Reduction of False Positives in Vessel Detection using HSV-based Thresholding

ANGELA CLARISSE C. CHUA, University of the Philippines- Diliman, Philippines

On instances where deep learning is not an option due to hardware limitations, vessel detection can be done using SVM. However, this does not guarantee a good precision as models trained with simple feature descriptors such as Histogram of Oriented Gradients (HOG) and Local Binary Pattern (LBP) are vulnerable to false positives. This research explores the idea of reducing these false positives by determining whether a ROI predicted by the SVM model contains a vessel or not using HSV-based thresholding.

CCS Concepts: • **Computing methodologies → Object detection**.

Additional Key Words and Phrases: ship detection, HSV colorspace, thresholding, segmentation

ACM Reference Format:
Angela Clarisse C. Chua. 2021. Reduction of False Positives in Vessel Detection using HSV-based Thresholding. In ,. ACM, New York, NY, USA, 8 pages. https://doi.org/10.1145/1122445.1122456

1 INTRODUCTION

Maritime surveillance is essential because it protects several elements of a country such as its security and its economy. Not only that but it is also important for disaster risk assessment and management. Vessel detection using machine learning applications is widely researched, whether it may be using classical machine learning methods such as SVM or deep learning. However, due to hardware, computational and data limitations, deep learning is not always a suitable option. Models like SVM can give quality predictions especially if it is trained on good datasets and feature descriptors. This can even be improved with pre-processing. However, all these steps require to be manually and carefully parameterized. Not only that but it is also at risk of false detection due to islands, heavy clouds, ocean waves, and the various and uncertain sea state conditions, like partial cloud cover, fog, wind, and swell [2]. Even in feature descriptors specifically used for ships such as S-HOG [7] can be vulnerable to false detection. Therefore, it is important to establish post-detection methods that can reduce the false positives predicted by the initial model.

2 METHODOLOGY

The methodology of this research is divided into three subsections. The first subsection describes the main idea of the whole research from pre-processing to post-processing. The second subsection, called *Model Selection*, discusses the process and baseline for selecting the proper feature descriptors which will give best possible initial detection. Lastly, the proposed *post-processing using HSV-based Thresholding* explores the concept and method behind the idea of using HSV values to determine whether a region of interest (ROI) contains a vessel(s) or not.

ACM ISBN 978-x-xxxx-xxxx-x/YY/MM. . . $15.00
https://doi.org/10.1145/1122445.1122456

2.1 Main Idea

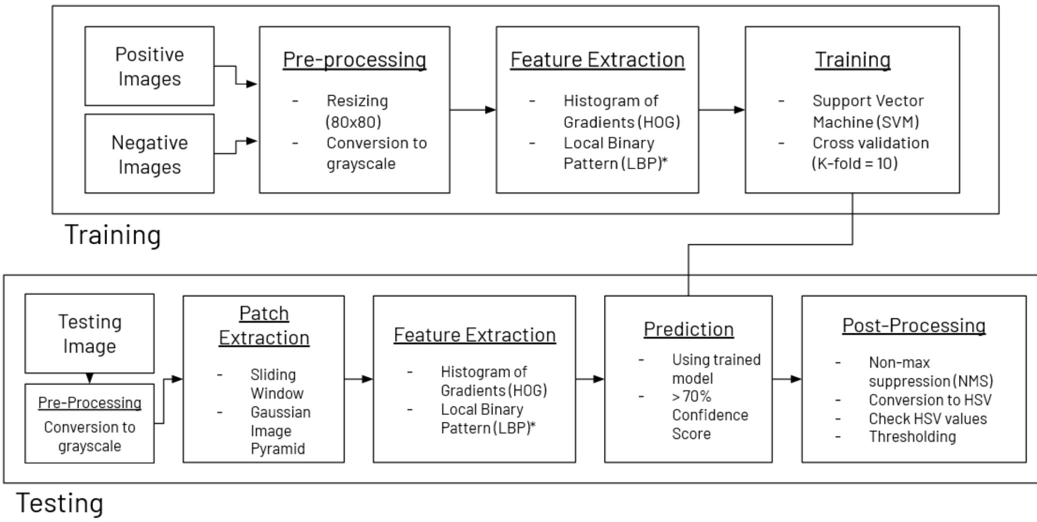


Fig. 1. Diagram of the proposed vessel detection system

Figure 1 shows the diagram of the system implementation of the proposed vessel detection system. It is divided into two stages. The training stage which trains the model using a support vector machine (SVM) and the testing stage which uses the trained model to predict the ROIs. The ROIs will then undergo a post-processing method to determine whether it actually contains a vessel or is just falsely identified as having one by the SVM model. This step reduces the number of false positives detected.

