UNIVERSITAS LAMBUNG MANGKURAT

Novitasari Novitasari <novitasari@ulm.ac.id>

Fwd: Paper Submission

YUSLENA SARI <yuzlena@ulm.ac.id> Fri, Jun 9, 2023 at 11:58 AM
To: Novitasari Novitasari <novitasari@ulm.ac.id>

Sent from my iPhone

Begin forwarded message:

From: editor@jsju.org
Date: 5 June 2023 21.03.34 GMT+8
To: yuzlena@ulm.ac.id
Subject: Re: Paper Submission

Dear Authors,

Greetings from Journal of Southwest Jiaotong University,
Thank you so much for being so interested in submitting your research to our Journal of Southwest Jiaotong University.

Two independent experts in the field have reviewed your manuscript.
We are pleased to inform you that the paper has been accepted for publication subject to minor revisions being made in response to the reviewers' comments given below.

You need to send the following to the editorial office via editor@jsju.org, if you want to publish in the Journal of Southwest Jiaotong University, Volume 58 (3), June 2023:
The article, strictly formatted according to the template recommendations for authors and correct the article on the reviewers'
recommendations, see http://jsju.org/JournalTemplate.docx, responses to reviewers, and proof of payment.

All revisions should be highlighted in the article in red. We hope you will do great.
Reviewers:
The topic is interesting and important. However, there are several key areas that need more work.
1 - In the introduction section, please highlight the purpose and novelty of this particular study.
2 - Authors should explain better the criteria for selecting the research object and the limitations of the application of the results obtained should be specified.
3 - Please cite sources under the tables. If they were developed by the authors, indicate "developed by the authors".
4 - The Conclusion section needs to justify the effectiveness of the approaches you propose.

Deadline for corrections and payment: June 12, 2023.
* Please write in the subject of the email: revised article and payment proof.

Journal of Southwest Jiaotong University charges Publication Fee at the rate of Euro 500.00. All articles published in our journal are open access and freely available online, immediately upon publication.
We recommend that the authors use the academic text editing service for the scientific articles, not just proofreading. Please use the American English option. We recommend the use of large, trusted companies with editors having a Ph.D. degree. You should also attach an editing certificate or use the editorial office services. Articles that native English speakers do not edit are not allowed for publication. The editorial team provides academic proofreading services for the authors at an additional cost - Euro 120 (for your text), if you don't attach an official English editing certificate. The fee's already included in the invoice.
We send you a link to pay by card:
https://invoice.stripe.com/i/acct_1Ja3OqJAnp8X2J3O/live_YWNjdF8xSmEzT3FKQW5wOFgySjNPLF9PMWdOR094NmF6MENOd1VHWExXNG5IM0hLYVIxVWk3LDc2NTEwODA10200Li9K07kY?s=db
The publication fee is 620 euros, the operator's fee for the transfer is 30 euros.

Banks fees must be paid by the customer for both payer and payee so that we can receive the full invoiced amount.

Southwest Jiaotong University:
Open access - free for readers, article processing fee (APC) paid by authors or institutions.
High visibility: indexed in many databases.
Subject Area and Category: multidisciplinary.
Rapid publication: manuscripts are peer-reviewed. The first decision is given to authors about 20 days after submission; acceptance for publication after revisions is done within 2-5 days (averages for articles published in this journal in the second half of 2020).
Recognition of reviewers: Reviewers who provide thorough review reports promptly receive vouchers that entitle them to a discount on the APC of their next publication.

Please read our published articles http://jsju.org/index.php/journal/issue/archive.

Sincerely yours,
Editorial Office of Journal of Southwest Jiaotong University
http://jsju.org/index.php/journal/index
该杂志的出版商 (Our Publisher): Science Press
网址 http://science-press.cn/西南交通大学学报/

-------- Original Message --------
Subject: Paper Submission
Date: 2023-04-28 15:16
From: JSJU <admin@jsju.org>
To: editor@jsju.org

Title of your paper: UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD
Corresponding Author's Email Address: yuzlena@ulm.ac.id
Author(s): Novitasari Novitasari, Yuslena Sari, Yudi Firmanul Arifin, Nurul Fathanah Mustamin, Erika Maulidiya
Keywords: Classification; class; CNN; GLCM; accuracy.
Abstract: Land cover is an important factor in geographic analysis, ranging from physical geography studies, approaches to sustainable planning to environmental analysis. Vegetation analysis according to the Indonesian National Standard (SNI 7645:2014) is classified based on density. The vegetation density index is divided into 4, namely non-vegetation, bare, medium and high. In the technical aspect to obtain information related to vegetation, this can be done using remote sensing. Remote sensing uses two data to obtain information, namely satellite data and UAV data. This study used UAV data with shooting locations in the Liang Anggang Protection Forest in classifying land cover. The method used was Convolutional Neural Network with feature extraction used in this study was GLCM. This research used the ShuffleNet v2 architecture on the CNN method. The findings of this study used two models, namely the CNN model without the GLCM process and compared to the CNN model with the addition of the GLCM process, resulting in a comparison that was quite far from the accuracy value obtained. The CNN model obtained an accuracy value of 80%, while the CNN model with GLCM using segmentation was 49.9% and without segmentation was 44.53%.

JournalTemplate (Done).docx
1166K

西南交通大学学报

第 57 卷 第 6 期
2022 年 12 月

JOURNAL OF SOUTHWEST JIAOTONG UNIVERSITY

Vol. 57 No. 6
Dec. 2022

ISSN: 0258-2724 **DOI：10.35741/issn.0258-2724.57.6.1**

Category
Please select the type of your manuscript
(see Author Guidelines)
Category.
Please select the most suitable field of scientific research
from the approved list of the journal categories

UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD

Novitasari Novitasari[a], Yuslena Sari[b,*], Yudi Firmanul Arifin[c], Nurul Fathanah Mustamin[d], Erika Maulidiya[e]

[a]Department of Civil Engineering, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, novitasari@ulm.ac.id
[b]Department of Information Technology, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, yuzlena@ulm.ac.id
[c]Faculty of Forestry, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, yudifirmanul@ulm.ac.id
[d]Department of Information Technology, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, nurul.mustamin@ulm.ac.id
[e]Department of Information Technology, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, 1810817220017@mhs.ulm.ac.id

Received: ▪ *Review:* ▪ *Accepted:* ▪ *Published*

Abstract

Land cover is an important factor in geographic analysis, ranging from physical geography studies, approaches to sustainable planning to environmental analysis. Vegetation analysis according to the Indonesian National Standard (SNI 7645:2014) is classified based on density. The vegetation density index is divided into 4, namely non-vegetation, bare, medium and high. In the technical aspect to obtain information related to vegetation, this can be done using remote sensing. Remote sensing uses two data to obtain information, namely satellite data and UAV data. This study used UAV data with shooting locations in the Liang Anggang Protection Forest in classifying land cover. The method used was Convolutional Neural Network with feature extraction used in this study was GLCM. This research used the ShuffleNet v2 architecture on the CNN method. The findings of this study used two models, namely the CNN model without the GLCM process and compared to the CNN model with the addition of the GLCM process, resulting in a comparison that was quite far from the accuracy value obtained. The CNN model obtained an accuracy value of 80%, while the CNN model with GLCM using segmentation was 49.9% and without segmentation was 44.53%.

Keywords: Classification; class; CNN; GLCM; accuracy.

摘要 The authors may not translate the abstract and keywords into Chinese themselves.

关键词:

I. INTRODUCTION

Peatland management in Indonesia has many challenges, due to peatland disaster, as flood and wildfire. Wildfire caused by human action is the biggest one. Many wildfires in Indonesia as intentional fires as part of residential developments. It is lead to change land cover. Land cover schange as one of internal factor as hydrological respon. From physical geography studies to approaches to sustainable planning to environmental analysis, land cover is an important factor in geographic analysis, especially in disaster mitigation in peatland. Environmental analysis needs surface vegetation-based land cover information [1]. The entire plant of an area that serves as a land cover is referred to as vegetation. Vegetation is the entire plant of an area that serves as a land cover. According to the Indonesian National Standard [2], vegetation analysis is classified based on density. Non-vegetation, bare, medium, and high vegetation density indexes are used [2]. In addition to determining the level of vegetation density on land cover, it is important to be able to distinguish vegetation density in the form of an image, which makes data processing easier. N.A. Harahap conducted research which provides an image of the classification of vegetation density classes based on the images shown in Figure 1.1.

(a) (b) (c) (d)

Figure 1. (a) Non Vegetation, (b) Bare Vegetation, (c) Medium Vegetation, (d) High Vegetation

Vegetation analysis is one method for studying the arrangement and composition of vegetation in terms of plant shape (structure). In terms of technology, remote sensing can be used to obtain information about vegetation. Remote sensing obtains information from two sources: satellite data and UAV data. Previous research that used remote sensing technology by utilizing satellite data resulted in data accuracy ranging from 63% - 85% using various methods [1], [3]. Because satellite data is a traditional format based on statistical reporting and sampling surveys, determining vegetation density is critical [3]. Remote sensing with satellite data has been widely used in the identification and classification of land cover patterns across a wide geographic coverage, but the use of satellite data, which has a high operating altitude and is easily influenced by weather, clouds, and other external factors, is being reconsidered. Remote sensing technology can quickly and precisely provide spatial information on the earth surface. The object being sensed, the sensor for recording the object, and the electronic waves emitted by the earth surface are the three main components of remote sensing.

Remote sensing technology can quickly and precisely provide spatial information on the earth surface. The object being sensed, the sensor for recording the object, and the electronic waves emitted by the earth surface are the three main components of remote sensing. As technology advances, remote sensing facilities such as the Unmanned Aerial Vehicle (UAV) become more practical and easier to implement. The emergence of UAV raises significant potential as a tool for environmental and ecological analysis, such as monitoring agricultural land, forest fires, arctic lichen distribution, and mapping of mangrove forests. The generation of spatial information based on aerial image data using drones has enormous potential for the advancement of remote sensing technology, such as area classification. The benefits of using a UAV include faster and more flexible data acquisition, results that are more real-time, and low and light operating and maintenance costs. Apart from the ability to fly through clouds and produce cloud-free images, it differs from satellite imagery, which is heavily influenced by atmospheric conditions. UAV imagery has a high resolution when compared to satellite imagery, reaching a spatial resolution of less than 1 cm, which is much more detailed than satellite (30cm) and aircraft (10cm) imagery [1]. Optimal results that can be obtained from the use of UAV in object classification and the appropriate method for processing data with UAV imagery.

Before being processed in a classification model, image data requires feature extraction techniques to determine certain characteristics possessed by images to aid in object identification (image analysis) [1], [4], [5]. The resulting features will be selected first in the feature extraction process to obtain features with a high influence as a reference for the classification process. The function of feature extraction is to extract the necessary information from an image. Shape, colour, and texture extraction are the three types of feature extraction. Images with a slight colour can benefit from feature extraction using the Gray Level Co-occurrence Matrix (GLCM) method, which is a second level statistical method that computes the frequency of pairs of pixels in an image that have the same gray level and applies the additional knowledge obtained through pixel spatial relationships [6]. Using edge information, the co-occurrence matrix embeds the distribution of grayscale transitions. The majority of the information required to calculate the threshold value in the GLCM technique is straightforward but efficient [7]. S. Karthikeyan and N. Rengarajan use the GLCM algorithm with up to 95% accuracy. Previous research has compared GLCM feature extraction to LBP, MI, CLBP, LBGLCM, and GLRLM, with the accuracy results proving that using feature extraction in classification using GLCM produces better results than using other methods. GLCM accuracy results range from 70% to 93% [7]–[9].

Visual interpretation methods, pixel-based digital classification methods, and object-based classification methods are used in land cover mapping based on remote sensing imagery. Land cover analysis researchers are interested in the use of data mining methods. Land classification, Machine Learning, and Deep Learning have all made extensive use of classification methods. Deep learning, which is included in the supervised classification, is developed and produced by the machine learning method. Deep learning methods are widely used in satellite image analysis because they are powerful and intelligent in image processing. Deep learning methods are still evolving, with the Convolutional Neural Network (CNN) deep learning method producing the most significant results in image recognition to date. Deep Learning has demonstrated that this architecture, particularly CNN, can learn human-level solutions to specific visual tasks. This method has been used extensively in remote sensing image analysis tasks such as object detection in images, image recording, scene classification, segmentation, object-based image analysis, and land use and land cover classification [10]. CNN is one of the most recent Deep Learning methods to emerge. This method has been shown to be useful for pattern recognition and object classification [1]. Previous research using the CNN method to classify land cover yielded satisfactory accuracy results ranging from 73% to 98% [1], [10], [11]

CNN has a variety of popular architectures, including LeNet5 (1998), AlexNet (2012), ZFNet (2013), GoogleNet (2014), ResNet (2015), FractalNet (2016), ShuffleNet (2018), and others. Previous research has compared the use of architecture on CNN in the field of classification. The compared architectures demonstrate the advantage and disadvantage of each, for architectures that are widely used in the field of image classification and are relatively new, and have been compared with several other architectures, ShuffleNet. ShuffleNet is a very efficient CNN architecture with fast accuracy. Research that has used the ShuffleNet architecture and has made comparisons with other architectures such as GoogleNet, DenseNet, MobileNet, Xception, IGCV2, EffNet V1, EffNet V2, IoTNet-3-5 and ResNet50 in the classification process states that the ShuffleNet architecture increases the accuracy of 82% - 98% with less memory usage and faster processing time [12]–[16].

The CNN method is widely used in the field of deep learning to conduct land cover classification. GLCM was used to extract features in this study. The ShuffleNet architecture on the CNN method will be used in this study. This research was carried out for a month in the Liang Anggang Protected Forest area, Banjarbaru block 1 area, with targeted data collection. The location for this study was chosen based on observations made during the observation and survey of the block 1 area, where, according to the 2017 Provincial Forestry Office, an area of 479 hectares of block 1 area is filled with land such as agriculture, plantations, roads and settlements, as well as 494 hectares of forest. In addition to being a peatland, the research site, particularly in block 1, meets the characteristics and suitability of the needs in collecting data for land cover classification in terms of vegetation density types (bare, medium, and high) that can be seen with the naked eye during observations and surveys. This study classified land cover, with a focus on vegetation density, and the research location was chosen in accordance with the data requirements. The objective of this study was to determine the results of the best deep learning methods in land cover classification based on vegetation density. This study created research updates by combining UAV data with shooting locations in the Liang Anggang Protected Forest.

II. RESEARCH METHODOLOGY

A. Research Site

This study was being conducted in the Liang Anggang Protected Forest in Banjarbaru City, South Kalimantan. This is the Tangi Timber KPHP's management area. The protected forest designation is based on Minister of Forestry Decree No. 672/Kpts-II/1991 and Kep Menhut No. 434/Kpts-II/1996 with a total area of 2,250 hectares divided into two protected forest blocks, namely block 1 covering an area of 960 hectares including Liang Anggang sub-district, Banjarbaru and block 2 covering an area of 1290 hectares including the Gambut District, Banjar Regency.

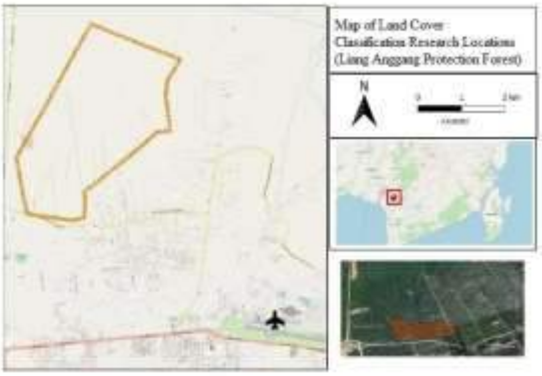


Figure 2. Map of the Liang Anggang Protected Forest Area

The study lasted one month, from November to December 2021, and focused on the Warning Area (lock signal area from the airport) that caused the drone to be unable to operate.

B. Research Procedure

This research was conducted in the Liang Anggang Protected Forest area by conducting a field survey to assess the state of the vegetation or areas within the Protected Forest area. This study collected image data using drones to capture images from a height of 20 meters over a one-month period. Land was assigned coordinates based on the goal of image data collection using Google Earth Pro tools. Land with coordinates was exported in .KML format and later imported into DroneDeploy (website) to make directing drone flights on land easier. Then, the imported KML file was configured for flight altitude and 2D or 3D image capture. An illustration of image capture is shown in Figure 3.

Figure 3. Illustration of Image Data Retrieval

Before proceeding to the next stage, image data that has been recorded and stored according to predetermined coordinate points was processed. To facilitate operation with the method that was used later, image data was labelled. The CNN method was used in this study. Image data that has already been processed was then fed into the classification process using the method used in this study. Image data was classified using each method, and the accuracy value was calculated using tools. The obtained accuracy value was then analysed and compared to draw conclusions. The flow of this research is shown in Figure 4.

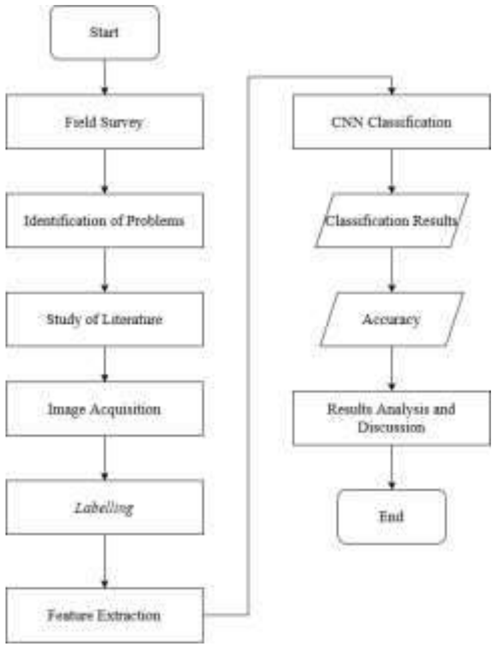


Figure 4. Diagram of Research Procedure

C. Feature Extraction

The purposes of feature extraction is to obtain the feature value of an object based on an image pixel intensity value relationship. The feature extraction process goal is to extract a special (unique) value from each image [17], [18]. This study used GLCM feature extraction with three primary features: correlation, homogeneity, and contrast. The feature extraction results created a GLCM version of the image using these three features. Figure 5 shows an illustration or description of the texture extraction results obtained with the GLCM feature.

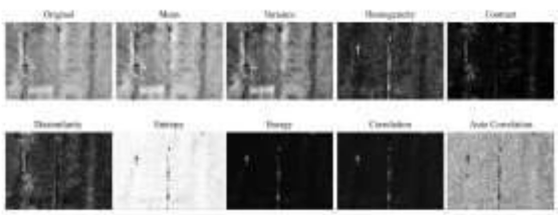


Figure 5. Result of Feature Extraction

The texture of an image was sought after by feature extracted images. The training data set consisted of 2400 images divided into three classes. This study applied 5 GLCM features to convert an input 2D image/image to an output 2D image/image to a gray level with a gray range of 0 to 1. The purpose of this step was to use gray level scaling to reduce the image volume to a more manageable size. Scaling to a grayscale level acted as a filter, removing some of the noise (de Mello, 2013). Figure 6 shows the scenario of the feature extraction test results with GLCM.

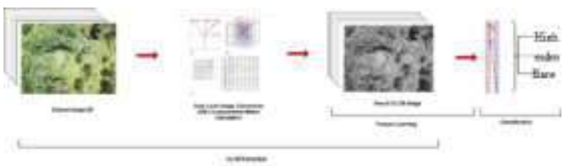


Figure 6. Ilustration of GLCM-CNN Feature Extraction

D. Classification of Convolutional Neural Networks

Only CNN neural networks can process grid structure data, such as two-dimensional images. The convolution layer is a linear algebra operation that generates a matrix of filters in the image to be processed. A convolution layer process is one of the many types of layers that can exist in a network. The image entered into the CNN classification model created during the fit model stage yielded an output calculated using the optimized weight. As a result, the classification model created should be able to classify the testing data into the correct class. This test was performed to calculate the accuracy value in the classification model that has been created. Figure 7 shows an illustration of the CNN classification process.

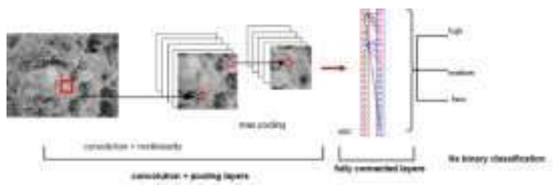


Figure 7. Illustration of CNN Classification Process

Figure 8 shows the classification flow using the CNN method

Figure 8. Stage of CNN Classification

E. Classification Analysis

The results of the UAV image classification using two methods were analysed and the level of accuracy was determined. This study applied accuracy testing with a confusion matrix in the form of overall accuracy (OA) and Kappa coefficient accuracy. Proceed with the analysis of the CNN method classification results to obtain accurate results from the use of the CNN method in land cover classification.

III. TESTING

A. Image Dataset

The dataset used in this study was divided into three categories: bare, medium, and high. The total number of images collected was 3000, with 1000 for each class type category. The classification of these three classes was based on the condition of the Liang Anggang Protected Forest where the research location, particularly block 1, meets the characteristics and suitability of the needs in collecting data for land cover classification in terms of vegetation density types (bare, medium, and high) that can be seen with the naked eye during observation and surveys. This study classified land cover, with a focus on vegetation density, and the research location was chosen in accordance with the data requirements. Table 1 shows the results of categorizing three classes of vegetation density in terms of images based on the division of the available dataset [19].

Table 1.
Image of Vegetation Density

Image	Type of Vegetation Density
	Bare
	Medium
	High

B. Image Cropping

Because the image data obtained with the drone was too large, the data was resized by cutting the image and selecting specific areas to be used as training data. Cropped image data aimed to facilitate the classification process, did not take up much space or memory, and the classification process was light, so it did not require a long time in the classification process later. The image data was cropped to 256 x 256 pixels, reducing the image size to 159 KB. The cropped image data was classified into three types: bare, medium, and dense/high [20], [21]. Figure 9 shows the cropping results of image data.

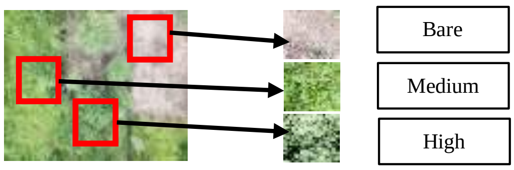


Figure 9. Image Data Cropping

C. Segmentation

Image segmentation was used to distinguish between objects and backgrounds [22], [23]. The separation process was designed to make classification and calculations easier. The image segmentation process was based on the difference in the image grayscale. To convert a colour image with r, g, and b matrix values into a grayscale image. The segmentation method, namely thresholding, can be used to change the colour image. The most basic method for segmenting was image development or image thresholding (de Mello, 2013). Thresholding was used to change the number of gray degrees in an image in order to create a binary image with pixel intensity values of 0 or 1.

D. Feature Extraction (GLCM)

In this study, GLCM was used for feature extraction, with three main features used: correlation, homogeneity, and contrast. This method was used to classify images, recognize textures, segment them, recognize objects, and analyse their colours. In the neighbourhood between pixels, GLCM had four angular directions: 0°, 45°, 90°, and 135°. When the angle was 0°, the pixel density was calculated by moving one distance to the right. Pixel adjacency was calculated using a 45° angle and 1 pixel distance to the top right. The angle is 90°, and the pixel density was calculated by a 1 pixel distance on top. A 135° angle was used, and neighbouring pixels were calculated by moving one pixel up [24]. The gray level of pixels was compared based on angle or neighbours at 0°, 45°, 90°, and 135° in this study. The feature extraction process was also compared the results of previously segmented images to those that have not been segmented.

