Object Based Remote Sensing Using Sentinel Data

Connor McLaughlin[1], Alan Woodley[12], Shlomo Geva[1], Timothy Chappell[1], Wayne Kelly[1], Wageeh Boles[1], Lance De Vine[1], Holly Hutson[1]

1 Science and Engeering Faulty, Queensland University of Technology
2 ARC Centre of Excellence for Mathematical & Statistical Frontiers
Brisbane, Australia
{c.mclaughlin, a.woodley, s.geva, timothy.chappell, w.kelly, w.boles, dimitri.perrin, l.devine, h.hutson}@qut.edu.au

Abstract— **Identifying changes on the Earth's surface is one of the most fundamental aspects of Earth observation from satellite images. Historically, the predominant form of analysis has measured change at a pixel level. Here, we present a new strategy that conducts the analysis based on objects. The objects are placed inside a random forest regressor. We have tested our approach in Queensland, Australia using Sentinel data. We find that the use of object-based approach either outperforms or is comparable to alternative approaches.**

Keywords—remote sensing, land change, object-based image analysis

I. INTRODUCTION

Identifying land change is an important application to remote sensing experts. Effective land change acquisition is important to environmental scientists, urban planners, and ecologists. Historically, the Queensland government have used their own strategy to identify land change. Despite the strength of this approach it has significant weaknesses such as only looking at individual pixels as well as investigating change only on an annual basis.

Here a different approach is described. First, we use object-based image analysis to identify likely areas of change. Second, we use a random forest regressor to quantify the area as change. Finally, we produce a rank of all the changed areas; a different evaluation approach to traditional approaches which use a set retrieval approach. This approach either outperforms or is comparable to these alternative approaches.

We begin by reviewing some of the literature. We then describe the approach in detail, describing how we are able to segment and classify each image. Finally, we describe the results of our experiment and compare it to the standard approach.

II. LITERATURE REVIEW

In remote sensing there are two ways to identify changes. First are pixel based approaches which operate on the reflectance values of a single pixel [1, 2]. In contrast, object based approaches focus on the identification of a homogeneous set of pixels [3, 4]. The advantage of object-based approaches is that they are less sensitive to outliers, such as registration errors, and focus on the entire object.

The success of an object-based approach requires accurate segmentation. This has been an active area of research, and a good summary is provided in [5]. One of the most effective methods is to start with seed points and expand them.

In terms of image acquisition there are two main alternatives. The first is from government satellites such as Landsat or Sentinel. NASA's Landsat has operated since the 1970s and is able to collect pixels in a 30m by 30m area. Sentinel, operated by the European Space Agency, is able to capture 10m by 10m pixels. An alternative is to use a private company such as planet which capture 3m by 3m pixels.

There are many land cover change detection studies specific to the Australian context. For example: the derivation of foliage projective cover (FPC) was derived from Landsat-5 TM and Landsat-7 ETM+ for the state of Queensland [6], the identification of heterogeneous patterns of fragmented forest and open woodland was mapped in the state of New South Wales using SPOT5 data [7] and the mapping of woody vegetation cover for the whole of Australia [8].

Detecting land change is important for environmental managers, since it allows them to estimate biodiversity, carbon dioxide sequestration, and impacts of natural events such as fire and floods [9, 10]. Often, one of the main sources of identifying land change is via detecting vegetative clearing. In Queensland, vegetation clearing is managed via a process which includes the Statewide Land and Trees Survey (SLATS), a survey which has been carried out over the last 20 years [11]. It uses an analytical workflow similar to other operational approaches and consists of the following stages:

1. For each area in Queensland, Foliage Project Cover (FPC) [6] is applied to two Landsat (now Sentinal-2) images captured approximately 12 months apart.
2. For each image pair, changes in FPC are calculated and compared with historical trends. Changes over a certain threshold are marked as candidate areas for further investigation.
3. Candidate areas are inspected on desktop software by an expert to determine if they are likely changes.
4. Likely changes are manually inspected by experts in the field.

SLATS is a highly recognized approach, however, it is labour intensive since it takes 12 months and requires analysis by experts. It also suffers from only registering pixels, therefore leading to isolated pixels being classified incorrectly. Here, we identify objects and rank them according to their level of change, allowing a user to search the tiles from top to bottom. This approach has previously worked on Planet tiles and now we apply it to Sentinel tiles [12].

