drone AI app on a smartphone or Apple device while the image is being captured or larvicidal application is taking place.

Key words: *Anopheles*, larval habitat, drones, artificial intelligence (AI), malaria, Uganda.

INTRODUCTION

Real-time, geospatial, predictive mapping from remote sensing can identify mosquito larval habitats in a large area that is difficult or near impossible to survey using conventional ground-based techniques. Recent advances that identify the reflective spectral signatures of active mosquito breeding sites, and their temporal evolution, have made predictive algorithms possible to search and identify previously unidentified larval habitats from an Unmanned Aerial Vehicle (UAV), and monitor their activity in real time. These aerial surveys can provide spatiotemporal data for targeting interventions to eliminate vectors before they become adult airborne biting mosquitoes, to reduce malaria transmission. Reference, capture point, larval habitats for *Anopheles gambiae* (*An. arabiensis* and *An. Funestus*), the main malaria mosquito vectors in Sub-Saharan Africa [www.who.int] can also be separately identified with the methodology.

In this research, we describe a process of real time, UAV based identification of new malaria mosquito habitats. Known larval habitats are first remotely processed through a breakdown in their spectral patterns of their changing red, green, and blue (RGB) wavelengths from repeated captured drone video, and on-the-ground, spectrometric, temporal measurements. RGB datasets constructed in ArcGIS provides an index case for how the reflectivity of vector arthropod, larval, habitat changes through seasons (Jacob et al., 2015). By modeling all local, capture point, wavelength, color, surface reflex as a combination of fully specified RGB values in ArcGIS, color fringing artifacts may be avoided while preserving sharp edges of georeferenced, gridded, land use land cover (LULC), habitat boundaries and their eco-geographic, seasonal classified, feature attributes (Jacob et al., 2011). This real time, cartographic methodology can aid in forecasting unknown, hyper-productive, aquatic and dry, *Anopheles*, larval, breeding sites using archived, time series, UAV, frequency-oriented, sample datasets. For remote identification of mosquito, vector arthropod, larval habitat and their respective, RGB, time series, capture point LULC signatures, the first step is often to construct a discrete tessellation of the region (Jacob et al., 2019).

A UAV, capture point, time series, RGB indexed, signature framework represents a vital component for retrieval systems where a vector control officer submits an query video and a real time system retrieves a ranked list of visually similar, classified, LULC, habitat types (e.g., a hyper-productive, flooded, *An. gambiae*, roadside ditch, aquatic foci) by differentially corrected GPS coordinate which has a positional accuracy of 0.178 m (Jacob et al.,

2015). The sensitivity and specificity of the RGB video signals at identifying multiple, grid-stratifiable, LULC classified, capture points can be subsequently evaluated by real time seasonal identification of the *Anopheles* breeding sites in natural settings (e.g., peri-urban, riceland, agricultural fields during pre-rain sample frames), followed by field verification (that is, "ground truthing") of the ArcGIS predictively mapped, aquatic, habitat sites. This data may then subsequently be fed into AI algorithms employing a real time app with associated software created to find other habitats with similar characteristics from new surveys of unknown terrain.

Our proposed computer vision approach for real time, LULC mapping, seasonal, anopheline, larval habitats is based on a Faster Region-based Convolutional Neural Network (Faster R-CNN) algorithm employing seasonally retrieved, RGB capture point, video analog datasets in a UAV real time platform. Faster R-CNN algorithm (Ren et al., 2015) is a state of the art artificial intelligence (AI) technique that not only classifies entities within an image, but can also localize where the entities of interest are within an image. Informally, the term "artificial intelligence" is typically used to describe machines (or computers or algorithms) that mimic cognitive functions that humans associate with the human mind, such as "learning" and "problem solving". In our proposed technique, we first learn the core features of objects in our drone image dataset using convolutional neural networks (CNNs), and then integrate manual ground truthing of larval habitats and their geolocations within the images, with learned feature maps from prior training in order to design a final region-based neural network for classification and localization. Subsequently, we assumed that at run-time, our network could classify unseen images as larval habitats as they are detected and localized where they are in the images. The technique also provides a confidence metric to the operator which can help remotely differentiate grid-stratifiable LULC aggregation and non-aggregation geolocations of georeferenced breeding site *Anopheles* foci based on like and unlike, neighboring feature habitat attributes (levels of intermittent canopy vegetation, catchment slope coefficients etc.).

Furthermore, we assumed a text-based real time retrieval system may be implementable in the future [e.g., where a vector control officer or researcher collaborator submits a textual description of an ecologically georeferenced, (henceforth eco-georeferenced), *Anopheles*, aquatic, breeding site, seasonal, capture point). The real time portal would retrieve a ranked list of

*Corresponding author. E-mail: bjacob1@usf.edu

relevant, unknown, habitat, capture point, grid-stratified, LULC locations stratified based on GPS ground coordinates. This real time portal, we assumed, could also facilitate this retrieval by finding the video clip of the unknown larval habitats that were manually assigned similar textual description (annotations) which may subsequently be employed for enabling the RGB signature framework through an iOS or Android application (app).

To summarize, the overall goal of this project is to develop a customized smartphone app that could identify the LULC geolocation of unknown, eco-georeferenceable, *Anopheles*, larval habitat capture points in agro-village, pasturelands from processed, real time, RGB, video images employing RGB signatures obtained from a drone aircraft seasonally flown in Akonyibedo village in Gulu District, Northern Uganda. Further, we assumed that once an Anopheline capture point, LULC, larval, habitat site, aquatic foci was identified and field validated, local village and entomological teams could be mobilized to the mapped habitat location employing the app and the GPS coordinates of the mapped site. Subsequently we assumed that an environmentally friendly tactic ("Seek and Destroy") could be used to bury the habitat using soil substrate and henceforth monitored for reapplication. Larger breeding habitats (e.g., rock-pit quarries, swamps) may be remotely real time targeted and treated with an environmental friendly larvicidal agent.

Interest in larval source management (LSM) as an adjunct intervention to control and eliminate malaria transmission in Uganda has recently increased mainly because long-lasting insecticidal nets (LLINs) and indoor residual spray (IRS) are ineffective against exophagic and exophilic mosquitoes. The urgent need to redesign vector control tools for mosquito populations resistant to current interventions may also require precise, seasonal, drone targeting of georeferenced, Anopheline habitats for increasing the relevance of LSM in Uganda. Therefore, knowledge of real time, drone sensed, time series, larval habitat, capture point, indexed, RGB, signature characterization for identification of unknown, productive, seasonal, positive, water bodies in endemic communities in Uganda would help to increase the impact of targeted larval mosquito control.

Here we develop a novel technique based on state-of-the-art AI enabled, real time, forecast vulnerability, grid-based, seasonal, model employing real time, geosampled, drone sensed, eco-georeferenced, *Anopheles (gambiae, funestus arabiensis)*, LULC, stratified, capture points for optimally identifying and subsequently forecasting geolocations of unknown, unsampled, seasonal, breeding site foci in Akonyibedo pastureland agro-village in Northern Uganda via a smartphone app. We develop multiple, real time, RGB, time series, habitat signatures so as to elucidate precise Anopheline, capture point geolocations. We develop binary bounding boxes (that is, 0-no habitat present and 1-habitat present) along with a

confidence metric on classified, drone sensed, gridded LULCs at the epi-entomological intervention site and develop a remote test procedure based on a real time resampling method so as to cartographically delineate unknown unsampled geolocations of potential, seasonal, hyper-productive, aquatic and non-aquatic, breeding site regions in the intervention study site. Thereafter, we conducted intense ground truthing exercises. Simulation studies in a real time UAV platform may be usable to generate a real time, drone sensing, RGB LULC, signature, iterative, interpolative methodology employing AI algorithms for optimally remotely identifying unknown, unsampled, eco-georeferenceable, hyper-productive, seasonal, aquatic and non-aquatic, *Anopheles (gambiae, funestus and arabiensis)* larval habitats throughout Uganda.