IV. Findings

A. Result of CNN Model

The formation of network architecture in the CNN algorithm can affect the results of model accuracy. In order to produce an optimal model, network architecture was used during the training process. This study applied an input image with a resolution of 256x256x3, with the aim for reducing image size so that the classification process took as little time as possible. This study applied the second version of the ShuffleNet architecture, which included one convolutional layer (Conv5), three stages (consisting of convolutional and shuffle units), one pooling layer (using Maxpool), and fc. The input image in the shuffleNet v2 model was 256 x 256 in size. The convolution and maximum pooling layers were added to the model's initial position to reduce the size of the feature graph. The convolution layer and pooling layer were replaced at the initial position by the convolutional layer (Conv1) with a 3 x 3 kernel, and the BN layer was added after Conv 1 and Conv 5. Figure 10 shows the flow of the proposed model.

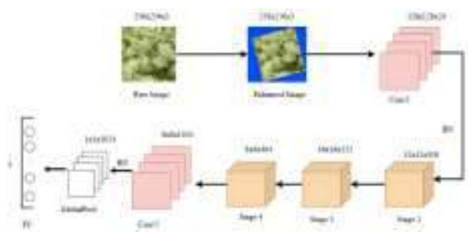


Figure 10. Flow of Model Shuffle Net v2 Model

The results of training and testing accuracy were obtained after going through several processes in the CNN algorithm. The value for training accuracy can be found in the "Accuracy" column, while the value for testing accuracy can be found in the "Validation Accuracy" column. Accuracy was the value calculated by calculating the accuracy of the training dataset and model predictions. Validation Accuracy is the value calculated by calculating the accuracy of the validation dataset and predictions from the model using validation dataset input data. This procedure used a total of 50 epochs. Previously, epoch comparisons were performed to determine the accuracy and validation results of each training with a different number of epochs. This epoch comparison was intended to find the best model. The number of epochs compared ranges between 25 and 100 [25]. The use of epoch had a significant impact on the resulting accuracy. Because epoch can improve accuracy and the resulting accuracy was stable, it was critical to use the correct epoch in training data to achieve maximum accuracy. The table below compares training results based on the number of epochs.

Table 2.
Comparison of Epoch

Epoch	**Accuracy**	**Loss**	**Accuracy Validation**	***Validation Loss***	**Time**
25	97.46%	0.0686	71%	1.9176	20minute
50	98.75%	0.0358	81.33%	1.2536	56minute
75	99.08%	0.0279	74%	2.4516	1hour 32minute
100	99.29%	0.0218	74.17%	1.1251	2hour 20minute

The training model's accuracy with a total of 50 epochs is 98.75% with a loss of 0.0218. The validation accuracy value for the 50 epochs is 81.33%, which is higher than the other epochs. According to the table, the closer to the highest epoch, the higher the accuracy obtained from the testing results. However, if more than 100 epochs are added, the accuracy value decreases because too many epochs can also affect the large number of datasets. The testing procedure used training data consisting of 2400 image data and 600 image data for each class, as well as 200 image data for each class. Table 3 shows the results of the confusion matrix.

Matrix		***Predict Class***		
		Bare	Medium	High
Actual Class	Bare	209	0	4
	Medium	9	92	88
	High	8	12	178

Table 3.
Confusion Matrix

Based on the results of table 3, the model's predictions on the new data testing data show promising results. Although the prediction of the Bare class is correctly classified as the Bare class, up to four miss classifications from the Bare image data input are classified as the High class. While the Medium class prediction is correctly classified as the Medium class, as many as 9 miss classifications from the input image data are classified as the Bare class. In addition, up to 88 misclassifications of input image data are classified as High. The High class prediction is correctly classified as the High class, but up to 8 miss classifications from the High image data input are classified as the Bare class. As many as 12 misclassifications of input image data were classified as Medium. The overall accuracy of the matrix and kappa accuracy are calculated as follows:

Overall Accuracy = 469/600 = 80%
Kappa = 70%

So, the model's accuracy with a 256x256 input image and a total of 600 image data obtained an accuracy value of 80% and a kappa accuracy of 70%.

B. Result of CNN Model with GLCM

The addition of three GLCM features, namely contrast, homogeneity, and correlation, is the result of the next training model. The procedure involved extracting 3,000 GLCM result image data and producing 9,000 image data that was processed by CNN. This study also compared the direction angles of 0°, 45°, 90°, and 135° to extract images per angle. The GLCM process used a total of 27,000 image data through the segmentation stage. This was done to determine how well each feature performed in the image classification process.

For the GLCM process with CNN going through the segmentation stage, the results of data training with the CNN model and each GLCM feature by going through the segmentation stage with 9,000 data for each angle, can be seen at an angle of 135° getting the highest validation accuracy value from other angles, namely 60.11% with a value validation loss of 0.8460. For each feature, this training procedure applied a total of 50 epochs. This training process took approximately 20-30 minutes per corner. Table 5 shows the results of the GLCM training data per corner.

Table 4.
Comparison of Training Per Angle

Angle	**Accu racy**	**Loss**	***Valid ation Accu racy***	***Valid ation Loss***	**Time**
0°	95.24 %	0.137 7	50.28 %	0.609 6	29 mnt 6 scnd
45°	94.99 %	0.134 6	50.39 %	0.552 6	31 mnt 20 scnd
90°	96.42 %	0.105 5	59.94 %	0.761 6	31 mnt 12 scnd
	96.26	0.117	60.11	0.846	31 mnt

Matrix		***Predict Class***		
		Bare	Medium	High
Actual Class	Bare	407	92	83
	Medium	54	419	137
	High	83	218	307

The training data is 9,000 images, and the test data is 1,800 images, with 3,000 images in each class. Table 5 shows the confusion matrix results for the CNN model process with GLCM that went through the segmentation stage.

Table 5.
Confusion Matrix

According to the results in table 5, the model's prediction results on new data testing data are poor. Although the prediction of the Bare class is correct, as many as 92 miss classifications from the Bare image data input are classified as Medium. In addition, up to 83 miss classifications from the Bare image data input are classified as High. While the Medium class prediction was correctly classified as the Medium class, as many as 54 miss classifications from the input image data were classified as the Bare class. In addition, 137 misclassifications of image data input Medium are classified as High. The High class prediction is correctly classified as the High class, but up to 83 miss classifications from the High image data input are classified as the Bare class. In addition, 218 miss classifications from the High image data input are classified as Medium. The overall accuracy of the matrix and kappa accuracy are calculated as follows:

Overall Accuracy = 1133/1800x100% = 62,99%

Kappa = 44,37%

So, with an input image of 256x256 pixels and a total of 1800 image data, the model produced an accuracy value of 62.99% and a kappa accuracy of 44.37%.

V. DISCUSSION

When the CNN model without the GLCM process was compared to the CNN model with the GLCM process, the comparison was quite far from the accuracy values obtained. The CNN model achieved an accuracy of 80%, while the CNN model with GLCM achieves 62.99% segmentation. This showed that the CNN model outperformed the GLCM process. According to the findings of the analysis, this occurred because the gray level in the image was leveled during the GLCM process, resulting in white and black colors in the image. The colours in the original image changed to white and black, resulting in a classification error. The GLCM process rendered the image colourless and rendered the entire image black.

During the testing of new data, there was a misclassification caused by nearly identical vegetation types. The input data for the CNN model was original image data with different types of vegetation, but based on the researcher's analysis, even though the texture between medium and high vegetation was different, the CNN model still had difficulty distinguishing and recognizing medium and high classes if the data simultaneously has the characteristics of an image that was filled with vegetation even though the type and texture of the vegetation was different. The CNN model with the GLCM method had a lot of misclassifications. The first reason was that the original image's colour had changed, making it

difficult for the model to distinguish between classes. The second issue was that the type and texture of the vegetation were not visible in the image, so when predicting with the CNN and GLCM models on prototypes, the bare class data was read as medium class. High class reads as medium class.

VI. CONCLUSION

The conclusion is that comparing the CNN model without the GLCM process to the CNN model with the GLCM process produces a comparison that is quite far from the accuracy value obtained. The CNN model achieves an accuracy of 80%, while the CNN model with GLCM achieves 62.99% segmentation. This demonstrates that the CNN model outperforms the GLCM process in land cover classification. This demonstrates that the image processing process has a significant impact on the classification and prediction stages.

ACKNOWLEDGMENT

This research was funded by DRTPM grant, National Basic Research Competition scheme, agreement number 113/E5/PG.02.00/PT/2023.

REFERENCES

[1] Z. Xu, K. Guan, N. Casler, B. Peng, and S. Wang, "A 3D convolutional neural network method for land cover classification using LiDAR and multi-temporal Landsat imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 144, pp. 423–434, 2018, doi: 10.1016/j.isprsjprs.2018.08.005.

[2] S. N. Indonesia and B. S. Nasional, "SNI: Klasifikasi penutup lahan," 2014.

[3] F. Zhao, X. Wu, and S. Wang, "Object-oriented Vegetation Classification Method based on UAV and Satellite Image Fusion," *Procedia Comput. Sci.*, vol. 174, no. 2019, pp. 609–615, 2020, doi: 10.1016/j.procs.2020.06.132.

[4] M. Alkaff, H. Khatimi, W. Puspita, and Y. Sari, "Modelling and predicting wetland rice production using support vector regression," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 17, no. 2, pp. 819–825, 2019, doi: 10.12928/TELKOMNIKA.V17I2.10145.

[5] Y. Sari, E. S. Wijaya, A. R. Baskara, and R. S. D. Kasanda, "PSO optimization on backpropagation for fish catch production prediction," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 18, no. 2, pp. 776–782, 2020, doi: 10.12928/TELKOMNIKA.V18I2.14826.

[6] Q. Wu, Y. Gan, B. Lin, Q. Zhang, and H. Chang, "An active contour model based on fused texture features for image segmentation," *Neurocomputing*, vol. 151, no. P3, pp. 1133–1141, 2015, doi: 10.1016/j.neucom.2014.04.085.

[7] Z. Xing and H. Jia, "Multilevel Color Image Segmentation Based on GLCM and Improved Salp Swarm Algorithm," *IEEE Access*, vol. 7, pp. 37672–37690, 2019, doi: 10.1109/ACCESS.2019.2904511.

[8] S. A. Alazawi, N. M. Shati, and A. H. Abbas, "Texture features extraction based on GLCM for face retrieval system," *Period. Eng. Nat. Sci.*, vol. 7, no. 3, pp. 1459–1467, 2019, doi: 10.21533/pen.v7i3.787.

[9] S. Ozturk and B. Akdemir, "Application of Feature Extraction and Classification Methods for Histopathological Image using GLCM, LBP, LBGLCM, GLRLM and SFTA," *Procedia Comput. Sci.*, vol. 132, no. Iccids, pp. 40–46, 2018, doi: 10.1016/j.procs.2018.05.057.

[10] O. Youme, T. Bayet, J. M. Dembele, and C. Cambier, "Deep Learning and Remote Sensing: Detection of Dumping Waste Using UAV," *Procedia Comput. Sci.*, vol. 185, no. June, pp. 361–369, 2021, doi: 10.1016/j.procs.2021.05.037.

[11] S. M. Hamylton *et al.*, "Evaluating techniques for mapping island vegetation from unmanned aerial vehicle (UAV) images: Pixel classification, visual interpretation and machine learning approaches," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 89, no. March, p. 102085, 2020, doi: 10.1016/j.jag.2020.102085.

[12] T. Lawrence and L. Zhang, "IoTNet:

An efficient and accurate convolutional neural network for IoT devices," *Sensors (Switzerland)*, vol. 19, no. 24, 2019, doi: 10.3390/s19245541.

[13] G. Liu *et al.*, "I3d-shufflenet based human action recognition," *Algorithms*, vol. 13, no. 11, 2020, doi: 10.3390/a13110301.

[14] G. Losapio *et al.*, "Lightweight and efficient convolutional neural networks for recognition of dolphin dorsal fins," pp. 68–72, 2020.

[15] Z. Wang and L. Ma, "SYOLO: An Efficient Pedestrian Detection," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 768, no. 7, 2020, doi: 10.1088/1757-899X/768/7/072067.

[16] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 6848–6856, 2018.

[17] Y. Sari, M. Alkaff, and R. A. Pramunendar, "Iris recognition based on distance similarity and PCA," *AIP Conf. Proc.*, vol. 1977, 2018, doi: 10.1063/1.5042900.

[18] Y. Sari, M. Alkaff, and M. Maulida, "Classification of Rice Leaf using Fuzzy Logic and Hue Saturation Value (HSV) to Determine Fertilizer Dosage," in *2020 Fifth International Conference on Informatics and Computing (ICIC)*, 2020, pp. 1–6. doi: 10.1109/ICIC50835.2020.9288585.

[19] Y. Sari, Y. F. Arifin, N. Novitasari, and M. R. Faisal, "Vegetation-Density Drone Dataset For Peatland Vegetation Classification," vol. 1, 2022, doi: 10.17632/TB26ZY2JST.1.

[20] Y. Sari, Y. Arifin, Novitasari, and M. Faisal, "Implementation of Deep Learning Based Semantic Segmentation Method To Determine Vegetation Density," *Eastern-European J. Enterp. Technol.*, vol. 5, no. 2–119, pp. 42–54, 2022, doi: 10.15587/1729-4061.2022.265807.

[21] Y. Sari, Y. F. Arifin, N. Novitasari, and M. R. Faisal, "Effect of Feature Engineering Technique for Determining Vegetation Density," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 7, pp. 655–661, 2022, doi: 10.14569/IJACSA.2022.0130776.

[22] Y. Sari, A. R. Baskara, and R. Wahyuni, "Classification of Chili Leaf Disease Using the Gray Level Co-occurrence Matrix (GLCM) and the Support Vector Machine (SVM) Methods," *2021 6th Int. Conf. Informatics Comput. ICIC 2021*, 2021, doi: 10.1109/ICIC54025.2021.9632920.

[23] Y. Sari, H. Suhud, A. R. Baskara, R. A. Pramunendar, and I. F. Radam, "Parking Lots Detection in Static Image Using Support Vector Machine Based on Genetic Algorithm," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 6, pp. 476–487, 2021, doi: 10.22266/ijies2021.1231.42.

[24] R. A. Pramunendar, D. P. Prabowo, D. Pergiwati, Y. Sari, P. N. Andono, and M. A. Soeleman, "New workflow for marine fish classification based on combination features and CLAHE enhancement technique," *Int. J. Intell. Eng. Syst.*, vol. 13, no. 4, pp. 293–304, 2020, doi: 10.22266/IJIES2020.0831.26.

[25] C. A. B. de Mello, "Image thresholding," *Digit. Doc. Anal. Process.*, vol. 2006, no. Snati, pp. 71–98, 2013, doi: 10.1201/9781003082224-3.

参考文:

COVER LETTER

UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD
Land cover is an important factor in geographic analysis, ranging from physical geography studies, approaches to sustainable planning to environmental analysis. Vegetation analysis according to the Indonesian National Standard (SNI 7645:2014) is classified based on density. The vegetation density index is divided into 4, namely non-vegetation, bare, medium and high. In the technical aspect to obtain information related to vegetation, this can be done using remote sensing. Remote sensing uses two data to obtain information, namely satellite data and UAV data. This study used UAV data with shooting locations in the Liang Anggang Protection Forest in classifying land cover. The method used was Convolutional Neural Network with feature extraction used in this study was GLCM. This research used the ShuffleNet v2 architecture on the CNN method. The findings of this study used two models, namely the CNN model without the GLCM process and compared to the CNN model with the addition of the GLCM process, resulting in a comparison that was quite far from the accuracy value obtained. The CNN model obtained an accuracy value of 80%, while the CNN model with GLCM using segmentation was 49.9% and without segmentation was 44.53%.
Classification; class; CNN; GLCM; accuracy.

Type of manuscript (please specify):

☐ Research article
☐ Review article
☐ Brief report
☐ Short communication
☐ Research note

Yuslena Sari Jl. Brig. Hasan Basry Kayutangi Banjarmasin, Indonesia, 70123	
Telephone# +6285247175500	Fax#
Email yuzlena@ulm.ac.id	

☐ I hereby confirm that the manuscript was prepared in accordance with the instructions for authors of scientific publications, and that the content of this manuscript, or most of it, was not published in the journal indicated, and the manuscript was not submitted for publication elsewhere.

☐ I hereby confirm my consent to pay the Article Processing Charges (APC) EUR 450 in the case of the manuscript acceptance for publication. I am aware and hereby confirm that the APC is non-refundable.

☐ I hereby confirm my consent to make the payment for English Language Editing Services (EUR 150).

Publication fee payer details

____

____

Signature of the Corresponding author **Date**

Copyright Agreement

Manuscript title: UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD

Full names of all authors: Novitasari Novitasari, Yuslena Sari, Yudi Firmanul Arifin, Nurul Fathanah Mustamin, Erika Maulidiya

Full name and address of the corresponding author:

Yuslena Sari

Telephone/Whatsap: +6285247175500 Fax: _________________ Email: yuzlena@ulm.ac.id_

License Agreement

(1) Authors own all the copyright rights for the paper.
(2) Submitted manuscript is an original paper.
(3) Authors hereby grant the Issues of Journal of SWJTU with an exclusive, royalty-free, worldwide license to email the paper to all who will ask for it.
(4) All authors have made a significant contribution to the research and are ready to assume joint responsibility for the paper.
(5) All authors have seen and approved the manuscript in the final form as it is submitted for publication.
(6) This manuscript has not been published and also has neither been submitted nor considered for publication elsewhere
(7) The text, illustrations and any other materials, included into the manuscript, do not infringe any existing intellectual property rights or other rights of any person or entity.
(8) The editors of the Issues of Journal of SWJTU, its personnel or the Editorial Board members accept no responsibility for the quality of the idea expressed in this publication.

I am the Corresponding author and have full authority to enter into this agreement.

Full name, affiliation and position: Yuslena Sari, Department of Information Technology, Universitas Lambung Mangkurat

Signature:____Date: 28 April 2023

Reviewers:

The topic is interesting and important. However, there are several key areas that need more work.

1 - In the introduction section, please highlight the purpose and novelty of this particular study.

2 - Authors should explain better the criteria for selecting the research object, and the limitations of the application of the results obtained should be specified.

3 - Please cite sources under the tables. If they were developed by the authors, indicate "developed by the authors".

4 - The Conclusion section needs to justify the effectiveness of the approaches you propose.

Authors response:

1. This manuscript's purpose and novelty are based on the only research on tropical peatland correlation with vegetation density in some peatland wildfires in Indonesia.
2. The limitations of this research focus on vegetation density in tropical peatlands. Many CNN models using in other classifications, but It is the first time using this model classified vegetation density.
3. Batasan penelitian dijabarkan (misalnya metode ini khusus utk apa?saya ga ngerti teori CNN nya
4. In conclusion, the authors add that this CNN model can be used to classify vegetation density in tropical peatlands.

西南交通大学学报

第 57 卷 第 6 期
2022 年 12 月

JOURNAL OF SOUTHWEST JIAOTONG UNIVERSITY

Vol. 57 No. 6
Dec. 2022

ISSN: 0258-2724 **DOI：10.35741/issn.0258-2724.57.6.1**

Category
Please select the type of your manuscript
(see Author Guidelines)
Category.
Please select the most suitable field of scientific research
from the approved list of the journal categories

UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD

Novitasari Novitasari[a], Yuslena Sari[b,]*, Yudi Firmanul Arifin[c], Nurul Fathanah Mustamin[d], Erika Maulidiya[e]

[a]Department of Civil Engineering, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, novitasari@ulm.ac.id
[b]Department of Information Technology, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, yuzlena@ulm.ac.id
[c]Faculty of Forestry, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, yudifirmanul@ulm.ac.id
[d]Department of Information Technology, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, nurul.mustamin@ulm.ac.id
[e]Department of Information Technology, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, 1810817220017@mhs.ulm.ac.id

Received: ▪ Review: ▪ Accepted: ▪ Published

Abstract

Land cover or vegetation density in tropical peatland is an essential factor in hydrology response in geographic analysis, ranging from physical geography studies and approaches to sustainable planning to environmental research. Vegetation analysis according to the Indonesian National Standard (SNI 7645:2014), is classified based on density. The vegetation density index is divided into 4, namely non-vegetation, bare, medium, and high. In the technical aspect, to obtain information related to vegetation, this can be done using remote sensing. Remote sensing uses two data to obtain information, namely satellite data and UAV data. This study used UAV data with shooting locations in the Liang Anggang Protection Forest in classifying land cover. The method used was Convolutional Neural Network with feature extraction used in this study was GLCM. This research used the ShuffleNet v2 architecture on the CNN method. The findings of this study used two models, namely the CNN model without the GLCM process and compared to the CNN model with the addition of the GLCM process, resulting in a comparison that was quite far from the accuracy value obtained. The CNN model obtained an accuracy value of 80%, while the CNN model with GLCM using segmentation was 49.9%, and without segmentation was 44.53%.

Keywords: Tropical Peatland, Vegetation Density, Classification; class; CNN; GLCM; accuracy.

摘要 The authors may not translate the abstract and keywords into Chinese themselves.

关键词:

I. INTRODUCTION

Peatland management in Indonesia has many challenges due to peatland disasters [1], such as floods and wildfires. Wildfire caused by human action is the biggest one. Many wildfires in Indonesia as intentional fires as part of residential developments [2]. It is leads to changes land cover as changes in vegetation density in tropical peatland. Land cover or vegetation density change is one of the internal factors of hydrological response. From physical geography studies to approaches to sustainable planning to environmental analysis, land cover is an essential factor in the geographic analysis, especially in disaster mitigation in tropical peatlands. Environmental analysis needs surface vegetation-based land cover information [3]. The entire plant of an area that serves as a land cover is referred to as vegetation. Vegetation is the entire plant of an area that serves as a land cover. According to the Indonesian National Standard [4], vegetation analysis is classified based on density. Non-vegetation, bare, medium, and high vegetation density indexes are used [4]. In addition to determining the level of vegetation density, it is important to be able to distinguish vegetation density in the form of an image, which makes data processing easier. N.A. Harahap (year) conducted research that provides an image of the classification of vegetation density classes based on the images shown in Figure 1.1.

(a) (b) (c) (d)

Figure 1. (a) Non Vegetation, (b) Bare Vegetation, (c) Medium Vegetation, (d) High Vegetation

Vegetation density analysis in tropical peatland is one method for studying the arrangement and composition of vegetation in terms of plant shape (structure). In terms of technology, remote sensing can be used to obtain information about vegetation. Remote sensing obtains information from two sources: satellite data and UAV data. Previous research that used remote sensing technology by utilizing satellite data resulted in data accuracy ranging from 63% - 85% using various methods [3], [5]. Because satellite data is a traditional format based on statistical reporting and sampling surveys, determining vegetation density is critical [5]. Remote sensing with satellite data has been widely used in the identification and classification of land cover patterns across a wide geographic coverage, but the use of satellite data, which has a high operating altitude and is easily influenced by weather, clouds, and other external factors, is being reconsidered. Remote sensing technology can quickly and precisely provide spatial information on the earth surface. The object being sensed, the sensor for recording the object, and the electronic waves emitted by the earth surface are the three main components of remote sensing.

Remote sensing technology can quickly and precisely provide spatial information on the earth surface. The object being sensed, the sensor for recording the object, and the electronic waves emitted by the earth surface are the three main components of remote sensing. As technology advances, remote sensing facilities such as the Unmanned Aerial Vehicle (UAV) become more practical and easier to implement. The emergence of UAV raises significant potential as a tool for environmental and ecological analysis, such as monitoring agricultural land, forest fires, arctic lichen distribution, and mapping of mangrove forests. The generation of spatial information based on aerial image data using drones has enormous potential for the advancement of remote sensing technology, such as area classification. The benefits of using a UAV include faster and more flexible data acquisition, results that are more real-time, and low and light operating and maintenance costs. Apart from the ability to fly through clouds and produce cloud-free images, it differs from satellite imagery, which is heavily influenced by atmospheric conditions. UAV imagery has a high resolution when compared to satellite imagery, reaching a spatial resolution of less than 1 cm, which is much more detailed than satellite (30cm) and aircraft (10cm) imagery [3]. Optimal results that can be obtained from the use of UAV in object classification and the appropriate method for processing data with UAV imagery.