III. METHOD

Here we present our overall approach for identifying change. It consists of 9 steps, presented in Figure 1, including segmentation, object-based analysis and random forest regression.
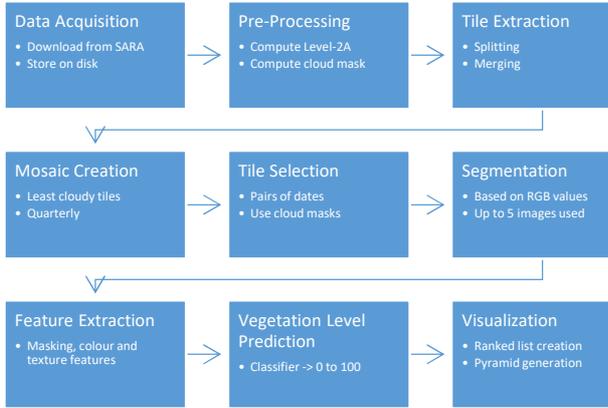

Fig. 1. The change detection pipeline described in this paper

A. Data Acquisition

Sentinel scenes were downloaded from the Sentinel Australasia Regional Access (SARA) service [13]. Where possible L2A products were downloaded, otherwise L1C products were downloaded. Level 2A products represent bottom-of-atmosphere reflectance applied via atmospheric correction and scene classification/correction. As shown in Figure 2, using the Level 2A product provides a better image for visualization and change detection purposes by increasing the contrast. Hence this improves the segmentation.

Fig. 2. Comparison of Level 1C and Level 2A images

B. Data Pre-Processing

Cloud obstruction was a significant issue. The Sentinel products do have a cloud mask included, however, we found that this mask missed significant cloudy areas, as well misclassifying some urban areas as cloud. As an alternative, we evaluated Fmask [14], an algorithm used for automatic detection of clouds, cloud shadows, snow, and water.

Cloud masks are generated from the L1C products at the same time that the L2A products are computed. Pixels containing either cloud or cloud shadows were excluded, as well as snow/ice. However, given the area of interest is the state of Queensland, it is unlikely that any snow/ice pixels were present.

Discarding the entire scene if clouds are detected would exclude many images, delaying detections until a completely clear scene for an area was obtained. As part of the Title Extraction process, we also split the cloud mask into tiles and exclude tiles based on the cloudy pixels, rather than the entire image. Figure 3 shows an example of these types of images as one of the images contains a few clouds. Discarding tiles enables us to continue to use the image for change detection purposes, while still avoiding mispredictions due to cloud or cover.
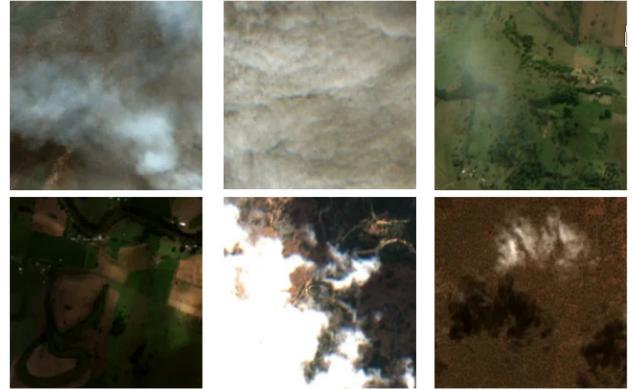

Fig. 3. Various forms of clouds in Sentinel scenes

While Fmask detects a large percentage of the clouds found in the Sentinel scenes, some clouds are still missed. As the vegetation level is often significantly changed by the presence of clouds, these tiles tend to show up higher in the rankings. However, we do not believe this to be a significant impediment to using the system, as most cloudy tiles are detected and those remaining can be discarded very quickly by a human examiner.

C. Tile Extraction

Sentinel scenes capture a large geographical area, where sizes vary from image to image. This presents a challenge for our change detection system, which assumes that the image being processed is square, and both before and after images cover the same area. Our solution to this problem is to take all the Sentinel scenes for a given day and extract or split the images into tiles each covering approximately 5km x 5km. As the images may not be perfectly square, multiple images from the same time period are merged, to ensure the lowest possible number of "no-data" pixels. "No-data" pixels are set to zero reflectance and are used as placeholders where no scene imagery covers the area.