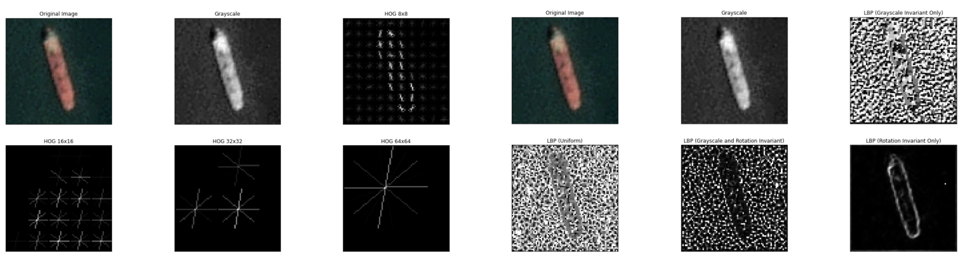
The training stage begins with a pre-processing method which will resize the training images to 80px x 80px and convert them to grayscale. The pre-processed images will go through feature extraction, which is dependent on the model selected. The extracted feature(s) will then undergo supervised training using a specific type of SVM, which is also dependent on the model selected. The classification is just binary, therefore, the labels are simply "vessel" or "not vessel" for positive and negative images respectively. This stage is done several times, where each iteration uses a different feature descriptor. The resulting models will then be assessed based on their results, specifically, the one with the highest F1-score will be chosen as the final model to be used for initial vessel detection.

The testing stage will take an input image and pre-process it by converting it to grayscale. Patches will then be extracted from the pre-processed image using the sliding window and Gaussian image pyramid method. Patch size is 80px x 80px, similar to the size of the training images. The same feature descriptor used in training the model will be used in these patches. Afterwards, the extracted features will be predicted by the trained model as "vessel" or "not vessel". Each prediction will have an accompanying confidence score. All patches > 70% confidence score will then be marked as a possible "vessel" and will be considered as a ROI.

The ROIs will then undergo a post-processing method, which will be further discussed under the subsection *Post-processing using HSV-based Thresholding.*

2.2 Model Selection

The initial detection is highly dependent of the trained model. Therefore, it is important to choose the model that will give the best result. In this research, two types of feature descriptors were chosen. The first one is Histogram of Oriented Gradients (HOG), HOG was proposed by Dalal and Triggs [4] as a feature descriptor to represent objects by the distribution of gradient intensities and orientations in spatially distributed regions, this has been widely acknowledged as one of the best features to capture the edge or local shape information of the objects. [3]. The second one is Local Binary Pattern by Ojala et al. [8], is a feature descriptor typically used for texture classification problems.



(a) HOG in various numbers of pixels per cell

(b) Different types of LBP

Fig. 2. The two feature descriptors considered for model selection

Different variations of HOG and LBP that were considered are seen in figure 2. Fig 2 (a) contains (from left to right) the original image, image in grayscale, HOG in 8x8 (pixels per cell), HOG 16x16, HOG 32x32 and HOG 64x64. In fig 2 (b), also from left to right, is the original image, image in grayscale, Grayscale Invariant LBP (Default LBP), Uniform LBP, Grayscale and Rotation Invariant LBP (RoR LBP) and Rotation Invariant LBP (Var). The last one is not included in testing due to computational limitations. It is important to note that the HOG and LBP features were normalized before being trained to maintain consistency and range among the values.

According to [1] and [6], HOG + LBP is a good combination for feature extraction especially in plant texture detection and face recognition. In [6], HOG + LBP was able to achieve a 91.25% accuracy as compared to HOG only (85.31%) or LBP only (40.6%) for plant texture detection. With this in mind, combination of HOG + LBP features will also be tested for vessel detection/identification and assess whether or not this combination will result to a higher F1-score as compared to training with just HOG or LBP only.

Tools used were OpenCV and skimage for image processing and scikit-learn for SVM.

