PAPER • **OPEN ACCESS**

Automated image analysis and improvisations to manage palm oil plantation

To cite this article: M N Akhtar *et al* 2020 *IOP Conf. Ser.: Mater. Sci. Eng.* **1007** 012082

IOP Conf. Series: Materials Science and Engineering 1007 (2020) 012082 doi:10.1088/1757-899X/1007/1/012082

Automated image analysis and improvisations to manage palm oil plantation

M N Akhtar[1], S. A. Khan[2], Mazlan Mohamed[3], Ayub Ahmed Janvekar[4, *]
[1]School of Aerospace Engineering, Universiti Sains Malaysia, 14300 Nibong Tebal, Penang, Malaysia
[2]Department of Mechanical Engineering, Faculty of Engineering, International Islamic University, 50728 Kuala Lumpur, Malaysia
[3]Advances Material Research Cluster, Faculty of Bioengineering and Technology, Universiti Malaysia Kelantan, Jeli Campus, 17600 Jeli, Kelantan, Malaysia
[4]School of Mechanical Engineering, VIT-Chennai Campus, Chennai, TN 600127, India

* ayubahmed.janvekar@vit.ac.in

Abstract. Palm oil industry plays an essential role in South-East Asian agricultural commodity sector as it contributes to the substantial gross domestic product of the country. However, with the advent of climate change and massive deforestation, the disease and malfunctioning in growth of palm tree has increased. Therefore, it has become essential to detect any form of disease in palm oil plantation which can hamper its productivity as it can cause a serious problem to the countries whose economic conditions are primarily dependent upon palm oil plantations. Hence, early detection of disease from the initial stage is crucial to the production of palm oil. In this regard, the proposed manuscript highlights the importance of image processing in detecting early disease in palm oil plantation using image segmentation and also proposes some improvisations in palm oil plantation which will be helpful in managing the palm oil commodity business.

1. Introduction
Some of the common occurring disease in plants are Bacterial disease leaf spots, leaf scorch and leaf blight with tattering[1]. Leaf scorch (also called leaf burn, leaf wilt, and sun scorch) results in browning of plant tissues, leaf margins and tips, and yellowing or darkening of veins which may lead to eventual wilting and abscission of the leaf [2]. Blight is a rapid and complete chlorosis which results in death of plant tissues such as leaves, branches, twigs, or floral organs[3]. Similar to other crops, oil palm also gets infected by numerous pathogens i.e., fungi which ultimately results in reduction of yield [4]. Palm oil producing countries have reported numerous diseases for which the most common one is Ganoderma. The basal stem rot or Ganoderma butt rot disease results in leaf blight and is caused by by Ganodermaboninense, is the most severe oil palm disease in South-East Asia[5]. Aluminium deficiencies in oil palms are acute in Sumatra [6]. Smith et al.[7] recorded over forty ailments of oil palms. They identified that the ailments are caused by effects from the various growth phases: Germination (with brownish germ ailments) and toddler stage growth. It is worth to be noted that seed decay and brown germ are acute in South-East Asia [6] which is likely due to potassium, phosphorus, boron and calcium deficiency [8]. Singh and Misra[9] described that essentially there are four steps in image segmentation of a diseased leaf, out of which, first one is used to convert the input leaf image in Red-Green-Blue (RGB) colorspace to Hue-Saturation-Intensity (HSI) colorspace. In the second step, a value for thresholding was set and green pixels were removed using the masking technique. For the third step, by using the value of threshold defined in

second step, diseased portion of the leaf was identified and extracted. Lastly in the fourth step, the segmentation was done.

Arya et al. [10] used image processing and genetic algorithm with adruino to detect diseased leaf in plants. For the diseased leaf the histogram matching was done detecting the edge. Sahoo et al. [11]used automated dead zone detection in 2D leaf's image. As per their work, segmentation of lesion-based region was carried out by defining a particular threshold value and in their final step, several diseases were categorized by computing the quotient of leaf area and lesion area. Bai et al. [12] and Ma et al. [13] came up with a leaf spot detection algorithm implemented using neighbourhood grayscale information. In their papers, process of diseased portion detection was done by comparing the effect of HSI, CIELAB, and YCbCr colour space.