METHODOLOGY

Study site

Uganda lies between the eastern and western sections of Africa's Great Rift Valley. The country shares borders with Sudan to the north, Kenya to the east, Lake Victoria to the southeast, Tanzania and Rwanda to the south and the Democratic Republic of Congo (DRC) to the west. Whilst the landscape is generally quite flat, most of the country is over 1,000 m (3,280 ft) in altitude. Mountainous regions include the Rwenzori Mountains that run along the border with the DRC, the Virunga Mountains on the border with Rwanda and the DRC, and Kigezi in the southwest of the country. An extinct volcano, Mount Elgon, straddles the border with Kenya. The capital city, Kampala, lies on the shores of Lake Victoria, the largest lake in Africa and second-largest freshwater inland body of water in the world. Jinja, located on the lake, is considered to be the start point of the River Nile, which traverses much of the country. The varied scenery includes tropical forest, a semi-desert area in the northeast, the arid plains of the Karamoja, the lush, heavily populated Buganda, the rolling savannah of Acholi, Bunyoro, Tororo and Ankole, tea plantations and the fertile cotton area of Teso.

Gulu District is a district in Northern Uganda. The district is named after its chief municipal, administrative and commercial center, the town of Gulu. The District is bordered by Lamwo District to the north, Pader District to the east, Oyam District to the south, Nwoya District to the southwest and Amuru District to the west. The district headquarters at Gulu are located approximately 340 kilometers (210 mi), by road, north of Uganda's capital city, Kampala. The coordinates of the district are: 02 45N, 32 00E.

Malaria background in Northern Uganda

The transmission intensity of malaria depends on: (1) Vector population or density which also depends on the presence of breeding sites and favorable temperatures, (2) Parasite-carrying individuals from whom mosquitoes pick the parasites; and, (3) Presence of a malaria susceptible population especially people with low immunity such as people migrating from areas of low malaria prevalence, pregnant women, children below 5 years and people living with HIV (www.who.gov). Once these individuals are bitten by infected mosquitoes, they develop clinical malaria after an incubation period of 1 to 2 weeks following an infective bite.

Malaria is caused by the *Plasmodium* parasite. Four human



Figure 1. (a) Malaria patients in hospital in Lira; (b) Treatment facility in Gulu district; (c) Malaria Health Clinic in Dokolo.

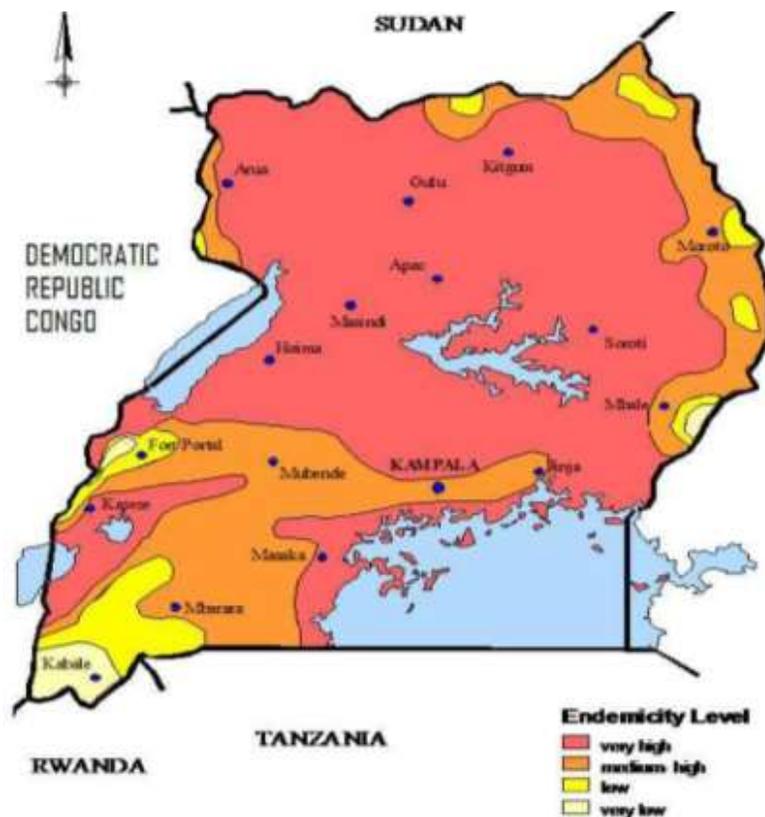


Figure 2. Malaria endemicity in Uganda.

species of *Plasmodium* (*malaria*, *vive*, *oval* and *falciparum*) occur in Northern Uganda although the predominant one is *P. falciparum* which accounts for 99% of the cases, according to Uganda Malaria Indicator Survey 2018. The malaria vectors in this region are mosquitoes of the *Anopheles* family, which breed in fresh water in temporary pools such as footprints and road cuts especially after rainfall and irrigation.

Anopheles gambiae, a highly efficient vector, along with *An. funestus* are the two main vectors; and morphometrically fewer *arabiensis* species are also present. These vectors are predominantly anthropophilic, endophilic and endophagic. The Ministry of Health (MoH), Uganda observed that using the Test, Treat and Track policy confirmed that malaria is the leading cause of morbidity and mortality; it accounts for 30-50% of outpatient visits at health facilities, 15-20% of all hospital admissions, up to 20% of all

hospital deaths and 27.2% of inpatient deaths (Figure 1a, b and c). Malaria transmission in Uganda (Figure 2) exhibits seasonality which follows the rainfall pattern. For example, in the northern region where there are two rainfall peaks, similar peak transmission periods occur that lag behind the rainy season peaks by about 4 weeks. These are associated with malaria morbidity which has been increasing in the recent decades.

The main malaria vector control method practiced in Northern Uganda is the use of Long Lasting Insecticide Treated Nets (LLINs) which has been continuously distributed to people but unfortunately has failed to curb down malaria transmission. This is due to the fact that many vectors are exophilic with the preferred biting time beginning in the early hours of the night when people are still outdoors. Secondly, majority of the local people go to bed late preferring to stay outdoors working while others socialize; hence, the



Figure 3. DJ 4 Phantom.



Figure 4. DJI 4K camera with a 8.8 mm/24 mm; f/2.8; 1" CMOS; 20 MP) for RGB capture point Anopheline habitat imaging in the Akonyibedo village epi-entomological, intervention, study site.

human vector contact in the early hours of the night has sustained malaria transmission. Also extensive agricultural insecticide and pesticides used on seeds, crops and horticultural gardens washed by rains into the running water have contaminated nearby breeding sites exposing immature stages of malaria vectors to these chemicals which have contributed to the resistance expressed in adults. Hence our assumption was that destruction of breeding sites would present a lasting solution to these challenges hampering malaria control in Northern Uganda.

