# Review of Methods for Per-Pixel Transforms for Convolutional Neural Networks to Identify Tree Boundaries

Honours Literature Review

Fergus Strangways-Dixon
Computer Science
University of Cape Town
strfer001@myuct.ac.za

## ABSTRACT

Segmentation of vegetation from the surrounding landscape is popular field of research, especially focused on analysing variations in vegetation that imply disease. Due to the vast size of land used to grow trees, drones are used to capture data. Being able to effectively identify individual trees from this data is the first step towards future research in the field and should be as efficient as possible as to not bottleneck future projects in this field. Some of the issues facing accurate segmentation of trees from aerial footage include the variation of sunlight intensity during capture, as well as the amount of visual noise from the environment those trees grow in. Due to the quantity of trees found on commercial farms, segmentation models should also be performant on large amounts of aerial data. This review seeks to evaluate the existing research on per-pixel engineered features and their impact on segmentation models in terms of both accuracy and processing time. Image data is assumed to be provided in RGB format, and this review explores the current methods used to convert this format into other industry standards, such as L*a*b*, HSI as well adding new dimensions, such as infrared overlays. The potential impact these transformations is discussed in the context of developing a Convolutional Neural Network for segmentation of trees, focused on improvements in accuracy and processing time. To understand the impact these transformations on pixel data will have on the performance of a particular Convolutional Neural Network, a review of existing architectures is performed, as well as the volume of training data required to bring them to industry-standard accuracy and performance. We establish from existing research that L*a*b* format results in the most accurate segmentation, however HSI provides easily separable colour channels that allow for reduced processing time without a noticeable loss of accuracy. The clear choice to evaluate the effect of these transformations is U-Nets, as it has the lowest barrier to entry and includes invaluable image manipulations tools that will allow for straightforward testing of future hypothesis on pixel transformations.

## KEYWORDS

Tree Segmentation, Colour Spaces, Pixel Transformations, Image Manipulation, Convolutional Neural Networks, U-Nets, Atrous

## 1 Introduction

There are many existing segmentation algorithms applied in the forestry industry, with the purpose of segmenting trees from the underlying landscape. Due to the nature of the industry, image data is usually captured by remote controlled drones capable of travelling the vast distances required. Once the images are captured, the images can be processed by a Convolutional Neural Network (CNN) to perform the segmentation. These CNNs often have the drawback of requiring vast amounts of training data to become accurate, and can suffer from overfitting, especially if the training data comes from the same time of day [7]. A robust solution should be able to reliably segment the images during any time of day, in any season. This may be achievable through vast amounts of training data captured at various times and seasons, however this will greatly increase the cost of creating the model and may not be guaranteed to work. This review will focus on alternative image representations that may positively impact the accuracy of the model, while keeping training costs to a minimum. While the architecture of the CNN will be important, a review of the existing research in various pixel-level transformations one can perform on an image is explored first, as well as how various segmentation strategies may interact with these newly engineered features of the image. Specifically, colour space manipulation will be the primary focus of per-pixel transformation research, as well as a brief review of infrared images overlaid on the original image and their potential impact on segmentation accuracy. This paper assumes the original images will be provided in the RGB colour space from our partner, and so will explore the potential gains obtained from transforming the images into the HSI, L*a*b* and other colour spaces. Much of the research highlights the benefits of using only certain channels of these colour spaces, to either highlight the pixel values most prevalent in trees, such as green-dominant channels, as well as removing unneeded noise, such as the Intensity component of HSI [14]. Finally, the theoretical impact these transforms will have on 3 proposed architecture of CNNs will be examined, namely Fully Convolutional Networks, Atrous Convolutional Neural Networks, and U-Nets.

## 2 Background

Since the development of digital image formats, researchers have manipulated image data for a variety of purposes. A large portion of this research was transforming the colour space of the existing image formats into new colour spaces for a wide variety of use cases. A detailed description of these colour spaces is defined further in this paper, as well as their potential advantages towards more accurate tree segmentation. Once a clear understanding of the existing research in colour spaces is achieved, this review will investigate the existing CNN models in existing research and their advantageous towards the task of tree segmentation.