Before being processed in a classification model, image data requires feature extraction techniques to determine certain characteristics possessed by images to aid in object identification (image analysis) [3], [6], [7]. The resulting features will be selected first in the feature extraction process to obtain features with a high influence as a reference for the classification process. The function of feature extraction is to extract the necessary information from an image. Shape, colour, and texture extraction are the three types of feature extraction. Images with a slight colour can benefit from feature extraction using the Gray Level Co-occurrence Matrix (GLCM) method, which is a second level statistical method that computes the frequency of pairs of pixels in an image that have the same gray level and applies the additional knowledge obtained through pixel spatial relationships [8]. Using edge information, the co-occurrence matrix embeds the distribution of grayscale transitions. The majority of the information required to calculate the threshold value in the GLCM technique is straightforward but efficient [9]. S. Karthikeyan and N. Rengarajan use the GLCM algorithm with up to 95% accuracy. Previous research has compared GLCM feature extraction to LBP, MI, CLBP, LBGLCM, and GLRLM, with the accuracy results proving that using feature extraction in classification using GLCM produces better results than using other methods. GLCM accuracy results range from 70% to 93% [9]–[11].

Visual interpretation methods, pixel-based digital classification methods, and object-based classification methods are used in land cover mapping based on remote sensing imagery. Land cover analysis researchers are interested in the use of data mining methods. Land classification, Machine Learning, and Deep Learning have all made extensive use of classification methods. Deep learning, which is included in the supervised classification, is developed and produced by the machine learning method. Deep learning methods are widely used in satellite image analysis because they are powerful and intelligent in image processing. Deep learning methods are still evolving, with the Convolutional Neural Network (CNN) deep learning method producing the most significant results in image recognition to date. Deep Learning has demonstrated that this architecture, particularly CNN, can learn human-level solutions to specific visual tasks. This method has been used extensively in remote sensing image analysis tasks such as object detection in images, image recording, scene classification, segmentation, object-based image analysis, and land use and land cover classification [12]. CNN is one of the most recent Deep Learning methods to emerge. This method has been shown to be useful for pattern recognition and object classification [3]. Previous research using the CNN method to classify land cover yielded satisfactory accuracy results ranging from 73% to 98% [3], [12], [13]

CNN has a variety of popular architectures, including LeNet5 (1998), AlexNet (2012), ZFNet (2013), GoogleNet (2014), ResNet (2015), FractalNet (2016), ShuffleNet (2018), and others. Previous research has compared the use of architecture on CNN in the field of classification. The compared architectures demonstrate the advantage and disadvantage of each, for architectures that are widely used in the field of image classification and are relatively new, and have been compared with several other architectures, ShuffleNet. ShuffleNet is a very efficient CNN architecture with fast accuracy. Research that has used the ShuffleNet architecture and has made comparisons with other architectures such as GoogleNet, DenseNet, MobileNet, Xception, IGCV2, EffNet V1, EffNet V2, IoTNet-3-5 and ResNet50 in the classification process states that the ShuffleNet architecture increases the accuracy of 82% - 98% with less memory usage and faster processing time [14]–[18].

The CNN method is widely used in the field of deep learning to conduct land cover classification. GLCM was used to extract features in this study. The ShuffleNet architecture on the CNN method will be used in this study. This research was carried out for a month in the Liang Anggang Protected Forest area, Banjarbaru block 1 area, with targeted data collection. The location for this study was chosen based on observations made during the observation and survey of the block 1 area, where, according to the 2017 Provincial Forestry Office, an area of 479 hectares of block 1 area is filled with land such as agriculture, plantations, roads and settlements, as well as 494 hectares of forest. In addition to being a peatland, the research site, particularly in block 1, meets the characteristics and suitability of the needs in collecting data for land cover classification in terms of vegetation density types (bare, medium, and high) that can be seen with the naked eye during observations and surveys. This study classified land cover, with a focus on vegetation density, and the research location was chosen in accordance with the data requirements. The objective of this study was to determine the results of the best deep learning methods in land cover classification based on vegetation density. This study created research updates by combining UAV data with shooting locations in the Liang Anggang Protected Forest.

II. RESEARCH METHODOLOGY

A. Research Site

This study was being conducted in the Liang Anggang Protected Forest in Banjarbaru City as the biggest wildfire in tropical peatland in South Kalimantan. This is the Tangi Timber KPHP's management area. The protected forest designation is based on Minister of Forestry Decree No. 672/Kpts-II/1991 and Kep Menhut No. 434/Kpts-II/1996 with a total area of 2,250 hectares divided into two protected forest blocks, namely block 1 covering an area of 960 hectares including Liang Anggang sub-district, Banjarbaru and block 2 covering an area of 1290 hectares including the Gambut District, Banjar Regency.

Figure 2. Map of the Liang Anggang Protected Forest Area

The study lasted one month, from November to December 2021, and focused on the Warning Area (lock signal area from the airport) that caused the drone to be unable to operate.

B. Research Procedure

This research was conducted in the Liang Anggang Protected Forest area by conducting a field survey to assess the state of the vegetation or areas within the Protected Forest area. This study collected image data using drones to capture images from a height of 20 meters over a one-month period. Land was assigned coordinates based on the goal of image data collection using Google Earth Pro tools. Land with coordinates was exported in .KML format and later imported into DroneDeploy (website) to make directing drone flights on land easier. Then, the imported KML file was configured for flight altitude and 2D or 3D image capture. An illustration of image capture is shown in Figure 3.

Figure 3. Illustration of Image Data Retrieval

Before proceeding to the next stage, image data that has been recorded and stored according to predetermined coordinate points was processed. To facilitate operation with the method that was used later, image data was labelled. The CNN method was used in this study. Image data that has already been processed was then fed into the classification process using the method used in this study. Image data was classified using each method, and the accuracy value was calculated using tools. The obtained accuracy value was then analysed and compared to draw conclusions. The flow of this research is shown in Figure 4.

Figure 4. Diagram of Research Procedure

C. Feature Extraction

The purposes of feature extraction is to obtain the feature value of an object based on an image pixel intensity value relationship. The feature extraction process goal is to extract a special (unique) value from each image [19], [20]. This study used GLCM feature extraction with three primary features: correlation, homogeneity, and contrast. The feature extraction results created a GLCM version of the image using these three features. Figure 5 shows an illustration or description of the texture extraction results obtained with the GLCM feature.

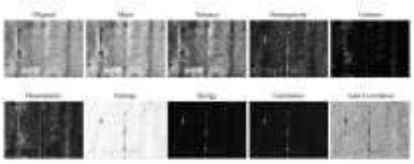


Figure 5. Result of Feature Extraction

The texture of an image was sought after by feature extracted images. The training data set consisted of 2400 images divided into three classes. This study applied 5 GLCM features to convert an input 2D image/image to an output 2D image/image to a gray level with a gray range of 0 to 1. The purpose of this step was to use gray level scaling to reduce the image volume to a more manageable size. Scaling to a grayscale level acted as a filter, removing some of the noise (de Mello, 2013). Figure 6 shows the scenario of the feature extraction test results with GLCM.

Figure 6. Ilustration of GLCM-CNN Feature Extraction

D. Classification of Convolutional Neural Networks

Only CNN neural networks can process grid structure data, such as two-dimensional images. The convolution layer is a linear algebra operation that generates a matrix of filters in the image to be processed. A convolution layer process is one of the many types of layers that can exist in a network. The image entered into the CNN classification model created during the fit model stage yielded an output calculated using the optimized weight. As a result, the classification model created should be able to classify the testing data into the correct class. This test was performed to calculate the accuracy value in the classification model that has been created. Figure 7 shows an illustration of the CNN classification process.

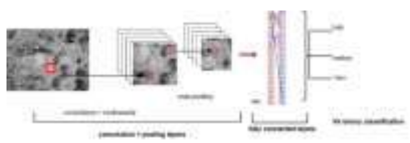


Figure 7. Illustration of CNN Classification Process

Figure 8 shows the classification flow using the CNN method

Figure 8. Stage of CNN Classification

E. Classification Analysis

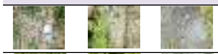
The results of the UAV image classification using two methods were analysed and the level of accuracy was determined. This study applied accuracy testing with a confusion matrix in the form of overall accuracy (OA) and Kappa coefficient accuracy. Proceed with the analysis of the CNN method classification results to obtain accurate results from the use of the CNN method in land cover classification.

III. TESTING

A. Image Dataset

The dataset used in this study was divided into three categories: bare, medium, and high. The total number of images collected was 3000, with 1000 for each class type category. The classification of these three classes was based on the condition of the Liang Anggang Protected Forest where the research location, particularly block 1, meets the characteristics and suitability of the needs in collecting data for land cover classification in terms of vegetation density types (bare, medium, and high) that can be seen with the naked eye during observation and surveys. This study classified land cover, with a focus on vegetation density, and the research location was chosen in accordance with the data requirements. Table 1 shows the results of categorizing three classes of vegetation density in terms of images based on the division of the available dataset [21].

Table 1.
Image of Vegetation Density

Image	Type of Vegetation Density
	Bare
	Medium
	High

Source:

B. Image Cropping

Because the image data obtained with the drone was too large, the data was resized by cutting the image and selecting specific areas to be used as training data. Cropped image data aimed to facilitate the classification process, did not take up much space or memory, and the classification process was light, so it did not require a long time in the classification process later. The image data was cropped to 256 x 256 pixels, reducing the image size to 159 KB. The cropped image data was classified into three types: bare, medium, and dense/high [22], [23]. Figure 9 shows the cropping results of image data.

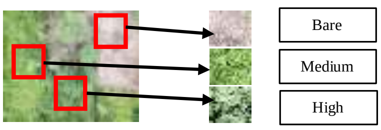


Figure 9. Image Data Cropping

C. Segmentation

Image segmentation was used to distinguish between objects and backgrounds [24], [25]. The separation process was designed to make classification and calculations easier. The image segmentation process was based on the difference in the image grayscale. To convert a colour image with r, g, and b matrix values into a grayscale image. The segmentation method, namely thresholding, can be used to change the colour image. The most basic method for segmenting was image development or image thresholding (de Mello, 2013). Thresholding was used to change the number of gray degrees in an image in order to create a binary image with pixel intensity values of 0 or 1.

D. Feature Extraction (GLCM)

In this study, GLCM was used for feature extraction, with three main features used: correlation, homogeneity, and contrast. This method was used to classify images, recognize textures, segment them, recognize objects, and analyse their colours. In the neighbourhood between pixels, GLCM had four angular directions: 0°, 45°, 90°, and 135°. When the angle was 0°, the pixel density was calculated by moving one distance to the right. Pixel adjacency was calculated using a 45° angle and 1 pixel distance to the top right. The angle is 90°, and the pixel density was calculated by a 1 pixel distance on top. A 135° angle was used, and neighbouring pixels were calculated by moving one pixel up [26]. The gray level of pixels was compared based on angle or neighbours at 0°, 45°, 90°, and 135° in this study. The feature extraction process was also compared the results of previously segmented images to those that have not been segmented.

IV. FINDINGS

A. Result of CNN Model

The formation of network architecture in the CNN algorithm can affect the results of model accuracy. In order to produce an optimal model, network architecture was used during the training process. This study applied an input image with a resolution of 256x256x3, with the aim for reducing image size so that the classification process took as little time as possible. This study applied the second version of the ShuffleNet architecture, which included one convolutional layer (Conv5), three stages (consisting of convolutional and shuffle units), one pooling layer (using Maxpool), and fc. The input image in the shuffleNet v2 model was 256 x 256 in size. The convolution and maximum pooling layers were added to the model's initial position to reduce the size of the feature graph. The convolution layer and pooling layer were replaced at the initial position by the convolutional layer (Conv1) with a 3 x 3 kernel, and the BN layer was added after Conv 1 and Conv 5. Figure 10 shows the flow of the proposed model.

Commented [A1]: Dibagi 4 kriteria atau 3, diintroduction dibagi 4

Figure 10. Flow of Model Shuffle Net v2 Model

The results of training and testing accuracy were obtained after going through several processes in the CNN algorithm. The value for training accuracy can be found in the "Accuracy" column, while the value for testing accuracy can be found in the "Validation Accuracy" column. Accuracy was the value calculated by calculating the accuracy of the training dataset and model predictions. Validation Accuracy is the value calculated by calculating the accuracy of the validation dataset and predictions from the model using validation dataset input data. This procedure used a total of 50 epochs. Previously, epoch comparisons were performed to determine the accuracy and validation results of each training with a different number of epochs. This epoch comparison was intended to find the best model. The number of epochs compared ranges between 25 and 100 [27]. The use of epoch had a significant impact on the resulting accuracy. Because epoch can improve accuracy and the resulting accuracy was stable, it was critical to use the correct epoch in training data to achieve maximum accuracy. The table below compares training results based on the number of epochs.

Table 2.
Comparison of Epoch

Epoch	**Accuracy**	**Loss**	**Accuracy Validation**	***Validation Loss***	**Time**
25	97.46%	0.0686	71%	1.9176	20minute
50	98.75%	0.0358	81.33%	1.2536	56minute
75	99.08%	0.0279	74%	2.4516	1hour 32minute
100	99.29%	0.0218	74.17%	1.1251	2hour 20minute

The training model's accuracy with a total of 50 epochs is 98.75% with a loss of 0.0218. The validation accuracy value for the 50 epochs is 81.33%, which is higher than the other epochs. According to the table, the closer to the highest epoch, the higher the accuracy obtained from the testing results. However, if more than 100 epochs are added, the accuracy value decreases because too many epochs can also affect the large number of datasets. The testing procedure used training data consisting of 2400 image data and 600 image data for each class, as well as 200 image data for each class. Table 3 shows the results of the confusion matrix.

Table 3.
Confusion Matrix

Matrix		***Predict Class***		
		Bare	Medium	High
Actual Class	Bare	209	0	4
	Medium	9	92	88
	High	8	12	178

Based on the results of table 3, the model's predictions on the new data testing data show promising results. Although the prediction of the Bare class is correctly classified as the Bare class, up to four miss classifications from the Bare image data input are classified as the High class. While the Medium class prediction is correctly classified as the Medium class, as many as 9 miss classifications from the input image data are classified as the Bare class. In addition, up to 88 misclassifications of input image data are classified as High. The High class prediction is correctly classified as the High class, but up to 8 miss classifications from the High image data input are classified as the Bare class. As many as 12 misclassifications of input image data were classified as Medium. The overall accuracy of the matrix and kappa accuracy are calculated as follows:

Overall Accuracy = 469/600 = 80%
Kappa = 70%

So, the model's accuracy with a 256x256 input image and a total of 600 image data obtained an accuracy value of 80% and a kappa accuracy of 70%.

B. Result of CNN Model with GLCM

The addition of three GLCM features, namely contrast, homogeneity, and correlation, is the result of the next training model. The procedure involved extracting 3,000 GLCM result image data and producing 9,000 image data that was processed by CNN. This study also compared the direction angles of 0°, 45°, 90°, and 135° to extract images per angle. The GLCM process used a total of 27,000 image data through the segmentation stage. This was done to determine how well each feature performed in the image classification process.

For the GLCM process with CNN going through the segmentation stage, the results of data training with the CNN model and each GLCM feature by going through the segmentation stage with 9,000 data for each angle, can be seen at an angle of 135° getting the highest validation accuracy value from other angles, namely 60.11% with a value validation loss of 0.8460. For each feature, this training procedure applied a total of 50 epochs. This training process took approximately 20-30 minutes per corner. Table 5 shows the results of the GLCM training data per corner.

Table 4.
Comparison of Training Per Angle

Angle	Accuracy	Loss	*Validation Accuracy*	*Validation Loss*	Time
0°	95.24 %	0.1377	50.28 %	0.6096	29 mnt 6 scnd
45°	94.99	0.134	50.39	0.552	31 mnt
	%	5	%	6	12 scnd
135°	96.26 %	0.1176	60.11 %	0.8460	31 mnt 11 scnd

The training data is 9,000 images, and the test data is 1,800 images, with 3,000 images in each class. Table 5 shows the confusion matrix results for the CNN model process with GLCM that went through the segmentation stage.

Table 5.
Confusion Matrix

Matrix		*Predict Class*		
		Bare	Medium	High
Actual Class	Bare	407	92	83
	Medium	54	419	137
	High	83	218	307

According to the results in table 5, the model's prediction results on new data testing data are poor. Although the prediction of the Bare class is correct, as many as 92 miss classifications from the Bare image data input are classified as Medium. In addition, up to 83 miss classifications from the Bare image data input are classified as High. While the Medium class prediction was correctly classified as the Medium class, as many as 54 miss classifications from the input image data were classified as the Bare class. In addition, 137 misclassifications of image data input Medium are classified as High. The High class prediction is correctly classified as the High class, but up to 83 miss classifications from the High image data input are classified as the Bare class. In addition, 218 miss classifications from the High image data input are classified as Medium. The overall accuracy of the matrix and kappa accuracy are calculated as follows:

Overall Accuracy = 1133/1800x100% = 62,99%

Kappa = 44,37%

So, with an input image of 256x256 pixels and a total of 1800 image data, the model produced an accuracy value of 62.99% and a kappa accuracy of 44.37%.

V. DISCUSSION

When the CNN model without the GLCM process was compared to the CNN model with the GLCM process, the comparison was quite far from the accuracy values obtained. The CNN model achieved an accuracy of 80%, while the CNN model with GLCM achieves 62.99% segmentation. This showed that the CNN model outperformed the GLCM process. According to the findings of the analysis, this occurred because the gray level in the image was leveled during the GLCM process, resulting in white and black colors in the image. The colours in the original image changed to white and black, resulting in a classification error. The GLCM process rendered the image colourless and rendered the entire image black.

During the testing of new data, there was a misclassification caused by nearly identical vegetation types. The input data for the CNN model was original image data with different types of vegetation, but based on the researcher's analysis, even though the texture between medium and high vegetation was different, the CNN model still had difficulty distinguishing and recognizing medium and high classes if the data simultaneously has the characteristics of an image that was filled with vegetation even though the type and texture of the vegetation was different. The CNN model with the GLCM method had a lot of misclassifications. The first reason was that the original image's colour had changed, making it

difficult for the model to distinguish between classes. The second issue was that the type and texture of the vegetation were not visible in the image, so when predicting with the CNN and GLCM models on prototypes, the bare class data was read as medium class. High class reads as medium class. This research is the only research to classified vegetation density in tropical peatland.

VI. Conclusion

The conclusion is that comparing the CNN model without the GLCM process to the CNN model with the GLCM process produces a comparison that is quite far from the accuracy value obtained. The CNN model achieves an accuracy of 80%, while the CNN model with GLCM achieves 62.99% segmentation. This demonstrates that the CNN model outperforms the GLCM process in land cover classification. This demonstrates that the image processing process has a significant impact on the classification and prediction stages in vegetation density in tropical peatland.

Acknowledgment

This research was funded by DRTPM grant, National Basic Research Competition scheme, agreement number 113/E5/PG.02.00/PT/2023.

References

[1] N. Novitasari, J. Sujono, S. Harto, A. Maas, and R. Jayadi, "Restoration of peat dome in ex-Mega rice project area in Central Kalimantan," *AIP Conf. Proc.*, vol. 1977, 2018.

[2] N. Novitasari, J. Sujono, S. Harto, A. Maas, and R. Jayadi, "Drought Index for Peatland Wildfire Management in Central Kalimantan, Indonesia During El Niño Phenomenon," *J. Disaster Res.*, vol. 14, no. 7, pp. 939–948, 2019.

[3] Z. Xu, K. Guan, N. Casler, B. Peng, and S. Wang, "A 3D convolutional neural network method for land cover classification using LiDAR and multi-temporal Landsat imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 144, pp. 423–434, 2018.

[4] S. N. Indonesia and B. S. Nasional, "SNI: Klasifikasi penutup lahan," 2014.

[5] F. Zhao, X. Wu, and S. Wang, "Object-oriented Vegetation Classification Method based on UAV and Satellite Image Fusion," *Procedia Comput. Sci.*, vol. 174, no. 2019, pp. 609–615, 2020.

[6] M. Alkaff, H. Khatimi, W. Puspita, and Y. Sari, "Modelling and predicting wetland rice production using support vector regression," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 17, no. 2, pp. 819–825, 2019.

[7] Y. Sari, E. S. Wijaya, A. R. Baskara, and R. S. D. Kasanda, "PSO optimization on backpropagation for fish catch production prediction," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 18, no. 2, pp. 776–782, 2020.

[8] Q. Wu, Y. Gan, B. Lin, Q. Zhang, and H. Chang, "An active contour model based on fused texture features for image segmentation," *Neurocomputing*, vol. 151, no. P3, pp. 1133–1141, 2015.

[9] Z. Xing and H. Jia, "Multilevel Color Image Segmentation Based on GLCM and Improved Salp Swarm Algorithm," *IEEE Access*, vol. 7, pp. 37672–37690, 2019.

[10] S. A. Alazawi, N. M. Shati, and A. H. Abbas, "Texture features extraction based on GLCM for face retrieval system," *Period. Eng. Nat. Sci.*, vol. 7, no. 3, pp. 1459–1467, 2019.

[11] S. Ozturk and B. Akdemir, "Application of Feature Extraction and Classification Methods for Histopathological Image using GLCM, LBP, LBGLCM, GLRLM and SFTA," *Procedia Comput. Sci.*, vol. 132, no. Iccids, pp. 40–46, 2018.

[12] O. Youme, T. Bayet, J. M. Dembele, and C. Cambier, "Deep Learning and Remote Sensing: Detection of Dumping Waste Using UAV," *Procedia Comput. Sci.*, vol. 185, no. June, pp. 361–369, 2021.

[13] S. M. Hamylton *et al.*, "Evaluating techniques for mapping island vegetation from unmanned aerial vehicle (UAV) images: Pixel classification, visual interpretation and machine learning approaches," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 89, no. March, p. 102085, 2020.

[14] T. Lawrence and L. Zhang, "IoTNet: An efficient and accurate convolutional neural network for IoT devices," *Sensors (Switzerland)*, vol. 19, no. 24, 2019.
[15] G. Liu *et al.*, "I3d-shufflenet based human action recognition," *Algorithms*, vol. 13, no. 11, 2020.
[16] G. Losapio *et al.*, "Lightweight and efficient convolutional neural networks for recognition of dolphin dorsal fins," pp. 68–72, 2020.
[17] Z. Wang and L. Ma, "SYOLO: An Efficient Pedestrian Detection," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 768, no. 7, 2020.
[18] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 6848–6856, 2018.
[19] Y. Sari, M. Alkaff, and R. A. Pramunendar, "Iris recognition based on distance similarity and PCA," *AIP Conf. Proc.*, vol. 1977, 2018.
[20] Y. Sari, M. Alkaff, and M. Maulida, "Classification of Rice Leaf using Fuzzy Logic and Hue Saturation Value (HSV) to Determine Fertilizer Dosage," in *2020 Fifth International Conference on Informatics and Computing (ICIC)*, 2020, pp. 1–6.
[21] Y. Sari, Y. F. Arifin, N. Novitasari, and M. R. Faisal, "Vegetation-Density Drone Dataset For Peatland Vegetation Classification," vol. 1, 2022.
[22] Y. Sari, Y. Arifin, Novitasari, and M. Faisal, "Implementation of Deep Learning Based Semantic Segmentation Method To Determine Vegetation Density," *Eastern-European J. Enterp. Technol.*, vol. 5, no. 2–119, pp. 42–54, 2022.
[23] Y. Sari, Y. F. Arifin, N. Novitasari, and M. R. Faisal, "Effect of Feature Engineering Technique for Determining Vegetation Density," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 7, pp. 655–661, 2022.
[24] Y. Sari, A. R. Baskara, and R. Wahyuni, "Classification of Chili Leaf Disease Using the Gray Level Co-occurrence Matrix (GLCM) and the Support Vector Machine (SVM) Methods," *2021 6th Int. Conf. Informatics Comput. ICIC 2021*, 2021.
[25] Y. Sari, H. Suhud, A. R. Baskara, R. A. Pramunendar, and I. F. Radam, "Parking Lots Detection in Static Image Using Support Vector Machine Based on Genetic Algorithm," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 6, pp. 476–487, 2021.
[26] R. A. Pramunendar, D. P. Prabowo, D. Pergiwati, Y. Sari, P. N. Andono, and M. A. Soeleman, "New workflow for marine fish classification based on combination features and CLAHE enhancement technique," *Int. J. Intell. Eng. Syst.*, vol. 13, no. 4, pp. 293–304, 2020.
[27] C. A. B. de Mello, "Image thresholding," *Digit. Doc. Anal. Process.*, vol. 2006, no. Snati, pp. 71–98, 2013.