D. Mosaic Creation

The Sentinel satellites do not cover the entirety of our search area every day, leading to "holes" in the images, or areas of "no-data", as seen in Figure 4. This can lead to reduced contextual information around the edges of these areas in our visualization tool, as seen below. Therefore, we compute and display our own mosaic. We generate 4 mosaics for each year of images, one for each quarter, selecting the date with the lowest amount of detected cloudy pixels for each tile. This mosaic is then displayed with a reduced opacity, as seen above, making it clear that those pixels should be considered "out of date", and could be up to or behind by 3 months, and should not be directly relied upon for determining change.

Fig. 4. Extraction of tiles from sentinel scenes

E. Tile Selection

For each date/tile combination, the number of pixels with identified clouds is compared against a configurable percentage threshold. In our testing, we required that for a tile to be considered for change detection, it cannot contain more than 30% cloud pixels.

For change detection, three options are available for selecting images to compare:

1) Day-to-day images: Change is computed on image pairs, for each date where a tile is present, the available dates are searched in reverse chronological order until a date is found where the number is cloudy pixels is below the threshold. This method provides the quickest turn-around for change detection results, however, detection ranking and accuracy is reduced as the amount of change between images is smaller.

2) Same-date previous-year: Change is computed using an image from at least N days in the past. For example, a setting of 365 days, and a date of August 8, 2018 would result in a previous image not after August 8, 2017.

3) Yearly change detection using reference dates: Change is computed against a per-tile reference image. This reference may be a known good date where no clouds are present. Detection can be performed against the reference dates from two years for a yearly change detection, or only using the reference as a "previous date" with new daily scenes as the "after" source.

F. Feature Selection

The image was segmented and for each segment, colour and texture features are computed, which forms the input to the vegetation level regressor. The process for computing these features is described below:

1) Texture Features: To compute texture features for each segment, a bounding box is computed, and all pixels contained are converted to greyscale using (1)

$$G = 0.216 . R + 0.7152 . G + .0722 . B \qquad (1)$$

Where:
G is the pixel's greyscale value
R is the value of the pixel's red value
G is the value of the pixel's green value
B is the value of the pixel's blue value

Each pixel value is then quantised to a number of grey-levels (default 8), and a GLCM (grey-level co-occurrence matrix) is computed [15]. The GLCM is a NxN matrix, where N is the number of grey-levels, with each element *i,j* representing the number of times grey-level *j* was a neighbour within the distance (default 3) of grey-level *i*. Masked pixels were not included in the GLCM computation process. Finally, the GLCM is normalized by computing the sum of all elements in the matrix and dividing each element by the sum.

For each segment, seven summary texture features (contrast, dissimilarity, homogeneity, ASM, IDM, energy and entropy) are computed, based on the expressions given in the scikit-image documentation. The values for these features are passed as-is into the classifier with no additional post-processing.

2) Vegetation Indicies: For each pixel, a series of vegetation indices are computed from the red, green, and blue bands. The 16-bit RGB values are first normalized to 0-1 from a range 0-2500, saturating over-bright pixels to 1. These indices are based on expressions by David and Ballado [16].

From the near-infrared band (Sentinel Band 8), two additional indices: the Normalized Difference Vegetation Index (NDVI) [17], and Enhanced Vegetation Index (EVI) [18] are computed. The band values are normalized to 0-1 using the quantification value from the Sentinel scene (usually 10000).

From the short-wave infra-red band (Sentinel Band 11), an additional index, the Normalized Difference Water Index (NDWI/NDII) is computed. As with the NIR band, this is also normalized to the quantification value.

As these features are per-pixel and we require feature vectors for each segment, summary values for each of these features are found by computing the mean and standard deviation across all pixels. Masked pixels are not considered in this calculation.

3) Colour Features: Finally, for each pixel a series of colour features are extracted. These features are computed by converting the pixel's RGB value to LAB and HSV space. The intensity in each band is also used as a feature (red, green, blue, near-infra-red, and short-wave infra-red). As with the vegetation indices, summary values are found by computing the mean and standard deviation of each colour space.