Segmentation is an essential process of object recognition, image compression, image database look-up and occlusion boundary estimation within stereo or motion system. The researchers these days are dealing with the problem of over segmentation of images which ultimately leads to inaccurate results and therefore, leaves a room for enhancing this problem [14]. Sharp changes in the intensity of image causes dissimilarity whereas similarity corresponds to the process of combining and matching the pixels with the neighbouring one based on its gray level pixel value match [15]. Some of the widely recognized techniques to implement image segmentation are; Otsu's threshold method for automated image segmentation, region growing and region merging technique, edge detection method, watershed transformation and histogram thresholding based algorithms [16]. Amongst all the techniques, Otsu's method is widely renowned method to carry out the process of image segmentation [17]. Since it is an automated process, therefore, it is easier to be applied on the bulk image data simultaneously. As the proposed research is dealing with image data, therefore, it is appropriate to use Open Computer Vision (OpenCV) library and it is also to be noted that Otsu's threshold technique has high degree compatibility with OpenCV [18]. With respect parallel processing to cut down the execution time, there is an Application Programming Interface (API) called Hadoop Interface for Image Processing (HIPI) which is an extensive set image processing framework and is only compatible with Hadoop Map-Reduce parallel programming model [19]. For the rest of the manuscript, Section 2 elaborates on segmentation model for diseased leaf detection followed by leaf blight analysis of palm tree using image processing algorithm in Section 3. Section 4 mentions some improvisations for the palm oil plantation and finally, Section 5 gives the conclusion of the proposed work.

2. Segmentation model for diseased leaf detection

Thresholding is considered to be an important technique for image segmentation which has got potential to identify and extract the target portion of an image from its actual background on the principal of distribution of gray levels in an image object. According to Otsu's method, an image is considered to be a two-dimensional grayscale intensity function which contains N pixels including gray levels ranging from 1 to L[20]. As per Otsu's analysis, the number of pixels having gray level 'i' is denoted by 'f_i'. Therefore, the probability function (P_i) of gray level 'i' in an image with N pixels could be written as (2)[21].

$$P_i = {}^{f_i}\!/_{N} \tag{1}$$

For the analysis of bi-level thresholding of an image, the pixels could be divided into two classes C_1 and C_2 respectively. C_1consists of first tier of gray level (1........,t) and C_2 consists

IOP Conf. Series: Materials Science and Engineering 1007 (2020) 012082 doi:10.1088/1757-899X/1007/1/012082

of second tier of gray level ($t+1$...........,L). Therefore, the gray level probability distribution for the two classes could be written as (2) and (3)[22].

$$C1 = P_1/\omega_1(t) \ldots \ldots . P_t/\omega_1(t) \quad (2)$$

$$C2 = P_{t+1}/\omega_2(t), P_{t+2}/\omega_2\,(t), \ldots . P_L/\omega_2(t) \quad (3)$$

Where $\omega_1(t) = \sum^{t}_{i=1} P_i$ and $\omega_2(t) = \sum^{L}_{i=t+1} P_i$

Above grey level probability distribution method could also be applied for M number of classes assuming that there are M-1 thresholds, $\{t_1, t_2 \ldots\ldots\ldots\ldots, t_{M-1}\}$ which divide the original image into M classes: C_1 for [1......,t_1], C_2 for [t_1+1........,t_2].......,C_i for [$t_{i-1}+1$.........,t_i] and C_m for [$t_{M-1}+1$..........,L] [22].Equation (4) represents a column vector.

$$x = \begin{matrix} x_1 \\ \vdots \\ x_n \end{matrix} \quad (4)$$

If the entered values in (4) are random pixel variables with a precise mean, then the segmented matrix [$seg(X_i, X_j)$] value $\sum$ is given by (5) [23].

$$\sum_{ij} = seg(X_i, X_j) = val[(X_i - \sigma_i)(X_j - \sigma_j)] \quad (5)$$

Where $\sigma_i = val(X_i)$ and $\sigma_j = val(X_j)$ are the assumed value of the i_{th} and the j_{th} entry in the vector X.

Now let us assume there are n such images to be segmented and if a single image is denoted by vector x, then the sample computed segmentation could be given by the formula in (6) [23].

$$Seg = \frac{1}{n}\sum_{i}^{n}(x_i - \bar{x})(x_i - \bar{x})^T = \frac{1}{n}\hat{X}\hat{X}^T \quad (6)$$

Where i = index for the set of n images, $\bar{x}$ = average of n image pixels
Equation (6) could also be rewritten in matrix form using $\hat{X}$ to denote the mean centred images $(x_i - \bar{x})$ in (7)

$$\begin{pmatrix} \vdots & & \vdots \\ \hat{x}_i & \ldots & \hat{x}_n \\ \vdots & & \vdots \end{pmatrix} * \begin{pmatrix} \cdots & \hat{x}_1 & \cdots \\ \cdots & \vdots & \ldots \\ \ldots & \hat{x}_n & \cdots \end{pmatrix} \quad (7)$$

Let us divide the image patches into v number of pixels based on their similarity. On similarity basis, let us categorize the set of pixels into different clusters i.e., C_1, C_2,.....,C_v.