参考文:

COVER LETTER

UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD
Land cover is an important factor in geographic analysis, ranging from physical geography studies, approaches to sustainable planning to environmental analysis. Vegetation analysis according to the Indonesian National Standard (SNI 7645:2014) is classified based on density. The vegetation density index is divided into 4, namely non-vegetation, bare, medium and high. In the technical aspect to obtain information related to vegetation, this can be done using remote sensing. Remote sensing uses two data to obtain information, namely satellite data and UAV data. This study used UAV data with shooting locations in the Liang Anggang Protection Forest in classifying land cover. The method used was Convolutional Neural Network with feature extraction used in this study was GLCM. This research used the ShuffleNet v2 architecture on the CNN method. The findings of this study used two models, namely the CNN model without the GLCM process and compared to the CNN model with the addition of the GLCM process, resulting in a comparison that was quite far from the accuracy value obtained. The CNN model obtained an accuracy value of 80%, while the CNN model with GLCM using segmentation was 49.9% and without segmentation was 44.53%.
Classification; class; CNN; GLCM; accuracy.

Type of manuscript (please specify):

☐ Research article
☐ Review article
☐ Brief report
☐ Short communication
☐ Research note

Yuslena Sari Jl. Brig. Hasan Basry Kayutangi Banjarmasin, Indonesia, 70123	
Telephone# +6285247175500	Fax#
Email yuzlena@ulm.ac.id	

☐ I hereby confirm that the manuscript was prepared in accordance with the instructions for authors of scientific publications, and that the content of this manuscript, or most of it, was not published in the journal indicated, and the manuscript was not submitted for publication elsewhere.

☐ I hereby confirm my consent to pay the Article Processing Charges (APC) EUR 450 in the case of the manuscript acceptance for publication. I am aware and hereby confirm that the APC is non-refundable.

☐ I hereby confirm my consent to make the payment for English Language Editing Services (EUR 150).

Publication fee payer details

____

____

Signature of the Corresponding author **Date**

Copyright Agreement

Manuscript title: UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD

Full names of all authors: Novitasari Novitasari, Yuslena Sari, Yudi Firmanul Arifin, Nurul Fathanah Mustamin, Erika Maulidiya

Full name and address of the corresponding author:

Yuslena Sari

Telephone/Whatsap: +6285247175500 Fax: _________________ Email: yuzlena@ulm.ac.id_

License Agreement

(1) Authors own all the copyright rights for the paper.
(2) Submitted manuscript is an original paper.
(3) Authors hereby grant the Issues of Journal of SWJTU with an exclusive, royalty-free, worldwide license to email the paper to all who will ask for it.
(4) All authors have made a significant contribution to the research and are ready to assume joint responsibility for the paper.
(5) All authors have seen and approved the manuscript in the final form as it is submitted for publication.
(6) This manuscript has not been published and also has neither been submitted nor considered for publication elsewhere
(7) The text, illustrations and any other materials, included into the manuscript, do not infringe any existing intellectual property rights or other rights of any person or entity.
(8) The editors of the Issues of Journal of SWJTU, its personnel or the Editorial Board members accept no responsibility for the quality of the idea expressed in this publication.

I am the Corresponding author and have full authority to enter into this agreement.

Full name, affiliation and position: Yuslena Sari, Department of Information Technology, Universitas Lambung Mangkurat

Signature:____Date: 28 April 2023

西 南 交 通 大 学 学 报

第 57 卷 第 6 期
2022 年 12 月

JOURNAL OF SOUTHWEST JIAOTONG UNIVERSITY

Vol. 57 No. 6
Dec. 2022

ISSN: 0258-2724 **DOI：10.35741/issn.0258-2724.57.6.1**

Category
Please select the type of your manuscript
(see Author Guidelines)
Category.
Please select the most suitable field of scientific research
from the approved list of the journal categories

UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD

Novitasari Novitasari[a], Yuslena Sari[b,*], Yudi Firmanul Arifin[c], Nurul Fathanah Mustamin[d], Erika Maulidiya[e]

[a]Department of Civil Engineering, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, novitasari@ulm.ac.id
[b]Department of Information Technology, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, yuzlena@ulm.ac.id
[c]Faculty of Forestry, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, yudifirmanul@ulm.ac.id
[d]Department of Information Technology, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, nurul.mustamin@ulm.ac.id
[e]Department of Information Technology, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, 1810817220017@mhs.ulm.ac.id

Received: ▪ *Review:* ▪ *Accepted:* ▪ *Published*

Abstract

Land cover or vegetation density in tropical peatland is an important factor as hydrology respon in geographic analysis, ranging from physical geography studies, approaches to sustainable planning to environmental analysis. Vegetation analysis according to the Indonesian National Standard (SNI 7645:2014) is classified based on density. The vegetation density index is divided into 4, namely non-vegetation, bare, medium and high. In the technical aspect to obtain information related to vegetation, this can be done using remote sensing. Remote sensing uses two data to obtain information, namely satellite data and UAV data. This study used UAV data with shooting locations in the Liang Anggang Protection Forest in classifying land cover. The method used was Convolutional Neural Network with feature extraction used in this study was GLCM. This research used the ShuffleNet v2 architecture on the CNN method. The findings of this study used two models, namely the CNN model without the GLCM process and compared to the CNN model with the addition of the GLCM process, resulting in a comparison that was quite far from the accuracy value obtained. The CNN model obtained an accuracy value of 80%, while the CNN model with GLCM using segmentation was 49.9% and without segmentation was 44.53%.

Keywords: Tropical Peatland, Vegetation Density, Classification; class; CNN; GLCM; accuracy.

摘要 The authors may not translate the abstract and keywords into Chinese themselves.

关键词:

I. INTRODUCTION

Peatland management in Indonesia has many challenges, due to peatland disaster [1], as flood and wildfire. Wildfire caused by human action is the biggest one. Many wildfires in Indonesia as intentional fires as part of residential developments [2]. It is lead to change land cover as change from vegetation density in tropical peatland. Land cover or vegetation density change as one of internal factor as hydrological respon. From physical geography studies to approaches to sustainable planning to environmental analysis, land cover is an important factor in geographic analysis, especially in disaster mitigation in tropical peatland. Environmental analysis needs surface vegetation-based land cover information [3]. The entire plant of an area that serves as a land cover is referred to as vegetation. Vegetation is the entire plant of an area that serves as a land cover. According to the Indonesian National Standard [4], vegetation analysis is classified based on density. Non-vegetation, bare, medium, and high vegetation density indexes are used [4]. In addition to determining the level of vegetation density, it is important to be able to distinguish vegetation density in the form of an image, which makes data processing easier. N.A. Harahap (year) conducted research which provides an image of the classification of vegetation density classes based on the images shown in Figure 1.1.

(a) (b) (c) (d)

Figure 1. (a) Non Vegetation, (b) Bare Vegetation, (c) Medium Vegetation, (d) High Vegetation

Vegetation density analysis in tropical peatland is one method for studying the arrangement and composition of vegetation in terms of plant shape (structure). In terms of technology, remote sensing can be used to obtain information about vegetation. Remote sensing obtains information from two sources: satellite data and UAV data. Previous research that used remote sensing technology by utilizing satellite data resulted in data accuracy ranging from 63% - 85% using various methods [3], [5]. Because satellite data is a traditional format based on statistical reporting and sampling surveys, determining vegetation density is critical [5]. Remote sensing with satellite data has been widely used in the identification and classification of land cover patterns across a wide geographic coverage, but the use of satellite data, which has a high operating altitude and is easily influenced by weather, clouds, and other external factors, is being reconsidered. Remote sensing technology can quickly and precisely provide spatial information on the earth surface. The object being sensed, the sensor for recording the object, and the electronic waves emitted by the earth surface are the three main components of remote sensing.

Remote sensing technology can quickly and precisely provide spatial information on the earth surface. The object being sensed, the sensor for recording the object, and the electronic waves emitted by the earth surface are the three main components of remote sensing. As technology advances, remote sensing facilities such as the Unmanned Aerial Vehicle (UAV) become more practical and easier to implement. The emergence of UAV raises significant potential as a tool for environmental and ecological analysis, such as monitoring agricultural land, forest fires, arctic lichen distribution, and mapping of mangrove forests. The generation of spatial information based on aerial image data using drones has enormous potential for the advancement of remote sensing technology, such as area classification. The benefits of using a UAV include faster and more flexible data acquisition, results that are more real-time, and low and light operating and maintenance costs. Apart from the ability to fly through clouds and produce cloud-free images, it differs from satellite imagery, which is heavily influenced by atmospheric conditions. UAV imagery has a high resolution when compared to satellite imagery, reaching a spatial resolution of less than 1 cm, which is much more detailed than satellite (30cm) and aircraft (10cm) imagery [3]. Optimal results that can be obtained from the use of UAV in object classification and the appropriate method for processing data with UAV imagery.

Before being processed in a classification model, image data requires feature extraction techniques to determine certain characteristics possessed by images to aid in object identification (image analysis) [3], [6], [7]. The resulting features will be selected first in the feature extraction process to obtain features with a high influence as a reference for the classification process. The function of feature extraction is to extract the necessary information from an image. Shape, colour, and texture extraction are the three types of feature extraction. Images with a slight colour can benefit from feature extraction using the Gray Level Co-occurrence Matrix (GLCM) method, which is a second level statistical method that computes the frequency of pairs of pixels in an image that have the same gray level and applies the additional knowledge obtained through pixel spatial relationships [8]. Using edge information, the co-occurrence matrix embeds the distribution of grayscale transitions. The majority of the information required to calculate the threshold value in the GLCM technique is straightforward but efficient [9]. S. Karthikeyan and N. Rengarajan use the GLCM algorithm with up to 95% accuracy. Previous research has compared GLCM feature extraction to LBP, MI, CLBP, LBGLCM, and GLRLM, with the accuracy results proving that using feature extraction in classification using GLCM produces better results than using other methods. GLCM accuracy results range from 70% to 93% [9]–[11].

Visual interpretation methods, pixel-based digital classification methods, and object-based classification methods are used in land cover mapping based on remote sensing imagery. Land cover analysis researchers are interested in the use of data mining methods. Land classification, Machine Learning, and Deep Learning have all made extensive use of classification methods. Deep learning, which is included in the supervised classification, is developed and produced by the machine learning method. Deep learning methods are widely used in satellite image analysis because they are powerful and intelligent in image processing. Deep learning methods are still evolving, with the Convolutional Neural Network (CNN) deep learning method producing the most significant results in image recognition to date. Deep Learning has demonstrated that this architecture, particularly CNN, can learn human-level solutions to specific visual tasks. This method has been used extensively in remote sensing image analysis tasks such as object detection in images, image recording, scene classification, segmentation, object-based image analysis, and land use and land cover classification [12]. CNN is one of the most recent Deep Learning methods to emerge. This method has been shown to be useful for pattern recognition and object classification [3]. Previous research using the CNN method to classify land cover yielded satisfactory accuracy results ranging from 73% to 98% [3], [12], [13]

CNN has a variety of popular architectures, including LeNet5 (1998), AlexNet (2012), ZFNet (2013), GoogleNet (2014), ResNet (2015), FractalNet (2016), ShuffleNet (2018), and others. Previous research has compared the use of architecture on CNN in the field of classification. The compared architectures demonstrate the advantage and disadvantage of each, for architectures that are widely used in the field of image classification and are relatively new, and have been compared with several other architectures, ShuffleNet. ShuffleNet is a very efficient CNN architecture with fast accuracy. Research that has used the ShuffleNet architecture and has made comparisons with other architectures such as GoogleNet, DenseNet, MobileNet, Xception, IGCV2, EffNet V1, EffNet V2, IoTNet-3-5 and ResNet50 in the classification process states that the ShuffleNet architecture increases the accuracy of 82% - 98% with less memory usage and faster processing time [14]–[18].

The CNN method is widely used in the field of deep learning to conduct land cover classification. GLCM was used to extract features in this study. The ShuffleNet architecture on the CNN method will be used in this study. This research was carried out for a month in the Liang Anggang Protected Forest area, Banjarbaru block 1 area, with targeted data collection. The location for this study was chosen based on observations made during the observation and survey of the block 1 area, where, according to the 2017 Provincial Forestry Office, an area of 479 hectares of block 1 area is filled with land such as agriculture, plantations, roads and settlements, as well as 494 hectares of forest. In addition to being a peatland, the research site, particularly in block 1, meets the characteristics and suitability of the needs in collecting data for land cover classification in terms of vegetation density types (bare, medium, and high) that can be seen with the naked eye during observations and surveys. This study classified land cover, with a focus on vegetation density, and the research location was chosen in accordance with the data requirements. The objective of this study was to determine the results of the best deep learning methods in land cover classification based on vegetation density. This study created research updates by combining UAV data with shooting locations in the Liang Anggang Protected Forest.

II. Research Methodology

A. Research Site

This study was being conducted in the Liang Anggang Protected Forest in Banjarbaru City as the biggest wildfire in tropical peatland in South Kalimantan. This is the Tangi Timber KPHP's management area. The protected forest designation is based on Minister of Forestry Decree No. 672/Kpts-II/1991 and Kep Menhut No. 434/Kpts-II/1996 with a total area of 2,250 hectares divided into two protected forest blocks, namely block 1 covering an area of 960 hectares including Liang Anggang sub-district, Banjarbaru and block 2 covering an area of 1290 hectares including the Gambut District, Banjar Regency.

Figure 2. Map of the Liang Anggang Protected Forest Area

The study lasted one month, from November to December 2021, and focused on the Warning Area (lock signal area from the airport) that caused the drone to be unable to operate.

B. Research Procedure

This research was conducted in the Liang Anggang Protected Forest area by conducting a field survey to assess the state of the vegetation or areas within the Protected Forest area. This study collected image data using drones to capture images from a height of 20 meters over a one-month period. Land was assigned coordinates based on the goal of image data collection using Google Earth Pro tools. Land with coordinates was exported in .KML format and later imported into DroneDeploy (website) to make directing drone flights on land easier. Then, the imported KML file was configured for flight altitude and 2D or 3D image capture. An illustration of image capture is shown in Figure 3.

Figure 3. Illustration of Image Data Retrieval

Before proceeding to the next stage, image data that has been recorded and stored according to predetermined coordinate points was processed. To facilitate operation with the method that was used later, image data was labelled. The CNN method was used in this study. Image data that has already been processed was then fed into the classification process using the method used in this study. Image data was classified using each method, and the accuracy value was calculated using tools. The obtained accuracy value was then analysed and compared to draw conclusions. The flow of this research is shown in Figure 4.

Figure 4. Diagram of Research Procedure

C. Feature Extraction

The purposes of feature extraction is to obtain the feature value of an object based on an image pixel intensity value relationship. The feature extraction process goal is to extract a special (unique) value from each image [19], [20]. This study used GLCM feature extraction with three primary features: correlation, homogeneity, and contrast. The feature extraction results created a GLCM version of the image using these three features. Figure 5 shows an illustration or description of the texture extraction results obtained with the GLCM feature.

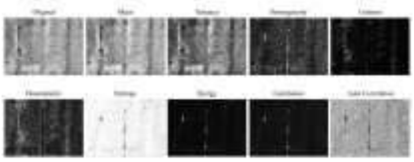


Figure 5. Result of Feature Extraction

The texture of an image was sought after by feature extracted images. The training data set consisted of 2400 images divided into three classes. This study applied 5 GLCM features to convert an input 2D image/image to an output 2D image/image to a gray level with a gray range of 0 to 1. The purpose of this step was to use gray level scaling to reduce the image volume to a more manageable size. Scaling to a grayscale level acted as a filter, removing some of the noise (de Mello, 2013). Figure 6 shows the scenario of the feature extraction test results with GLCM.

Figure 6. Ilustration of GLCM-CNN Feature Extraction

D. Classification of Convolutional Neural Networks

Only CNN neural networks can process grid structure data, such as two-dimensional images. The convolution layer is a linear algebra operation that generates a matrix of filters in the image to be processed. A convolution layer process is one of the many types of layers that can exist in a network. The image entered into the CNN classification model created during the fit model stage yielded an output calculated using the optimized weight. As a result, the classification model created should be able to classify the testing data into the correct class. This test was performed to calculate the accuracy value in the classification model that has been created. Figure 7 shows an illustration of the CNN classification process.

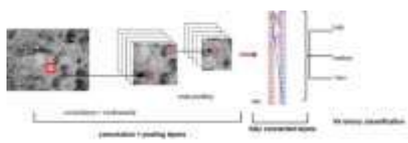


Figure 7. Illustration of CNN Classification Process

Figure 8 shows the classification flow using the CNN method

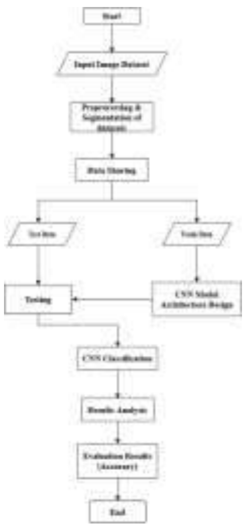


Figure 8. Stage of CNN Classification

E. Classification Analysis

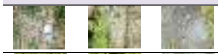
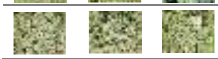
The results of the UAV image classification using two methods were analysed and the level of accuracy was determined. This study applied accuracy testing with a confusion matrix in the form of overall accuracy (OA) and Kappa coefficient accuracy. Proceed with the analysis of the CNN method classification results to obtain accurate results from the use of the CNN method in land cover classification.

III. TESTING

A. Image Dataset

The dataset used in this study was divided into three categories: bare, medium, and high. The total number of images collected was 3000, with 1000 for each class type category. The classification of these three classes was based on the condition of the Liang Anggang Protected Forest where the research location, particularly block 1, meets the characteristics and suitability of the needs in collecting data for land cover classification in terms of vegetation density types (bare, medium, and high) that can be seen with the naked eye during observation and surveys. This study classified land cover, with a focus on vegetation density, and the research location was chosen in accordance with the data requirements. Table 1 shows the results of categorizing three classes of vegetation density in terms of images based on the division of the available dataset [21].

Table 1.
Image of Vegetation Density

Image	Type of Vegetation Density
	Bare
	Medium
	High

Source:

B. Image Cropping

Because the image data obtained with the drone was too large, the data was resized by cutting the image and selecting specific areas to be used as training data. Cropped image data aimed to facilitate the classification process, did not take up much space or memory, and the classification process was light, so it did not require a long time in the classification process later. The image data was cropped to 256 x 256 pixels, reducing the image size to 159 KB. The cropped image data was classified into three types: bare, medium, and dense/high [22], [23]. Figure 9 shows the cropping results of image data.

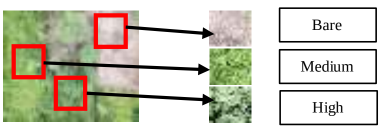


Figure 9. Image Data Cropping

C. Segmentation

Image segmentation was used to distinguish between objects and backgrounds [24], [25]. The separation process was designed to make classification and calculations easier. The image segmentation process was based on the difference in the image grayscale. To convert a colour image with r, g, and b matrix values into a grayscale image. The segmentation method, namely thresholding, can be used to change the colour image. The most basic method for segmenting was image development or image thresholding (de Mello, 2013). Thresholding was used to change the number of gray degrees in an image in order to create a binary image with pixel intensity values of 0 or 1.

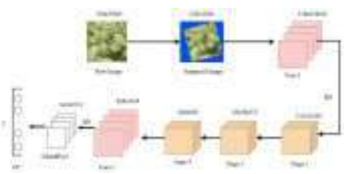
D. Feature Extraction (GLCM)

In this study, GLCM was used for feature extraction, with three main features used: correlation, homogeneity, and contrast. This method was used to classify images, recognize textures, segment them, recognize objects, and analyse their colours. In the neighbourhood between pixels, GLCM had four angular directions: 0°, 45°, 90°, and 135°. When the angle was 0°, the pixel density was calculated by moving one distance to the right. Pixel adjacency was calculated using a 45° angle and 1 pixel distance to the top right. The angle is 90°, and the pixel density was calculated by a 1 pixel distance on top. A 135° angle was used, and neighbouring pixels were calculated by moving one pixel up [26]. The gray level of pixels was compared based on angle or neighbours at 0°, 45°, 90°, and 135° in this study. The feature extraction process was also compared the results of previously segmented images to those that have not been segmented.

IV. FINDINGS

A. Result of CNN Model

The formation of network architecture in the CNN algorithm can affect the results of model accuracy. In order to produce an optimal model, network architecture was used during the training process. This study applied an input image with a resolution of 256x256x3, with the aim for reducing image size so that the classification process took as little time as possible. This study applied the second version of the ShuffleNet architecture, which included one convolutional layer (Conv5), three stages (consisting of convolutional and shuffle units), one pooling layer (using Maxpool), and fc. The input image in the shuffleNet v2 model was 256 x 256 in size. The convolution and maximum pooling layers were added to the model's initial position to reduce the size of the feature graph. The convolution layer and pooling layer were replaced at the initial position by the convolutional layer (Conv1) with a 3 x 3 kernel, and the BN layer was added after Conv 1 and Conv 5. Figure 10 shows the flow of the proposed model.



Commented [A1]: Dibagi 4 kriteria atau 3, diintroduction dibagi 4

Figure 10. Flow of Model Shuffle Net v2 Model

The results of training and testing accuracy were obtained after going through several processes in the CNN algorithm. The value for training accuracy can be found in the "Accuracy" column, while the value for testing accuracy can be found in the "Validation Accuracy" column. Accuracy was the value calculated by calculating the accuracy of the training dataset and model predictions. Validation Accuracy is the value calculated by calculating the accuracy of the validation dataset and predictions from the model using validation dataset input data. This procedure used a total of 50 epochs. Previously, epoch comparisons were performed to determine the accuracy and validation results of each training with a different number of epochs. This epoch comparison was intended to find the best model. The number of epochs compared ranges between 25 and 100 [27]. The use of epoch had a significant impact on the resulting accuracy. Because epoch can improve accuracy and the resulting accuracy was stable, it was critical to use the correct epoch in training data to achieve maximum accuracy. The table below compares training results based on the number of epochs.

Table 2.
Comparison of Epoch

Epoch	**Accuracy**	**Loss**	**Accuracy Validation**	***Validation Loss***	**Time**
25	97.46%	0.0686	71%	1.9176	20minute
50	98.75%	0.0358	81.33%	1.2536	56minute
75	99.08%	0.0279	74%	2.4516	1hour 32minute
100	99.29%	0.0218	74.17%	1.1251	2hour 20minute

The training model's accuracy with a total of 50 epochs is 98.75% with a loss of 0.0218. The validation accuracy value for the 50 epochs is 81.33%, which is higher than the other epochs. According to the table, the closer to the highest epoch, the higher the accuracy obtained from the testing results. However, if more than 100 epochs are added, the accuracy value decreases because too many epochs can also affect the large number of datasets. The testing procedure used training data consisting of 2400 image data and 600 image data for each class, as well as 200 image data for each class. Table 3 shows the results of the confusion matrix.