2.3 Post-processing using HSV-based Thresholding

The main concept of the proposed post-processing method is to segregate objects such as vessels and land to water. This method assumes that land-to-water ratio (or “not vessel”) is greater than vessel-to-water ratio per patch or ROI. The ratio can be calculated by simply getting the average pixel value of the masked image. If the average pixel value of the ROI is 0, this means its mask is pure black or simply "water". If a vessel is present, then the avarage pixel value must be greater than 0 but less than a certain threshold. Therefore, anything higher than a certain threshold shall be tagged as “not vessel” and will be disregarded.

Thresholding is usually not a preferred method for segmentation due to it being region size-dependent and also due to the varying individual image characteristics, it is also time consuming [9]. However, it can be realized that by accessing the HSV values of the image, it is possible to do a better masking/thresholding, thus, resulting in an improved segmentation.

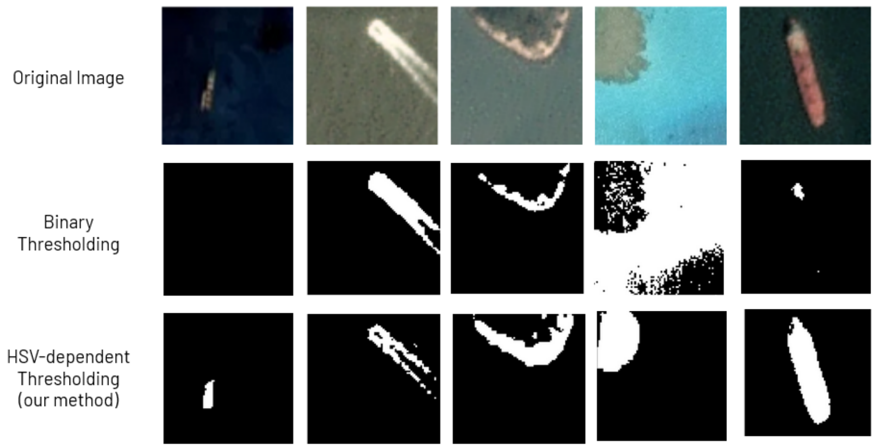


Fig. 3. Comparison between Binary Thresholding and HSV-based Thresholding

Figure 3 shows the difference between binary thresholding and the proposed HSV-based thresholding. The HSV-based thresholding can mask the vessel even at varying image conditions. For example, the vessel in the first image disappears when masked using binary thresholding, this is due to the contrast of the original image. Using the proposed HSV-based thresholding, the vessel is now recognized as a separate object. Another example is the fourth image, where it shows an inverted mask when using binary thresholding, that is, the water is recognized as "land" (white) and the land is recognized as "water" (black). The HSV-based thresholding, on the other hand, was able to acknowledge the initial image condition and mask the image the proper way.

The concept behind HSV-based thresholding is by relating the color range of the background (water) and object (vessel/land/others) to the HSV values of the ROI. By pure observation, it can be noticed that at certain hue, saturation or value, there is a distinct color range that separates the background from the object. Table 1 shows the relationship between saturation value and background/object color range, table 2 shows the same concept for hue values. The color range is normalized in a way that 0 and 1 both correspond to the color red, the colors between 0 and 1 follows the sequence of color of the visible spectrum.

Table 1. Relationship between Saturation Value and Background/Object Color Range

Saturation Value	Background Color Range	Object Color Range
Low Saturation (<0.25)	0.10-0.20	0-0.05
Medium Saturation 1 (0.25-0.30)	0.20-0.30	0-0.15
Medium Saturation 2 (0.30-0.45)	0.30-0.50	0-0.20
High Saturation (0.45-0.60)	0.50-0.70	0-0.35
Extremely High Saturation (>0.60)	0.70-1.00	0-0.50

Table 2. Relationship between Hue Value and Background/Object Color Range

Hue Value	Background Color Range	Object Color Range
Low Hue (<0.20)	0-0.10	0.30-0.70
Medium Hue (0.20-0.40)	0.20-0.30	0.40-0.70
High Hue (>0.40)	0.30-0.50	0.55-0.70

Using this relationship, the objects within the ROI can be identified just by masking out the said object color range. However, it is important to note that some ranges have their drawbacks and a simple masking using these relationships does not give the best segmentation. For example, the low saturation has a small range of colors to be masked, the medium 2 and high saturation, on the other hand, are not always coherent, this means that the objects within the image are usually hollow and that their borders are usually the only one within the object color range (not the whole object itself). Lastly, the extremely high saturation value overlaps the color red between the background and the object itself. Therefore, the only feasible range to use is the medium saturation 1 range. An image or ROI that is not within this range cannot be masked right away and shall therefore undergo hue check. The full process of the HSV-based thresholding method is show in figure 4.