Now let us define the set group of every unsigned pixel which at least borders one of the clusters as defined in (8) [24].

$$S = \{x \notin \bigcup_{i=1}^{v} C_i \bigwedge \exists k : N(x) \bigcap C_k \neq \emptyset\} \quad (8)$$

Here, x is the pixel to be assigned, where N(x) denotes the current neighbouring pixel of point *x* which is a part of cluster C_k. As per (8),*x* does not lie the cluster C_i and *k* belongs to pixel *x* such that *N(x)* is a part of cluster C_k (Cluster with *k* pixels).
Now let us denoteδas the difference of measurebetween the pixels as defined in (9) [24].

$$\delta(x, C_i) = \left| l(x) - mean_{y \in C_i}[l(y)] \right| \tag{9}$$

Where *l(x)* denotes the pixel value of point *x* and *I* denotes the index of the cluster such that *N(x)* intersect C_i. *l(y)* denotes the pixel value of point *y*.

Now to select whether $q \in S$ and cluster C_j where $j \in [1,n]$ such that:

$$\delta(q, C_j) = \min_{x \in s, k \in (1,n)} \{(x, C_k)\} \tag{10}$$

Where *S* is defined in (8)

Now if $\delta(q, C_j)$ is lesser than the predefined threshold point t_p set by the programmer, the pixel is assigned to cluster C_j, else it must be assigned to another most considerable cluster *C* such that:

$$C = arg \min_{C_k} \{\delta(Z, C_k)\} \tag{11}$$

Now if $\delta(q, C_n) < t_p$, then the pixel is allocated to C_n. If neither of the condition is satisfied, then the formation of new cluster C_{n+1}takes place.
After the pixel has been allocated to the cluster, the mean pixel value of the cluster must be updated. According to Gedraite and Hadad[25], the function which is used to generate the kernel is a Gaussian function comprising of 2 dimensions and could be defined using (12):

$$f(q, r) = a.e^{-\left\{\frac{(q-q_0)^2}{2\delta q^2} + \frac{(r-r_0)^2}{2\delta r^2}\right\}} \tag{12}$$

Where *q* and *r* are the vectors, *a* is the amplitude, (q_0, r_0) is the centre, δq and δr is the standard deviation in *q* and *r* direction. Filter is defined using the variance of the Gaussian distribution. This parameter drastically affects the filtering results. The quality factor (*Q*) function defined for segmented image using Gaussian blur is given by the (13) [25]:

$$Q(s, t) = \frac{\sigma_{s.t}}{\sigma_s{}^2 . \sigma_t{}^2} . 2 . \frac{\bar{s}. \bar{t}}{s^2 + t^2} . \frac{2 . \sigma_s{}^2 . \ \sigma_t{}^2}{\sigma_s{}^2 + \ \sigma_t{}^2} \tag{13}$$

Where, *t* is the image without noise and *s* is the filtered image,$\sigma_{s.t}$ is the covariance between two images, $\sigma_s{}^2$ is the variance of filtered image and $\sigma_t{}^2$ is the variance of source image without noise.$\bar{s}$ and $\bar{t}$ are the mean of images *s* and *t*. This quality factor determines the covariance between two images, the distribution in the contrast and distortion in luminance.

3. Analysis of leaf blight in palm leaf

For the pre-processing of image samples, it is required to convert the acquired input image from its original form to bilateral blur form which acts as a kernel matrix in order to remove the noise from the image. Post noise removal, the transformed bilateral blurred RGB image is converted to HSV colour space format to analyze the diseased regions within the plant leafs. Once the

IOP Conf. Series: Materials Science and Engineering 1007 (2020) 012082 doi:10.1088/1757-899X/1007/1/012082

analysis of various regions is done, then it is required to apply image thresholding to segment the diseased portion area of the entire plant leafs captured in the image in the form of non-zero pixels. Post diseased portion estimation, the contour is drawn in order to verify the boundary of the diseased portion detection. Figure 1 shows the analysis of the diseased palm leafs caused due to leaf blight. The image was taken from 1-meter distance and has the resolution of 4000x3000. For analysing the leaf blight in palm leaf sample, it is suitable to use Hue-Saturation-Value (HSV) colorspace to separate colour components from intensity for various reasons, such as robustness to lighting changes, or removing shadows. It is worth to be noted that each classified diseased leaf will have its own colorspace. Post HSV threshold analysis, area of the diseased portion could be computed using the count of non-zero pixels. The pseudocode of the image segmentation algorithm is shown in Algorithm 1.