Malaria transmission occurs year-round with two peaks from May to June and from November to December following distinct rainy seasons in Northern Uganda. In addition to climate and altitude, other factors that influence malaria in the country include high human concentration near vector habitats (e.g., agro-villages and boarding schools in proximity to marshlands or rice fields), population movement (especially from areas of low to high transmission), irrigation schemes (especially in the eastern and southern parts of the country), and cross-border movement of people (especially in the eastern and southeastern parts of the country).

UAV tactics

Drone surveys were carried out using a DJI Phantom 4 Pro

quadcopter (Figure 3) fitted with a DJI 4K camera (Figure 4) for seasonal, RGB, capture point, Anopheline, larval habitat, imagery collection. The camera was composed of single-band cameras [Green, Red, Red Edge and Near Infrared—(NIR)] of 1.2 MP for multispectral imagery collection.

The wayward flight plan over the agro-pastureland, epi-entomological, intervention study site in Northern Uganda was programmed with Pix4D Capture app in an iPad Mini 4 (Apple, California, US) (Figure 5). Pix4Dcapture automatically imaged the Anopheline, larval habitat, RGB, multispectral, capture image, LULC data. We processed post-flight images on the cloud applications which produced georeferenced maps and models that were tailored for ground truthing. The connection between the controller and DJI Phantom 4 Pro and 3DR Solo was set up using DJI GO 4 app (Figure 6).

For RGB capture point of Anopheline larval habitat and real time, LULC, 2-D and 3-D mapping in Akonyibedo village, the DJI Phantom 4 Pro drone was flown to an altitude of approximately 100 m, initially which gave a ground sampling distance (GSD) or spatial resolution of 0.1 m/pixel. Grids of 500 m x 500 m were drawn in Pix4D. Households and a buffer of at least 250 m were covered using several grids for imaging multiple, georeferenced, *Anopheles*, larval habitat, capture points. In each grid, 50 LULC waypoints were automatically calculated to ensure an overlap of at least 70% between neighboring images, necessary to generate an orthomosaic image. The flight plan was preloaded onto the DJI Phantom 4 Pro



Figure 5. Pix4D Capture app in an iPad Mini 4 for real time UAV signature RGB seasonal mapping Anopheline larval habitat capture points in Akonyibedo agro-village.



Figure 6. The DJI GO 4 app.

drone and the flight path was followed automatically. A flying time of ~30 min without a change of battery was required to complete the survey in each grid (Figure 7). The optimal flight height was 25-30 ft (7.5-9 m) above the capture point. We monitored the flight height in real time. Multispectral mapping was conducted over 7 randomly

sampled water bodies. In each water body, the drone was flown to an altitude of approximately 6 ft to 25 ft, which assured a GSD of 0.02 m/pixel. A grid of 270 m × 270 m was drawn in Pix4D and the RGB multispectral camera was set up to take an image each second during the 20-min flight time of the drone.

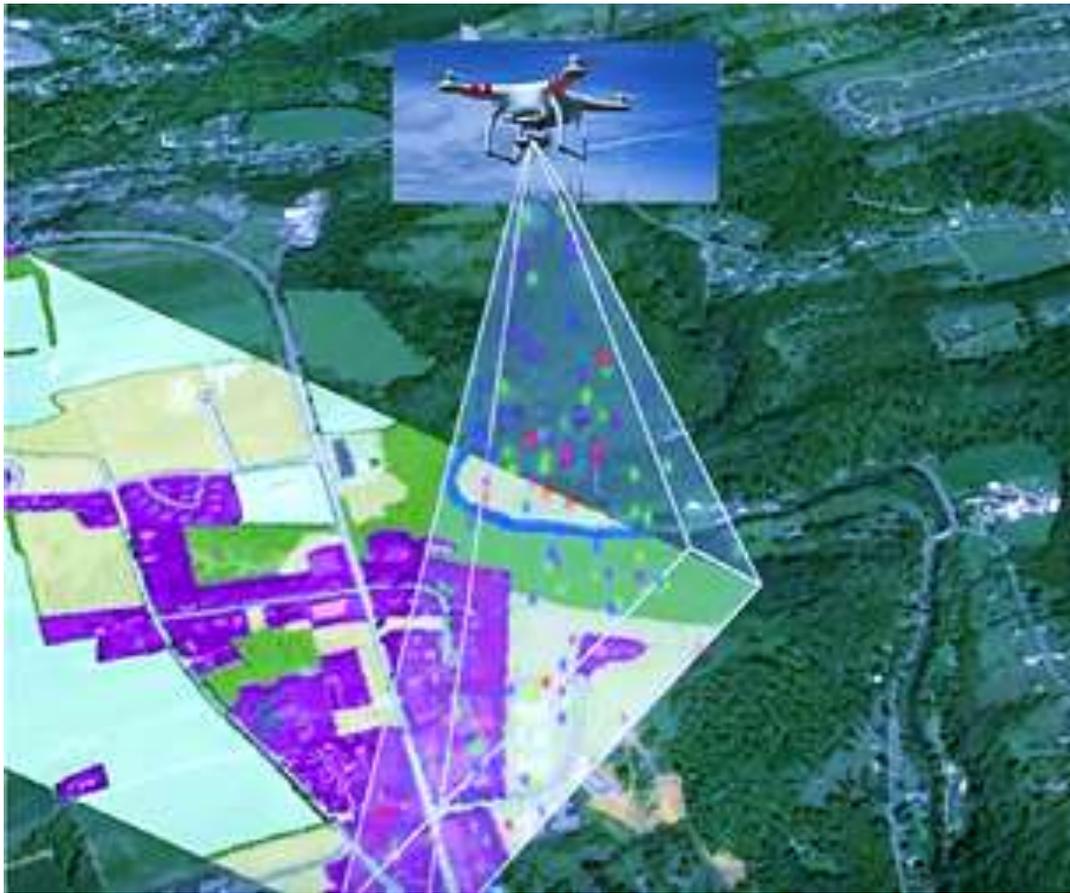


Figure 7. Simulated image of capture point RGB mapping.

Orthomosaic construction

The photogrammetric processing (gridded, LULC surface, capture point, habitat measurements based on photographs) was conducted in AgiSoft Photoscan Pro (www.agisoft.com). The resulting real time UAV imagery was imported into Photoscan and processed to construct an orthomosaic (that is, mosaic of overlapped LULC images) which included correction for topographic, capture point, signature distortions) for each georeferenced, Anopheline, larval habitat.

Three approaches were employed for conducting the spectro-temporal explicit LULC classification: (1) a classifier with particular focus on identifying water bodies placing the orthomosaic images into five groups: low vegetation, high vegetation, bare soil, urban and water bodies; (2) a classifier with a particular focus on differentiating water bodies with presence or absence of *Anopheles* larvae, classifying the orthomosaics into six groups: low vegetation, high vegetation, bare soil, urban, water bodies positive for *Anopheles* and water bodies negative for *Anopheles* and (3) a classifier with a particular focus on differentiating water bodies as positive or negative for *Anopheles* classifying only the water bodies drone detected into two groups: water bodies positive and negative for *Anopheles*.

In order to measure the statistical separability between positive (aquatic habitats consistently harboring *Anopheles* >50% of the time)-and negative (aquatic habitats consistently harboring

Anopheles < 50% of the time) - water body classes in approaches 2 and 3, an interclass separability analysis was conducted using the Jeffries Matusita (JM) distance. Briefly, JM is a measure of the average difference between two-class (positive and negative water body) density functions by pair-wise comparison and ranges between 0 and 2. A JM distance of 0 implied no separation and 2 for full separation between LULC classes geosampled in Akonyibedo village (Figure 8).

