**SEMI-SUPERVISED PIXEL CLASSIFICATION IN SATALITE IMAGES**

SINGLE ELEMENT IDENTIFICATION OF DOWNED AIRCRAFT WITH R

John Camp

Machine Learning

University of Georgia

**SEMI-SUPERVISED PIXEL CLASSIFICATION IN SATALITE IMAGES**

Locating Downed Aircraft with R

The challenge which drives satellite image analytics is the detection and classification of objects of interests over large geographical areas. Existing methods unfortunately have proved inadequate for the location and classification of small heterogeneous objects. Several pipelines utilizing convolutional neural networks have proven adept at recognizing people, cars, etc. in cell phone photos. (Ciregan, Meier and Schmidhuber 2012) In one form or another these are all sliding window approaches. (Szegedy, Toshev and Erhan 2013) A bounding box with dimensions similar to those of the object of interest is placed at various positions and in varying scales in the image. Row by row and pixel by pixel the frame is examined for objects that fit within the box. These are “brute force” methods which require a powerful computer and a lot of time. The methods are effective, in part, because the models have been trained to recognize cars, people, signs, and other urban objects. Scale then is rarely an issue. (Wojek and Schiele 2005)

These methods are not optimized for the detection of small objects in large images and have fared poorly when applied to such images. (Lu and Weng 2007) Satellite images have varying scale, so depending on the distance between the satellite and the object, and the power of the satellite’s equipment, the object of interest can range from one pixel, to most of the image. The use of Light Detection and Ranging (LIDAR) to accurately calculate the distance from the camera to the earth’s surface holds promise, and will hopefully soon address the issue of shifting scale, but that research is only just beginning. (Reutebuch, Andersen and McGaughey 2005)

There are methods optimized for satellite image analysis. These methods are not concerned with the search and classification with heterogeneous objects however, but instead with the classification of ground cover: forest, lakes, erosion, urban areas, etc. along with other features such as pipelines, rivers, roads, and rail systems. These algorithms are very complex. They use multiple image bands, thermal, and LIDAR images to identify varying types of ground cover and train the model for further analysis. (Bork and Su 2007) Still these methods are not suitable for small object classification.

In this project I use pixel classification, an older method predating low orbiting satellites, to locate downed aircraft in remote, single feature terrain. Each image of a downed aircraft is a data set. The single feature terrain is the training set and the aircraft is the test. I plot a binarized image, based on color Z-scores, representing the reflectivity of first the single feature terrain and second the aircraft.

**Why Search for Downed Aircraft?**

There are two primary motivations for mapping the locations of these planes. The first reason is the recovery and repatriation of the airmen’s remains. Some of these airmen’s spouses are still living, but even their children are nearing the ends of their lives and deserve, if possible, answers and closure. The second reason is that these planes are very rare. There are very few WWII planes remaining and even the parts from demolished aircraft can be very valuable. Consider the Rolls-Royce Merlin Engine that powered most British aircraft. A rebuildable core is worth $20,000.00[[1]](#footnote-1), while the Daimler-Benz DB 600 series engines used in many of the German planes is worth ten times that. An intact airframe might be worth more than a million dollars. [[2]](#footnote-2) Many of these planes were lost in remote areas and have not been disturbed since.

For instance, in 2012 oil workers walked up on an intact P40 Kittyhawk, a WWII fighter produced in America by Curtis-Wright and delivered to British Commonwealth forces fighting in North Africa as a component of Lend lease. The serial number read: 87-21-705. The whereabouts of that plane was last known on June 28, 1942, when Flt Sargent Dennis Copping of the Royal Air Force was ordered to fly the damaged fighter from a frontline airfield in Libya to a British airfield in Egypt for repairs. For reasons unknown, Sargent Copping apparently crash landed. It is known that he survived the landing, but died soon after as his remains were found only a few miles from the plane. (Corbley 2013)

**A P40 with its familiar shark’s teeth Sargent Copping’s P40**

That plane sat on the desert floor unnoticed for 70 years. It has happened before. In 1960 a British Petroleum exploration team found the B24 Liberator “Lady Be Good” in the Libyan Desert. (Romito 2016) In 1994 an oil team wandered up on a British Commando truck with still functioning machineguns. In 1992 UN relief workers in Ethiopia found 5 British and 3 Swedish fighters. (Cooper 2015)

The point is that the Sahara Desert is really big, and in most areas so inhospitable that people do not go there at all. (Leslie 1998) There were also a lot of aircraft that landed and were never recovered, many simply disappeared, and thousands were destroyed in combat. Most of those planes are still there. This does not apply to just the desert. There are thousands of WWII planes that crashed or crash-landed in remote areas, including The Yukon Territory, Alaska, Siberia, Greenland, Iceland, The Karelian Peninsula, and Scandinavia.

**Heterogeneous Object Identification**

The challenges of identifying heterogeneous objects in remote sensing data are many. As mentioned, the area to be searched is massive. Another challenge is that airplanes are rather small objects. As an example, consider this image of the P40 found in 2012. The image was taken by the WorldView-2 Satellite with eight-band multispectral color imagery at 1.5-m resolution. (Corbley 2013) This is the highest definition image available.