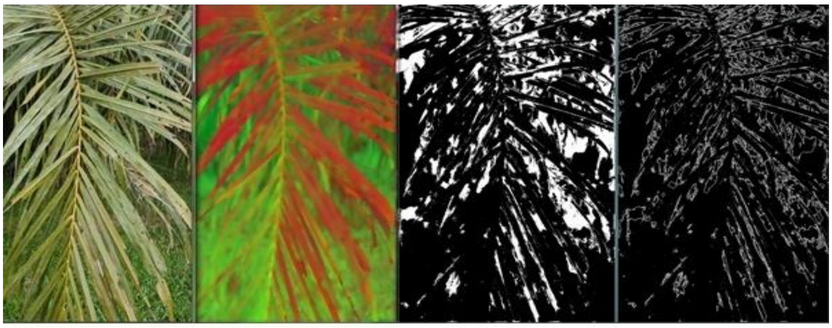

Figure 1. Analysis of leaf blight and tattering in palm tree (original resolution: 4032x3024)

ALGORITHM 1
Requirements: Bundled input images

1. Input image JPG format
2. Resize the image to 500x500 // resize (img, image,size
3. Apply filter (Bilateral Blur) // Bilater
4. Filter (image, dst,Kernel_lenghth,Kernel_lenghth/2,Kernel_length*2)
5. Convert RGB colorspace to HSV colorspace // cvtColor (dst, HSV, CV_BGR2HSV)
6. Apply Threshold adjustments by adjusting the HSV trackbars
7. SegmentThresholded pixels // (0[Black],255[White] Pixels)
8. Draw contours
9. Store segmented pixels // white/black pixels
10. Emit segmented pixels variable //write (new IntWritable(1),newIntWritable (region of interest [ROI] pixels));

Table 1 shows the accuracy detection for leaf tattering in 20 test samples using three different colorspace. The training and the testing sets for tattered palm leafs along with their detection accuracy is shown. From the results it can be seen that the detection accuracy is enhanced by HSV implemented using standardized image segmentation algorithm compared to other approaches. From the results it can be seen that detection accuracy using HSV is 100%. In the second phase and the third phase, classification were done using CIELAB and YcbCr colorspace classifiers for which the accuracy diminished up to 9 % - 10%. The results clearly conclude that for detecting palm leaf blight, image segmentation implemented using HSV colorspace is the most accurate option.

Table 1. Accuracy Detect Using Different Colorspace

Diseased Samples	No. Of images (Train Data)	No. Of images (Test Data)	Detection accuracy of proposed algorithm using different colorspace (%)		
			HSV	CIELAB	YCbCr
Palm	25	20	100% (count of non-zero pixel: 58485)	91%(count of non-zero pixel: 53221)	90% (count of non-zero pixel: 52636)

4. Improvisations required in Palm Oil Industry

The technologies and tools connected with precision farming have drawn the attention of researchers from the oil palm sector in South-East Asia. These resources and technology give a chance to comprehend and capitalise on the variabilities in the areas that have been recognised from the planters, however little emphasis has been applied until today. The following sub-sections gives a brief overview of some of the deemed necessary steps which should be taken by the Malaysian Pam Oil Board (MPOB) to increase the production of palm oil in the country.

4.1 Yield Monitoring Using UAV

Quantification of FFB from UAV stream pictures for return map creation is the first step in practicing precision agriculture in oil palm plantations. One of the advantages of utilizing autonomous UAV is their cheap price and lower price per mission flight which makes them appropriate for research purpose in the applicability of yield tracking. The main motive is to assess the feasibility of getting artificial intelligence based-UAV that could fly over oil palm plantations and also collect high-resolution images from various angles for automatic creation of return maps. These maps may tell farmers/planters the precise location where they can apply best quantity of fertilizer, pesticide and water to ensure sustainability. Naturally, mobile robots with sensors and camera mounted on top of UAV can also be utilized for such program; yet as mentioned earlier, we intend to propose an idea that involves a fleet of artificial intelligence-based UAV to monitor the palm oil plantation. By using different sensor-based imaging and measurement methods on every UAV, a real-time image processing system could be created for precise identification of the quantity of fresh fruit bunch.