The position of the drone at the time of image capture for each photo was recorded automatically by the on-board GPS; thus, an orthomosaic was georeferenced without the need of Ground Control Points (GCP). A 3-D digital elevation model (DEM) was built in the real time portal using WGS 84 resolution of 0.1 m per pixel for the RGB and multispectral, capture point, LULC, georeferenced, Anopheline habitat, UAV images.

A normalized difference vegetation index (NDVI) was calculated for each capture point LULC based on the bands from the drone camera using the following formula: $NDVI = \frac{NIR - Red}{NIR + Red}$. The NDVI is a simple graphical indicator that can be used to analyze remote sensing measurements, from a space platform, and assess whether an Anopheline, larval habitat, capture point being observed contains live green vegetation or not (Jacob et al., 2015) (Figure 9). For each georeferenced capture point in Akonyibedo village orthomosaics were constructed: (1) a 3-band RGB image from the DJI 4K camera and an 8-band composite image (Table 1).

The image classification was conducted in Google Earth Engine



Figure 8. (A) Partially shade-canopied Anopheline larval habitat capture point in Akonyibedo agro-village complex; (B) Sampling for immature *Anopheles* in a road side ditch habitat.

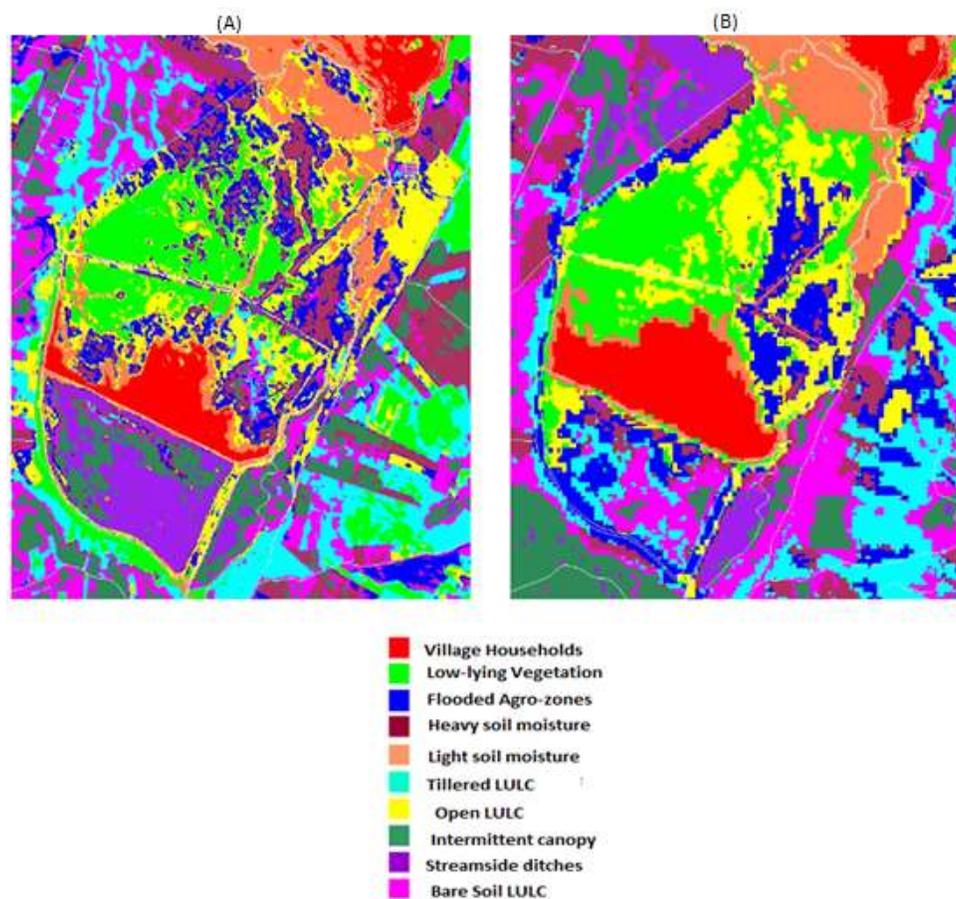


Figure 9(a and b). Two seasonal NDVI maps of varying LULCs in Akonyibedo agro-village.

Table 1. RGB and multispectral bands used as features in the classification.

	Spectral band	Label	Wavelength range (nm)	Resolution (m)
RGB bands	Blue	Blue	492 to 455	0.1
	Green	Green	577 to 492	
	Red	Red	780 to 622	
Multispectral bands	Green	green_m	550	0.02
	Red	red_m	660	
	Red Edge	edge_red	735	
	Near Infrared	nir	790	
Other	NDVI	ndvi	-1 to 1	0.02

Table 2. Relevant details on our dataset of breeding habitats.

Total duration of all videos	32 min
Total number of frames	1,889
Total number of frames with breeding habitats	1,100
Total number of frames without breeding habitats	789

(GEE). GEE is a cloud-based platform for planetary-scale geospatial analysis. [<https://earthengine.google.com>]. All classification analyses were conducted in the online Integrated Development Environment (IDE) at <https://code.earthengine.google.com> (repositories for data and code available in Supplementary information). All 8-band multispectral orthomosaics were uploaded to GEE assets and a supervised classification was performed using AI algorithms on RGB data.

Technical details

Data preparation

We collected around 32 min video of the whole village using DJI4 Drone (GPS enabled) during three seasonal collection frames (dry, pre-rain and rain) throughout 2019. Essentially four data gathering experiment were conducted, and each experiment was approximately 8 min long. Relevant details on our dataset are shown in Table 2. The data were used to train and validate our AI algorithm. We state clearly here that the height of the drone was chosen between 6ft to 25ft during these experiments.

After the process of data collection, and in order to process the videos, we first extracted each frame within each video as one image. For the entire duration of 32 min the number of frames (that is, images) extracted was 1,889. Out of these, the images which contained potential larval habitats (that is, sources of standing water) were 1,100 that were subsequently annotated (localized and labeled) using *labelling* tool (GitHub-Tzutalin, 2019). The others did not contain any source of stagnant water. Annotated images were further verified by an expert researcher for ground truth validity. In our efforts in designing an AI algorithm for classification, we employed 70% of annotated images towards training a model, and the rest images were used for validation. The total number of training and validation images generated after the aforementioned split were 770, and 330 respectively (for the class of images containing a potential larval habitat), and the split was similar for the other

classes.

We point out that the resolution of each original image extracted was 4,096 × 2,160 pixels, where each pixel was a representation of RGB color space. This was a relatively large size hence we assumed that there would be slow down training time. In order to accelerate training, without compromising accuracy, we reduced the image size by a factor of 4. This resulted in an image of size 1,024 × 524 (we observed no loss in model accuracy with this reduced size). After that, we normalized the RGB value of each pixel in an image by dividing it by 255. This aided in avoiding poor contrast from the image. Further, to increase the training images, we randomly zoomed in/ zoomed out each image between 0.5 and 1.5. Doing so, helped to add more robustness to our model for operating on unseen entomological and LULC data. Via this procedure, the total number of training images (original and scaled) generated was 1,540. We point out that the DJI Phantom drone provided GPS extraction capability using the notion of .SRT files containing the GPS information for each frame. Naturally, these were also extracted per frame, and they were stored in a .CSV file.