Table 3.
Confusion Matrix

Matrix		***Predict Class***		
		Bare	Medium	High
Actual Class	Bare	209	0	4
	Medium	9	92	88
	High	8	12	178

Based on the results of table 3, the model's predictions on the new data testing data show promising results. Although the prediction of the Bare class is correctly classified as the Bare class, up to four miss classifications from the Bare image data input are classified as the High class. While the Medium class prediction is correctly classified as the Medium class, as many as 9 miss classifications from the input image data are classified as the Bare class. In addition, up to 88 misclassifications of input image data are classified as High. The High class prediction is correctly classified as the High class, but up to 8 miss classifications from the High image data input are classified as the Bare class. As many as 12 misclassifications of input image data were classified as Medium. The overall accuracy of the matrix and kappa accuracy are calculated as follows:

Overall Accuracy = 469/600 = 80%
Kappa = 70%

So, the model's accuracy with a 256x256 input image and a total of 600 image data obtained an accuracy value of 80% and a kappa accuracy of 70%.

B. Result of CNN Model with GLCM

The addition of three GLCM features, namely contrast, homogeneity, and correlation, is the result of the next training model. The procedure involved extracting 3,000 GLCM result image data and producing 9,000 image data that was processed by CNN. This study also compared the direction angles of 0°, 45°, 90°, and 135° to extract images per angle. The GLCM process used a total of 27,000 image data through the segmentation stage. This was done to determine how well each feature performed in the image classification process.

For the GLCM process with CNN going through the segmentation stage, the results of data training with the CNN model and each GLCM feature by going through the segmentation stage with 9,000 data for each angle, can be seen at an angle of 135° getting the highest validation accuracy value from other angles, namely 60.11% with a value validation loss of 0.8460. For each feature, this training procedure applied a total of 50 epochs. This training process took approximately 20-30 minutes per corner. Table 5 shows the results of the GLCM training data per corner.

Table 4.
Comparison of Training Per Angle

Angle	Accuracy	Loss	*Validation Accuracy*	*Validation Loss*	Time
0°	95.24 %	0.1377	50.28 %	0.6096	29 mnt 6 scnd
45°	94.99 [illegible]	0.134[illegible]	50.39 [illegible]	0.552[illegible]	31 mnt [illegible]
[illegible]	[illegible] %	[illegible]5	[illegible] %	[illegible]6	[illegible] 12 scnd
135°	96.26 %	0.1176	60.11 %	0.8460	31 mnt 11 scnd

Matrix		*Predict Class*		
		Bare	Medium	High
Actual Class	Bare	407	92	83
	Medium	54	419	137
	High	83	218	307

The training data is 9,000 images, and the test data is 1,800 images, with 3,000 images in each class. Table 5 shows the confusion matrix results for the CNN model process with GLCM that went through the segmentation stage.

Table 5.
Confusion Matrix

According to the results in table 5, the model's prediction results on new data testing data are poor. Although the prediction of the Bare class is correct, as many as 92 miss classifications from the Bare image data input are classified as Medium. In addition, up to 83 miss classifications from the Bare image data input are classified as High. While the Medium class prediction was correctly classified as the Medium class, as many as 54 miss classifications from the input image data were classified as the Bare class. In addition, 137 misclassifications of image data input Medium are classified as High. The High class prediction is correctly classified as the High class, but up to 83 miss classifications from the High image data input are classified as the Bare class. In addition, 218 miss classifications from the High image data input are classified as Medium. The overall accuracy of the matrix and kappa accuracy are calculated as follows:

Overall Accuracy = 1133/1800x100% = 62,99%

Kappa = 44,37%

So, with an input image of 256x256 pixels and a total of 1800 image data, the model produced an accuracy value of 62.99% and a kappa accuracy of 44.37%.

V. Discussion

When the CNN model without the GLCM process was compared to the CNN model with the GLCM process, the comparison was quite far from the accuracy values obtained. The CNN model achieved an accuracy of 80%, while the CNN model with GLCM achieves 62.99% segmentation. This showed that the CNN model outperformed the GLCM process. According to the findings of the analysis, this occurred because the gray level in the image was leveled during the GLCM process, resulting in white and black colors in the image. The colours in the original image changed to white and black, resulting in a classification error. The GLCM process rendered the image colourless and rendered the entire image black.

During the testing of new data, there was a misclassification caused by nearly identical vegetation types. The input data for the CNN model was original image data with different types of vegetation, but based on the researcher's analysis, even though the texture between medium and high vegetation was different, the CNN model still had difficulty distinguishing and recognizing medium and high classes if the data simultaneously has the characteristics of an image that was filled with vegetation even though the type and texture of the vegetation was different. The CNN model with the GLCM method had a lot of misclassifications. The first reason was that the original image's colour had changed, making it

difficult for the model to distinguish between classes. The second issue was that the type and texture of the vegetation were not visible in the image, so when predicting with the CNN and GLCM models on prototypes, the bare class data was read as medium class. High class reads as medium class. This research is the only research to classified vegetation density in tropical peatland.

VI. CONCLUSION

The conclusion is that comparing the CNN model without the GLCM process to the CNN model with the GLCM process produces a comparison that is quite far from the accuracy value obtained. The CNN model achieves an accuracy of 80%, while the CNN model with GLCM achieves 62.99% segmentation. This demonstrates that the CNN model outperforms the GLCM process in land cover classification. This demonstrates that the image processing process has a significant impact on the classification and prediction stages in vegetation density in tropical peatland.

ACKNOWLEDGMENT

This research was funded by DRTPM grant, National Basic Research Competition scheme, agreement number 113/E5/PG.02.00/PT/2023.

REFERENCES

[1] N. Novitasari, J. Sujono, S. Harto, A. Maas, and R. Jayadi, "Restoration of peat dome in ex-Mega rice project area in Central Kalimantan," *AIP Conf. Proc.*, vol. 1977, 2018.

[2] N. Novitasari, J. Sujono, S. Harto, A. Maas, and R. Jayadi, "Drought Index for Peatland Wildfire Management in Central Kalimantan, Indonesia During El Niño Phenomenon," *J. Disaster Res.*, vol. 14, no. 7, pp. 939–948, 2019.

[3] Z. Xu, K. Guan, N. Casler, B. Peng, and S. Wang, "A 3D convolutional neural network method for land cover classification using LiDAR and multi-temporal Landsat imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 144, pp. 423–434, 2018.

[4] S. N. Indonesia and B. S. Nasional, "SNI: Klasifikasi penutup lahan," 2014.

[5] F. Zhao, X. Wu, and S. Wang, "Object-oriented Vegetation Classification Method based on UAV and Satellite Image Fusion," *Procedia Comput. Sci.*, vol. 174, no. 2019, pp. 609–615, 2020.

[6] M. Alkaff, H. Khatimi, W. Puspita, and Y. Sari, "Modelling and predicting wetland rice production using support vector regression," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 17, no. 2, pp. 819–825, 2019.

[7] Y. Sari, E. S. Wijaya, A. R. Baskara, and R. S. D. Kasanda, "PSO optimization on backpropagation for fish catch production prediction," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 18, no. 2, pp. 776–782, 2020.

[8] Q. Wu, Y. Gan, B. Lin, Q. Zhang, and H. Chang, "An active contour model based on fused texture features for image segmentation," *Neurocomputing*, vol. 151, no. P3, pp. 1133–1141, 2015.

[9] Z. Xing and H. Jia, "Multilevel Color Image Segmentation Based on GLCM and Improved Salp Swarm Algorithm," *IEEE Access*, vol. 7, pp. 37672–37690, 2019.

[10] S. A. Alazawi, N. M. Shati, and A. H. Abbas, "Texture features extraction based on GLCM for face retrieval system," *Period. Eng. Nat. Sci.*, vol. 7, no. 3, pp. 1459–1467, 2019.

[11] S. Ozturk and B. Akdemir, "Application of Feature Extraction and Classification Methods for Histopathological Image using GLCM, LBP, LBGLCM, GLRLM and SFTA," *Procedia Comput. Sci.*, vol. 132, no. Iccids, pp. 40–46, 2018.

[12] O. Youme, T. Bayet, J. M. Dembele, and C. Cambier, "Deep Learning and Remote Sensing: Detection of Dumping Waste Using UAV," *Procedia Comput. Sci.*, vol. 185, no. June, pp. 361–369, 2021.

[13] S. M. Hamylton *et al.*, "Evaluating techniques for mapping island vegetation from unmanned aerial vehicle (UAV) images: Pixel classification, visual interpretation and machine learning approaches," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 89, no. March, p. 102085, 2020.

[14] T. Lawrence and L. Zhang, "IoTNet: An efficient and accurate convolutional neural network for IoT devices," *Sensors (Switzerland)*, vol. 19, no. 24, 2019.

[15] G. Liu *et al.*, "I3d-shufflenet based human action recognition," *Algorithms*, vol. 13, no. 11, 2020.

[16] G. Losapio *et al.*, "Lightweight and efficient convolutional neural networks for recognition of dolphin dorsal fins," pp. 68–72, 2020.

[17] Z. Wang and L. Ma, "SYOLO: An Efficient Pedestrian Detection," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 768, no. 7, 2020.

[18] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 6848–6856, 2018.

[19] Y. Sari, M. Alkaff, and R. A. Pramunendar, "Iris recognition based on distance similarity and PCA," *AIP Conf. Proc.*, vol. 1977, 2018.

[20] Y. Sari, M. Alkaff, and M. Maulida, "Classification of Rice Leaf using Fuzzy Logic and Hue Saturation Value (HSV) to Determine Fertilizer Dosage," in *2020 Fifth International Conference on Informatics and Computing (ICIC)*, 2020, pp. 1–6.

[21] Y. Sari, Y. F. Arifin, N. Novitasari, and M. R. Faisal, "Vegetation-Density Drone Dataset For Peatland Vegetation Classification," vol. 1, 2022.

[22] Y. Sari, Y. Arifin, Novitasari, and M. Faisal, "Implementation of Deep Learning Based Semantic Segmentation Method To Determine Vegetation Density," *Eastern-European J. Enterp. Technol.*, vol. 5, no. 2–119, pp. 42–54, 2022.

[23] Y. Sari, Y. F. Arifin, N. Novitasari, and M. R. Faisal, "Effect of Feature Engineering Technique for Determining Vegetation Density," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 7, pp. 655–661, 2022.

[24] Y. Sari, A. R. Baskara, and R. Wahyuni, "Classification of Chili Leaf Disease Using the Gray Level Co-occurrence Matrix (GLCM) and the Support Vector Machine (SVM) Methods," *2021 6th Int. Conf. Informatics Comput. ICIC 2021*, 2021.

[25] Y. Sari, H. Suhud, A. R. Baskara, R. A. Pramunendar, and I. F. Radam, "Parking Lots Detection in Static Image Using Support Vector Machine Based on Genetic Algorithm," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 6, pp. 476–487, 2021.

[26] R. A. Pramunendar, D. P. Prabowo, D. Pergiwati, Y. Sari, P. N. Andono, and M. A. Soeleman, "New workflow for marine fish classification based on combination features and CLAHE enhancement technique," *Int. J. Intell. Eng. Syst.*, vol. 13, no. 4, pp. 293–304, 2020.

[27] C. A. B. de Mello, "Image thresholding," *Digit. Doc. Anal. Process.*, vol. 2006, no. Snati, pp. 71–98, 2013.

参考文:

COVER LETTER

UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD
Land cover is an important factor in geographic analysis, ranging from physical geography studies, approaches to sustainable planning to environmental analysis. Vegetation analysis according to the Indonesian National Standard (SNI 7645:2014) is classified based on density. The vegetation density index is divided into 4, namely non-vegetation, bare, medium and high. In the technical aspect to obtain information related to vegetation, this can be done using remote sensing. Remote sensing uses two data to obtain information, namely satellite data and UAV data. This study used UAV data with shooting locations in the Liang Anggang Protection Forest in classifying land cover. The method used was Convolutional Neural Network with feature extraction used in this study was GLCM. This research used the ShuffleNet v2 architecture on the CNN method. The findings of this study used two models, namely the CNN model without the GLCM process and compared to the CNN model with the addition of the GLCM process, resulting in a comparison that was quite far from the accuracy value obtained. The CNN model obtained an accuracy value of 80%, while the CNN model with GLCM using segmentation was 49.9% and without segmentation was 44.53%.
Classification; class; CNN; GLCM; accuracy.

Type of manuscript (please specify):

- ☐ Research article
- ☐ Review article
- ☐ Brief report
- ☐ Short communication
- ☐ Research note

Yuslena Sari Jl. Brig. Hasan Basry Kayutangi Banjarmasin, Indonesia, 70123	
Telephone# +6285247175500	Fax#
Email yuzlena@ulm.ac.id	

☐ I hereby confirm that the manuscript was prepared in accordance with the instructions for authors of scientific publications, and that the content of this manuscript, or most of it, was not published in the journal indicated, and the manuscript was not submitted for publication elsewhere.

☐ I hereby confirm my consent to pay the Article Processing Charges (APC) EUR 450 in the case of the manuscript acceptance for publication. I am aware and hereby confirm that the APC is non-refundable.

☐ I hereby confirm my consent to make the payment for English Language Editing Services (EUR 150).

Publication fee payer details

____

____

Signature of the Corresponding author **Date**

Copyright Agreement

Manuscript title: UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD

Full names of all authors: Novitasari Novitasari, Yuslena Sari, Yudi Firmanul Arifin, Nurul Fathanah Mustamin, Erika Maulidiya

Full name and address of the corresponding author:

Yuslena Sari

Telephone/Whatsap: +6285247175500 Fax: _________________ Email: yuzlena@ulm.ac.id_

License Agreement

(1) Authors own all the copyright rights for the paper.
(2) Submitted manuscript is an original paper.
(3) Authors hereby grant the Issues of Journal of SWJTU with an exclusive, royalty-free, worldwide license to email the paper to all who will ask for it.
(4) All authors have made a significant contribution to the research and are ready to assume joint responsibility for the paper.
(5) All authors have seen and approved the manuscript in the final form as it is submitted for publication.
(6) This manuscript has not been published and also has neither been submitted nor considered for publication elsewhere
(7) The text, illustrations and any other materials, included into the manuscript, do not infringe any existing intellectual property rights or other rights of any person or entity.
(8) The editors of the Issues of Journal of SWJTU, its personnel or the Editorial Board members accept no responsibility for the quality of the idea expressed in this publication.

I am the Corresponding author and have full authority to enter into this agreement.

Full name, affiliation and position: Yuslena Sari, Department of Information Technology, Universitas Lambung Mangkurat

Signature:____Date: 28 April 2023

UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD

by Novitasari Novitasari

西南交通大学学报

第 57 卷 第 6 期
2022 年 12 月

JOURNAL OF SOUTHWEST JIAOTONG UNIVERSITY

Vol. 57 No. 6
Dec. 2022

ISSN: 0258-2724 DOI：10.35741/issn.0258-2724.57.6.1

Category
Please select the type of your manuscript
(see Author Guidelines)
Category.
Please select the most suitable field of scientific research
from the approved list of the journal categories

UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD

Novitasari Novitasari[a], Yuslena Sari[b,*], Yudi Firmanul Arifin[c], Nurul Fathanah Mustamin[d], Erika Maulidiya[e]

[a]Department of Civil Engineering, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, novitasari@ulm.ac.id
[b]Department of Information Technology, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, yuzlena@ulm.ac.id
[c]Faculty of Forestry, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, yudifirmanul@ulm.ac.id
[d]Department of Information Technology, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, nurul.mustamin@ulm.ac.id
[e]Department of Information Technology, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, 1810817220017@mhs.ulm.ac.id

Received: • Review: • Accepted: • Published

Abstract

Land cover is an important factor in geographic analysis, ranging from physical geography studies, approaches to sustainable planning to environmental analysis. Vegetation analysis according to the Indonesian National Standard (SNI 7645:2014) is classified based on density. The vegetation density index is divided into 4, namely non-vegetation, bare, medium and high. In the technical aspect to obtain information related to vegetation, this can be done using remote sensing. Remote sensing uses two data to obtain information, namely satellite data and UAV data. This study used UAV data with shooting locations in the Liang Anggang Protection Forest in classifying land cover. The method used was Convolutional Neural Network with feature extraction used in this study was GLCM. This research used the ShuffleNet v2 architecture on the CNN method. The findings of this study used two models, namely the CNN model without the GLCM process and compared to the CNN model with the addition of the GLCM process, resulting in a comparison that was quite far from the accuracy value obtained. The CNN model obtained an accuracy value of 80%, while the CNN model with GLCM using segmentation was 49.9% and without segmentation was 44.53%.

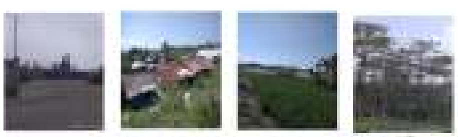
Keywords: Classification; class; CNN; GLCM; accuracy.

摘要 The authors may not translate the abstract and keywords into Chinese themselves.

关键词:

I. INTRODUCTION

From physical geography studies to approaches to sustainable planning to environmental analysis, land cover is an important factor in geographic analysis. Environmental analysis needs surface vegetation-based land cover information [1]. The entire plant of an area that serves as a land cover is referred to as vegetation. Vegetation is the entire plant of an area that serves as a land cover. According to the Indonesian National Standard [2], vegetation analysis is classified based on density. Non-vegetation, bare, medium, and high vegetation density indexes are used [2]. In addition to determining the level of vegetation density on land cover, it is important to be able to distinguish vegetation density in the form of an image, which makes data processing easier. N.A. Harahap conducted research which provides an image of the classification of vegetation density classes based on the images shown in Figure 1.1.



(a) (b) (c) (d)

Figure 1. (a) Non Vegetation, (b) Bare Vegetation, (c) Medium Vegetation, (d) High Vegetation

Vegetation analysis is one method for studying the arrangement and composition of vegetation in terms of plant shape (structure). In terms of technology, remote sensing can be used to obtain information about vegetation. Remote sensing obtains information from two sources: satellite data and UAV data. Previous research that used remote sensing technology by utilizing satellite data resulted in data accuracy ranging from 63% - 85% using various methods [1], [3]. Because satellite data is a traditional format based on statistical reporting and sampling surveys, determining vegetation density is critical [3]. Remote sensing with satellite data has been widely used in the identification and classification of land cover patterns across a wide geographic coverage, but the use of satellite data, which has a high operating altitude and is easily influenced by weather, clouds, and other external factors, is being reconsidered. Remote sensing technology can quickly and precisely provide spatial information on the earth surface. The object being sensed, the sensor for recording the object, and the electronic waves emitted by the earth surface are the three main components of remote sensing.

Remote sensing technology can quickly and precisely provide spatial information on the earth surface. The object being sensed, the sensor for recording the object, and the electronic waves emitted by the earth surface are the three main components of remote sensing. As technology advances, remote sensing facilities such as the Unmanned Aerial Vehicle (UAV) become more practical and easier to implement. The emergence of UAV raises significant potential as a tool for environmental and ecological analysis, such as monitoring agricultural land, forest fires, arctic lichen distribution, and mapping of mangrove forests. The generation of spatial information based on aerial image data using drones has enormous potential for the advancement of remote sensing technology, such as area classification. The benefits of using a UAV include faster and more flexible data acquisition, results that are more real-time, and low and light operating and maintenance costs. Apart from the ability to fly through clouds and produce cloud-free images, it differs from satellite imagery, which is heavily influenced by atmospheric conditions. UAV imagery has a high resolution when compared to satellite imagery, reaching a spatial resolution of less than 1 cm, which is much more detailed than satellite (30cm) and aircraft (10cm) imagery [1]. Optimal results that can be obtained from the use of UAV in object classification and the appropriate method for processing data with UAV imagery.

Before being processed in a classification model, image data requires feature extraction techniques to determine certain characteristics possessed by images to aid in object identification (image analysis) [1], [4], [5]. The resulting features will be selected first in the feature extraction process to obtain features with a high influence as a reference for the classification process. The function of feature extraction is to extract the necessary information from an image. Shape, colour, and texture extraction are the three types of feature extraction. Images with a slight colour can benefit from feature extraction using the Gray Level Co-occurrence Matrix (GLCM) method, which is a second level statistical method that computes the frequency of pairs of pixels in an image that have the same gray level and applies the additional knowledge obtained through pixel spatial relationships [6]. Using edge information, the co-occurrence matrix embeds the distribution of grayscale transitions. The majority of the information required to calculate the threshold value in the GLCM technique is straightforward but efficient [7]. S. Karthikeyan and N. Rengarajan use the GLCM algorithm with up to 95% accuracy. Previous research has compared GLCM feature extraction to LBP, MI, CLBP, LBGLCM, and GLRLM, with the accuracy results proving that using feature extraction in classification using GLCM produces better results than using other methods. GLCM accuracy results range from 70% to 93% [7]–[9].

Visual interpretation methods, pixel-based digital classification methods, and object-based classification methods are used in land cover mapping based on remote sensing imagery. Land cover analysis researchers are interested in the use of data mining methods. Land classification, Machine Learning, and Deep Learning have all made extensive use of classification methods. Deep learning, which is included in the supervised classification, is developed and produced by the machine learning method. Deep learning methods are widely used in satellite image analysis because they are powerful and intelligent in image processing. Deep learning methods are still evolving, with the Convolutional Neural Network (CNN) deep learning method producing the most significant results in image recognition to date. Deep Learning has demonstrated that this architecture, particularly CNN, can learn human-level solutions to specific visual tasks. This method has been used extensively in remote sensing image analysis tasks such as object detection in images, image recording, scene classification, segmentation, object-based image analysis, and land use and land cover classification [10]. CNN is one of the most recent Deep Learning methods to emerge. This method has been shown to be useful for pattern recognition and object classification [1]. Previous research using the CNN method to classify land cover yielded satisfactory accuracy results ranging from 73% to 98% [1], [10], [11].

CNN has a variety of popular architectures, including LeNet5 (1998), AlexNet (2012), ZFNet (2013), GoogleNet (2014), ResNet (2015), FractalNet (2016), ShuffleNet (2018), and others. Previous research has compared the use of architecture on CNN in the field of classification. The compared architectures demonstrate the advantage and disadvantage of each, for architectures that are widely used in the field of image classification and are relatively new, and have been compared with several other architectures, ShuffleNet. ShuffleNet is a very efficient CNN architecture with fast accuracy. Research that has used the ShuffleNet architecture and has made comparisons with other architectures such as GoogleNet, DenseNet, MobileNet, Xception, IGCV2, EffNet V1, EffNet V2, IoTNet-3-5 and ResNet50 in the classification process states that the ShuffleNet architecture increases the accuracy of 82% - 98% with less memory usage and faster processing time [12]–[16].

The CNN method is widely used in the field of deep learning to conduct land cover classification. GLCM was used to extract features in this study. The ShuffleNet architecture on the CNN method will be used in this study. This research was carried out for a month in the Liang Anggang Protected Forest area, Banjarbaru block 1 area, with targeted data collection. The location for this study was chosen based on observations made during the observation and survey of the block 1 area, where, according to the 2017 Provincial Forestry Office, an area of 479 hectares of block 1 area is filled with land such as agriculture, plantations, roads and settlements, as well as 494 hectares of forest. In addition to being a peatland, the research site, particularly in block 1, meets the characteristics and suitability of the needs in collecting data for land cover classification in terms of vegetation density types (bare, medium, and high) that can be seen with the naked eye during observations and surveys. This study classified land cover, with a focus on vegetation density, and the research location was chosen in accordance with the data requirements. The objective of this study was to determine the results of the best deep learning methods in land cover classification based on vegetation density. This study created research updates by combining UAV data with shooting locations in the Liang Anggang Protected Forest.

II. Research Methodology

A. Research Site

This study was being conducted in the Liang Anggang Protected Forest in Banjarbaru City, South Kalimantan. This is the Tangi Timber KPHP's management area. The protected forest designation is based on Minister of Forestry Decree No. 672/Kpts-II/1991 and Kep Menhut No. 434/Kpts-II/1996 with a total area of 2,250 hectares divided into two protected forest blocks, namely block 1 covering an area of 960 hectares including Liang Anggang sub-district, Banjarbaru and block 2 covering an area of 1290 hectares including the Gambut District, Banjar Regency.

