International Journal of Computing and Digital Systems
ISSN (2210-142X)
Int. J. Com. Dig. Sys. 16, No.1 (Jul-24)

http://dx.doi.org/10.12785/ijcds/160116

An Instance Segmentation Method for Nesting Green Sea Turtle's Carapace using Mask R-CNN

Mohamad Syahiran Soria[1], Khalif Amir Zakry[2], Irwandi Hipiny[3], Hamimah Ujir[4], Ruhana Hassan[5] and Alphonsus Ligori Jerry[6]

[1,2,3,4]*Faculty of Computer Science and Information Technology, Universiti Malaysia Sarawak, Sarawak, Malaysia*
[5]*Faculty of Resource Science and Technology, Universiti Malaysia Sarawak, Sarawak, Malaysia*
[6]*Sarawak Forestry Corporation, Sarawak, Malaysia*

Received 17 Oct. 2023, Revised 14 Apr. 2024, Accepted 16 Apr. 2024, Published 1 Jul. 2024

Abstract: This research presents an improved instance segmentation method using Mask Region-based Convolutional Neural Network (Mask R-CNN) on nesting green sea turtles' images. The goal is to achieve precise segmentation to produce a dataset fit for future re-identification tasks. Using this method, we can skip the labour-intensive and tedious task of manual segmentation by automatically extracting the carapace as the Region-of-Interest (RoI). The task is non-trivial as the image dataset contains noise, blurry edges, and low contrast between the target object and background. These image defects are due to several factors, including jittering footage due to camera motion, the nesting event occurring during a low-light environment, and the inherent limitation of the Complementary Metal-Oxide-Semiconductor (CMOS) sensor used in the camera during our data collection. The CMOS sensor produces a high level of noise, which can manifest as random variations in pixel brightness or colour, especially in low-light conditions. These factors contribute to the degradation of image quality, causing difficulties when performing RoI segmentation of the carapaces. To address these challenges, this research proposes including Contrast-Limited Adaptive Histogram Equalization (CLAHE) as the data pre-processing step to train the model. CLAHE enhances contrast and increases differentiation between the carapace structure and the background elements. Our research findings demonstrate the effectiveness of Mask R-CNN when combined with CLAHE as the data pre-processing step. With CLAHE technique, there is an average increase of 1.55% in Intersection over Union (IoU) value compared to using Mask R-CNN alone. The optimal configuration managed an IoU value of 93.35%.

Keywords: Computer Vision, Instance Segmentation, Mask R-CNN, CLAHE, Deep Learning

1. Introduction

Chelonia Mydas is one of the seven species of sea turtles in the world. Despite being a commonly-sighted species worldwide, the green sea turtles is classified as an endangered animal, with its population declining tremendously over the years [1]. As part of the conservation initiative, researchers perform biometric re-identification to identify an individual from a biometric sample of the carapace scute pattern using learned image descriptors [2]. This task is done practically using computer vision and deep learning techniques with scute pattern imagery as input. Nevertheless, obtaining an accurate classification requires a ton of training data. During the training phase, the neural network will learn all the features inside the image. As a result, this can impact longer training time and require expensive computing resources to train the model to differentiate between the target object vs. the unnecessary image background. Significantly, the model might be trained on unneeded features affecting the classification prediction. It has been proven that the locations or backgrounds during training will impact the classification system because the model cannot generalize well to new locations [3]. Moreover, manually cropping a target object is a cumbersome and laborious task [4]. Therefore, there is a need to segment the carapace, focusing on the scute pattern, to obtain an accurate segmentation of the target object in an automated manner.

Different methods and approaches have been done to leverage the performance of Mask R-CNN to tackle noisy and low contrast images. For the first time, we proposed the advantages of CLAHE as a data preprocessing technique to nesting green sea turtles' images. By using CLAHE, we could improve the local contrast and enhance the visibility of the subject against the background which is clinical to produce better image segmentation.

E-mail address: 22020014@siswa.unimas.my, 21020463@siswa.unimas.my, mhihipni@unimas.my

The dataset has a few inherent image challenges that need to be solved to enhance Mask R-CNN's performance in predicting the turtle's carapace segmentation. In this paper, we evaluated Mask R-CNN performance using images extracted from videos of sea turtles captured during the nesting season. As we wish to reduce interferences with the animals, we used a HIK Vision camera to capture videos in low-light conditions, as shown in Figure 1. The camera model is a closed-circuit television (CCTV) type camera equipped with a CMOS sensor. CMOS is an image sensor technology commonly used in consumer-level digital cameras and other imaging devices. This sensor technology exhibits higher levels of noise that can manifest as random variations in pixel brightness or colour, especially in low-light conditions [5]. Moreover, our dataset images also contain a loose boundary between objects due to various factors, including image noise, blurring, low image resolution, and similar intensity values between the carapace against the flipper of the turtle as shown in Figure 1. These factors introduce uncertainty and ambiguity, leading to challenges in image segmentation, object detection, and boundary extraction tasks [6] [7].