Convolutional neural network based object-detection algorithms to localize the breeding habitats

The state of the art techniques in image recognition relies on the notion of convolutional neural networks (CNNs) (Krizhevsky et al., 2012). We provide a brief overview here. There are two components here - feature extraction part and classification. During feature extraction, the network performed a series of convolutions and pooling to get the critical LULC features in the gridded image that aided classification. The convolution layers extracted features from the images by performing convolution operation with the use of filters on the input images to generate the feature map. Each convolution layer had 3 dimensions (width, height and depth). Typically the image contained n filters where the filter size was (a, b, c) and where a and b was the width and height of the filter and c represented the number of color channels of an image. Each filter independently

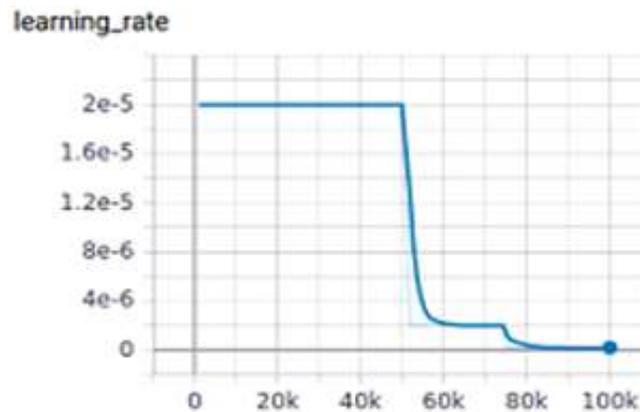


Figure 10. Learning rate for each iteration.

convolved on the input image that was subsequently followed by pooling to generate a feature map. The pooling layer aided in reducing the size of the feature map. After the filter size was chosen, stride size needed to be chosen. Stride size is the size of the step by which filter moves across an image. The process of convolution worked by computing a dot product between the filter and the local region of the image on which the filter was mounted. Since deep convolutional neural networks contain several convolution layers, each layer employed different filter sizes. As such, different feature maps were integrated together at the end and this acted as the final output of the convolution layer.

Subsequently, after the convolution layers, we added a few dense layers for classification. The neurons (essentially a non-linear function that takes multiple inputs and renders out a single output) in the dense layers were fully connected to all the neurons of the previous layers. The last dense layers consisted of neurons equal to the number of classes. We aimed to classify, and render the probability of each class present in the image. The one with the highest probability was predicted as the correct class for the particular image when presented for classification.

While classification of potential larval habitats in an image may alone suffice in most cases, we wanted to add another feature in our design to also highlight where the predicted larval habitat was within the image (that is, localization of the object within the image). When the operator viewed the habitat localized (bounded within a solid box) more details about the size of the larval habitat were inferred. Furthermore, the number of habitats present in the image was accessible. In addition, we generated a confidence metric to the operator for each bounding box which was emplaced in an image. Our solution to do so was based on the notion of Faster R-CNN.

In this Region CNN approach, we executed several steps which were challenging. First, we trained the drone sensed, seasonal, capture point, gridded LULC, image datasets using a pre-trained convolutional neural network model and extracted the convolutional feature map from the last layer of the trained network. This step enabled the neural network to understand the key features within the image that separates multiple classes of objects within the image. To train towards feature recognition, the right neural network must be used and adapted. For this study, we used Inception V2 (Szegedy et al., 2016) as the base pre-trained convolutional network for extracting the feature maps. Inception V2 is a complex deep learning architecture which employs smart factorization methods to make the convolutions efficient in terms of computational complexity. This architecture helped in achieving the same for our problem.

After extracting the feature maps, in order to localize objects of interest in the image, we employed the notion of Faster Region-based CNNs (Faster R-CNNs) (Ren et al., 2015). To do so, a few steps needed to be executed. First, we predefined anchor aspect ratios and scaled in our images. Anchors are the rectangular boxes that are used to scan objects in the image. These were emplaced during training the neural network. For the case of our drone imagery, we set the base anchor size as (256, 256) pixels, with scaling ratios and aspect ratios as (0.25, 0.5, 0.75, 1.0, 1.50, 2.0, 2.5) and (0.25, 0.5, 0.75, 1.0, 1.50, 2.0, 2.5) respectively. The width and height of each anchor was set as

$$\begin{aligned} \text{width_anchor} &= \text{scale}[i] * \text{sqrt}(\text{aspect_ratio}[i]) * \text{base_anchor}[0], \\ \text{height_anchor} &= \text{scale}[i] * \text{base_anchor}[1] / \text{sqrt}(\text{aspect_ratio}[i]) \end{aligned}$$

where i was the index of the matrices of scales and aspect ratios. In total, 300 anchors were generated per image for training. Then, we fed our anchors and feature maps to the region proposal network (RPN). Here, the task was to train the network to identify those boxes in the image that were indicative of Anopheles larval habitats. To do so, we manually ground truthed each box prior to training as a potential larval habitat and the classified background. The RPN consists of small convolutional layers, was trained to identify the anchors having the relevant objects of interest (in our case, a potential, eco-georeferenceable, seasonal, Anopheline larval habitat, capture point aquatic foci) with an "objectness" score and return the ones that most likely contain objects within the image based on the score along with the center coordinates, width and height of all anchors [Ren et al. 2015].

Finally, we resize the feature maps of the anchors (that is, anchors which contains objects of interest) learned from the above step into fixed sized feature maps which serve as input to two branches to: a) classifier to identify potential larval habitats within the anchor; and b) regression to tighten the anchor for improved accuracy. All two steps are done in parallel. We fine-tuned the classifier by varying learning rate. We used stochastic gradient descent (SGD) solver for 50,000 iterations with a base learning rate of 2e-5, then another 25,000 iterations by reducing the base learning 2e -6 and the rest 25,000 with 2e-7 for faster convergence. The learning rate for each iteration is shown in Figure 10.

During training, one important criterion for correctness is the loss function (Ren et al., 2015). Briefly, the loss function measures the learning ability of a neural network architecture. Typically, it ranges from 0 to 1 where 0 means perfect learning and 1 means no