Figure 2. Map of the Liang Anggang Protected Forest Area

The study lasted one month, from November to December 2021, and focused on the Warning Area (lock signal area from the airport) that caused the drone to be unable to operate.

B. Research Procedure

This research was conducted in the Liang Anggang Protected Forest area by conducting a field survey to assess the state of the vegetation or areas within the Protected Forest area. This study collected image data using drones to capture images from a height of 20 meters over a one-month period. Land was assigned coordinates based on the goal of image data collection using Google Earth Pro tools. Land with coordinates was exported in .KML format and later imported into DroneDeploy (website) to make directing drone flights on land easier. Then, the imported KML file was configured for flight altitude and 2D or 3D image capture. An illustration of image capture is shown in Figure 3.

Figure 3. Illustration of Image Data Retrieval

Before proceeding to the next stage, image data that has been recorded and stored according to predetermined coordinate points was processed. To facilitate operation with the method that was used later, image data was labelled. The CNN method was used in this study. Image data that has already been processed was then fed into the classification process using the method used in this study. Image data was classified using each method, and the accuracy value was calculated using tools. The obtained accuracy value was then analysed and compared to draw conclusions. The flow of this research is shown in Figure 4.

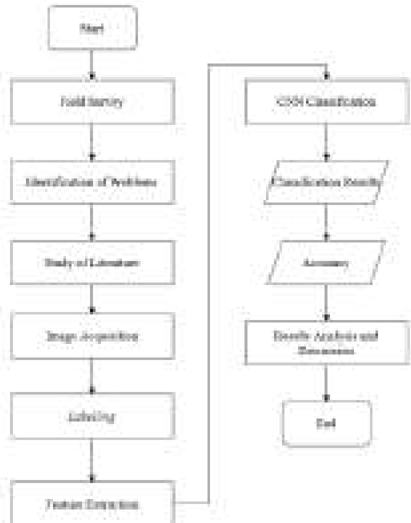


Figure 4. Diagram of Research Procedure

C. Feature Extraction

The purposes of feature extraction is to obtain the feature value of an object based on an image pixel intensity value relationship. The feature extraction process goal is to extract a special (unique) value from each image [17], [18]. This study used GLCM feature extraction with three primary features: correlation, homogeneity, and contrast. The feature extraction results created a GLCM version of the image using these three features. Figure 5 shows an illustration or description of the texture extraction results obtained with the GLCM feature.

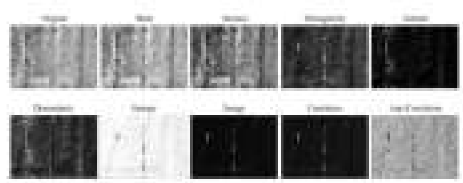


Figure 5. Result of Feature Extraction

The texture of an image was sought after by feature extracted images. The training data set consisted of 2400 images divided into three classes. This study applied 5 GLCM features to convert an input 2D image/image to an output 2D image/image to a gray level with a gray range of 0 to 1. The purpose of this step was to use gray level scaling to reduce the image volume to a more manageable size. Scaling to a grayscale level acted as a filter, removing some of the noise (de Mello, 2013). Figure 6 shows the scenario of the feature extraction test results with GLCM.

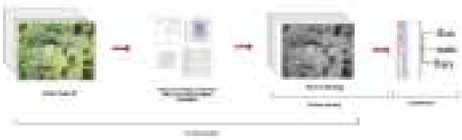


Figure 6. Ilustration of GLCM-CNN Feature Extraction

D. Classification of Convolutional Neural Networks

Only CNN neural networks can process grid structure data, such as two-dimensional images. The convolution layer is a linear algebra operation that generates a matrix of filters in the image to be processed. A convolution layer process is one of the many types of layers that can exist in a network. The image entered into the CNN classification model created during the fit model stage yielded an output calculated using the optimized weight. As a result, the classification model created should be able to classify the testing data into the correct class. This test was performed to calculate the accuracy value in the classification model that has been created. Figure 7 shows an illustration of the CNN classification process.

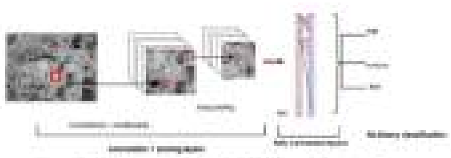


Figure 7. Illustration of CNN Classification Process

Figure 8 shows the classification flow using the CNN method

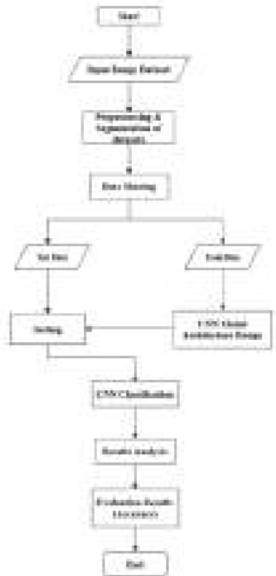


Figure 8. Stage of CNN Classification

E. Classification Analysis

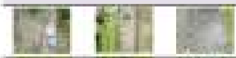
The results of the UAV image classification using two methods were analysed and the level of accuracy was determined. This study applied accuracy testing with a confusion matrix in the form of overall accuracy (OA) and Kappa coefficient accuracy. Proceed with the analysis of the CNN method classification results to obtain accurate results from the use of the CNN method in land cover classification.

III. TESTING

A. Image Dataset

The dataset used in this study was divided into three categories: bare, medium, and high. The total number of images collected was 3000, with 1000 for each class type category. The classification of these three classes was based on the condition of the Liang Anggang Protected Forest where the research location, particularly block 1, meets the characteristics and suitability of the needs in collecting data for land cover classification in terms of vegetation density types (bare, medium, and high) that can be seen with the naked eye during observation and surveys. This study classified land cover, with a focus on vegetation density, and the research location was chosen in accordance with the data requirements. Table 1 shows the results of categorizing three classes of vegetation density in terms of images based on the division of the available dataset [19].

Table 1.
Image of Vegetation Density

Image	Type of Vegetation Density
	Bare
	Medium
	High

B. Image Cropping

Because the image data obtained with the drone was too large, the data was resized by cutting the image and selecting specific areas to be used as training data. Cropped image data aimed to facilitate the classification process, did not take up much space or memory, and the classification process was light, so it did not require a long time in the classification process later. The image data was cropped to 256 x 256 pixels, reducing the image size to 159 KB. The cropped image data was classified into three types: bare, medium, and dense/high [20], [21]. Figure 9 shows the cropping results of image data.

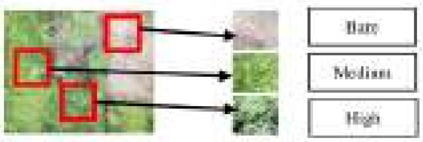


Figure 9. Image Data Cropping

C. Segmentation

Image segmentation was used to distinguish between objects and backgrounds [22], [23]. The separation process was designed to make classification and calculations easier. The image segmentation process was based on the difference in the image gray scale. To convert a colour image with r, g, and b matrix values into a grayscale image. The segmentation method, namely thresholding, can be used to change the colour image. The most basic method for segmenting was image development or image thresholding (de Mello, 2013). Thresholding was used to change the number of gray degrees in an image in order to create a binary image with pixel intensity values of 0 or 1.

D. Feature Extraction (GLCM)

In this study, GLCM was used for feature extraction, with three main features used: correlation, homogeneity, and contrast. This method was used to classify images, recognize textures, segment them, recognize objects, and analyse their colours. In the neighbourhood between pixels, GLCM had four angular directions: 0°, 45°, 90°, and 135°. When the angle was 0°, the pixel density was calculated by moving one distance to the right. Pixel adjacency was calculated using a 45° angle and 1 pixel distance to the top right. The angle is 90°, and the pixel density was calculated by a 1 pixel distance on top. A 135° angle was used, and neighbouring pixels were calculated by moving one pixel up [24]. The gray level of pixels was compared based on angle or neighbours at 0°, 45°, 90°, and 135° in this study. The feature extraction process was also compared the results of previously segmented images to those that have not been segmented.

IV. FINDINGS

A. Result of CNN Model

The formation of network architecture in the CNN algorithm can affect the results of model accuracy. In order to produce an optimal model, network architecture was used during the training process. This study applied an input image with a resolution of 256x256x3, with the aim for reducing image size so that the classification process took as little time as possible. This study applied the second version of the ShuffleNet architecture, which included one convolutional layer (Conv5), three stages (consisting of convolutional and shuffle units), one pooling layer (using Maxpool), and fc. The input image in the shuffleNet v2 model was 256 x 256 in size. The convolution and maximum pooling layers were added to the model's initial position to reduce the size of the feature graph. The convolution layer and pooling layer were replaced at the initial position by the convolutional layer (Conv1) with a 3 x 3 kernel, and the BN layer was added after Conv 1 and Conv 5. Figure 10 shows the flow of the proposed model.

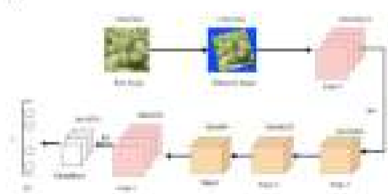


Figure 10. Flow of Model Shuffle Net v2 Model

The results of training and testing accuracy were obtained after going through several processes in the CNN algorithm. The value for training accuracy can be found in the "Accuracy" column, while the value for testing accuracy can be found in the "Validation Accuracy" column.

Accuracy was the value calculated by calculating the accuracy of the training dataset and model predictions. Validation Accuracy is the value calculated by calculating the accuracy of the validation dataset and predictions from the model using validation dataset input data. This procedure used a total of 50 epochs. Previously, epoch comparisons were performed to determine the accuracy and validation results of each training with a different number of epochs. This epoch comparison was intended to find the best model. The number of epochs compared ranges between 25 and 100 [25]. The use of epoch had a significant impact on the resulting accuracy. Because epoch can improve accuracy and the resulting accuracy was stable, it was critical to use the correct epoch in training data to achieve maximum accuracy. The table below compares training results based on the number of epochs.

Table 2.
Comparison of Epoch

Epoch	**Accuracy**	**Loss**	**Accuracy Validation**	***Validation Loss***	**Time**
25	97.46%	0.0686	71%	1.9176	20minute
50	98.75%	0.0358	81.33%	1.2536	56minute
75	99.08%	0.0279	74%	2.4516	1hour 32minute
100	99.29%	0.0218	74.17%	1.1251	2hour 20minute

The training model's accuracy with a total of 50 epochs is 98.75% with a loss of 0.0218. The validation accuracy value for the 50 epochs is 81.33%, which is higher than the other epochs. According to the table, the closer to the highest epoch, the higher the accuracy obtained from the testing results. However, if more than 100 epochs are added, the accuracy value decreases because too many epochs can also affect the large number of datasets. The testing procedure used training data consisting of 2400 image data and 600 image data for each class, as well as 200 image data for each class. Table 3 shows the results of the confusion matrix.

Table 3.
Confusion Matrix

Matrix		***Predict Class***		
		Bare	Medium	High
Actual Class	Bare	200	0	4
	Medium	9	92	88
	High	8	12	178

Based on the results of table 3, the model's predictions on the new data testing data show promising results. Although the prediction of the Bare class is correctly classified as the Bare class, up to four miss classifications from the Bare image data input are classified as the High class. While the Medium class prediction is correctly classified as the Medium class, as many as 9 miss classifications from the input image data are classified as the Bare class. In addition, up to 88 misclassifications of input image data are classified as High. The High class prediction is correctly classified as the High class, but up to 8 miss classifications from the High image data input are classified as the Bare class. As many as 12 misclassifications of input image data were classified as Medium. The overall accuracy of the matrix and kappa accuracy are calculated as follows:

Overall Accuracy = 469/600 = 80%
Kappa = 70%

So, the model's accuracy with a 256x256 input image and a total of 600 image data obtained an accuracy value of 80% and a kappa accuracy of 70%.

B. Result of CNN Model with GLCM

The addition of three GLCM features, namely contrast, homogeneity, and correlation, is the result of the next training model. The procedure involved extracting 3,000 GLCM result image data and producing 9,000 image data that was processed by CNN. This study also compared the direction angles of 0°, 45°, 90°, and 135° to extract images per angle. The GLCM process used a total of 27,000 image data through the segmentation stage. This was done to determine how well each feature performed in the image classification process.

For the GLCM process with CNN going through the segmentation stage, the results of data training with the CNN model and each GLCM feature by going through the segmentation stage with 9,000 data for each angle, can be seen at an angle of 135° getting the highest validation accuracy value from other angles, namely 60.11% with a value validation loss of 0.8460. For each feature, this training procedure applied a total of 50 epochs. This training process took approximately 20-30 minutes per corner. Table 5 shows the results of the GLCM training data per corner.

Table 4.
Comparison of Training Per Angle

Angle	Accuracy	Loss	*Validation Accuracy*	*Validation Loss*	Time
0°	95.24 %	0.1377	50.28 %	0.6096	29 mnt 6 scnd
45°	94.99 %	0.1346	50.39 %	0.5526	31 mnt 20 scnd
90°	96.42 %	0.1055	59.94 %	0.7616	31 mnt 12 scnd
135°	96.26 %	0.1176	60.11 %	0.8460	31 mnt 11 scnd

The training data is 9,000 images, and the test data is 1,800 images, with 3,000 images in each class. Table 5 shows the confusion matrix results for the CNN model process with GLCM that went through the segmentation stage.

Table 5.
Confusion Matrix

Matrix		*Predict Class*		
		Bare	Medium	High
Actual Class	Bare	407	92	83
	Medium	54	419	137
	High	83	218	307

According to the results in table 5, the model's prediction results on new data testing data are poor. Although the prediction of the Bare class is correct, as many as 92 miss classifications from the Bare image data input are classified as Medium. In addition, up to 83 miss classifications from the Bare image data input are classified as High. While the Medium class prediction was correctly classified as the Medium class, as many as 54 miss classifications from the input image data were classified as the Bare class. In addition, 137 misclassifications of image data input Medium are classified as High. The High class prediction is correctly classified as the High class, but up to 83 miss classifications from the High image data input are classified as the Bare class. In addition, 218 miss classifications from the High image data input are classified as Medium. The overall accuracy of the matrix and kappa accuracy are calculated as follows:

Overall Accuracy = 1133/1800x100% = 62,99%

Kappa = 44,37%

So, with an input image of 256x256 pixels and a total of 1800 image data, the model produced an accuracy value of 62.99% and a kappa accuracy of 44.37%.

V. Discussion

When the CNN model without the GLCM process was compared to the CNN model with the GLCM process, the comparison was quite far from the accuracy values obtained. The CNN model achieved an accuracy of 80%, while the CNN model with GLCM achieves 62.99% segmentation. This showed that the CNN model outperformed the GLCM process. According to the findings of the analysis, this occurred because the gray level in the image was leveled during the GLCM process, resulting in white and black colors in the image. The colours in the original image changed to white and black, resulting in a classification error. The GLCM process rendered the image colourless and rendered the entire image black.

During the testing of new data, there was a misclassification caused by nearly identical vegetation types. The input data for the CNN model was original image data with different types of vegetation, but based on the researcher's analysis, even though the texture between medium and high vegetation was different, the CNN model still had difficulty distinguishing and recognizing medium and high classes if the data simultaneously has the characteristics of an image that was filled with vegetation even though the type and texture of the vegetation was different. The CNN model with the GLCM method had a lot of misclassifications. The first reason was that the original image's colour had changed, making it difficult for the model to distinguish between classes. The second issue was that the type and texture of the vegetation were not visible in the image, so when predicting with the CNN and GLCM models on prototypes, the bare class data was read as medium class. High class reads as medium class.

VI. Conclusion

The conclusion is that comparing the CNN model without the GLCM process to the CNN model with the GLCM process produces a comparison that is quite far from the accuracy value obtained. The CNN model achieves an accuracy of 80%, while the CNN model with GLCM achieves 62.99% segmentation. This demonstrates that the CNN model outperforms the GLCM process in land cover classification. This demonstrates that the image processing process has a significant impact on the classification and prediction stages.

Acknowledgment

This research was funded by DRTPM grant, National Basic Research Competition scheme, agreement number 113/E5/PG.02.00/PT/2023.

References

[1] Z. Xu, K. Guan, N. Casler, B. Peng, and S. Wang, "A 3D convolutional neural network method for land cover classification using LiDAR and multi-temporal Landsat imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 144, pp. 423–434, 2018, doi: 10.1016/j.isprsjprs.2018.08.005.

[2] S. N. Indonesia and B. S. Nasional, "SNI: Klasifikasi penutup lahan," 2014.

[3] F. Zhao, X. Wu, and S. Wang, "Object-oriented Vegetation Classification Method based on UAV and Satellite Image Fusion," *Procedia Comput. Sci.*, vol. 174, no. 2019, pp. 609–615, 2020, doi: 10.1016/j.procs.2020.06.132.

[4] M. Alkaff, H. Khatimi, W. Puspita, and Y. Sari, "Modelling and predicting wetland rice production using support vector regression," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 17, no. 2, pp. 819–825, 2019, doi: 10.12928/TELKOMNIKA.V17I2.10145.

[5] Y. Sari, E. S. Wijaya, A. R. Baskara, and R. S. D. Kasanda, "PSO optimization on backpropagation for fish catch production prediction," *Telkomnika (Telecommunication Comput. Electron. Control.*, vol. 18, no. 2, pp. 776–782, 2020, doi: 10.12928/TELKOMNIKA.V18I2.14826.

[6] Q. Wu, Y. Gan, B. Lin, Q. Zhang, and H. Chang, "An active contour model based on fused texture features for image segmentation," *Neurocomputing*, vol. 151, no. P3, pp. 1133–1141, 2015, doi: 10.1016/j.neucom.2014.04.085.

[7] Z. Xing and H. Jia, "Multilevel Color Image Segmentation Based on GLCM and Improved Salp Swarm Algorithm," *IEEE Access*, vol. 7, pp. 37672–37690, 2019, doi: 10.1109/ACCESS.2019.2904511.

[8] S. A. Alazawi, N. M. Shati, and A. H. Abbas, "Texture features extraction based on GLCM for face retrieval system," *Period. Eng. Nat. Sci.*, vol. 7, no. 3, pp. 1459–1467, 2019, doi: 10.21533/pen.v7i3.787.

[9] S. Ozturk and B. Akdemir, "Application of Feature Extraction and Classification Methods for Histopathological Image using GLCM, LBP, LBGLCM, GLRLM and SFTA," *Procedia Comput. Sci.*, vol. 132, no. Iccids, pp. 40–46, 2018, doi: 10.1016/j.procs.2018.05.057.

[10] O. Youme, T. Bayet, J. M. Dembele, and C. Cambier, "Deep Learning and Remote Sensing: Detection of Dumping Waste Using UAV," *Procedia Comput. Sci.*, vol. 185, no. June, pp. 361–369, 2021, doi: 10.1016/j.procs.2021.05.037.

[11] S. M. Hamylton *et al.*, "Evaluating techniques for mapping island vegetation from unmanned aerial vehicle (UAV) images: Pixel classification, visual interpretation and machine learning approaches," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 89, no. March, p. 102085, 2020, doi: 10.1016/j.jag.2020.102085.

[12] T. Lawrence and L. Zhang, "IoTNet: An efficient and accurate convolutional neural network for IoT devices," *Sensors (Switzerland)*, vol. 19, no. 24, 2019, doi: 10.3390/s19245541.

[13] G. Liu *et al.*, "I3d-shufflenet based human action recognition," *Algorithms*, vol. 13, no. 11, 2020, doi:

10.3390/a13110301.

[14] G. Losapio *et al.*, "Lightweight and efficient convolutional neural networks for recognition of dolphin dorsal fins," pp. 68–72, 2020.

[15] Z. Wang and L. Ma, "SYOLO: An Efficient Pedestrian Detection," *IOP Conf. Ser. Mater. Sci. Eng.*, vol. 768, no. 7, 2020, doi: 10.1088/1757-899X/768/7/072067.

[16] X. Zhang, X. Zhou, M. Lin, and J. Sun, "Shufflenet: An extremely efficient convolutional neural network for mobile devices," *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, pp. 6848–6856, 2018.

[17] Y. Sari, M. Alkaff, and R. A. Pramunendar, "Iris recognition based on distance similarity and PCA," *AIP Conf. Proc.*, vol. 1977, 2018, doi: 10.1063/1.5042900.

[18] Y. Sari, M. Alkaff, and M. Maulida, "Classification of Rice Leaf using Fuzzy Logic and Hue Saturation Value (HSV) to Determine Fertilizer Dosage," in *2020 Fifth International Conference on Informatics and Computing (ICIC)*, 2020, pp. 1–6. doi: 10.1109/ICIC50835.2020.9288585.

[19] Y. Sari, Y. F. Arifin, N. Novitasari, and M. R. Faisal, "Vegetation-Density Drone Dataset For Peatland Vegetation Classification," vol. 1, 2022, doi: 10.17632/TB26ZY2JST.1.

[20] Y. Sari, Y. Arifin, Novitasari, and M. Faisal, "Implementation of Deep Learning Based Semantic Segmentation Method To Determine Vegetation Density," *Eastern-European J. Enterp. Technol.*, vol. 5, no. 2–119, pp. 42–54, 2022, doi: 10.15587/1729-4061.2022.265807.

[21] Y. Sari, Y. F. Arifin, N. Novitasari, and M. R. Faisal, "Effect of Feature Engineering Technique for Determining Vegetation Density," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 7, pp. 655–661, 2022, doi: 10.14569/IJACSA.2022.0130776.

[22] Y. Sari, A. R. Baskara, and R. Wahyuni, "Classification of Chili Leaf Disease Using the Gray Level Co-occurrence Matrix (GLCM) and the Support Vector Machine (SVM) Methods," *2021 6th Int. Conf. Informatics Comput. ICIC 2021*, 2021, doi: 10.1109/ICIC54025.2021.9632920.

[23] Y. Sari, H. Suhud, A. R. Baskara, R. A. Pramunendar, and I. F. Radam, "Parking Lots Detection in Static Image Using Support Vector Machine Based on Genetic Algorithm," *Int. J. Intell. Eng. Syst.*, vol. 14, no. 6, pp. 476–487, 2021, doi: 10.22266/ijies2021.1231.42.

[24] R. A. Pramunendar, D. P. Prabowo, D. Pergiwati, Y. Sari, P. N. Andono, and M. A. Soeleman, "New workflow for marine fish classification based on combination features and CLAHE enhancement technique," *Int. J. Intell. Eng. Syst.*, vol. 13, no. 4, pp. 293–304, 2020, doi: 10.22266/IJIES2020.0831.26.

[25] C. A. B. de Mello, "Image thresholding," *Digit. Doc. Anal. Process.*, vol. 2006, no. Snati, pp. 71–98, 2013, doi: 10.1201/9781003082224-3.

参考文:

COVER LETTER

UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD
Land cover is an important factor in geographic analysis, ranging from physical geography studies, approaches to sustainable planning to environmental analysis. Vegetation analysis according to the Indonesian National Standard (SNI 7645:2014) is classified based on density. The vegetation density index is divided into 4, namely non-vegetation, bare, medium and high. In the technical aspect to obtain information related to vegetation, this can be done using remote sensing. Remote sensing uses two data to obtain information, namely satellite data and UAV data. This study used UAV data with shooting locations in the Liang Anggang Protection Forest in classifying land cover. The method used was Convolutional Neural Network with feature extraction used in this study was GLCM. This research used the ShuffleNet v2 architecture on the CNN method. The findings of this study used two models, namely the CNN model without the GLCM process and compared to the CNN model with the addition of the GLCM process, resulting in a comparison that was quite far from the accuracy value obtained. The CNN model obtained an accuracy value of 80%, while the CNN model with GLCM using segmentation was 49.9% and without segmentation was 44.53%.
Classification; class; CNN; GLCM; accuracy.