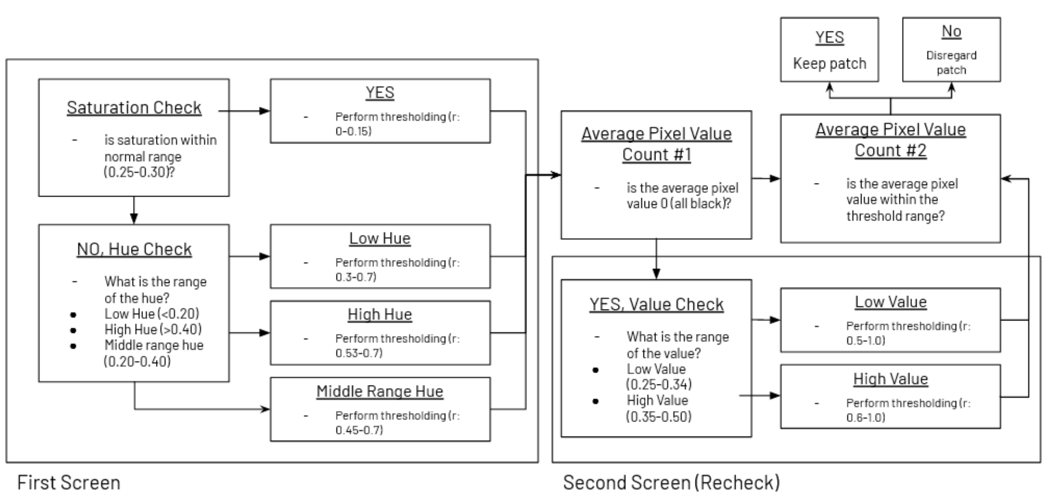


Fig. 4. HSV-based Thresholding method

The ROI will first undergo saturation check, as mentioned in the previous paragraph, if the ROI's saturation value is within the medium saturation 1 range, it will proceed to masking, if not, it will undergo hue check. This step determines the proper color range needed to be masked. After the masking process, the masked ROI will undergo the first *average pixel value count*. If the result is 0, this means that the mask is pure black and that the ROI is identified as water. However, this is not always accurate, sometimes, just by using the hue-based masking, small objects can come undetected. In this case, the value (v in HSV) shall be accessed. The same concept with the value and color range relationship shall be applied. This step is considered as the *second screening*. This step is succeeded by another average pixel value count which will determine whether a ROI contains a vessel or not.

3 RESULTS

3.1 Model Selection

The dataset is taken from [5], which contains 1520 positive (vessel) images and 1951 negative (not vessel) images. The SVM model used to test the different types of HOG and LBP is a polynomial SVM with a degree of 7 as it gave the best result for each test.

Table 3. Quantitative results of model trained with various HOG and LBP feature descriptors

Feature Descriptor	Precision	Recall	F1-Score
HOG (8x8)	96%	82%	88%
HOG (16x16)	**96%**	94%	**95%**
HOG (32x32)	87%	92%	89%
HOG (64x64)	80%	83%	82%
LBP (Default)	83%	93%	88%
LBP (Uniform)	76%	94%	84%
LBP (RoR)	76%	93%	84%

Table 4. Quantitative results of different SVM models trained with HOG (16x16) + LBP (Default)

SVM Model	Precision	Recall	F1-Score
Linear SVM	93%	91%	92%
Radical Basis (RBF) SVM	92%	95%	94%
Polynomial SVM (d=7)	94%	**96%**	**95%**
Sigmoid SVM	78%	74%	76%

Table 3 shows that HOG (16x16) performed the best with an F1-score of 95%. It also gave the highest precision with 96% (together with HOG 8x8) and highest recall with 94% (together with Uniform LBP). However, the F1-score was chosen as the basis because it describes the balance between the precision and recall. Therefore, the HOG and LBP type with the highest F1-score are considered as the best performing type.