### 2.1 Pixel Transforms

Single pixel feature transforms are most often colour space transformations. These transforms convert image data from a source, usually RGB, into various other formats for different use cases.

The most widely spread colour space is RGB, as this is the format computer screens require display data. This format works well for describing light as components of 3 LEDs embedded in display pixels, however is not a direct representation of how a human eye perceives light. This colour space can be useful for the task of image segmentation as the Green component can be weighted as the most important value of the pixel in the CNN, using the Red and Blue as auxiliary components to help isolate background data.

The L*a*b* colour space is closer to human vision as it includes a Lightness channel, as well as a* and b* chrominance channels. This allows us to separate the light factor from the colour values themselves, not possible in RGB. Hernández-Hernández et al. [6] tests a wide variety of colour spaces in the context of weed detection in agriculture, and found their technique resulted in 99.2% accuracy in the L*a*b* colour space. Xiaosong Wang, Xinyuan Huang, and Hui Fu [13] found that the hues of tree canopies are predominantly found in the negative end of the a-channel, they then performed 2-dimensional OTSU segmentation of automatic threshold in the a-channel in order to reliably segment trees from an image. Xiao-Song Wang, Xin-Yuan Huang, and Hui Fu [12] corroborate these findings, stating that L*a*b* is a perceptually uniform orthogonal Cartesian coordinate system, which results in the difference in 2 pixels representing the same difference in the human eyes' visual system and resulting in more accurate segmentation.

Another industry standard for a digital format of human-like vision representation is the HSI colour space. The Hue component represents the dominant wavelength of the colour, the Saturation component is the relative purity of the Hue, and Intensity of zero describes pure white, and one describes pure black. L. Tang, L. Tian, and B.L. Steward [11] show that it is possible to decouple the intensity component to better represent how human vision perceives light in a digital manner, they also provide criticism of the HSI space by describing issues with sensor noise and minor reflectance variations leading to instabilities in the converted images, which can have serious negative effects on single dimension segmentation algorithms, however they find through normalization of the pixel values that the HSI colour space that it can theoretically lead to higher accuracy in higher-dimensional segmentation strategies. G. Ruiz-Ruiz, J. Gómez-Gil, and L.M. Navas-Gracia [9] found that due to the separable nature of the HSI colour space, they were able to reduce their computational time of their clustering process and Bayesian classifier as they required only 2 and 1 components of the HSI space respectively. The removal of the Intensity component also increased accuracy for images with varying illumination levels present in real farm fields. This strategy resulted in a 25 times improvement in segmentation time compared to the original RGB segmentation strategy without a significant loss in accuracy. Dianyuan Han and Xinyuan Huang [5] used an interesting approach to manipulate the HSI format for segmenting based on a dominant colour, green in the case of vegetation segmentation; "the hue of trees is estimated or collected, and then the distance between the hue of every pixels of image and tree hue was calculated and then normalized, which can avoid classifying the regions with large deference in colour from trees as trees and achieve good results." [5]. As with the L*a*b* space, Liying Zheng, Jingtao Zhang, and Qianyu Wang [14] found that they could isolate green in the HSI colour space by the Hue component being roughly $120^{\circ}$. They then used this in a mean-shift segmentation strategy discussed further in this paper to achieve better segmentation results for green vegetation.

A fairly new approach to vegetation detection and segmentation is the use of infrared image overlays over original data, giving the segmentation an additional attribute to segment on. Adimara Bentivoglio Colturato et al. [15] use drone mounted FLIR cameras to detect potential diseases in vegetation, specifically tree trunks. This additional spectrum of light provided to the CNN can aid in both traditional segmentations, especially in detecting trees from background vegetation, as well as the additional feature of being able to assess the overall health of an area when compared to past data.