G. Classifier Training

The vegetation level predictor/regressor was trained on a set of 236 labelled polygons. Each of these polygons was given a vegetation level between 0 and 100. These labels were distributed across all of Queensland. To train the regressor the following process was used:

1. A list of overlapping tiles was produced for the entire set of labelled polygons.

2. For each date where the tile is present, a segment was created using the same parameters as the change detection process.

3. Features are extracted for the tile in the same manner as the change detection process.

4. Each segment was tested against each labelled polygon. If a segment intersects with a labelled polygon, and the label covers a configurable threshold of the segment (90%), the segment is labelled with the polygon's vegetation level, and added to the training set. If multiple labelled polygons intersect the segment, the label with the greatest coverage area was used.

5. Segments without a label are discarded.

Across our 236 labelled polygons, this resulted in a training set of approximately 100,000 labelled segments. To evaluate the initial performance of the vegetation level predictor, an Extremely Randomized Trees (Extra-Trees) regressor was used, with the number of estimators/trees set to 100. To optimize feature pre-processing, selection and hyper-parameters, the training set was passed to the TPOT AutoML tool, and optimized using 100 generations and a population size of 50. K-Fold cross-validation was used to create test/train sets, with 5 splits, using negative mean squared error as the metric.

H. Vegetation Level Prediction

The vegetation label prediction was similar to previous work [19]. The following process is used to select the segmentation images for a before/after date pair:

- An arbitrary maximum number of images is chosen for segmentation (N). In our testing, we used 5 segmentation images.
- Up to N images are found for both the before/after dates, where the tile is present in the date, and the number of cloudy pixels is below the threshold.
- Any "holes" or "no-data" pixels in the tile are replaced with pixels from the closest date.
- The "after" date is added to the segmentation image set.
- Alternating between the before/after images to ensure the segmentation is not biased in either direction, images are added until the total number equals N.

In some cases, there may be insufficient non-cloudy images to satisfy the maximum segmentation image constraint. In these cases, only the images which are found are used. At a minimum, this will always be at least two: the "before" image and the "after" image.

After the segmentation was computed, a feature vector is extracted for each segment in both the "before" and "after" images. The trained regressor is used to predict a vegetation level for the segment. Each of these vegetation levels is saved, along with the segmentation for ranked list creation. A sample of the segmentation and predicted vegetation levels can be seen in Figure 5.

Fig. 5. Segmentation and predicted vegetation levels

I. Ranked List Creation

After all segments in a time point have been classified, a ranked list of tiles is generated to enable examiners to focus on the areas with the most significant change. A list is precomputed for each change magnitude (from 0 to 100) by examining the vegetation levels for each segment. The predicate used to filter these levels is shown in (2):

$$C = (V_{Before} - V_{After}) > T \qquad (2)$$

Where:
  C is the level of change
  V_{Before} is the previous vegetation level
  V_{After} is the current vegetation level
  T is the threshold (0-100)

Therefore, only negative change (higher vegetation transitioning to lower vegetation) will be considered. If positive change is desired, it can be found by swapping the greater than operator for less than. In addition, for each tile, the number of changed segments is computed and stored. If a segment contains one or more masked pixels in either segment, it is not included in the change count regardless of the vegetation levels; however, we still track it for statistical purposes.

For each tile, we compute a "weighted change" value, which is used as the basis for ordering the ranked list. The total area (in meters) of change is also tracked but is not currently used in the list rankings. The "weighted change" is shown in (3) to (5):

$$M = \frac{\sum_{i=1}^{N} C_i}{S} \qquad (3)$$

Where:
  M is the mean amount of change per tile
  N is the number of change segments in the tile
  C_i is the amounts if change for a given tile
  S is the number if changed segments

$$P = \frac{C}{T} \times 100 \qquad (4)$$

Where:
    P is the percentage of the change area
    C is the size of the changed area
    T is the total area

$$W = P \times \left(\frac{M}{100}\right) \qquad (5)$$

Where:
    W is the weighted change
    P is the percentage of the change area
    M is the mean amount of change per tile