4.2 Management Zone Optimization

The recent technology, management and information degree in the oil palm plantation does not provide the vision for exact direction of oil palm agriculture. Consequently, production of management fields inside the farm predicated on palm era, agronomic and land infrastructure information remains the best method to rekindle inputs. The problem here would be to choose the size or scale of each management field, that is likely too large today at 10 to 100 ha. Research carried out previously proves that the spatial version of fresh fruit bunch return is isotropic using a selection of approximately 3 palm distant space [26]. This usually means that the best size of management field is 32 palm patterned using triangular spacing. However, by means of the frequent street spacing of 20 palm space, the minimal, functional field management size ought to be 160 palms (8 palms rows x 20 palms each row) approximately 1.2 ha. Additional work is necessary to determine the mentioned specifications.

4.3 Mechanized Fertilization

The primary agronomic limitation to higher yield production is generally insufficient soil nutrient source which however can be corrected by substantial number of fertilisers. Therefore, fertiliser is considered to bear most significant expense item in the creation of oil palms in South-East Asia. It represents approximately 60 - 70 percent of the area that upkeeps price of oil palms. Incorrect fertilisation techniques might lead to high financial losses during reduction of harvest or excess fertilisation which might lead to dangers of elevated nutrient losses from run-off, contraceptive as well as other nutritional supplement reduction mechanisms. Precision farming seems to provide an accurate solution to these issues. These variations might be partly reduced by maintaining management zones for fertilization.

5. Conclusion

The demand of ever-increasing population toward research efforts should be steered toward enhancing the performance of the palm oil industry and reducing the negative environmental impact it causes. Image processing has an essential role in palm oil industry monitoring which also takes into consideration its economical as well as environmental aspect. It is also useful for the evaluation of ecological as well as harvest condition. The palm oil industry provides a room for dynamic shift which could be tailored with machine automation and methods of precision agriculture in order to bring a decrement in the labour cost and thus improve productivity with quality. Numerous palm oil plantation companies have initiated employing the latest technology and they also have a dedicated work unit for GIS/remote sensing. The use of remote sensing-based image processing aids businesses to gather vital details which eventually impacts their future business perspective. Some techniques have been already employed but is not highlighted in public or are unpublished as their main focus is not associated to academic publication. However, since the current era is observing swift shift from Industry 3.0 to Industry 4.0, therefore, knowledge transfer of latest precision agriculture methods must get more focus in order to advance in mutual advantage for agricultural and industrial sector.

6. References

[1] Torres G, Sarria G, Martinez G, Varon F, Drenth A, Guest D 2016 Bud rot caused by Phytophthora palmivora: a destructive emerging disease of oil palm. Phytopathology, 106, 320-329.

[2] De Assis Costa, O.Y, Tupinambá, D.D, Bergmann, J.C, Barreto, C.C, Quirino, B.F 2018 Fungal diversity in oil palm leaves showing symptoms of Fatal Yellowing disease. PloS one 13, e0191884.

[3] Khaled, A.Y, Aziz, S.A., Bejo, S.K., Nawi, N.M, Seman, I.A. 2018 Spectral features selection and classification of oil palm leaves infected by Basal stem rot (BSR) disease using dielectric spectroscopy. Computers and Electronics in Agriculture, 144, 297-309.

[4] Nurdiansyah F Clough Y Wiegand, K Tscharntke, T Local, 2016 and Landscape Management Effects on Pests, Diseases, Weeds and Biocontrol in Oil Palm Plantations-