Table 4 shows the results of the combination of the two best performing HOG and LBP types, which are HOG (16x16) and LBP (Default) trained in various SVM models. It can be seen that the polynomial SVM (degree = 7) gave the highest F1-score. In general, when using the F1-score as a basis, the HOG+LBP combination did not give an improvement, however, it was able to increase the recall from 94% to 96%. The precision, on the other hand, gave a lower score.

3.2 Post-processing using HSV-based Thresholding

Using the model trained with the HOG (16x16) feature descriptor and 32 images from Planet.com, National Aeronautics and Space Administration (NASA) and Asia Maritime Transparency Initiative (AMTI), the initial detection was conducted. The initial detection gave an average precision of 52%, this means that the model predicted a lot of false positives or ROIs that do not really contain a vessel(s) but were predicted as having one. This precision shall be called as the *raw precision*. The generated ROIs from the initial detection then underwent the proposed post-processing step using

HSV-based thresholding and was able increase the average precision to 89%. This gave a 71.15% increase as compared to the raw precision. The post-processing step also got a recall score of 92%, this means the “loss” (the actual vessels that were initially detected and later disregarded by the post-processing method) is low.

Fig. 5. (Left) Initial detection of vessels in San Francisco Bay, USA (Planet.com). The green boxes are the ROIs that will undergo post-processing. (Right) Red boxes are the detection after the proposed post-processing method.

An example can be seen in figure 5. The left image shows the initial detection of the trained SVM model in green boxes, these boxes are also the ones considered as ROIs that will undergo the proposed post-processing method. The right image shows the detection after post-processing in red boxes. It can be noticed that the false positives initially detected by the trained model significantly decreased, leading to a more precise prediction.

4 CONCLUSION

The research shows that combining HOG (16x16) + LBP (Default) gives the same F1-score as HOG (16x16) alone. However, the recall increased from the previous highest of 94% to 96%. This is not the case for precision, which is lower than the highest one attained by the HOG (16x16) alone.

The proposed post-processing method, on the other hand, was able to notably decrease the false positives detected by the initial detection, specifically, it was able to increase the precision by 71.15% of the 32 images tested with only a small loss. This post-processing method can be useful for vessel detection systems based on classical machine learning methods such as SVM, which feature extraction step is manually done and carefully parameterized to get a more precise detection.

Moreover, the research was able to present a novel thresholding method using the relationship between HSV values and color ranges of the background and objects. This can be useful not only as a post-processing method for reduction of false positives in vessel detection but also for general segmentation of vessels in varying image conditions.

REFERENCES

[1] V. F. Arguedas. 2015. Texture-based vessel classifier for electro-optical satellite imagery. *2015 IEEE International Conference on Image Processing (ICIP)* (2015), 3866–3870.

[2] J. Liu C. Dong and F. Xu. 2018. Ship Detection in Optical Remote Sensing Images Based on Saliency and a Rotation-Invariant Descriptor. *Remote Sensing.* 10, 3 (2018).

[3] S. Charchut. 2020. *Ship Detection Feature Analysis in Optical Satellite Imagery through Machine Learning Applications through Machine Learning Applications.* Master’s thesis. University of New Orleans.

[4] N. Dalal and B. Triggs. 2005. Histograms of oriented gradients for human detection. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)* (2005), 886–893.
[5] Kaggle. 2017. Ships in Satellite Imagery. https://www.kaggle.com/rhammell/ships-in-satellite-imagery
[6] M. Billah M. Islam and M. Yousuf. 2019. Automatic Plant Detection Using HOG and LBP Features With SVM. *International Journal of Computer (IJC)* 33 (04 2019), 26–38.
[7] J. Lin Y. Li S. Qi, J. Ma and J. Tian. 2015. Unsupervised Ship Detection Based on Saliency and S-HOG Descriptor From Optical Satellite Images. *IEEE Geoscience and Remote Sensing Letters* 12, 7 (2015), 1451–1455.
[8] M. Pietikäinen T. Ojala and D. Harwood. 1994. Performance Evaluation of Texture Measures with Classification Based on Kullback Discrimination of Distributions. (1994).
[9] O.I. Singh T. Sinam T.R. Singh, S. Roy and K.M. Singh. 2011. A New Local Adaptive Thresholding Technique in Binarization. *IJCSI International Journal of Computer Science Issue* 8, 6 (2011), 271–277.