The plane is in the center of the image and it should be clear that the sliding window method would be a poor choice to locate this object. The sliding window classification method is a supervised classification in which a shape similar to the object of interest is “slid” over the image searching for objects that “fit” within the “window.” The shape of an airplane is clearly present, so a model could be trained to classify objects of this shape. However the P40, which has a wing span of 30 ft, is tiny when compared to the image. It is true that there are many planes, trucks, tanks, etc. in the Sahara, but the area containing these relics is more than 3.5 million square miles. Placing a bounded box over every pixel in millions of images would be inefficient to say the least, especially when nearly all of the pixels are devoid of any man made material.

Ground cover classification would also be a poor choice. Ground cover classification is a supervised classification method usually conducted on high orbit images that survey geographical regions like deltas and mountain ranges. The purpose of ground cover classification is to color code varied types of ground cover, and the challenge is to accurately separate types, like forests from urban areas. The inadequacy of ground cover classification was highlighted during the search for both Malaysia Airlines Flight 370 and the Indian Air Force An-32 that disappeared in 2016. In both cases machine classifications were abandoned and replaced by human examination of images.

**Supervised and Unsupervised Pixel Classification**

Object classification algorithms are viable only on high resolution images, hence their popularity identifying objects in digital photos. Object classification on satellite images is now more popular since high resolution low orbit images have become readily available. When analysis was done on data collected from the older low resolution and high orbit satellites, pixel classification was the method of choice.

Pixel classification is a clustering method that can be either unsupervised or supervised. Pixels are clustered based on their level of reflectance. Unsupervised, the user simply identifies how many clusters to generate. The user then examines the images and labels the clusters according to professional judgment. The most popular algorithm for unsupervised classification is K-means clustering. (Theiler and Gisler 1997)

Supervised, the user trains the model with areas of the image judged as typical of each land cover class. The classification of land cover is based on the spectral signature defined in the training set. The most common supervised classification algorithms are maximum likelihood and minimum-distance classification. (Mather and Tso 2009)

**Single Feature Heterogeneous Object Detection**

There is a distinct advantage to searching for objects in most remote areas, and that is the general homogeneity of the terrain. The desert is mostly sand, the icecap is nearly all glacier, and much of the sub-arctic lands are covered in tundra. From a land cover perspective, these are single feature terrains and any cluster analysis would reveal one primary cluster that would encompass nearly all of the data. This means that downed aircraft could be identified by only one feature, if that one feature were sufficiently different from the surrounding terrain.

Objects have seven features that can, to some extent, be observed through satellite imagery. These include tone (relative brightness), hue (Red, Green, Blue Color), texture, pattern, shape, size, and with the aid of LIDAR height. Texture, pattern, and shape are much more suited to object classification methods, which would still be too inefficient even considering one feature. Without a uniform scale, size is not an option. Height may be useful in the future as 3D images of remote locations become available but it is not an option now. That leaves tone and hue. The human eye can distinguish around 10 million of the 16.8 million unique colors, but can only detect around 30 of the 256 shades of brightness. (Riaz, et al. 2008) This would make tone the best candidate.

Satellite images are not photographs. They are recordings of varied reflectance. Reflectance is the fraction of incoming solar radiation that is not absorbed by the Earth. Instead it is reflected back out. Satellite images approximate sensor readings just above the Earth's surface. Clay, sand, water, forest, etc. all have different levels of spectral reflectance, each with signature wavelengths. (Feng 2012) The search area then should have one predominate wavelength releflecting from the surface. If the relectivity of the metal in the planes is different enough from the ground cover in the search area, then the aircraft will be an anomaly, and its location identified. Reflectance is measured by surface albedo. The surface albedo of the Western Desert in Egypt is estimated at between 30% and 33%, while the surface albedo of aluminum ranges from 40% to nearly 100% depending on corrosion level. (Pinker and Karnieli 1995) This is a scientific way of saying that aluminum is shinier than sand.

**Method**

One difficulty in building a supervised model for aircraft classification is finding the data to train the algorithm. This is certainly the case here. There simply are not enough images of downed aircraft in the varied remote locations. This is why I have termed the method semi-supervised, while it probably is really a supervised method of classification.