This research proposed an instance segmentation method of a sea turtle's carapace in an automated manner using a Mask R-CNN-based method. We implemented Contrast-Limited Adaptive Histogram Equalization (CLAHE) [8] method as the data pre-processing step. In many cases, CLAHE improves local contrast while enhancing and preserving details inside the image. Application examples include CT scans [9], retinal fundus [10] and animals [11]. In Section 3, We performed sensitivity analysis on a few internal parameters of CLAHE to show their impacts towards mask segmentation prediction by comparing the segmented results against our manually annotated ground truth.

Previous works have incorporated different image enhancement techniques for data preprocessing with Mask R-CNN. Jiangping et al. [12] addressed the problem of protein crystallization in macromolecular crystallography and the challenges posed by image quality and classification algorithms. In their work, CLAHE was used as a data preprocessing step to enhance the visibility of protein crystals in images. This inclusion improved the classification accuracy of the Mask R-CNN model, resulting in a 42% improvement in mean average precision. However, the network sometimes misidentifies non-crystal spots outside the droplet as crystals, highlighting the need for enhance segmentation accuracy. Naufal et al. [13] addressed the problem of face mask detection under low-light conditions during the Coronavirus disease (COVID-19) pandemic. The low-light condition can make image detection more difficult due to the presence of high-noise images, poor illumination, and reflectance. To overcome this issue, the images need to be processed using two different approaches, such as CLAHE and Gamma Correction, before training the model to provide visual quality by fine-tuning brightness and

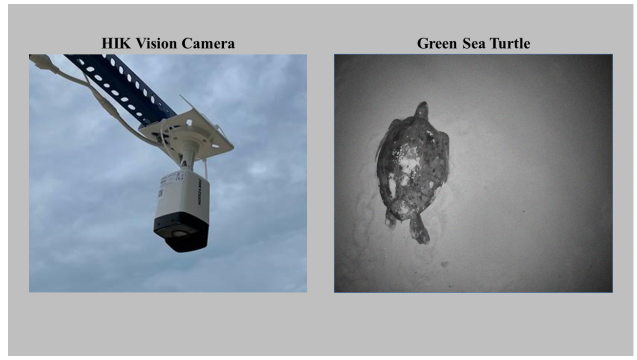

Figure 1. Hik Vision Camera used in the data collection and noisy turtle images captured

contrast levels. The methods improved the detection of face masks under low-light conditions by 98.13% accuracy using the Gamma correction method and 97.86% for the CLAHE method. The proposed method shows better result compared with the data trained using pre-trained models such as MobileNetV2, VGG16, and VGG19. However, the paper also suggests finding more alternative methods that can improve performance and classification for low-light images. Xu et al. [14] adopted CLAHE techniques for data preprocessing as a preliminary step in preparing the input data for their tuberculosis detection model. Given that chest radiograph images often exhibit high noise levels and low contrast, CLAHE was utilized to extract essential features for training the model to detect tuberculosis lesions within the images. Consequently, this method effectively enhanced the contrast between pulmonary tuberculosis regions and the background.

Unlike previous studies, we have developed an image segmentation model using Mask R-CNN combined with the CLAHE algorithm to segment biometric parts of wildlife animals in extremely low-light conditions at night. Thus, the CLAHE segmentation technique was utilized as a data preprocessing step to overcome the problem of low contrast in the images. This method improves the local contrast visibility of fine features and enhances the overall image clarity. This enhancement is crucial for achieving accurate segmentation with Mask R-CNN, particularly under challenging lighting conditions.

2. Method

To train our dataset with Mask R-CNN, we collected the dataset, annotated it, enhanced the images using CLAHE, and evaluated the mask performance based on Intersection over Union (IoU). This section provides a detailed discussion of the methodology followed.

A. Data Acquisition

For this research, we collected video data at Talang Besar Island, located in Sarawak, Malaysia. Each video data is around 2 - 3 minutes in duration and contains video footage of a slowly moving green sea turtles. The dataset