Type of manuscript (please specify):

- ☐ Research article
- ☐ Review article
- ☐ Brief report
- ☐ Short communication
- ☐ Research note

Yuslena Sari Jl. Brig. Hasan Basry Kayutangi Banjarmasin, Indonesia, 70123	
Telephone# +6285247175500	Fax#
Email yuzlena@ulm.ac.id	

☐ I hereby confirm that the manuscript was prepared in accordance with the instructions for authors of scientific publications, and that the content of this manuscript, or most of it, was not published in the journal indicated, and the manuscript was not submitted for publication elsewhere.

☐ I hereby confirm my consent to pay the Article Processing Charges (APC) EUR 450 in the case of the manuscript acceptance for publication. I am aware and hereby confirm that the APC is non-refundable.

☐ I hereby confirm my consent to make the payment for English Language Editing Services (EUR 150).

Publication fee payer details

____

____

Signature of the Corresponding author Date

Copyright Agreement

Manuscript title: UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD

Full names of all authors: Novitasari Novitasari, Yuslena Sari, Yudi Firmanul Arifin, Nurul Fathanah Mustamin, Erika Maulidiya

Full name and address of the corresponding author:

Yuslena Sari

Telephone/Whatsap: +6285247175500 Fax: ________________ Email: yuzlena@ulm.ac.id_

License Agreement

(1) Authors own all the copyright rights for the paper.
(2) Submitted manuscript is an original paper.
(3) Authors hereby grant the Issues of Journal of SWJTU with an exclusive, royalty-free, worldwide license to email the paper to all who will ask for it.
(4) All authors have made a significant contribution to the research and are ready to assume joint responsibility for the paper.
(5) All authors have seen and approved the manuscript in the final form as it is submitted for publication.
(6) This manuscript has not been published and also has neither been submitted nor considered for publication elsewhere
(7) The text, illustrations and any other materials, included into the manuscript, do not infringe any existing intellectual property rights or other rights of any person or entity.
(8) The editors of the Issues of Journal of SWJTU, its personnel or the Editorial Board members accept no responsibility for the quality of the idea expressed in this publication.

I am the Corresponding author and have full authority to enter into this agreement.

Full name, affiliation and position: Yuslena Sari, Department of Information Technology, Universitas Lambung Mangkurat

Signature:____Date: 28 April 2023

UTILIZATION OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD

ORIGINALITY REPORT

18%	14%	16%	%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

1%

★ web.archive.org

Internet Source

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		

西南交通大学学报

第 58 卷 第 3 期
2023 年 6 月

JOURNAL OF SOUTHWEST JIAOTONG UNIVERSITY

Vol. 58 No. 3
June 2023

ISSN: 0258-2724

DOI：10.35741/issn.0258-2724.58.3.63

Research article

Computer and Information Science

USE OF UAV IMAGES FOR PEATLAND COVER CLASSIFICATION USING THE CONVOLUTIONAL NEURAL NETWORK METHOD

使用卷积神经网络方法将无人机图像用于泥炭地覆盖分类

Novitasari Novitasari [a], Yuslena Sari [b,]*, Yudi Firmanul Arifin [c], Nurul Fathanah Mustamin [b], Erika Maulidiya [b]

[a] Department of Civil Engineering, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, novitasari@ulm.ac.id

[b] Department of Information Technology, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, yuzlena@ulm.ac.id, nurul.mustamin@ulm.ac.id, 1810817220017@mhs.ulm.ac.id

[c] Faculty of Forestry, Universitas Lambung Mangkurat
Jl. Brig. Hasan Basry, Banjarmasin, Indonesia, yudifirmanul@ulm.ac.id

Received: March 8, 2023 ▪ Reviewed: April 24, 2023
▪ Accepted: May 26, 2023 ▪ Published: June 30, 2023

Abstract

Land cover or vegetation density in tropical peatland is an essential factor in hydrology response in geographic analysis, ranging from physical geography studies and approaches to sustainable planning to environmental research. Vegetation analysis according to the Indonesian National Standard (SNI 7645:2014) is classified on the basis of density. The vegetation density index is divided into four categories: non-vegetation, bare, medium, and high. In the technical aspect, to obtain information related to vegetation, this can be done using remote sensing. Remote sensing uses two types of data to obtain information: satellite data and UAV data. This study used UAV data with shooting locations in the Liang Anggang Protection Forest for classifying land cover. The method used was convolutional neural network with feature extraction used in this study was GLCM. This research used the ShuffleNet v2 architecture for the CNN method. The findings of this study used two models: the CNN model without GLCM process and compared to the CNN model with the addition of GLCM process, resulting in a comparison that was quite far from the accuracy value obtained. The CNN model obtained an accuracy value of 80%, while the CNN model with GLCM using segmentation gained 49.9% and without segmentation - 44.53%.

Keywords: Tropical Peatland, Vegetation Density, Classification, Class, Convolutional Neural Network, Gray Level Co-Occurrence Matrix, Accuracy

摘要 热带泥炭地的土地覆盖或植被密度是地理分析中水文响应的一个重要因素，范围从自然地理

研究和可持续规划方法到环境研究。根据印度尼西亚国家标准(SNI 7645:2014)的植被分析是根据密度进行分类的。植被密度指数分为无植被、裸露、中、高四类。在技术方面，要获得与植被有关的信息，可以使用遥感来完成。遥感使用两种类型的数据来获取信息：卫星数据和无人机数据。本研究使用带有梁安岗防护林拍摄地点的无人机数据对土地覆盖进行分类。使用的方法是卷积神经网络，本研究中使用的特征提取是 GLCM。本研究将洗牌网络 v2 架构用于美国有线电视新闻网方法。这项研究的结果使用了两个模型：没有 GLCM 过程的美国有线电视新闻网模型和添加了 GLCM 过程的美国有线电视新闻网模型进行比较，导致比较结果与获得的精度值相去甚远。美国有线电视新闻网模型获得了 80%的准确度值，而使用 GLCM 使用分割的美国有线电视新闻网模型获得了 49.9%，没有分割- 44.53%。

关键词: 热带泥炭地，植被密度，分类，类，卷积神经网络，灰度共生矩阵，精度

I. INTRODUCTION

Peatland management in Indonesia has many challenges due to peatland disasters [1], such as floods and wildfires. Wildfire caused by human action is the biggest. Many wildfires in Indonesia are intentional fires as part of residential developments [2]. They lead to change in land cover as changes in vegetation density in tropical peatland. Land cover or vegetation density change is one of the internal factors of hydrological response. From physical geography studies to approaches to sustainable planning to environmental analysis, land cover is an essential factor in geographic analysis, especially in disaster mitigation in tropical peatlands. Environmental analysis needs surface vegetation-based land cover information [3]. The entire plant of an area that serves as a land cover is referred to as vegetation. Vegetation is the entire plant of an area that serves as a land cover. According to the Indonesian National Standard [4], vegetation analysis is classified based on density. Non-vegetation, bare, medium, and high vegetation density indexes are used [4]. In addition to determining the level of vegetation density, it is important to be able to distinguish vegetation density in the form of an image, which makes data processing easier. Vegetation density in tropical peatland is classified [5]–[8] is based on Figure 1.

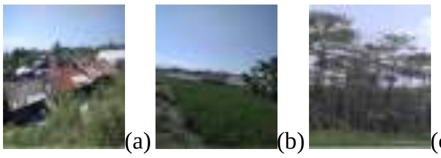

(a) (b) (c)

Figure 1. (a) Bare vegetation, (b) medium vegetation, (c) high vegetation

Vegetation density analysis in tropical peatland is one method for studying the arrangement and composition of vegetation in terms of plant shape (structure). In terms of technology, remote sensing can be used to obtain information about vegetation. Remote sensing obtains information from two sources: satellite data and UAV data. Previous research that used remote sensing technology by using satellite data resulted in data accuracy ranging from 63-85% using various methods [3], [9]. Because satellite data is a traditional format based on statistical reporting and sampling surveys, determining vegetation density is critical [9]. Remote sensing with satellite data has been widely used in the identification and classification of land cover patterns across a wide geographic coverage, but the use of satellite data, which has a high operating altitude and is easily influenced by weather, clouds, and other external factors, is being reconsidered. Remote sensing technology can quickly and precisely provide spatial information on the earth surface. The object being sensed, the sensor for recording the object, and the electronic waves emitted by the earth surface are the three main components of remote sensing.

Remote sensing technology can quickly and precisely provide spatial information on the earth surface. The object being sensed, the sensor for recording the object, and the electronic waves emitted by the earth surface are the three main components of remote sensing. As technology advances, remote sensing facilities such as the unmanned aerial vehicle (UAV) become more practical and easier to implement. The emergence of UAVs raises significant potential as a tool for environmental and ecological analysis, such as monitoring agricultural land, forest fires, Arctic lichen distribution, and mapping mangrove forests. The generation of spatial information based on aerial image data using drones has enormous potential for the advancement of remote sensing technology, such as area classification. The benefits of using a UAV include faster and more flexible data acquisition, more real-time results, and low and light operating and maintenance costs. Apart from the ability to fly through clouds and produce cloud-free images, it differs from satellite imagery, which is heavily influenced by atmospheric conditions. UAV imagery has a high resolution when compared to satellite imagery, reaching a spatial resolution of less than 1 cm, which is much more detailed than satellite (30 cm) and aircraft (10 cm) imagery [3]. Optimal results can be obtained from the use of UAV in object classification and the appropriate method for processing data with UAV imagery.

Before being processed in a classification model, image data requires feature extraction techniques to determine certain characteristics possessed by images to aid in object identification (image analysis) [3], [10], [11]. The resulting features will be selected first in the feature extraction process to obtain features with a high influence as a reference for the classification process. The function of feature extraction is to extract the necessary information from an image. Shape, colour, and texture extraction are the three types of feature extraction. Images with a slight colour can benefit from feature extraction using the gray level co-occurrence matrix (GLCM) method, which is a second-level statistical method that computes the frequency of pairs of pixels in an image that have the same gray level and applies the additional knowledge obtained through pixel spatial relationships [12]. Using edge information, the co-occurrence matrix embeds the distribution of grayscale transitions. Most of the information required to calculate the threshold value in the GLCM technique is straightforward but efficient [13]. S. Karthikeyan and N. Rengarajan used the GLCM algorithm with up to 95% accuracy. Previous research has compared GLCM feature extraction to LBP, MI, CLBP, LBGLCM, and GLRLM, with the accuracy results proving that using feature extraction in classification using GLCM produces better results than using other methods. GLCM accuracy results range from 70% to 93% [13]–[15].

Visual interpretation methods, pixel-based digital classification methods, and object-based classification methods are used in land cover mapping based on remote sensing imagery. Land cover analysis researchers are interested in the use of data mining methods. Land classification, machine learning, and deep learning have all made extensive use of classification methods. Deep learning, which is included in supervised classification, is developed and produced using the machine learning method. Deep learning methods are widely used in satellite image analysis because they are powerful and intelligent in image processing. Deep learning methods are still evolving, with the convolutional neural network (CNN) deep learning method producing the most significant results in image recognition to date. Deep Learning has demonstrated that this architecture, particularly CNN, can learn human-level solutions to specific visual tasks. This method has been used extensively in remote sensing image analysis tasks such as object detection in images, image recording, scene classification, segmentation, object-based image analysis, and land use and land cover classification [16]. CNN is one of the most recent deep learning methods to emerge. This method is useful for pattern recognition and object classification [3]. Previous research using the CNN method to classify land cover yielded satisfactory accuracy results ranging from 73% to 98% [3], [16], [17].

CNN has a variety of popular architectures, including LeNet5 (1998), AlexNet (2012), ZFNet (2013), GoogleNet (2014), ResNet (2015), FractalNet (2016), and ShuffleNet (2018). Previous research has compared the use of architecture on CNN in the field of classification. The compared architectures demonstrate their advantages and disadvantages, architectures that are widely used in the field of image classification and are relatively new, and have been compared with several other architectures. ShuffleNet is a very efficient CNN architecture with fast accuracy. Research that has used the ShuffleNet architecture and has made comparisons with other architectures such as GoogleNet, DenseNet, MobileNet, Xception, IGCV2, EffNet V1, EffNet V2, IoTNet-3-5 and ResNet50 in the classification process states that the ShuffleNet architecture increases the accuracy of 82%-98% with less memory usage and faster processing time [18]–[22].

The CNN method is widely used in the field of deep learning to conduct land cover classification. GLCM was used to extract features in this study. The ShuffleNet architecture on the CNN method will be used in this study. CNN is improved by using GLCM feature extraction to address the limitations of CNN. The high complexity is a limitation of CNN in feature extraction. This research was carried out for a month in the Liang Anggang Protected Forest, Banjarbaru Block 1, with the targeted data collection. The location for this study was chosen based on observations made during the observation and survey of the block 1 area, where, according to the 2017 Provincial Forestry Office, an area of 479 hectares of block 1 area is filled with land such as agriculture, plantations, roads and settlements, as well as 494 hectares of forest. In addition to being a peatland, the research site, particularly in Block 1, meets the characteristics and suitability of the needs in collecting data for land cover classification in terms of vegetation density types (bare, medium, and high) that can be seen with the naked eye during observations and surveys. This study classified land cover with a focus on vegetation density, and the research location was chosen in accordance with the data requirements. The objective of this study was to determine the results of the best deep learning methods in land cover classification based on vegetation density. This study created research updates by combining UAV data with shooting locations in the Liang Anggang Protected Forest.

II. RESEARCH METHODOLOGY

A. Research Site

This study was conducted in the Liang Anggang Protected Forest in Banjarbaru as the biggest wildfire in tropical peatland in South Kalimantan. This is Kayu Tangi KPHP management area. The protected forest designation is based on Minister of Forestry Decree No. 672/Kpts-II/1991 and Kep Menhut No. 434/Kpts-II/1996 with a total area of 2,250 hectares divided into two protected forest blocks, namely block 1 covering an area of 960 hectares including Liang Anggang sub-district, Banjarbaru and block 2 covering an area of 1290 hectares including Gambut District, Banjar Regency.

Figure 2. Map of the Liang Anggang Protected Forest

The study lasted for one month, from November to December 2021, and focused on the warning area (lock signal area from the airport) that caused the drone to be unable to operate.

B. Research Procedure

This research was conducted in the Liang Anggang Protected Forest area by conducting a field survey to assess the state of vegetation or areas within the protected forest area. This study collected image data using drones to capture images from a height of 20 m over a one-month period. Land was assigned coordinates based on the goal of image data collection using Google Earth Pro tools. Land with coordinates was exported in .KML format and later imported into DroneDeploy (website) to make directing drone flights on land easier. Then, the imported KML file was configured for flight altitude and 2D or 3D image capture. An illustration of image capture is shown in Figure 3.

Figure 3. Illustration of the image data retrieval

Before proceeding to the next stage, image data that had been recorded and stored according to predetermined coordinate points were processed. To facilitate the operation with the method that was used later, the image data were labeled. The CNN method was used in this study. Image data that had already been processed were then fed into the classification process using the method used in this study. The image data were classified using each method, and the accuracy value was calculated using the tools. The obtained accuracy value was then analyzed and compared to draw conclusions. The flow of this research is shown in Figure 4.

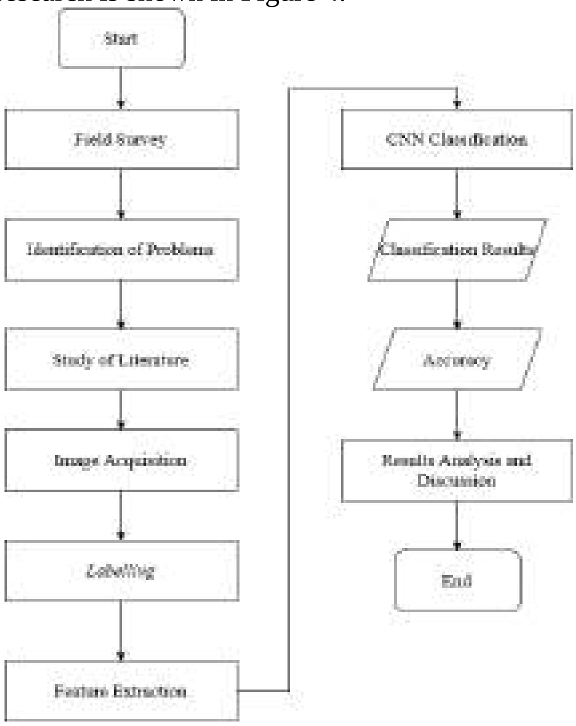


Figure 4. Diagram of the research procedure

C. Feature Extraction

The purpose of feature extraction is to obtain the feature value of an object based on an image pixel intensity value relationship. The goal of the feature extraction process is to extract a special (unique) value from each image [23], [24]. This study used GLCM feature extraction with three primary features: correlation, homogeneity, and contrast. The feature extraction results created a GLCM version of the image using these three features. Figure 5 shows an illustration or description of the texture extraction results obtained using the GLCM feature.

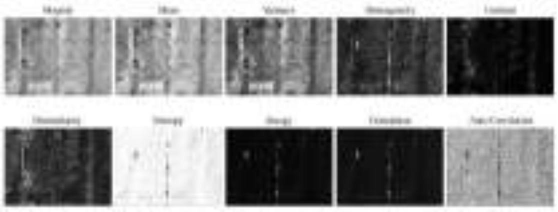


Figure 5. Results of the feature extraction

The texture of an image was sought after by feature-extracted images. The training data set consisted of 2400 images divided into three classes. This study applied 5 GLCM features to convert an input 2D image/image to an output 2D image/image to a gray level with a gray range of 0 to 1. The purpose of this step was to use gray-level scaling to reduce the image volume to a more manageable size. Scaling to a grayscale level acted as a filter, removing some of the noise [28]. Figure 6 shows the scenario of the feature extraction test results with GLCM.

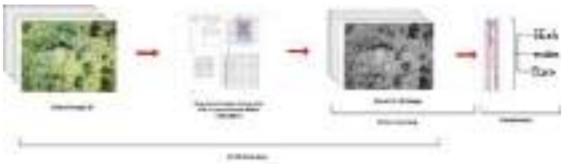


Figure 6. Illustration of GLCM-CNN feature extraction

D. Classification of Convolutional Neural Networks

Only CNN neural networks can process grid structure data, such as two-dimensional images. The convolution layer is a linear algebra operation that generates a matrix of filters in the image to be processed. A convolution layer process is one of the many types of layers that can exist in a network. The image entered into the CNN classification model created during the fit model stage yielded an output calculated using the optimized weight. As a result, the classification model created should be able to classify the testing data into the correct class. This test was performed to calculate the accuracy value in the classification model created. Figure 7 illustrates the CNN classification process.

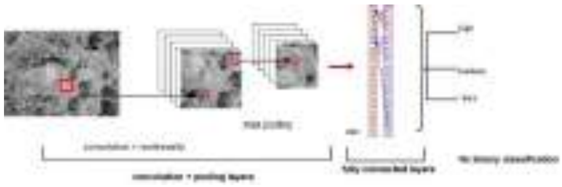


Figure 7. Illustration of CNN classification process

Figure 8 shows the classification flow using the CNN method.

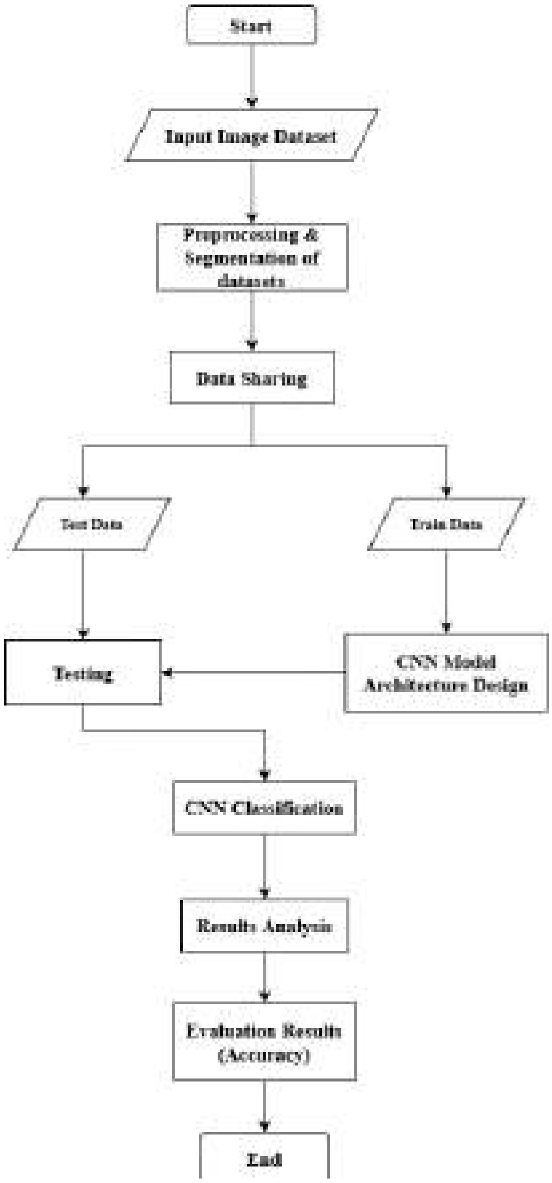


Figure 8. CNN classification stage

E. Classification Analysis

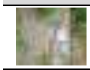
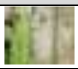
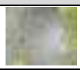
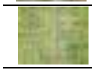
The results of UAV image classification using two methods were analyzed, and the level of accuracy was determined. This study applied accuracy testing with a confusion matrix in the form of overall accuracy (OA) and Kappa coefficient accuracy. Proceed with the analysis of the CNN method classification results to obtain accurate results from the use of the CNN method in the land cover classification.

III. TESTING

A. Image Dataset

The dataset used in this study was divided into three categories: bare, medium, and high. The total number of images collected was 3000, with 1000 for each class type category. The classification of these three classes was based on the condition of the Liang Anggang Protected Forest where the research location, particularly block 1, meets the characteristics and suitability of the needs in collecting data for land cover classification in terms of vegetation density types (bare, medium, and high) that can be seen with the naked eye during observation and surveys. This study classified land cover with a focus on vegetation density, and the research location was chosen in accordance with the data requirements. Table 1 shows the results of categorizing three classes of vegetation density in terms of images based on the division of the available dataset [6].

Table 1.
The images of vegetation density [5]–[8]

Image			Type of Vegetation Density
			Bare
			Medium
			High

B. Image Cropping

Because the image data obtained with the drone was too large, the data was resized by cutting the image and selecting specific areas to be used as training data. Cropped image data aimed to facilitate the classification process, did not take up much space or memory, and the classification process was light, so it did not require a long time in the classification process later. The image data was cropped to 256 x 256 pixels, reducing the image size to 159 KB. The cropped image data were classified into three types: bare, medium, and dense/high [5], [7]. Figure 9 shows the cropping results of the image data.

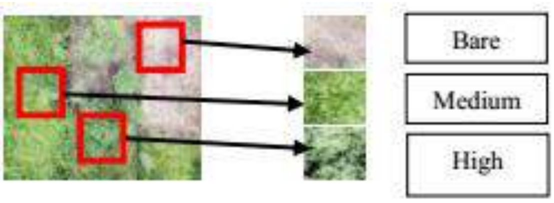


Figure 9. Image data cropping

C. Segmentation

Image segmentation was used to distinguish between objects and backgrounds [25], [26]. The separation process was designed to make classification and calculations easier. The image segmentation process was based on the difference in the image grayscale. To convert a colour image with r, g, and b matrix values into a grayscale image, the segmentation method, namely thresholding, can be used to change the colour image. The most basic method for segmenting is image development or image thresholding [28]. Thresholding was used to change the number of gray degrees in an image to create a binary image with pixel intensity values of 0 or 1.