IV. RESULTS

A. Comparision with SLATS 17/18

To evaluate the performance of the Sentinel-based change detection system, we used the public SLATS report/shapefile from 2017/2018 as a baseline [20]. SLATS Landsat image date footprints were used to determine the dates of the scenes used for both the "before" (2017) and "after" (2018) images. However, as Landsat and Sentinel are produced from different satellites, it was not always possible to use the exact same date as SLATS. In this case, for the "before" image, we selected a Sentinel scene no later than the Landsat date from 2017, with a maximum of 25% cloud cover. The "after" image was selected from a date no earlier than the Landsat date from 2018, with the same cloud cover requirement. Cloud percentage was computed by dividing the total number of pixels by the number of pixels with any type of cloud/shadow detected. In most cases, the images we used were no more than a week before/after the Landsat scenes. For these results, all bands supported by the change detection system were used (red, green, blue, near-infra-red, short-wave infra-red).

We converted the SLATS polygons to a list of tiles. Any tile with pixels intersecting a SLATS polygon was included. This resulted in approximately 12,000 tiles in the "ground truth" or "relevant document" set, from a total of approximately 120,000 tiles. Recall measures which fraction of results were returned compared to the relevant document list. In this case, the relevant document list is the tiles which contain one or more polygons present in the SLATS dataset. Figure 6 shows that at thresholds up to approximately 20, all relevant tiles are returned. However, as the vegetation predictions tend to vary between images even when no change is present, thresholds lower than 10 should not be considered accurate as this may just be noise showing up as change. The recall drops as the change threshold increases, as not all change in the SLATS dataset is "significant". At thresholds closer to 100, almost no relevant tiles are returned, as this would require a vegetation level in the high 90s to reduce to 0, and we found that even dense forest areas were predicted to be between 90 and 100.

Precision @ 1000 measures how many of the first 1000 results are present in the relevant document (or tile) list. As the time required to determine whether to exclude an image is a false positive is minimal, we believe it is not unrealistic to expect a compliance officer to examine 1000 documents in a short time; for example, a week. Figure 7 shows that our

precision at thresholds is around 27% at the lowest thresholds, gradually increasing as the threshold is also increased, up to a peak of near 40%. At higher thresholds, while the precision improves, this is a side-effect of having fewer tiles returned overall due to excluding less significant change.
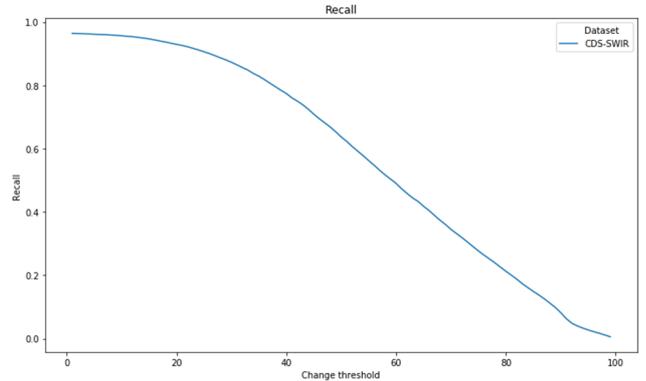

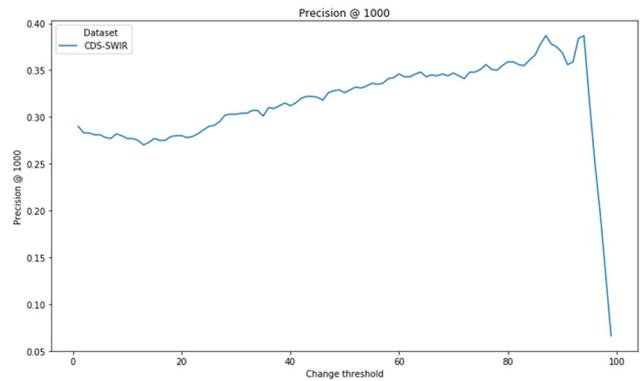Fig. 6. Recall across thresholds compared to SLATS 17/18

Fig. 7. Precision @ 1000 across thresholds compared to SLATS 17/18

As we were unable to get information on the false positive rates of the SLATS system – only an approximate number of 90% – it is difficult for us to clearly state how well our system performs in comparison. If the quoted number of approximately 90% is correct, this would suggest that our system outperforms SLATS in precision. However, we are also only looking at two points in time across a year. We expect examining daily images would result in greater noise and additional false positives due the short time frame between images.