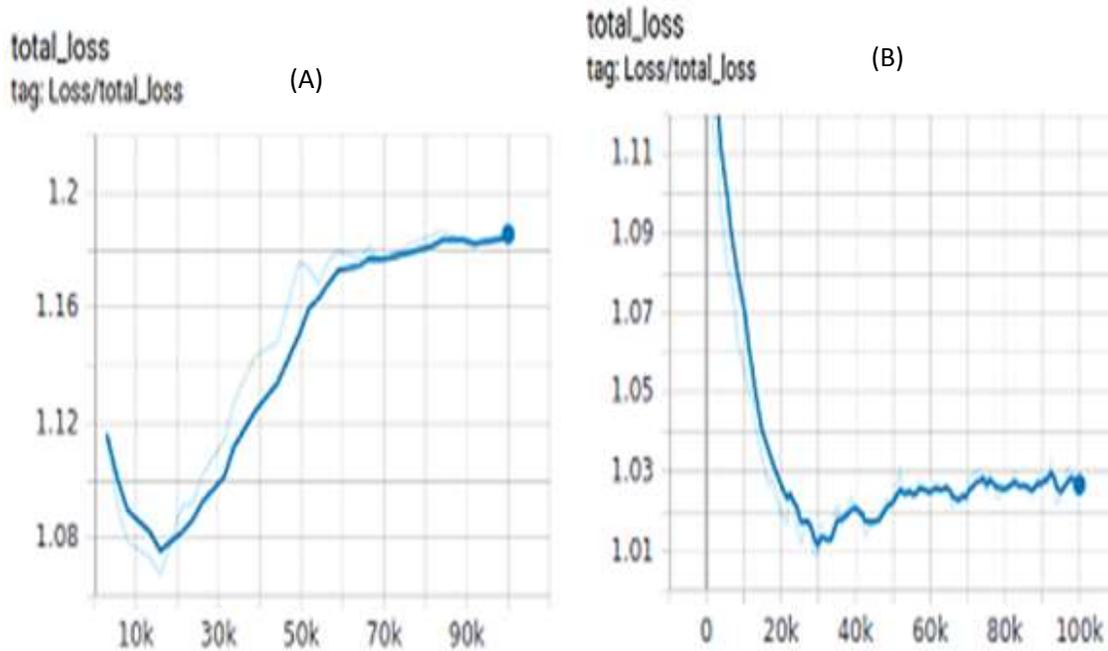


Figure 11. Total loss graph (a) without and (b) with focal loss.

learning. The goal of training the anopheline larval habitat data was to minimize the loss of training data. We noticed that the default binary cross entropy loss function (which is standard) had increased after 15,000 iterations. To minimize the loss, we applied a novel focal loss function which penalizes instances of false negatives which in our case was an actual larval habitat classified as background (Lin et al., 2017). We noted that the number of anchors containing breeding habitats was lesser in comparison to anchors that contained the background class. Hence, the classifier was biased towards the background class which initially resulted in biased learning. Focal loss is an improvement over the more standard cross entropy loss, since it operates by lowering the loss of well classified cases and emphasizing the misclassified ones (Lin et al., 2017). We considered the following equation,

$$Q = \begin{cases} p & \text{if } y = +1 \\ 1 - p & \text{otherwise.} \end{cases}$$

Here y denoted the ground-truth of the class (+1 for larval habitat and -1 for background), and $p \in [0,1]$, which was the model estimated probability for the class label $y = +1$. The Focal loss (FL) for each anchor k is defined as

$$FL(k) = -(1-Q)^\lambda \log(Q).$$

Here λ is a tunable parameter. In this research, when a potential *Anopheles* larval habitat was misclassified as background and Q was small, the modulating factor $(1-Q)^\lambda$, was close to 1 and this did not affect loss. When Q tends to 1, the modulating factor is close to 0 and loss for well classified examples is down-weighted. $\lambda=2$ performed optimal. After applying focal loss, we were able to determine that overall minimal loss was achieved in the model.

Essentially, when the loss in training data is similar to the loss in validation data (recall that validation data is not used to train the model), and when they do not decrease any further, the process of training is complete. The process of training and validation of the model was an iterative process. The final set of parameters of the neural net architecture that met our (loss) criteria are presented (Figure 11a and b).

We defined a user-understandable metric for revealing the quality of our trained and validated neural network. Our metric was mAP over Intersection over Union (IoU) threshold. The IoU metric measures the correctness of a given bounding box (Ren et al., 2015). It is formally, the area of the intersection of the predicted box and ground truth box divided by the area of union of the predicted box and the ground truth box. It is illustrated in Figure 12 where Green denotes the Ground Truth Box, and Red denotes a Predicted Box. In our design, the IoU threshold was set as 0.7. When the IoU is 1, then a perfect classification and emplacement of the bounding box occurred. Lower IoU values indicate poorer performance (Ren et al., 2015).

By computing the True Positives, False Positives and False Negatives for classification in the validation set, we were able to integrate a final metric called mean Average Precision or mAP. Denoting AP as the Average Precision (AP) for finding the area under the precision-recall curve of each class, the mean Average Precision or mAP score was calculated by taking the mean AP over all classes over IoU threshold. Note that Precision and Recall are standard metrics in binary classification problems. The Precision was defined as the ratio of True Positives to the Sum of True Positives and False Positives. The Recall was subsequently defined as the ratio of True Positives to the Sum of True Positives and False Negatives. A high Precision and Recall indicate a more accurate classifier (Powers, 2011).

The mean Average Precision (mAP) of all images in our validation set was determined to 0.87 for an IoU of 0.7. Note that if when we set the IoU with lower thresholds (that is, less than 0.7), the mAP

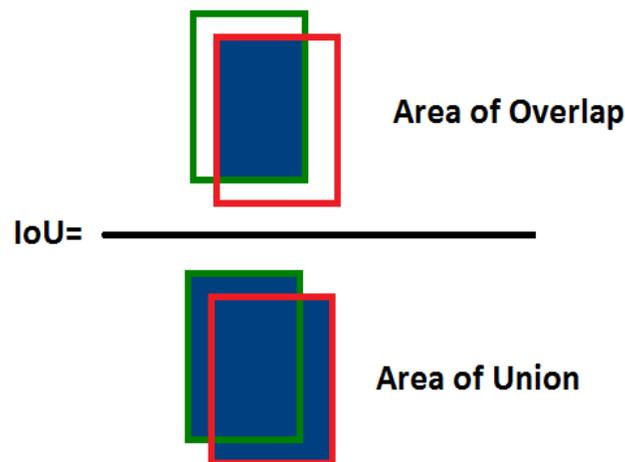


Figure 12. The mAP of the images in validation set.



Figure 13. An Instance of a Predicted *Anopheles* larval habitat Akonyibedo by our AI Algorithm.

was higher. We trained our Faster R-CNN model using the annotated images to detect and locate sources of breeding habitats (Figure 12). It took close to 10 hours test and validate the model using a graphic processing unit (GPU) cluster. The cluster has 4 nodes of GeForce GTX TITAN X each with 12 GB memory.

RESULTS

In order to test the correctness (or accuracy) of our neural network model, we tested it with completely unseen images retrieved from the drone spectral library. By unseen, we mean that the images that were fed into the neural network for classification and localization of larval habitats which were never used during training or

validation (details of which are elaborated above). Our algorithm performed optimally here, and we were confident that our model was one of high fidelity. For testing we flew the UAV over neighboring unsampled locations at the peri-entomological, agro-pastureland, intervention site. During these wayward flights, 7 videos with a total of 15 min were collected. The total number of frames extracted was 959 with 60% of them containing at least one potential larval habitat. In some frames georeferenced breeding site, *Anopheles* foci were repeated

Subsequently, each frame was fed into our model for classification and localization. We derived robust outcomes. Figures 13 and 14 reveal breeding habitats



Figure 14. An instance of a predicted anopheles larval habitat Akonyibedo by our AI algorithm.

identified accurately. The probabilities of detection by our neural network (that is, the confidence it had in making the prediction) were close to 0.99 in most images. This signifies that every potential source of stagnant water (that is, larval *Anopheles* habitat) was correctly classified and localized. There was a minute number of False Positives (<10%) present. It may be expected that our model accuracy would increase with increased training data.

For practical deployment purposes, once we predicted a frame to have at-least one georeferenced, anopheline, larval habitat, it was extracted based on its GPS indexed geolocation from the associated .csv file which was provided by the drone. The final output to the operator was provided as an image with bounding boxes and its GPS location was able to geolocate the anopheline habitats in the form of a simple smartphone app. These habitat sites were subsequently classified as productive or not for prioritizing seasonal, breeding sites, for implementing larval control strategies by overlaying a georeferenced Anopheline, RGB analog, video larval habitat seasonal capture point, signature over the UAV, real time, geosampled, gridded, LULC images.