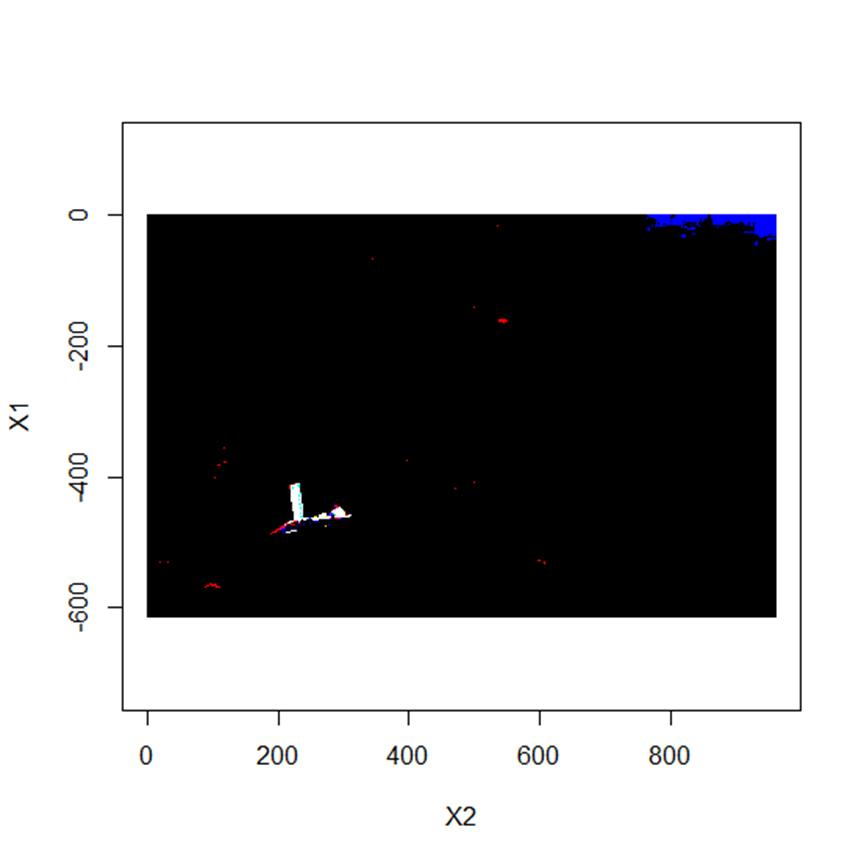
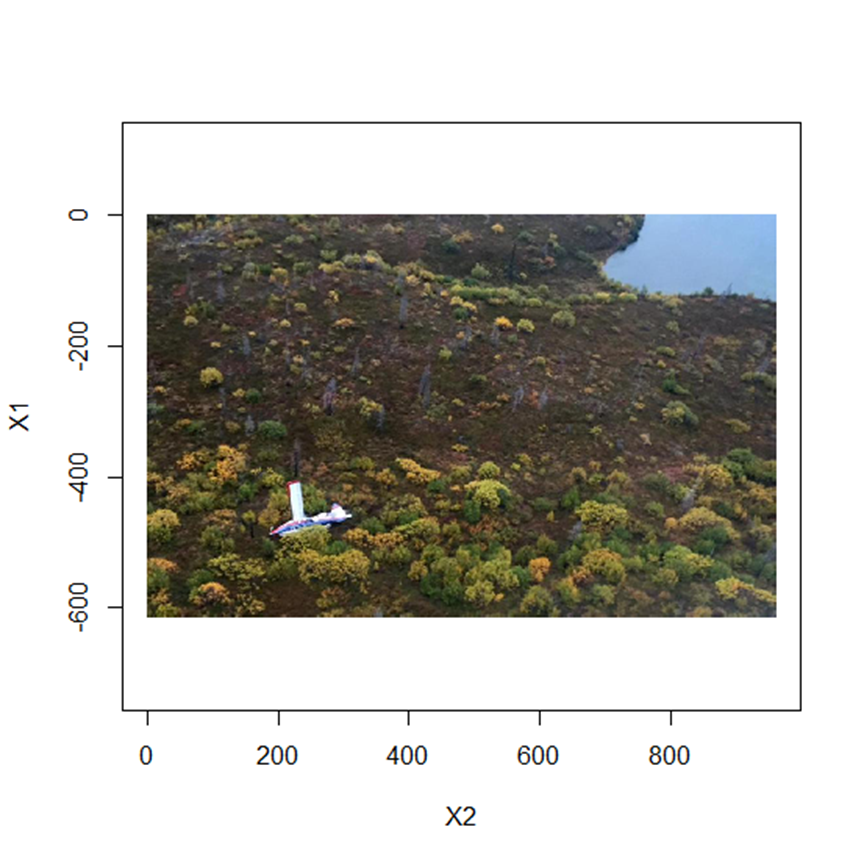
I first attempted several clustering techniques, including K-means and nearest neighbor, but none of the clustering methods fully separated the aircraft from the ground cover. As predicted, there was one cluster encompassing most of the data. There were several smaller clusters but the number was set at two for obvious reasons. Unfortunately, the algorithms classified a small percentage of sand with the aircraft, resulting in a salt and pepper noise pattern that completely obscured the planes. Therefore, none of these methods successfully identified any of the aircraft. I then decided to try a much simpler method that would deliver a binarized image, hopefully of the ground cover and the aircraft.