Another approach to pixel transformations is compression. This can be an effective transform for image pre-processing as it can remove unneeded detail from an image before submitting it for segmentation. Wavelet transforms aim to mimic the masking effect of the human eye by removing tiny details at high resolution [1]. R.N. Strickland and Hee Il Hahn [10] use a modified version on wavelet transforms to improve segmentation of microcalcifications in mammograms.

### 2.2 Segmentation Methods

Segmentation methods are techniques used by computer programs to isolate instances in an image, these methods may or may not form part of a neural network. One of these methods is the Watershed Segmentation method. This involves choosing low lying areas in an image, such as dark patches, and using the pixel values as gradient. The algorithm begins at these lower points and begins to flood the image. When these basins converge, segmentation boundaries are discovered [4]. Min Bai and Raquel Urtasun [2] proposed using this as the segmentation method in a CNN. Their model generates a modified energy landscape to extract object instances. Their experimental results showed a double in performance from state-of-the-art in the Cityscapes Instance segmentation. Their model was used primarily to extract objects such as pedestrians and cars from images captured in a city, with a focus on complex images containing objects at varying distances, and the modified watershed transform performed well in this context. This technique may be modified to segment trees in a similar way, especially if the drone footage is of high enough quality to recognize the height of the tree. V. Grau, A.u.j. Mewes, M. Alcaniz, R. Kikinis, and S.k. Warfield [4] recognize a number of drawbacks of the Watershed Segmentation method that may be relevant to use on tree segmentation; the algorithm is sensitive to noise, and so local variations, such as shrubs surrounding the canopy, or shadows from images captured early or late in the day may have serious effects on the result of the segmentation. They propose anisotropic filters to mitigate this. The algorithm can also suffer from over segmentation if the gradient of colour is not substantial, resulting in many small catchment basins and a useless result. They propose using a marker overlay to identify start positions to reduce the number of minima in the image.

Mean-shift segmentation involves iteratively shifting every pixel in an image to the average of its neighbours. Once these begin to converge, segmentation is done by isolating regions of similar pixels [14]. This method may be applicable because trees are very rarely the same colour as their surrounding environment and are grouped together in similar colours already.

## 2.3 Convolutional Neural Network Models

Convolutional Neural Networks are machine learning models that use layers of connected neurons to perform an extremely wide variety of tasks, especially image classification [7]. Some of the sub-types of CNNs this paper shall examine for the purpose of tree segmentation with the aforementioned per-pixel transformations are Fully Convolutional Neural Networks (FCNN), U-Nets and Atrous Convolutional Neural Networks.

FCNNs for image segmentation proposed by Jonathan Long, Evan Shelhamer, and Trevor Darrell [7] differ from an original CNN by improving the architecture with multi-resolution layer classifiers. Their approach allows a model to not just classify one label from an image, but segment various parts of the image with different labels through up-sampling. This would be essential to the task of tree segmentation.

U-Nets are a novel architecture proposed by Olaf Ronneberger, Philipp Fischer, and Thomas Brox [8] for use in the medical field. Their focus was segmenting cells from medical images, however the topography of a cell may translate well into the topography of a forested area, if their existing model can be finetuned successfully.

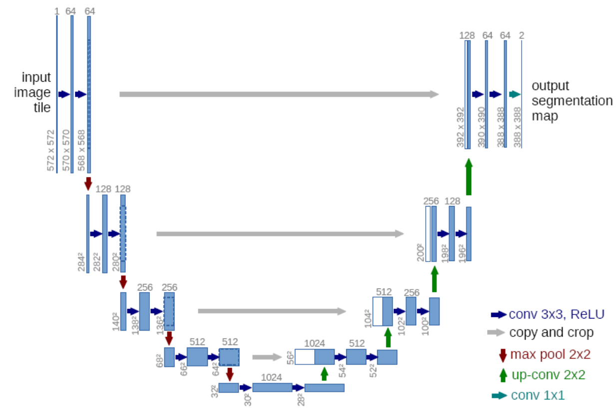


**Figure 1. U-net architecture (example for 32x32 pixels in the lowest resolution). [8]**

Figure 1 shows the architecture of a U-Net model, due to the unique connection and filter style, the model requires far fewer training images and results in higher accuracy on segmentation than existing FCNNs [8].