The sensitivity and specificity of the RGB drone signals at identifying positive and dry habitats were subsequently evaluated in blinded experimental studies in a natural setting followed by extensive ground-truthing of the UAV sampled, real time, LULC classified, model outputs employing ground coordinates which had a positional accuracy 0.178 m. All or 100% of the capture points from 90 (unique) larval habitats forecasted by the model in Akonyibedo village were identified, and were found to contain *Anopheles* larvae. Pyrethrium Spray Catch (PSC) statistics were subsequently developed.

DISCUSSION

We employed a real time, UAV-differentially corrected

GPS-AI platform to prognosticatively delineate georeferenceable geolocations of probable unknown, unsampled, *Anopheles*, breeding site, oviposition, capture point, seasonal foci in various landscapes (e.g., grassy, streamside, irrigation ditches, vernal roadside pools etc.) in an peri-urban agro-village, pastureland complex (Akonyibedo) in Northern Uganda. We employed a DJ4 Phantom drone which had a RGB camera, with a 1-inch, 20-megapixel sensor that shot 4K video at 60 frames per second (fps) of different LULCs in the study site agro-village complex employing a variety of automatic flight modes including Draw, ActiveTrack, TapFly, Gesture, and Tripod settings. The UAV had ample internal storage and battery life (up to 128GB via microSD and 30 minutes, respectively). Two flight altitudes (6-25 m) with two flight modes (stop and cruising modes) were employed for acquiring precise capture point, ground control point, indexed, landscape measurements. We eco-georeferenced, multiple, *An. gambiae funestus* and *arabiensis*, breeding site, LULC, seasonal, capture points in a variety of different peri-urban, pastureland, agro-village, land cover classifications in Akonyibedo village. These gridded images were assessed to identify characteristics unique to the habitats.

We analyzed the influence of real time, multispectral, radiometric, UAV, real time, classified, spectral RGB signature, bi-directional reflux, wavelength, emissivity parameters for remote discrimination of multiple seasonal hyperproductive, aquatic and dry, seasonal, *Anopheles*, capture point, LULC images. The reflectance spectrum of multiple, agro-village, agro-pastureland anopheline habitats was established in the semi-autonomous, drone aircraft dashboard as a plot of the fraction of radiation capture point reflected which was a function of unmixed, RGB indexed, incident wavelengths in the peri-urban pastureland epi-entomological Ugandan intervention study site.

The AI component of our technique was the successful

training of a Faster R-CNN (Region-based Convolutional Neural Network) model using all annotated RGB images from the drone video to detect and locate sources of Anopheline aquatic breeding geolocations at the intervention, epi-entomological site. The R-CNN model essentially performed two tasks, classification and regression. The classification component found the object of interest (in our case, a potential larval habitat) within an image, and the regression components emplaced a tight bounding box. Our final metric to assess accuracy was mean Average Precision (mAP), which is a function of another metric called IoU, essentially compared the overlap amongst the ground truth datasets and predicted bounding boxes after classification and localization. For a relatively high IoU threshold of 0.7, the mAP value during validation was 0.87, which was a high number for a complex problem like ours. With more training data, the accuracy will only improve further.

The sensitivity and specificity of the RGB capture point, LULC signals at identifying the targeted seasonal Anopheline habitats was then evaluated in blinded experimental studies by real time drone identification of breeding sites in a natural setting followed by extensive ground-truthing of the real time model outputs. All hyperproductive, seasonal aquatic and dry habitats were identified. Once optimized the models were incorporated into a mobile app that analyzed the density of breeding site pixels. The app was employed to remotely identify LULC properties in which individual breeding sites and clusters of habitats (that is, "hot spots") were observed representing a potential, Anopheline, larval habitat, hyperproductive, aquatic and non-aquatic, seasonal, intervention site. The app was employed to remotely identify LULC properties in which individual breeding sites and clusters of habitats ("hot spots") were observed. These Anopheline, larval habitat, hyperproductive, aquatic foci were then prioritized based on field sampled entomological data (larval counts, Euclidean distance to a homestead etc.) and then catalogued as a seasonal, intervention site in the app. The app in the real time platform recorded the GPS location of the aquatic breeding site as a pin of an unknown habitat on a Google Earth map.

We noted that the real time platform performed all the necessary system checks (e.g., battery levels, differential correction of GCPs, camera storage, etc.) while autonomously flying a wayward mission in Akonyibedo village with the proper ortho-overlapping for real time imaging an experimental or predicted, georeferenceable, capture point, Anopheline larval, habitat, LULC site. The platform had cloud image processing, single-click collaboration, export, and integration options with the capability of embedding object-based classification methods (e.g., capture point, RGB wavelength separation) for, optimal land cover mapping. The real time platform allowed measuring un-aggregated, RGB, wavelength,

reflectance, capture point, 2-D and 3-D, LULC reflux variations in the UAV, geosampled empirical datasets of seasonal, hyper-productive, aquatic and non-aquatic, habitat pixel frequencies, where probability of a habitat containing larvae or pupae was measureable.

Elucidating first-order, differential, eco-georeferenceable, capture point, LULC surface reflux through dimensionless, radiation-based, discontinuously vegetated, seasonal, habitat canopies employing off-nadir, NIR and red wavelength, radiance in the Akonyibedo agrovillage, pastureland, epi-entomological, intervention, study site, approximated unobserved, isoline convergence and soil-perturbed, LULC responses in the sub-meter resolution, Anopheline habitat, spectral, signatures in the drone dashboard real time machine learning module. The UAV allowed the mapping of the vegetation LULC at very high spatial resolution. For reflectance measurements and vegetation indices (Vis) to be comparable between seasonal, eco-georeferenceable, capture point, larval habitat sites and over time, careful flight planning and robust radiometric calibration procedures is required (Jacob et al., 2019).

Two sources of uncertainty that require attention for future, anopheline, larval habitat, signature mapping are illumination geometry and the effect of flying height. This study developed methods to quantify and visualize effects in imagery from the Parrot Sequoia, a UAV-mounted multispectral sensor. Changes in illumination geometry over one day had visible effects on both individual Anopheline habitat, capture point LULC images and orthomosaics. Average NIR reflectance and NDVI in regions of interest were slightly lower around solar noon, and the contrast between shadowed and well-illuminated areas increased over the day in all multispectral bands. Per-pixel differences in NDVI maps were spatially variable, and much larger than average differences in some classified LULC areas at the epi-entomological intervention study site. Results relating to flying height indicated that 6-25 though small increases in NIR reflectance with height were observed. These results underline the need to consider illumination geometry when carrying out UAV vegetation surveys for targeting eco-georeferenceable, hyper-productive, aquatic and dry, seasonal, Anopheline larval habitat, capture points.