The method was tested on the images of seven downed or crashed aircraft. Each image was treated as a different data set. The test images include one aircraft in the Sahara Desert, one in the Mojave Desert, one in the Namibia Desert, one in the Nevada desert, one on the tundra in Alaska, one in the Alaskan wilderness, and one on the ice cap in Greenland. In each case a portion of the image excluding the aircraft was used as a training set. Using the R package JPEG, the mean and standard deviation of the test set were calculated and recorded. Brightness Z scores were then calculated by subtracting the test set mean from the dataset and then dividing by the standard deviation. Density plots of the brightness scores were made and used to determine the reflectivity threshold of the model. The results were then plotted.

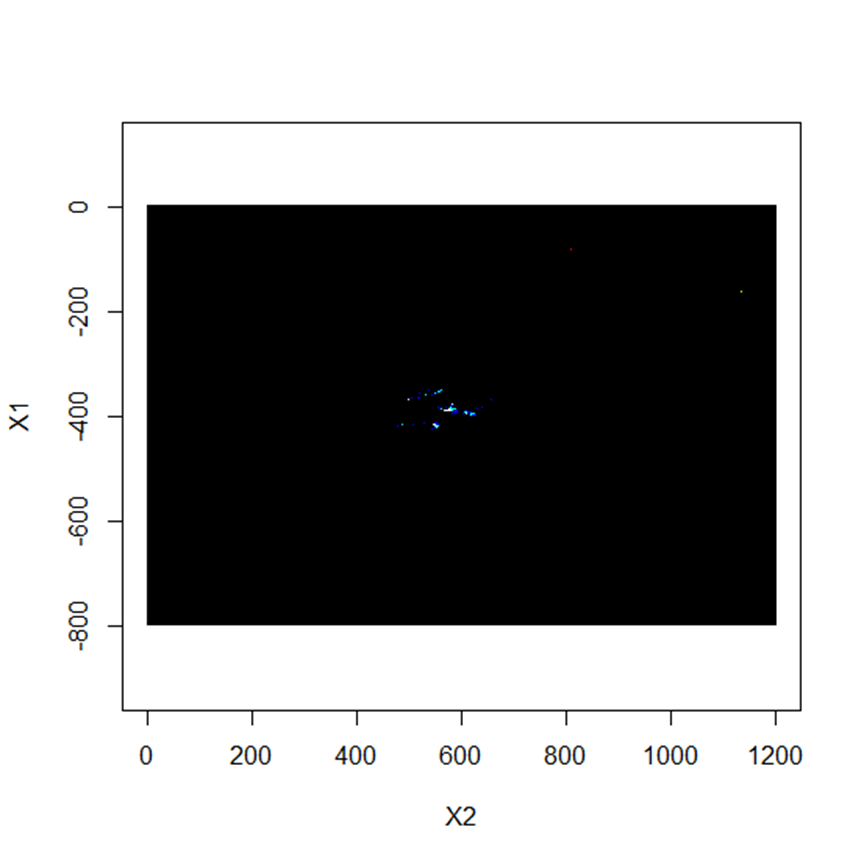
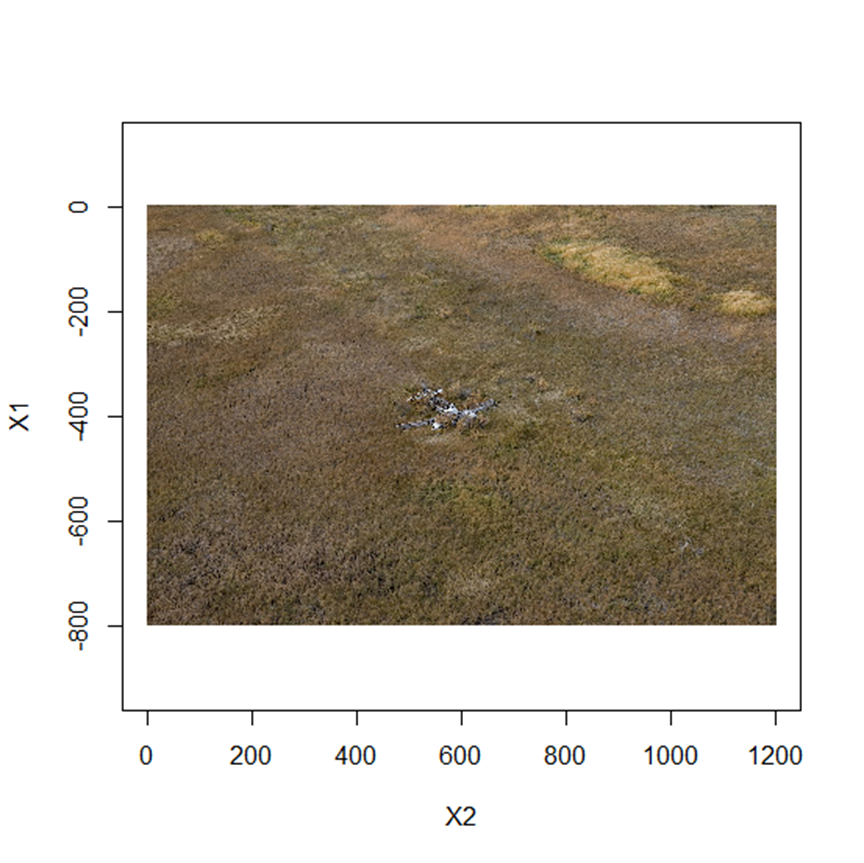
**Results**

The image on the left is the original jpeg after conversion to a matrix. The one on the right is an image showing black for ground cover and blue for the identified object, except for the plot of the P38 on the Greenland Ice Cap. Because ice is more reflective than aluminum, the plot is of the lower z scores and as a result the ground cover is white.