Atrous CNNs are designed for segmentation where detailed spatial information is desired [3]. Due to their unique architecture, Atrous CNNs excel at segmenting based on the depth of the image, accurately segmenting objects from the background noise [3].

**Figure 2. Visualization of Atrous CNN segmenting complex deep images [3]**

# 3 Per-Pixel Transformations

The primary focus of manipulating pixel values in images is changing the format of the colour they attempt to represent, as well as isolating the components those formats consist of. This review seeks to identify potential improvements in accuracy, training time and segmentation time in the discussed CNN architectures by performing the aforementioned per-pixel transformations on the provided image data before passing it to a

CNN for segmentation and investigating the viability of doing these transformations within the network itself.

### 3.1 HSI Colour Transforms

The HSI colour space is reflective of how humans perceive light, using hue, saturation, and intensity to define colours. When converting from RGB to HSI, the RGB values are normalized from 0-255 to 0-1. The following formulae are then used:

$$I = \frac{R+G+B}{3}$$

$$S = 1 - \frac{3}{R+G+B}[\min(R,G,B)]$$

$$H = \cos^{-1}\left\{\frac{(R-G)+(R-B)}{2\sqrt{(R-G)^2+(R-B)(G-B)}}\right\}$$

$$R \neq G, R \neq B$$

$$if \quad B > G, H = 2\pi - H$$

**Figure 3. RGB to HSI conversion [5]**

Figure 3 defines component Intensity as directly proportional to the total amount of colour in an RGB pixel, however is not representative of the colour itself. Hue describes the position on the colour wheel as an angle between $0^{\circ}$-$360^{\circ}$ and Saturation is representative of how polluted that colour is with white light [14]. Together, Hue and Saturation closely resemble the way a human eye perceives colour. For the purpose of segmenting vegetation, the HSI colour space allows us to isolate greens at the $120^{\circ}$ Hue value. Although this is possible with isolating the Green component from the RGB colour space, the HSI colour space allows the removal of light pollution by discarding the Intensity channel. This may lead to greater accuracy in a CNN model when the sunlight intensity of images varies greatly and will also lead to faster processing times due to only working with a 2-channel image [14].

### 3.2 Lab Colour Transforms

The L*a*b* colour space is a perceptually uniform orthogonal Cartesian coordinate system. This means the difference between pixels should be intuitive to a human observing them and allow them to quantify the difference [12, 13]. To convert from RGB to the L*a*b* colour space, the formulae in figure 4 are used by Xiao-Song Wang, Xin-Yuan Huang, and Hui Fu [12]:

$$\begin{cases} L^* = 116 f(\frac{Y}{Y_n}) - 16 \\ a^* = 500[f(\frac{X}{X_n}) - f(\frac{Y}{Y_n})] \\ b^* = 200[f(\frac{Y}{Y_n}) - f(\frac{Z}{Z_n})] \end{cases}$$

X, Y, Z, $X_n$, $Y_n$, and $Z_n$ are the coordinates of CIEXYZ color space. The solution to convert digital images from the RGB space to the CIEXYZ color space is as the following formula.

$$\begin{pmatrix} X \\ Y \\ Z \end{pmatrix} = \begin{pmatrix} 0.608 & 0.174 & 0.201 \\ 0.299 & 0.587 & 0.114 \\ 0.000 & 0.066 & 1.117 \end{pmatrix} \cdot \begin{pmatrix} R \\ G \\ B \end{pmatrix}$$

$X_n$, $Y_n$, and $Z_n$ are respectively corresponding to the white value of the parameter.

$$f(x) = \begin{cases} x^{\frac{1}{3}} & x > 0.008856 \\ 7.787x + \frac{16}{116} & x \leq 0.008856 \end{cases}$$

**Figure 4. RGB to L*a*b colour space formulae [12]**

In the L*a*b* colour space, a negative direction of the a-channel represents the pixel's tendency towards green [13]. A CNN may be able to utilize this better than a standard RGB issue by either only processing the a-channel, or processing 2 or all 3 channels and heavily weighting the a-channel.