D. Feature Extraction (GLCM)

In this study, GLCM was used for feature extraction with three main features: correlation, homogeneity, and contrast. This method was used to classify images, recognize textures, segment them, recognize objects, and analyse their colours. In the neighborhood between pixels, GLCM had four angular directions: 0°, 45°, 90°, and 135°. When the angle was 0°, the pixel density was calculated by moving one distance to the right. Pixel adjacency was calculated using a 45° angle and 1-pixel distance to the top right. The angle was 90°, and the pixel density was calculated by a 1-pixel distance on the top. A 135° angle was used, and neighboring pixels were calculated by moving one pixel up [27]. The gray level of pixels was compared based on the angle or neighbors at 0°, 45°, 90°, and 135° in this study. The feature extraction process also compared the results of previously segmented images to those that had not been segmented.

IV. FINDINGS

A. Result of CNN Model

The formation of network architecture in the CNN algorithm can affect the results of model accuracy. To produce an optimal model, network architecture was used during the training process. This study applied an input image with a resolution of 256x256x3, with the aim of reducing image size so that the classification process took as little time as possible. This study applied the second version of the ShuffleNet architecture, which included one convolutional layer (Conv5), three stages (consisting of convolutional and shuffle units), one pooling layer (using Maxpool), and fc. The input image in the ShuffleNet v2 model was 256 x 256 in size. The convolution and maximum pooling layers were added to the model's initial position to reduce the size of the feature graph. The convolution layer and pooling layer were replaced at the initial position by the convolutional layer (Conv1) with a 3 x 3 kernel, and the BN layer was added after Conv 1 and Conv 5. Figure 10 shows the flow of the proposed model.

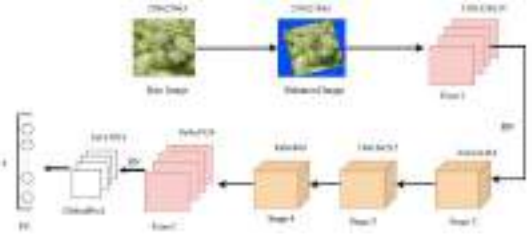


Figure 10. The ShuffleNet v2 model flow

The results of training and testing accuracy were obtained after going through several processes in the CNN algorithm. The value for training accuracy can be found in the "Accuracy" column, while the value for testing accuracy can be found in the "Validation Accuracy" column. Accuracy was the value calculated by calculating the accuracy of the training dataset and model predictions. Validation accuracy is the value calculated by calculating the accuracy of the validation dataset and predictions from the model using validation dataset input data. This procedure used 50 epochs. Previously, epoch comparisons were performed to determine the accuracy and validation results of each training with a different number of epochs. This epoch comparison was intended to find the best model. The number of epochs compared ranges between 25 and 100 [28]. The use of the epoch had a significant impact on the resulting accuracy. Because an epoch can improve accuracy, and the resulting accuracy was stable, it was critical to use the correct epoch in training data to achieve maximum accuracy. The table below compares training results based on the number of epochs.

Table 2.
Comparison of epochs

Epoch	Accuracy	Loss	Accuracy Validation	Validation Loss	Time
25	97.46%	0.0686	71%	1.9176	20 min
50	98.75%	0.0358	81.33%	1.2536	56 min
75	99.08%	0.0279	74%	2.4516	1 h 32 min
100	99.29%	0.0218	74.17%	1.1251	2 h 20 min

The training model's accuracy with 50 epochs is 98.75% with a loss of 0.0218. The validation accuracy value for the 50 epochs is 81.33%, which is higher than that for the other epochs. According to the table, the closer to the highest epoch, the higher the accuracy obtained from the testing results. However, if more than 100 epochs are added, the accuracy value decreases because too many epochs can also affect the large number of datasets. The testing procedure used training data consisting of 2400 image data, 600 image data for each class, and 200 image data for each class. Table 3 shows the results of the confusion matrix.

Table 3.
Confusion matrix

Matrix	Predict Class			
Actual Class		Bare	Medium	High
	Bare	209	0	4
	Medium	9	92	88
	High	8	12	178

Based on the results of Table 3, the model's predictions on the new data testing data show promising results. Although the prediction of the bare class is correctly classified as the bare class, up to four miss classifications from the bare image data input are classified as the high class. While the medium class prediction is correctly classified as the medium class, as many as 9 miss classifications from the input image data are classified as the bare class. In addition, up to 88 misclassifications of input image data are classified as high. The high class prediction is correctly classified as the high class, but up to 8 miss classifications from the high image data input are classified as the bare class. As many as 12 misclassifications of input image data were classified as medium. The overall accuracy of the matrix and kappa accuracy are calculated as follows:

Overall accuracy = 469/600 = 80%

Kappa = 70%

Thus, the model's accuracy with a 256x256 input image and 600 image data obtained an accuracy value of 80% and a kappa accuracy of 70%.

B. Result of CNN Model with GLCM

The addition of three GLCM features, namely contrast, homogeneity, and correlation, is the result of the next training model. The procedure involved extracting 3,000 GLCM result image data and producing 9,000 image data that were processed by the CNN. This study also compared the direction angles of 0°, 45°, 90°, and 135° to extract images per angle. The GLCM process used a total of 27,000 image data through the segmentation stage. This was done to determine how well each feature performed in the image classification process.

For the GLCM process with CNN going through the segmentation stage, the results of data training with the CNN model and each

GLCM feature by going through the segmentation stage with 9,000 data for each angle, can be seen at an angle of 135° getting the highest validation accuracy value from other angles, namely 60.11% with a value validation loss of 0.8460. For each feature, this training procedure applied 50 epochs. This training process took approximately 20-30 minutes per corner. Table 5 shows the results of the GLCM training data per corner.

Table 4.
The comparison of training per angle

Angle	Accuracy	Loss	Validation Accuracy	Validation Loss	Time
0°	95.24%	0.1377	50.28%	0.6096	29 min 6 sec
45°	94.99%	0.1346	50.39%	0.5526	31 min 20 sec
90°	96.42%	0.1055	59.94%	0.7616	31 min 12 sec
135°	96.26%	0.1176	60.11%	0.8460	31 min 11 sec

The training data is 9,000 images, and the test data is 1,800 images, with 3,000 images in each class. Table 5 shows the confusion matrix results for the CNN model process with GLCM that went through the segmentation stage.

Table 5.
Confusion matrix

Matrix	Predict Class			
Actual Class		Bare	Medium	High
	Bare	407	92	83
	Medium	54	419	137
	High	83	218	307

According to the results in Table 5, the model's prediction results on new data testing data are poor. Although the prediction of the bare class is correct, as many as 92 miss classifications from the bare image data input are classified as medium. In addition, up to 83 miss classifications from the bare image data input are classified as high. While the medium class prediction was correctly classified as the medium class, as many as 54 miss classifications from the input image data were classified as the bare class. In addition, 137 misclassifications of medium image data input were classified as high. The high class prediction was correctly classified as the high class, but up to 83 miss classifications from the high image data input were classified as the bare class. In addition, 218 miss classifications from the high image data input were classified as medium. The overall accuracy of the matrix and kappa accuracy are calculated as follows:

Overall Accuracy = 1133/1800x100% = 62,99%

Kappa = 44,37%

Thus, with an input image of 256x256 pixels and 1800 image data, the model produced an accuracy value of 62.99% and a kappa accuracy of 44.37%.

V. DISCUSSION

When the CNN model without the GLCM process was compared to the CNN model with the GLCM process, the comparison was quite far from the accuracy values obtained. The CNN model achieved an accuracy of 80%, while the CNN model with GLCM achieved 62.99% segmentation. This showed that the CNN model outperformed the GLCM process. According to the findings of the analysis, this occurred because the gray level in the image was leveled during the GLCM process, resulting in white and black colors in the image. The colors in the original image changed to white and black, resulting in a classification error. The GLCM process rendered the image colorless and rendered the entire image black.

During the testing of new data, there was a misclassification caused by nearly identical vegetation types. The input data for the CNN model was original image data with different types of vegetation, but based on the researcher's analysis, even though the texture between medium and high vegetation was different, the CNN model still had difficulty distinguishing and recognizing medium and high classes if the data simultaneously has the characteristics of an image that was filled with vegetation even though the type and texture of the vegetation was different. The CNN model with the GLCM method had many misclassifications. The first reason was that the original image's colour had changed, making it difficult for the model to distinguish between classes. The second issue was that the type and texture of the vegetation were not visible in the image, so when predicting with the CNN and GLCM models on prototypes, the bare class data was read as medium class. High class reads as medium class. This research is the only to classify vegetation density in tropical peatland.

VI. CONCLUSION

The conclusion is that comparing the CNN model without the GLCM process to the CNN model with the GLCM process produces a comparison that is quite far from the accuracy value obtained. The CNN model achieves an accuracy of 80%, while the CNN model with GLCM achieves 62.99% segmentation. This demonstrates that the CNN model outperforms the GLCM process in the land cover classification. This demonstrates that the image processing process has a significant impact on the stages of classification and prediction of vegetation density in tropical peatland.

ACKNOWLEDGMENT

This research was funded by DRTPM grant, National Basic Research Competition scheme, Agreement No. 113/E5/PG.02.00/PT/2023.

REFERENCES

[1] NOVITASARI, N., SUJONO, J., HARTO, S., MAAS, A., and JAYADI, R. (2018) Restoration of peat dome in ex-Mega rice project area in Central Kalimantan. *AIP Conference Proceedings*, 1977, 040008.

[2] NOVITASARI, N., SUJONO, J., HARTO, S., MAAS, A., and JAYADI, R. (2019) Drought Index for Peatland Wildfire Management in Central Kalimantan, Indonesia During El Niño Phenomenon. *Journal of Disaster Research*, 14 (7), pp. 939–948.

[3] XU, Z., GUAN, K., CASLER, N., PENG, B., and WANG, S. (2018) A 3D convolutional neural network method for land cover classification using LiDAR and multi-temporal Landsat imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 144, pp. 423–434.

[4] BADAN STANDAR NASIONAL INDONESIA (2014) SNI Klasifikasi penutup lahan.

[5] SARI, Y., ARIFIN, Y., NOVITASARI, and FAISAL, M. (2022) Implementation of Deep Learning Based Semantic Segmentation Method To Determine Vegetation Density. *Eastern-European Journal of Enterprise Technologies*, 5 (2(119)), pp. 42–54.

[6] SARI, Y., ARIFIN, Y.F., NOVITASARI, N., and FAISAL, M.R. (2022) *Vegetation-Density Drone Dataset for Peatland Vegetation Classification*. Mendeley Data. https://doi.org/10.17632/tb26zy2jst.1

[7] SARI, Y., ARIFIN, Y.F., NOVITASARI, N., and FAISAL, M.R. (2022) Effect of Feature Engineering Technique for Determining Vegetation Density. *International Journal of Advanced Computer Science and Applications*, 13 (7), pp. 655–661.

[8] SARI, Y., ARIFIN, Y.F., NOVITASARI, and FAISAL, M.R. (2022) The Effect of Batch Size and Epoch on Performance of ShuffleNet-CNN Architecture for Vegetation Density Classification. In: *Proceedings of the 7th International Conference on Sustainable Information Engineering and Technology, Malang, November 2022*. New York: Association for Computing Machinery, pp. 39–46.

[9] ZHAO, F., WU, X., and WANG, S. (2020) Object-Oriented Vegetation Classification Method Based on UAV and Satellite Image Fusion. *Procedia Computer Science*, 174, pp. 609–615.

[10] ALKAFF, M., KHATIMI, H., PUSPITA, W., and SARI, Y. (2019) Modelling and predicting wetland rice production using support vector regression. *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, 17 (2), pp. 819–825.

[11] SARI, Y., WIJAYA, E.S., BASKARA, A.R., and KASANDA, R.S.D. (2020) PSO optimization on backpropagation for fish catch production prediction. *TELKOMNIKA (Telecommunication, Computing, Electronics and Control)*, 18 (2), pp. 776–782.

[12] WU, Q., GAN, Y., LIN, B., ZHANG, Q., and CHANG, H. (2015) An active contour model based on fused texture features for image segmentation. *Neurocomputing*, 151 (Part 3), pp. 1133–1141.

[13] XING, Z. and JIA, H. (2019) Multilevel Color Image Segmentation Based on GLCM and Improved Salp Swarm Algorithm. *IEEE Access*, 7, pp. 37672–37690.

[14] ALAZAWI, S.A., SHATI, N.M., and ABBAS, A.H. (2019) Texture features extraction based on GLCM for face retrieval system. *Periodicals of Engineering and Natural Sciences*, 7 (3), pp. 1459–1467.

[15] OZTURK, S. and AKDEMIR, B. (2018) Application of Feature Extraction and Classification Methods for Histopathological Image using GLCM, LBP, LBGLCM, GLRLM and SFTA. *Procedia Computer Science*, 132, pp. 40–46.

[16] YOUME, O., BAYET, T., DEMBELE, J.M., and CAMBIER, C. (2021) Deep Learning and Remote Sensing: Detection of Dumping Waste Using UAV. *Procedia*

Computer Science, 185, pp. 361–369.

[17] HAMYLTON, S.M., MORRIS, R.H., CARVALHO, R.C., RODER, N., BARLOW, P., MILLS, K., and WANG, L. (2020) Evaluating techniques for mapping island vegetation from unmanned aerial vehicle (UAV) images: Pixel classification, visual interpretation and machine learning approaches. *International Journal of Applied Earth Observation and Geoinformation*, 89, 102085.

[18] LAWRENCE, T. and ZHANG, L. (2019) IoTNet: An efficient and accurate convolutional neural network for IoT devices. *Sensors*, 19 (24), 5541.

[19] LIU, G., ZHANG, C., XU, Q., CHENG, R., SONG, Y., YUAN, X., and SUN, J. (2020) I3D-Shufflenet Based Human Action Recognition. *Algorithms*, 13 (11), 301.

[20] LOSAPIO, G., MAGLIETTA, R., POLITI, T., STELLA, E., FANIZZA, C., HARTMAN, K., CARLUCCI, R., DIMAURO, G., and RENÒ, V. (2020) Lightweight and efficient convolutional neural networks for recognition of dolphin dorsal fins. In: *Proceedings of the IMEKO TC-19 International Workshop on Metrology for the Sea, Naples, October 2020,* pp. 68–72. Available from https://www.imeko.org/publications/tc19-Metrosea-2020/IMEKO-TC19-MetroSea-2020-13.pdf.

[21] WANG, Z. and MA, L. (2020) SYOLO: An Efficient Pedestrian Detection. *IOP Conference Series: Materials Science and Engineering*, 768 (7), 072067.

[22] ZHANG, X., ZHOU, X., LIN, M., and SUN, J. (2018) ShuffleNet: An Extremely Efficient Convolutional Neural Network for Mobile Devices. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, Utah, June 2018.* Manhattan, New York: Institute of Electrical and Electronics Engineers, pp. 6848–6856.

[23] SARI, Y., ALKAFF, M., and PRAMUNENDAR, R.A. (2018) Iris recognition based on distance similarity and PCA. *AIP Conference Proceedings*, 1977, 020044.

[24] SARI, Y., ALKAFF, M., and MAULIDA, M. (2020) Classification of Rice Leaf Using Fuzzy Logic and Hue Saturation Value (HSV) to Determine Fertilizer Dosage. In: *Proceedings of the 5th International Conference on Informatics and Computing, Gorontalo, November 2020*. Manhattan, New York: Institute of Electrical and Electronics Engineers, pp. 1–6.

[25] SARI, Y., BASKARA, A.R., and WAHYUNI, R. (2021) Classification of Chili Leaf Disease Using the Gray Level Co-Occurrence Matrix (GLCM) and the Support Vector Machine (SVM) Methods. In: *Proceedings of the 6th International Conference on Informatics and Computing, Jakarta, November 2021*. Manhattan, New York: Institute of Electrical and Electronics Engineers, pp. 1-4.

[26] SARI, Y., SUHUD, H., BASKARA, A. R., PRAMUNENDAR, R.A., and RADAM, I.F. (2021) Parking Lots Detection in Static Image Using Support Vector Machine Based on Genetic Algorithm. *International Journal of Intelligent Engineering and Systems*, 14 (6), pp. 476–487.

[27] PRAMUNENDAR, R.A., PRABOWO, D.P., PERGIWATI, D., SARI, Y., ANDONO, P.N., and SOELEMAN, M.A. (2020) New workflow for marine fish classification based on combination features and CLAHE enhancement technique. *International Journal of Intelligent Engineering and Systems*, 13 (4), pp. 293–304.

[28] RAJINIKANTH, V., PRIYA, E., LIN, H., and LIN, F. (2021) Image thresholding. In: *Hybrid Image Processing Methods for Medical Image Examination*. Boca Raton, Florida: CRC Press, pp. 69-102.

参考文:

[1] NOVITASARI, N.、SUJONO, J.、HARTO, S.、MAAS, A. 和 JAYADI, R. (2018) 加里曼丹中部前巨型水稻项目区泥炭丘修复. AIP 会议论文集，1977，040008。

[2] NOVITASARI, N.、SUJONO, J.、HARTO, S.、MAAS, A. 和 JAYADI, R.（2019）厄尔尼诺现象期间印度尼西亚加里曼丹中部泥炭地野火管理的干旱指数。灾害研究杂志，14 (7)，第 939–948 页。

[3] XU, Z., GUAN, K., CASLER, N., PENG, B., 和 WANG, S. (2018) 一种使用激光雷达和多时相陆地卫星图像进行土地覆盖分类的 3 丁卷积神经网络方法。ISPRS 摄影测量与遥感杂志，144，第 423–434 页。

[4] 巴丹标准印度尼西亚国民银行 (2014) SNI 土地覆盖分类。

[5] SARI, Y., ARIFIN, Y., NOVITASARI, 和 FAISAL, M. (2022) 实施基于深度学习的语义分割方法来确定植被密度。东欧企业技术杂志，5 (2(119))，第 42-54 页。

[6] SARI, Y., ARIFIN, Y.F., NOVITASARI, N., 和 FAISAL, M.R. (2022) 用于泥炭地植被分类的植被密度无人机数据集。门德利数据。https://doi.org/10.17632/tb26zy2jst.1

[7] SARI, Y., ARIFIN, Y.F., NOVITASARI, N., 和 FAISAL, M.R. (2022) 特征工程技术对确定植被密度的影响。国际高级计算机科学与应用杂志，13 (7)，第 655–661 页.

[8] SARI, Y., ARIFIN, Y.F., NOVITASARI, 和 FAISAL, M.R. (2022) 批量大小和时期对洗牌网-美国有线电视新闻网架构植被密度分类性能的影响. 载于：第七届可持续信息工程与技术国际会议论文集，玛琅，2022 年 11 月。纽约：计算机协会，第 39-46 页。

[9] ZHAO, F., WU, X., 和 WANG, S. (2020) 基于无人机和卫星图像融合的面向对象植被分类方法。程序计算机科学，174，第 609-615 页。

[10] ALKAFF, M.、KHATIMI, H.、PUSPITA, W. 和 SARI, Y. (2019) 使用支持向量回归建模和预测湿地水稻产量。电信公司（电信、计算、电子和控制），17 (2)，第 819–825 页。

[11] SARI, Y.、WIJAYA, E.S.、BASKARA, A.R. 和 KASANDA, R.S.D. (2020) 用于鱼类产量预测的反向传播的 PSO 优化。电信公司（电信、计算、电子和控制），18 (2)，第 776–782 页。

[12] WU, Q., GAN, Y., LIN, B., ZHANG, Q., 和 CHANG, H. (2015) 基于融合纹理特征的主动轮廓模型用于图像分割。神经计算，151（第 3 部分），第 1133–1141 页.

[13] XING, Z. 和 JIA, H. (2019) 基于 GLCM 和改进的萨尔普群的多级彩色图像分割算法。IEEE 访问，7，页数 37672–37690。

[14] ALAZAWI, S.A., SHATI, N.M., 和 ABBAS, A.H. (2019) 基于 GLCM 的人脸检索系统纹理特征提取。工程与自然科学期刊，7 (3)，第 1459–1467 页。

[15] OZTURK, S. 和 AKDEMIR, B. (2018) 使用 GLCM、腰痛、LBGLCM、GLRLM 和自由贸易协定对组织病理学图像进行特征提取和分类方法的应用。程序计算机科学，132，第 40-46 页。

[16] YOUME, O., BAYET, T., DEMBELE, J.M., 和 CAMBIER, C. (2021) 深度学习和遥感：使用无人机检测倾倒垃圾。程序计算机科学，185，第 361-369 页。

[17] HAMYLTON, S.M., MORRIS, R.H., CARVALHO, R.C., RODER, N., BARLOW, P., MILLS, K., 和 WANG, L. (2020) 从无人机（无人机）测绘岛屿植被的评估技术) 图像：像素分类、视觉解释和机器学习方法。国际应用地球观测与地理信息杂志, 89, 102085.

[18] LAWRENCE, T. 和 ZHANG, L. (2019) 物联网：一种用于物联网设备的高效准确的卷积神经网络。传感器，19 (24)，5541.

[19] LIU, G., ZHANG, C., XU, Q., CHENG, R., SONG, Y., YUAN, X., 和 SUN, J. (2020) 基于三维-洗牌网的人类动作识别。算法，13 (11), 301。

[20] LOSAPIO, G.、MAGLIETTA, R.、POLITI, T.、STELLA, E.、FANIZZA, C.、HARTMAN, K.、CARLUCCI, R.、DIMAURO, G. 和 RENÒ, V.（2020) 用于识别海豚背鳍的轻量级高效卷积神经网络。在：伊梅科 TC-19 国际海洋计量研讨会论文集，那不勒斯，2020 年 10 月，第 68-72 页。可从 https://www.imeko.org/publications/tc19-Metrosea-2020/IMEKO-TC19-MetroSea-2020-13.pdf 获得。

[21] WANG, Z. 和 MA, L. (2020) 奏乐：一种高效的行人检测。眼压会议系列：材料科学与工程，768（7），072067。

[22] ZHANG, X., ZHOU, X., LIN, M., 和 SUN, J. (2018) 洗牌网：用于移动设备的极其高效的卷积神经网络. 在：IEEE/CVF 计算机视觉和模式识别会议记录，犹他州盐湖城，2018 年 6 月。纽约曼哈顿：电气和

电子工程师协会，第 6848-6856 页。

[23] SARI, Y. 、 ALKAFF, M. 和 PRAMUNENDAR, R.A. (2018) 基于距离相似度和主成分分析的虹膜识别。AIP 会议论文集，1977，020044。

[24] SARI, Y., ALKAFF, M. 和 MAULIDA, M. (2020) 使用模糊逻辑和色相饱和值(单纯疱疹病毒)对稻叶进行分类以确定肥料用量。载于：第五届信息学与计算国际会议论文集，哥伦打洛，2020 年 11 月。纽约曼哈顿：电气和电子工程师协会，第 1-6 页。

[25] SARI, Y., BASKARA, A.R. 和 WAHYUNI, R. (2021) 使用灰度共生矩阵(GLCM)和支持向量机(支持向量机)方法对辣椒叶病进行分类。载于：第六届信息学与计算国际会议论文集，雅加达，2021 年 11 月。纽约曼哈顿：电气和电子工程师协会，第 1-4 页。

[26] SARI, Y.、SUHUD, H.、BASKARA, A. R. 、 PRAMUNENDAR, R.A. 和 RADAM, I.F. (2021) 使用基于遗传算法的支持向量机检测静态图像中的停车场。国际智能工程与系统杂志，14 (6)，第 476–487 页。

[27] PRAMUNENDAR, R.A., PRABOWO, D.P., PERGIWATI, D., SARI, Y., ANDONO, P.N., 和 SOELEMAN, M.A. (2020) 基于组合特征和克拉赫增强技术的海洋鱼类分类新工作流程。国际智能工程与系统杂志，13 (4)，第 293–304 页。

[28] RAJINIKANTH, V. 、 PRIYA, E. 、 LIN, H. 和 LIN, F. (2021) 图像阈值处理。在：医学图像检查的混合图像处理方法。佛罗里达州博卡拉顿：钢筋混凝土出版社，第 69-102 页。