**Alaskan Wilderness**

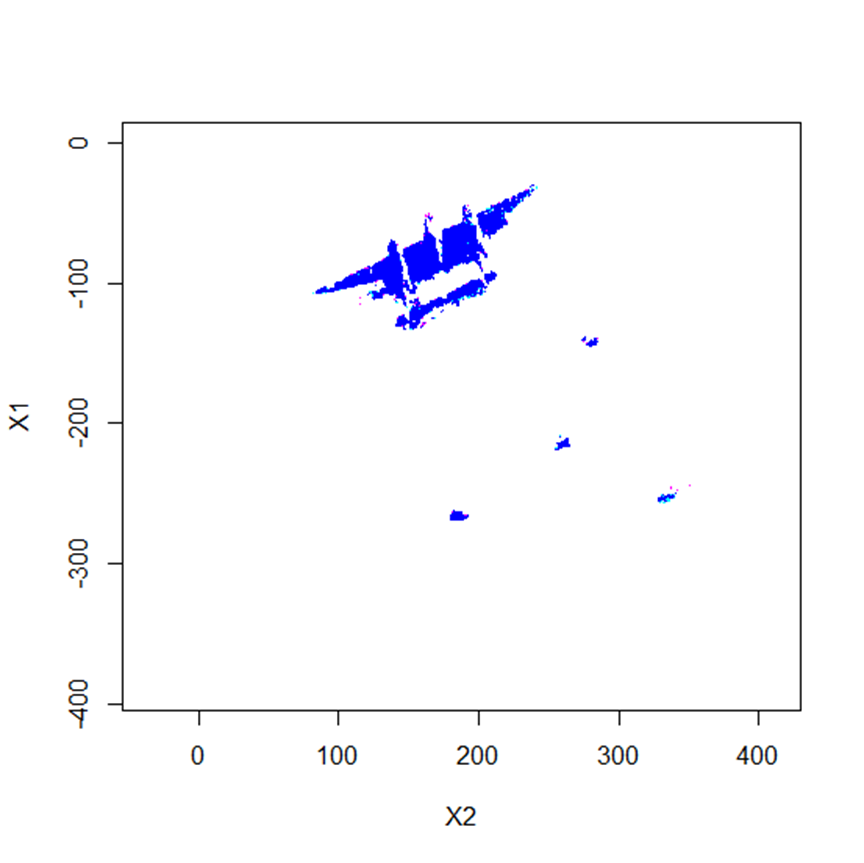
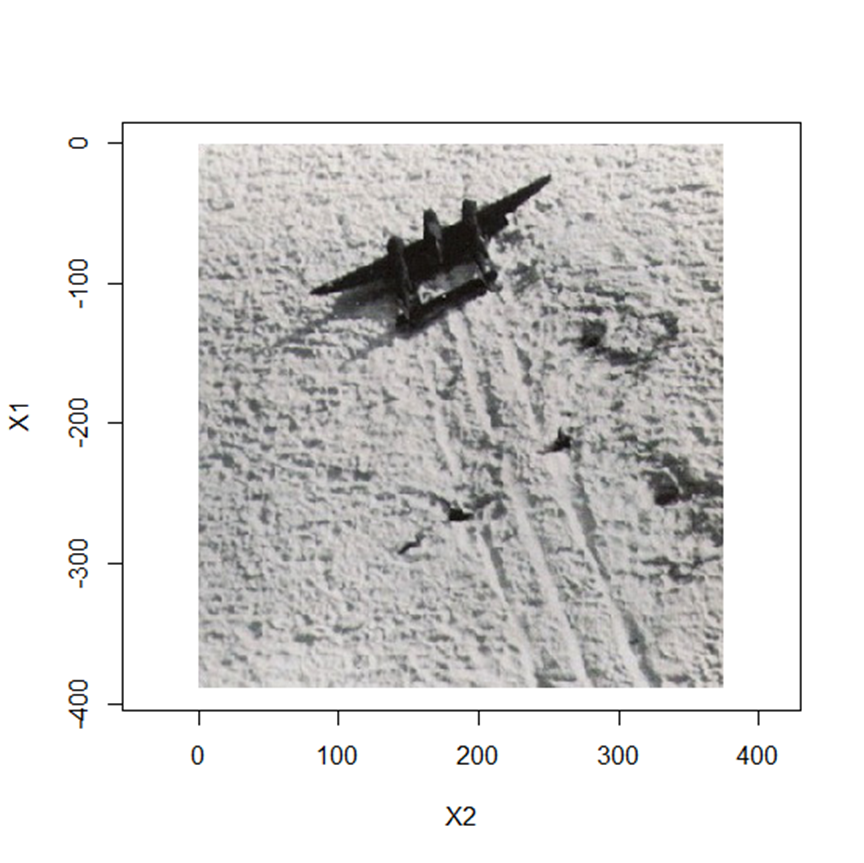