### 3.3 Infrared Imagery Overlays

Infrared (IR) cameras are still lacking in resolution compared to normal cameras, however they can still be of use to the issue of segmenting trees from drone footage. If a drone could be fit with dual sensors (one for normal light, one for IR), the resulting IR image can be up-sampled and layered over the existing high-resolution image already processed by one of the previously mentioned per-pixel transforms. This would provide a CNN an additional feature to learn and perform segmentation on. Adimara Bentivoglio Colturato et al. [15] found that small shrubs and leaves tend not to appear on IR imagery, however trees, especially their trunks, register clearly from drone level footage. This may allow a CNN to use the IR layer to identify tree centres and discard additional vegetation that may confuse the model around the base of the tree.

## 4 Convolutional Neural Network Models

Although not the primary focus of this paper, CNN architecture will play a large role in the success of the suggested pixel transformations. Fully Convolutional Neural Networks contain a large variety of sub-architectures useful for various tasks and would require a large amount of research into which style would suit tree segmentation best. U-Nets are a more definitive architectural style and are supported by pre-existing tools such as a plugin to Fiji – an image processing program. This allows for easy initial set up and provides tools for creating overlays and performing pixel transformations on images. The team behind U-nets has also provided the community with pre-trained models focused on segmenting cells in medical imagery, which can be easily cross trained into tree segmentation. It does not seem the U-Nets project has an easily accessible command line interface, and so may be hard to automate with vast amounts of training data. However, due to the previous points discussed on U-Nets, the architecture drastically reduces the amount of training data required to perform at industry standard. U-Nets are also designed to work with minimal colour channels, as the nature of images used for micro-scale segmentation in medical use tend to be on a grey scale, and so U-Nets should be more receptive to the methods discussed around isolating colour channels. Atrous Networks excel at depth perception, and so if the quality of drone footage provided by Aerobotics are of high enough quality to render depth appropriately, as well as the effect of overlaying IR data on the original data, may prove Atrous networks the positively impacted by the effect of the pixel transformations.

## 5 Discussion

It is evident the use of colour in segmentation has been widely researched, with many papers focused around using colour spaces that closely resemble the way a human perceives natural light. This may be because we seek to develop algorithms that closely resemble our own segmentation technique. However, it is clear that these pixel values can be manipulated further for specific use cases. For this paper's purpose of investigating the segmentation of trees, the focus will be on isolating green-sensitive colour channels from the various colour spaces as discussed previously in order to both improve the accuracy of segmentation by removing noise from the image, as well as reducing processing time. This will reduce the number of colour channels the chosen CNN will receive. The papers examined in this review that advocated the use of either the HSI or L*a*b* colour spaces did not substantiate their choice well, they seemed to accept one of these as industry-standard and did not investigate the impact of other colour spaces on their segmentation algorithms. J.L. Hernández-Hernández, G. García-Mateos, J.M. González-Esquiva, D. Escarabajal-Henarejos, A. Ruiz-Canales, and J.M. Molina-Martínez [6] proposed an algorithm for automatically selecting the best colour space for segmentation of plants and soil, comparing 11 colour spaces on various images from agriculture. They found L*a*b* to be the most commonly selected and resulted in 99.2% accuracy of vegetation segmentation. Their approach to colour space selection was far more thorough and definitive, both admitting that no colour space will be the best in every situation due to various camera equipment and lighting variations, however through thorough experimentation they concede L*a*b* is the clear choice for current segmentation algorithms. Papers that advocate the use of HSI colour space focus on the ability to safely remove the Intensity component without losing the actual colour data, removing a colour channel and reducing processing time of segmentation [5, 9, 11]. Depending on the volume of data used for tree segmentation by Aerobotics, this factor may give it a clear advantage. When developing new segmentation procedures, performance is often a secondary concern as without accuracy, the process is not useful, and so L*a*b* is hypothesized to provide the best results.