Real time, UAV-based, seasonal, habitat time series mapping could have a wider range of applications for precisely determining geolocations of Anopheline, larval habitat, LULC capture points. For example, high resolution DEMs generated in a real time portal may be useful tools to analyze watersheds and small streams favorable to *Anopheles* breeding sites that are shaped by intermittent heavy rain. Moreover, these DEMs may support the identification of seasonally flooded areas, common in peri-urban agro-pastureland, epi-entomological, intervention sites that possibly increase human-mosquito contact and therefore are associated



Figure 15. Another more complex instance of how our algorithms classify and localize larval habitats with a confidence metric in Akonyibedo village in Uganda.

with a higher risk of malaria. Also, these seasonal, RGB, video, analog signatures may be input into a real time framework in ArcGIS for efficient visual, larval habitat, capture point, similarity-based matching for scaling up to a larger epi-entomological intervention site (e.g., agro-irrigated, pastureland, agro-village geolocation to a district level sub-county) in a real time UAV platform. All drone sampled real time data were streamed into ground stations via WiFi where control personnel viewed the live footage using a multi-directional, mobile, hand held device (e.g., Android technology, iPhones). Sub-county, local, mosquito control personnel in Akonyibedo village measured the real time UAV captured, mosquito, capture point, signatures (that is, georeferenced breeding site, habitat targets).

Sub-county control personnel at the intervention subsequently visited the field verified tagged properties and applied an environmental friendly treatment for controlling Anopheline habitats in the epi-entomological, intervention study site. This treatment included entirely burying the habitat with soil substrate. We all imitated a real time treatment for large habitats such as swamps and rock pit quarry habitats whereby the drone was used to deliver an environmental friendly insecticide to these foci within a period of 30 days there was a drastic reduction of airborne Anopheline mosquitoes based on Pyrethrium knockdown exercises conducted at random households at

100% active breeding sites as at baseline, on average, over 8 female anopheles mosquitoes were found per house spayed. On destruction of 65% breeding sites, a monitoring PSC showed a vector reduction to 1 female *Anopheles* mosquitoes per household in 30 days. This means that if all the breeding sites were destroyed, there would be an effective vector reduction to zero indoor resting female *Anopheles* mosquitoes per house hold.

In this research we noted that many active breeding sites in Akonyibedo village were potholes created by delivery trucks on marram roads; we also noted that these potholes were constantly stirred up by moving trucks, greatly destabilizing the young *Anopheles* mosquito development. This allowed “Seek and Destroy” to focus on these small positive breeding sites. Ironically, these reports have indicated that the locals believe that filling road potholes is the job of the government, and so these potholes have remained the main breeding sites for *Anopheles* mosquitoes in these agro villages. Reapplication of soil substrates in peri-urban anopheline breeding sites may be required due to precipitation and other environmental changes (e.g., agricultural land cover changes such as agro-irrigation flooding to post-tillering to pre-harvest rice seasonal cycle (approximately 120 to 150 days) . Hence these treated habitats should be bi-weekly monitored. These sites may then then re-treated as necessary (Figure 15).

Real time technology in a UAV offers the potential to identify and treat large water bodies such as a rock pit quarry habitats. For example, real time drone-based imagery has the potential to provide ancillary information for planning of logistics for treating large habitats: that is, location and nature of access capture points/routes to the swamps to direct field vector control teams to initiate real time seasonal larvicidal treatment. One of the greatest advantages of real time drone systems is their flexibility for real time larvicidal treatment for large habitats *Anopheles*; insecticide may be targeted which can be very cost-effective. Blanket treatment protocols are very expensive (Jacob et al., 2015). Although drones cannot be flown in the rain, they are not reliant on clear sky conditions (as they are flown at low altitudes, below clouds, unlike optical satellites) which can be very efficient for treating large seasonal Anopheline habitats. Additionally, drone, real time imagery can be used to establish and monitor links between environmental factors (e.g., low lying vegetation, soil moisture) and malaria, disease transmission, such as changes in land cover and the emergence of new vector habitats.

Capturing data multiple longitudinal entomological surveys throughout Uganda would provide the tools to study Anophelinae breeding site dynamics for optimally employing “Seek and Destroy” larval control for small commercial road ditch habitats and real time drone larviciding of larger rockpit quarry and swamp breeding habitats for lowering malaria transmission. For instance, the adaptation to more permanent anthropogenic larval habitats on LULC change sites (e.g., rice tillers to flooded pre-harvest habitat), eco-georeferenceable geolocations has been hypothesized to be the cause of a resident population of *An. arabiensis* in rice schemes leading to a perennial presence of this species, hence promoting and maintaining continual *Plasmodium* transmission. “Seek and Destroy” and season real time drone larviciding may be cost efficient and timely implemented in these LULC seasonal change sites in Uganda.

Overall the most important methodological caveat in this study is the definition of positive and negative, Anopheline, capture point water bodies. Although we sampled only 7 negative and 7 positive water bodies capture points for the presence of immature *Anopheles*, we achieved a 100% sensitivity and specificity during field verification exercises. Future research work in Uganda should consider more frequent, seasonal, drone sensed, real time surveillance of additional water bodies from other sub-county, epi-entomological, intervention communities and additional drone flights over the survey localities at different times of the day and under various atmospheric conditions.

In conclusion, real time, drone sensed, capture point, time series, *Anopheles*, breeding site, RGB, signature interpolation using machine learning algorithms can optimally geolocate unknown, unsampled, seasonal,

larval/pupal habitat, capture point, LULC, seasonal maps and 2-D live maps in a UAV dashboard. This data may be live streamed to mobile hand-held devices (i-phone, android technology) for immediate mapping of unknown, unsampled foci using an Android or IOS app. Thereafter environmental friendly Seek and Destroy and drone larviciding may be implemented for reducing immature anophelines and hence reducing malaria transmission throughout Uganda.

CONFLICT OF INTERESTS

The authors have not declared any conflict of interests.

REFERENCES

- GitHub-Tzutalin/LabelImg: LabelImg is a graphical image annotation tool and label object bounding boxes in image. Accessed December 21, 2019. <https://github.com/tzutalin/labelimg>.
- Jacob B, Mwangangi JM, Mbogo CB, Novak RJ (2011). A Taxonomy of Unmixing Algorithms Using Decomposing a QuickBird Visible and Near Infra-red Pixel of an *Anopheles arabiensis* Habitat. *Open Remote Sensing* 17(3):11-24.
- Jacob B, Novak RJ (2015). Integrating a Trimble Recon X 400 MHz Intel PXA255 Xscale CPU® Mobile Field Data Collection System Using Differentially Corrected Global Positioning System Technology and a Real-Time Bidirectional Actionable Platform within an ArcGIS Cyberenvironment for Implementing Malaria Mosquito Control. *Advances in Remote Sensing* 3(3):141-196.
- Jacob B (2019). Robustifying content-based real time video retrieval of RGB analog signatures by scalability and genericness indexed with time sequence frames for optimally delineating unsampled eco-georeferenceable capture point *Anopheline gambiae* s.l. aquatic foci based on similarity frame sequencing in a non-compressed geospectrotemporal domain. *Journal of Photogrammetry and Remote Sensing* (In press 2019).
- Krizhevsky A, Sutskever I, Hinton GE (2012). Imagenet classification with deep convolutional neural networks. In *Advances in Neural Information Processing Systems*, pp. 1097-1105.
- Lin TY, Goyal P, Girshick R, He K, Dollár P (2017). Focal loss for dense object detection. In *Proceedings of the IEEE International Conference on Computer Vision*, pp. 2980-2988.
- Powers DM (2011). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation.
- Ren S, He K, Girshick R, Sun J (2015). Faster r-cnn: Towards real-time object detection with region proposal networks. In *Advances in Neural Information Processing Systems*, pp. 91-99.
- Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z (2016). Rethinking the inception architecture for computer vision. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 2818-2826.