**Alaskan Tundra**

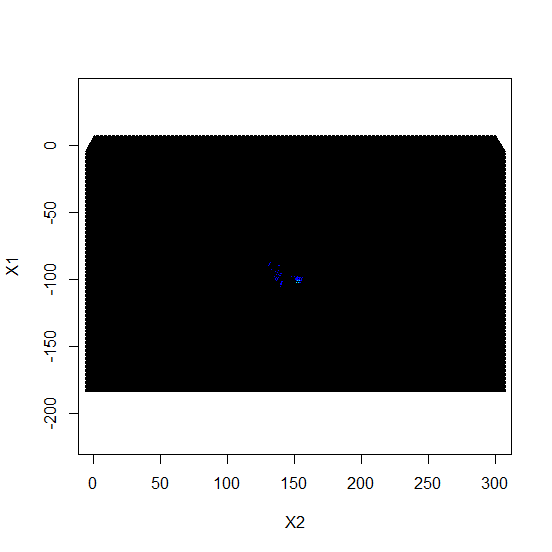
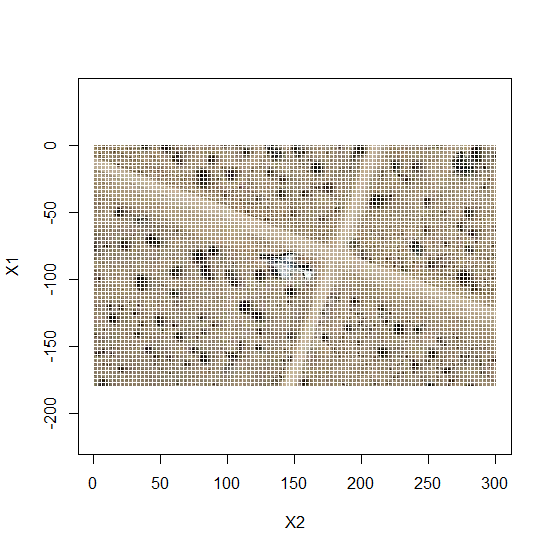
****

**Greenland Ice Cap**

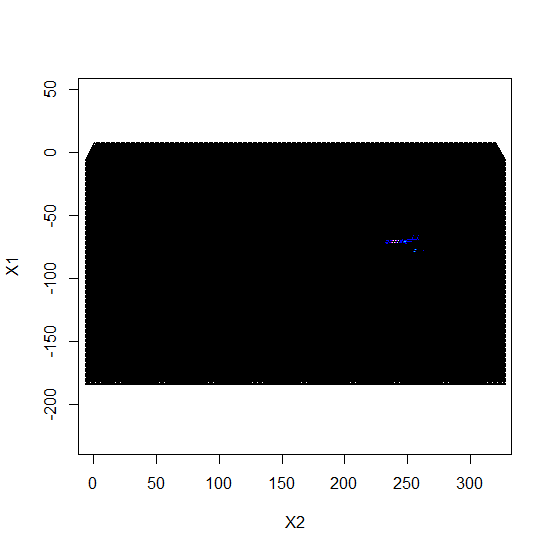
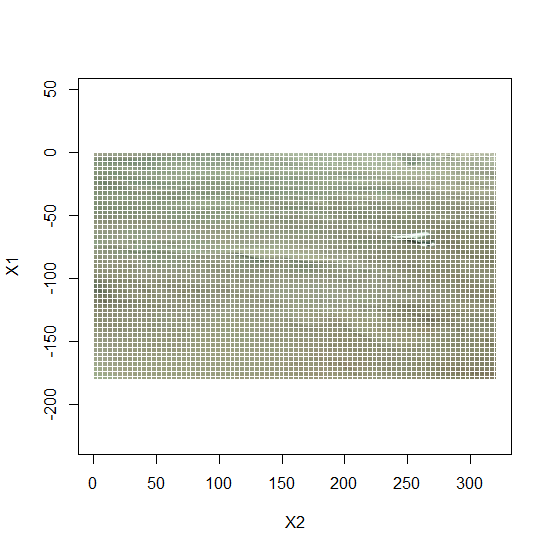
**P38**