In terms of CNN architectures, of the 3 examined none have a particular affinity to the aforementioned pixel transformations discovered in the readings. The U-Nets platform is a clear choice for experimentation with the discussed transformations as it has a low barrier to entry, with an existing tool chain for processing images provided, as well as pre-trained and fine-tuneable models [8]. Atrous Networks are interesting as they excel at depth perception, certainly an advantage when analysing topographical image of tree canopies. However, there are no clear existing models and tools to easily experiment with pixel transformation using Atrous networks, and so may be an additional architecture type examined if time allows. Choosing a Fully Convolutional Network type for tree segmentation may be a research topic in itself and presents no clear advantage over U-Nets or Atrous networks. Another restraint on this research is that of quantity of correctly annotated training data for the CNN model, and U-Nets was designed for creating state-of-the-art performing models from very little training data, again making it the clear choice to begin experimentation with these pixel transformations and the impact they have on accuracy and performance of the model.

## 6 Conclusions

There has been vast amount of research into colour spaces, and the effects they have on segmenting image data. RGB is considered poor for segmentation tasks, and researchers have determined that transforming pixel values from RGB into either HSI or L*a*b* colour spaces to better represent how the human eye perceives colour. The papers examined in this review also emphasise the ability to separate the colour channels present in these formats to highlight particular colours based on the use case, in this case separating greens. In HSI, Hue values at roughly 120º or negative values in the L*a*b* a-channel are responsible for greens in the final image, and so will be the primary data used by the CNN model to segment trees with. Due to this separation, the papers reviewed also highlight that because the pixel colour channels are reduced, it leads to a significant improvement in training and segmentation time, without a loss in accuracy if performed correctly. A hypothesis that will be explored using both HSI and L*a*b* colour transforms will be reducing the loss of accuracy due to the time of day an image was captured. As mentioned earlier it is possible to remove components of images in these formats to theoretically make the colour of vegetation less dependent on the brightness of the sunlight present, thus improving the segmentation accuracy of the model. The researchers that developed the transformations discussed in this review did not apply these to CNN models, only traditional segmentation techniques such as the Watershed approach. With modern advances in CNN architectures, exploring these transformations for use in CNN models is an interesting research field.

None of the explored CNN architectures seem to display a natural affinity towards the discussed transformations, and so U-Nets will be the initial model used to test these techniques due to the existing framework for processing images as well as fine-tuning models, as well as the architecture's inherent attribute of requiring small data sets to achieve industry-standard accuracy and segmentation time. The existing models provided with the U-Nets architecture are also trained on segmenting topographical images of cells, which may be similar enough to tree canopies from drone footage to allow for easy fine-tuning for our use case. If time allows, an Atrous CNN model may respond well to infrared overlays on the original imagery, however Atrous networks seem to have a higher barrier to entry getting the model functional and may not have a significant impact on accuracy. Adding infrared channel data to pixels will also negatively impact training and segmentation time, potentially without a tangible impact on performance. This may be an interesting transformation to experiment with, should be a lower priority to the aforementioned techniques.

## REFERENCES


[1] Marc Antonini, Pierre Mathieu, and Ingrid Daubechies. 1992. Image Coding Using Wavelet Transform. *IEEE Transactions on Image Processing* 1, 2 (April 1992), 205–220.