Overall precision gives an indication of how many false positives are in the entire results list. As seen in Figure 8, excluding the lowest 10 thresholds due to noise, our true positive rate is around 17%, peaking at around 45% at the highest thresholds. This translates into approximately ~38,000 tiles at a threshold of 40. However, due to us ranking the results based on the magnitude and size of the change, an examiner can spend their time more efficiently as the entire list does not need to be examined due to the greatest areas of interest being sorted to the top of the list.
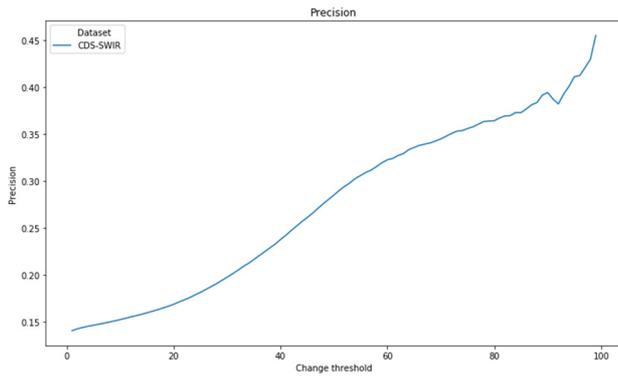
Fig. 8. Overall precision across thresholds compared to SLATS 17/18.

Fig. 10. Precision @ 1000 against SLATS 17/18 with band restriction applied.

B. Impact of Additional Bands

In addition to overall results, we also examined the impact of restricting the change detection system to only RGB and RGB+NIR bands. The recall and precision @ 1000 can be found in Figure 9. The usage of additional bands did not significantly affect recall at the lower thresholds, with a measurable difference of around 5-10% at higher thresholds. The precision for the first 1000 results was more significantly affected by the additional bands, as seen in Figure 10 with 20% for RGB only vs 30% for NIR/SWIR bands at a threshold of 20. This would suggest that using only the RGB bands produces a greater number of false positives, thereby showing that the accuracy of the level predictor is highly dependent on the spectral image data.
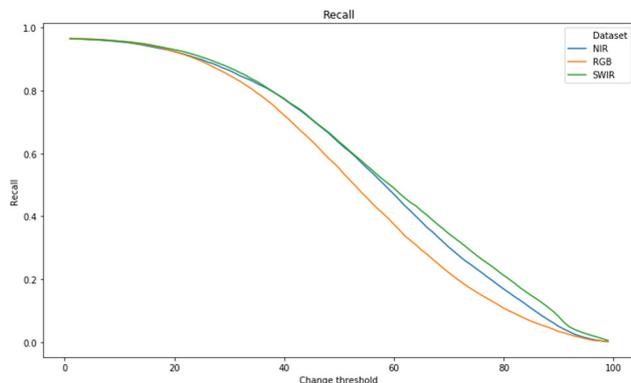

Fig. 9. Recall against SLATS 17/18 with band restriction applied.

V. CONCLUSION

Here, we have described our system, which uses object-based image analysis to find land change. It uses a random forest regression to classify images that identify vegetation land change. Our results show that the results are comparable to other approaches and in some cases outperform state of the art approaches.
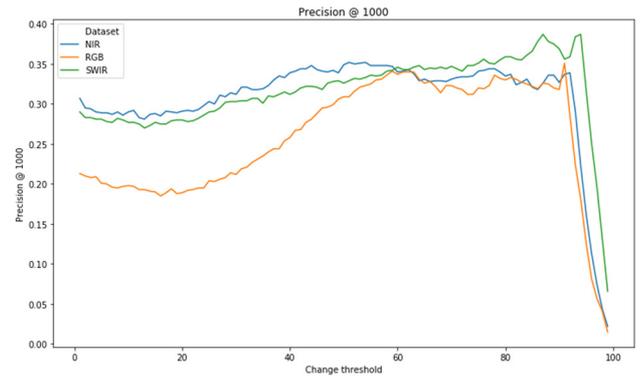
REFERENCES

[1] M. D. Hussain, A. Chen, H. W. Cheng, and D. Stanley, "Change detection from remotely sensed images: From pixelbased to object-based approaches," *ISPRS Journal of Photogrammetry and Remote Sensing* vol. 80, pp. 91-106, 2013.