****

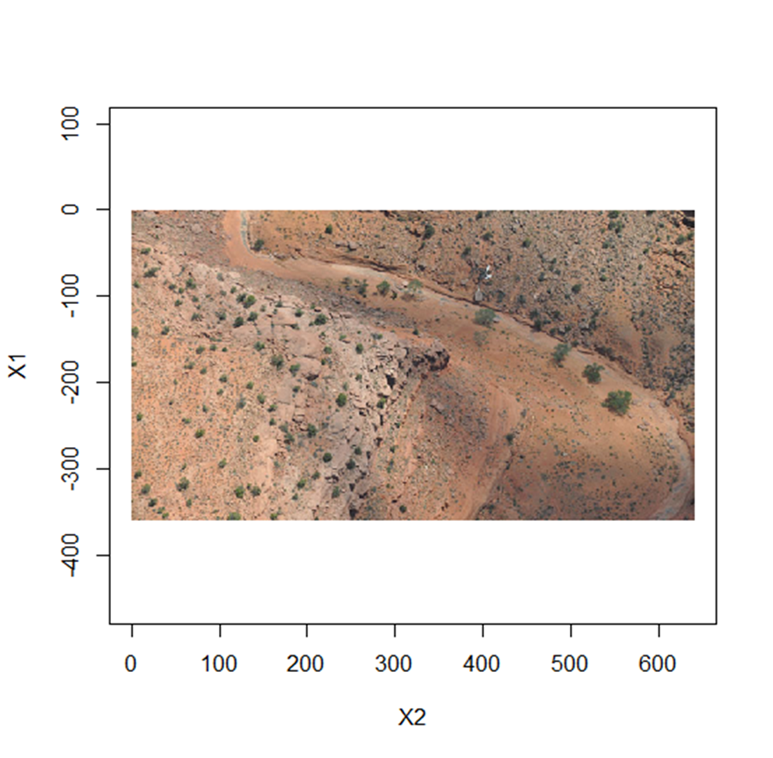
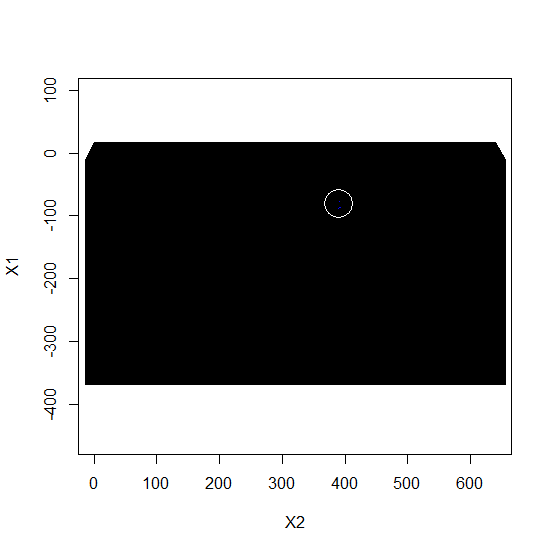
**Nevada Desert**