A Review. Local and Landscape Management of Biological Pest Control in Oil Palm Plantations, 14.
[5] Izzuddin M., Nisfariza M., Ezzati B., Idris A., Steven M., Boyd, D. 2018 Analysis Of Airborne Hyperspectral Image Using Vegetation Indices, Red Edge Position And Continuum Removal For Detection Of Ganoderma Disease In Oil Palm. Journal of Oil Palm Research, 30, 416-428.
[6] Olafisoye O, Oguntibeju O, Osibote O 2017 Trace elements and radionuclides in palm oil, soil, water, and leaves from oil palm plantations: A review. Critical reviews in food science and nutrition, 57, 1295-1315.
[7] Smith T, Evers S, Yule C.M, Gan J, 2018, In situ tropical peatland fire emission factors and their variability, as determined by field measurements in Peninsula Malaysia. Global Biogeochemical Cycles , 32, 18-31.
[8] Woittiez L.S, van Wijk M.T, Slingerland M, van Noordwijk M, Giller K.E., 2017, Yield gaps in oil palm: A quantitative review of contributing factors. European Journal of Agronomy, 83, 57-77.
[9] Singh V, Misra A.K, 2017, Detection of plant leaf diseases using image segmentation and soft computing techniques. Information Processing in Agriculture, 4, 41-49.
[10] Arya M, Anjali K, Unni D, 2018, In Detection of unhealthy plant leaves using image processing and genetic algorithm with Arduino, 2018 International Conference on Power, Signals, Control and Computation (EPSCICON), 2018; IEEE: ,pp 1-5.
[11] Sahoo R.K, Panda R, Barik R.C, Panda S.N, Automatic Dead Zone Detection in 2-D Leaf Image Using Clustering and Segmentation Technique. International Journal of Image, Graphics & Signal Processing , 10.
[12] Bai X, Li X, Fu Z, L v X, Zhang L, A fuzzy clustering segmentation method based on neighborhood grayscale information for defining cucumber leaf spot disease images. Computers and Electronics in Agriculture 2017, 136, 157-165.
[13] Ma J, Du K, Zheng F, Zhang L, Gong Z, Sun Z 2018,A recognition method for cucumber diseases using leaf symptom images based on deep convolutional neural network. Computers and Electronics in Agriculture 2018, 154, 18-24.
[14] Jolliffe I.T, Cadima J 2016,Principal component analysis: a review and recent developments. Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences , 374, 20150202.
[15] Demšar U, Harris P, Brunsdon C, Fotheringham A.S, McLoone S 2013, Principal component analysis on spatial data: an overview. Annals of the Association of American Geographers , 103, 106-128.
[16] Firdousi R, Parveen S 2014, Local Thresholding Techniques in Image Binarization. International Journal Of Engineering And Computer Science, 3.
[17] Wang W, Duan L, Wang Y 2017, Fast Image Segmentation Using Two-Dimensional Otsu Based on Estimation of Distribution Algorithm. Journal of Electrical and Computer Engineering, 2017.
[18] Ali M, Siarry P, Pant M 2017, Multi-level image thresholding based on hybrid differential evolution algorithm. application on medical images. In Metaheuristics for Medicine and Biology, Springer, pp 23-36.
[19] Akhtar M N, Saleh JM, Awais H and Bakar E A 2020 Map-Reduce based tipping point scheduler for parallel image processing Expert Systems With Applications 139 112848 doi:10.1016/j. eswa2019.112848

[20] Hancock P.J, Baddeley R.J, Smith L.S 1992 ,The principal components of natural images. Network: computation in neural systems, 3, 61-70.
[21] Tremeau A, Borel N 1997, A region growing and merging algorithm to color segmentation. Pattern Recognition, 30, 1191-1203.
[22] Yin P.-Y, Wu T.-H 2017, Multi-objective and multi-level image thresholding based on dominance and diversity criteria. Applied Soft Computing, 54, 62-73.
[23] Prasetyo E, Adityo R.D, Suciati N, Fatichah C 2017, In Mango leaf image segmentation on HSV and YCbCr color spaces using Otsu thresholding, Science and Technology-Computer (ICST), 2017 3rd International Conference on, 2017; IEEE: , pp 99-103.
[24] Gu S, Zuo W, Xie Q, Meng D, Feng X, Zhang L 2015, In Convolutional sparse coding for image super-resolution, Proceedings of the IEEE International Conference on Computer Vision;pp 1823-1831.
[25] Gedraite E.S, Hadad M 2011, Investigation on the effect of a Gaussian Blur in image filtering and segmentation. ELMAR Proceedings, 393-396.
[26] Georgi C, Spengler D, Itzerott S, Kleinschmit B 2018, Automatic delineation algorithm for site-specific management zones based on satellite remote sensing data. Precision Agriculture, 19, 684-707.

Acknowledgments
The authors would like to acknowledge School of Aerospace Engineering, Universiti Sains Malaysia for providing research facilities to conduct this work.