[2] Min Bai and Raquel Urtasun. 2017. Deep Watershed Transform for Instance Segmentation. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2017), 5221–5229. DOI:http://dx.doi.org/10.1109/cvpr.2017.305

[3] Liang-Chieh Chen, George Papandreou, Florian Schroff, and Hartwig Adam. 2017. Rethinking Atrous Convolution for Semantic Image Segmentation. *CoRR* abs/1706.05587 (2017).

[4] V. Grau, A.u.j. Mewes, M. Alcaniz, R. Kikinis, and S.k. Warfield. 2004. Improved Watershed Transform for Medical Image Segmentation Using Prior Information. *IEEE Transactions on Medical Imaging* 23, 4 (April 2004), 447–458. DOI:http://dx.doi.org/10.1109/tmi.2004.824224

[5] Dianyuan Han and Xinyuan Huang. 2010. A Tree Image Segmentation Method Based on 2-D OTSU in HSI Colour Space. *2010 International Conference on Computational Intelligence and Software Engineering* (2010). DOI:http://dx.doi.org/10.1109/wicom.2010.5600669

[6] J.L. Hernández-Hernández, G. García-Mateos, J.M. González-Esquiva, D. Escarabajal-Henarejos, A. Ruiz-Canales, and J.M. Molina-Martínez. 2016. Optimal colour space selection method for plant/soil segmentation in agriculture. *Computers and Electronics in Agriculture* 122 (2016), 124–132. DOI:http://dx.doi.org/10.1016/j.compag.2016.01.020

[7] Jonathan Long, Evan Shelhamer, and Trevor Darrell. 2015. Fully convolutional networks for semantic segmentation. *2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)* (2015), 3431–3440. DOI:http://dx.doi.org/10.1109/cvpr.2015.7298965

[8] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. 2015. U-Net Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015* 9351 (2015), 234–241. DOI:http://dx.doi.org/10.1007/978-3-319-24574-4_28

[9] G. Ruiz-Ruiz, J. Gómez-Gil, and L.M. Navas-Gracia. 2009. Testing different colour spaces based on hue for the environmentally adaptive segmentation algorithm (EASA). *Computers and Electronics in Agriculture* 68, 1 (2009), 88–96. DOI:http://dx.doi.org/10.1016/j.compag.2009.04.009

[10] R.N. Strickland and Hee Il Hahn. 1996. Wavelet transforms for detecting microcalcifications in mammograms. *IEEE Transactions on Medical Imaging* 15, 2 (April 1996), 218–229. DOI:http://dx.doi.org/10.1109/42.491423

[11] L. Tang, L. Tian, and B.L. Steward. 2000. Colour Image Segmentation With Genetic Algorithm For In-Field Weed Sensing. Transactions of the ASAE 43, 4 (2000), 1019–1027. DOI:http://dx.doi.org/10.13031/2013.2970

[12] Xiao-Song Wang, Xin-Yuan Huang, and Hui Fu. 2009. The Study of Colour Tree Image Segmentation. *2009 Second International Workshop on Computer Science and Engineering* (2009), 303–307. DOI:http://dx.doi.org/10.1109/wcse.2009.818

[13] Xiaosong Wang, Xinyuan Huang, and Hui Fu. 2010. A Colour-Texture Segmentation Method to Extract Tree Image in Complex Scene. *2010 International Conference on Machine Vision and Human-machine Interface* (2010). DOI:http://dx.doi.org/10.1109/mvhi.2010.138

[14] Liying Zheng, Jingtao Zhang, and Qianyu Wang. 2009. Mean-shift-based colour segmentation of images containing green vegetation. *Computers and Electronics in Agriculture* 65, 1 (2009), 93–98. DOI:http://dx.doi.org/10.1016/j.compag.2008.08.002

[15] Adimara Bentivoglio Colturato et al. 2013. Pattern Recognition in Thermal Images of Plants Pine Using Artificial Neural Networks. *Engineering Applications of Neural Networks Communications in Computer and Information Science* (September 2013), 406–413. DOI:http://dx.doi.org/10.1007/978-3-642-41013-0_42