**Namibia Desert**

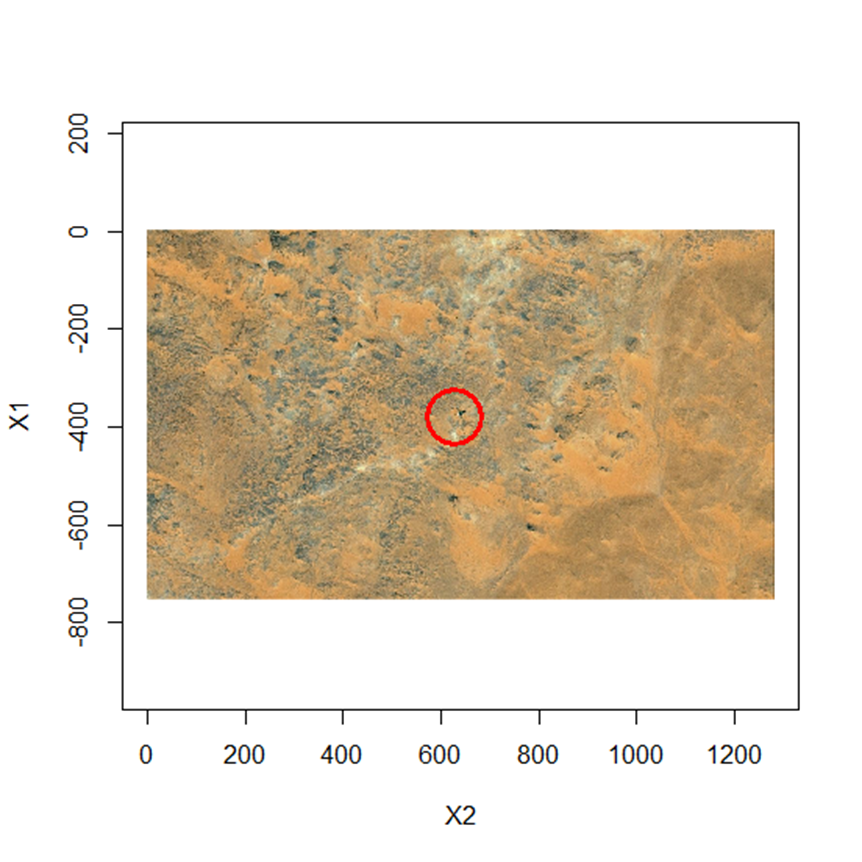


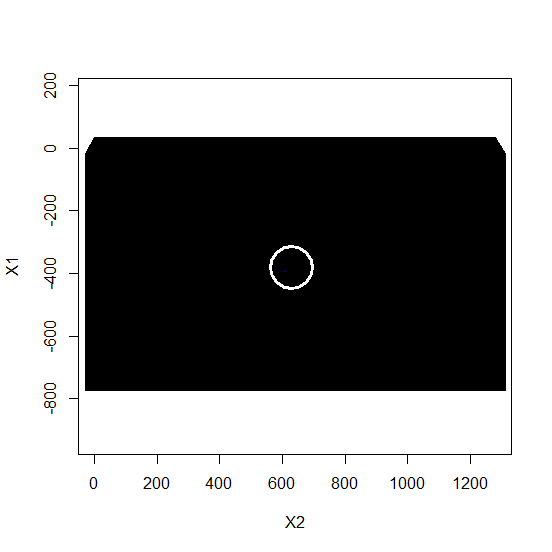
**Mojave Desert**

****

**Sahara Desert**

**Sargent Copping’s P40**

****



In all seven images the aircraft was identified. The two most difficult to classify were the images of the small plane in the Mojave Desert and Sergeant Copping’s P40 in Egypt. In both cases the aircraft are very small in the image, with the Mojave Desert plane being debris from a wreck. In both cases only two small blue dots registered, which I assume are one pixel each. Nevertheless, it identified the wrecks.

**Conclusion**

The method used here is very simple, and it is very possible that a more sophisticated model measuring multiple features, or employing some sort of clustering algorithm might perform better. The method could certainly be improved with the inclusion of LIDAR to assign distance values to objects. However this method works, so the next task is to take advantage of several R packages that allow R to interface with Google Earth, with the ultimate goal being an algorithm that will “crawl” through Landsat images when I am not around.

**Bibliography**

Bork, Edward W., and Jason G. Su. 2007. "Integrating LIDAR data and multispectral imagery for enhanced classification of rangeland vegetation: A meta analysis." *Remote Sensing of Environment* 11-24.

Ciregan, Dan, Ueli Meier, and Jürgen Schmidhuber. 2012. "Multi-column deep neural networks for image classification." *In Computer Vision and Pattern Recognition .* 3642-3649.

Cooper, Tom. 2015. *Wings over Ogaden: The Ethiopian–Somali War, 1978–1979. .* UK: Helion and Company.

Corbley, Kevin. 2013. "Lost WWII fighter in Egypt offers lesson in varied uses of satellite imagery." *The Leading Edge* 186-188.

Corbley, Kevin. 2013. "Lost WWII fighter in Egypt offers lesson in varied uses of satellite imagery." *The Leading Edge* 186-188.

Feng, Min. 2012. "An approach for quickly labeling land cover types for multiple epochs at globally selected locations." *Geoscience and Remote Sensing Symposium* (Geoscience and Remote Sensing Symposium) 6203-6206.

Leslie, Edward E. 1998. *Desperate journeys, abandoned souls: true stories of castaways and other survivors.* Boston: Houghton Mifflin Harcourt.

Lu, Dengsheng, and Qihao Weng. 2007. "A survey of image classification methods and techniques for improving classification performance." *International journal of Remote sensing* 823-870.

Mather, Paul, and Brandt Tso. 2009. *Classification methods for remotely sensed data. : , 2009.* Boca Raton: CRC press.

Pinker, R. T., and A. Karnieli. 1995. "Characteristic spectral reflectance of a semi-arid environment." *International journal of remote sensing* 1341-1363.

Reutebuch, Stephen E., Hans-Erik Andersen, and Robert J. McGaughey. 2005. "Light detection and ranging (LIDAR): an emerging tool for multiple resource inventory." *Journal of Forestry* 286-292.

Riaz, Muhammad, Youngbae, Kim Gwangwon Kang, Pan Sungbum, and Jongan Park. 2008. "Efficient image retrieval using adaptive segmentation of HSV color space." *Computational Sciences and Its Applications.* International Conference on. 491-496.

Romito, Joseph. 2016. "Flights of No Return: Aviation History's Most Infamous One-Way Tickets to Immortality." *Air Power History* 53-54.

Szegedy, Christian, Alexander Toshev, and Dumitru Erhan. 2013. "Deep neural networks for object detection." *Advances in Neural Information Processing Systems* 2553-2561.

Theiler, James P., and Galen Gisler. 1997. "Contiguity-enhanced k-means clustering algorithm for unsupervised multispectral image segmentation." In *Optical Science, Engineering and Instrumentation*, 108-118. International Society for Optics and Photonics.

Wojek, Christian, and Bernt Schiele. 2005. "A performance evaluation of single and multi-feature people detection." *Pattern Recognition* 82-91.

1. https://www.bonhams.com/auctions/17257/lot/121A/ [↑](#footnote-ref-1)
2. http://www.platinumfighters.com/p-51d-n38227 [↑](#footnote-ref-2)