[2] A. Woodley, L.-X. Tang, S. Geva, R. Nayak, and T. Chappell, "Parallel K-Tree: A multicore, multinode solution to extreme clustering," *Future Generation Computer Systems,* vol. 99, pp. 333-345, 2019.

[3] M. Lu, J. Chen, H. Tang, Y. Rao, P. Yang, and W. Wu, "Land cover change detection by integrating object-based data blending model of Landsat and MODIS," *Remote Sensing of Environment* vol. 184, pp. 374-386, 2016.

[4] B. Desclee, P. Bogaert, and P. Defourny, "Forest change detection by statistical object-based method," *Remote Sensing of Environment* vol. 102, pp. 1-11, 2006.

[5] S. R. Vantaram and E. Saber, "Survey of contemporary trends in color image segmentation," *Journal of Electronic Imaging,* vol. 21, 2012.

[6] J. D. Armston, R. J. Denham, T. J. Danaher, P. F. Scarth, and T. Moffiet, "Prediction and validation of foliage projective cover from Landsat-5 TM and Landsat-7 ETM+ imagery for Queensland, Australia," *Journal of Applied Remote Sensing,* vol. 3, 2008.

[7] A. Fisher, M. Day, T. Gill, A. Roff, T. Danaher, and N. Flood, "Large area, high-resolution tree cover mapping with multi-temporal SPOT5 imagery, New South Wales, Australia " *Remote Sensing,* vol. 8, 2016.

[8] T. Gill, K. Johansen, S. Phinn, R. Trevithick, P. Scarth, and J. Armston, "A method for mapping Australian woody vegetation cover by linking continental-scale field data and long-term Landsat time series," *International Journal of Remote Sensing* vol. 38, pp. 679-705, 2017.

[9] C. Sarker, L. Mejias Alvarez, F. Maire, and A. Woodley, "Flood mapping with convolutional neural networks using spatio-contextual pixel information," *Remote Sensing,,* vol. 11, pp. 23-31, 2019.

[10] R. I. Jony, A. Woodley, A. Raj, and D. Perrin, "Ensemble classification technique for water detection in satellite images," in *2018 Digital Image Computing: Techniques and Applications (DICTA)*, Canberra, Australia, 2018, pp. 1-8.

[11] Remote Sensing Centre, "Statewide Landcover and Trees Study: Overview of Methods," Department of Environment and Science, The State of Queensland, Brisbane, Queensland,2018.

[12] A. Woodley, C. McLaughlin, H. Hutson, S. Geva, T. Chappell, W. Kelly*, et al.*, "High resolution change detection using Planet mosaic," in *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, 2019, pp. 6578-6581.

[13] Australian Regional Copernicus Data Hub. (2019). *Home - Copernicus Australia*. Available: https://copernicus.nci.org.au/

[14] Z. Zhu, S. Wang, and C. E. Woodcock, "Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images," *Remote Sensing of Environment,* vol. 159, pp. 269-277, 2015.

[15] R. Haralick, "Statistical and Structural Approaches to Texture," *Proceedings of the IEEE* vol. 67, pp. 786-804, 1979.

[16] L. C. G. David and A. H. Ballado, "Vegetation indices and textures in object-based weed detection from UAV imagery," presented at the 2016 6th IEEE International Conference on Control System, Computing and Engineering (ICCSCE), 2016.

[17] R. S. Defries and J. R. G. Townshend, "NDVI-derived land cover classifications at a global scale," *International Journal of Remote Sensing,* vol. 15, pp. 3567-3586, 1994.

[18] Z. Jiang, A. R. Huete, K. Didan, and T. Miura, "Development of a two-band enhanced vegetation index without a blue band," *Remote Sensing of Environment,* vol. 112, pp. 3833-3845, 2008.

[19] C. McLaughlin, H. Hutson, L. De Vine, A. Woodley, S. Geva, T. Chappell*, et al.*, "Change detection over the state of Queensland using high resolution Planet satellite mosaics," in *2019 Digital Image Computing: Techniques and Applications (DICTA)*, 2019, pp. 1-8.

[20] Queensland Government, "Land cover change in Queensland, Statewide Landcover and Trees Study Summary Report: 2016–17 and 2017–18," Brisbane, Australia 2019.