An enhanced MobileNetV2-191h architecture for accurate crescent moon detection using semantic segmentation

Muhammad Habib Yuhandri [a,1,*], Yuhandri [a,2], Sumijan [a,3]

[a] Universitas Putra Indonesia YPTK, Lubuk Begalung Street No. 1, Padang 25221, Indonesia
[1] mhabibyuhandri@upiyptk.ac.id*; [2] yuyu@upiyptk.ac.id; [3] sumijan@upiyptk.ac.id
* corresponding author

ARTICLE INFO

Article history
Received March 9, 2026
Revised May 27, 2026
Accepted May 29, 2026
Available online August 31, 2026

Keywords
Crescent moon detection
Semantic segmentation
Convolutional neural network
MobileNetV2
Deep learning

ABSTRACT

This study proposes an enhanced lightweight semantic segmentation framework, called MobileNetV2-191H, for accurate crescent moon detection in astronomical observation images. Crescent moon detection is a challenging task because the crescent moon typically appears as a very thin and low-contrast object that is highly affected by atmospheric interference, cloud cover, image noise, and illumination variation near the horizon. The dataset used in this study was obtained from crescent moon observation videos provided by the Indonesian Agency for Meteorology, Climatology, and Geophysics (BMKG), resulting in 6,283 extracted image frames that were divided into training and testing datasets. The proposed MobileNetV2-191H framework was developed by adapting the original MobileNetV2 architecture through the integration of additional segmentation-oriented layers, including transposed convolution, feature refinement, and pixel classification modules, to improve pixel-level segmentation capability for thin crescent moon objects. Experimental evaluation was conducted using semantic segmentation metrics, including precision, recall, F1-score, and Intersection over Union (IoU). The proposed model achieved 84.81% precision, 85.32% recall, 84.07% F1-score, and 72.64% IoU, outperforming several baseline segmentation models such as U-Net, DeepLabV3, SegNet, and EfficientNet. These results demonstrate that the proposed lightweight framework effectively improves crescent moon segmentation performance while maintaining computational efficiency for real-time astronomical observation applications.

© 2026 The Author(s).
This is an open access article under the CC–BY-SA license.

1. Introduction

The start of a month in the Islamic lunar calendar relies heavily on tracking down the first glimpse of the new crescent moon right after its conjunction phase [1], [2]. Spotting this initial sliver of light is absolutely vital for Muslims across the globe, as it marks the official beginning of sacred months like Ramadan, Shawwal, and Dhu al-Hijjah. Unfortunately, trying to catch sight of it with the naked eye presents major obstacles because the young crescent is incredibly thin, lacks strong contrast, and is easily obscured by atmospheric issues like heavy cloud cover, blurry air turbulence, and visual interference near the horizon line. Because these harsh natural conditions make it exceptionally tough to detect the moon visually, experts increasingly turn to advanced digital image processing methods to boost its visibility and ensure much higher precision during these critical observations [3]–[5]. With the advancement of computer vision technology, artificial intelligence–based approaches, particularly deep learning methods, have been increasingly employed for the analysis of astronomical images [6], [7]. One of the most widely used methods is the Convolutional Neural Network (CNN), which is capable of automatically extracting

image features through hierarchical learning processes [8]–[11]. CNN has been proven effective across various image segmentation and object detection applications involving complex visual characteristics, including medical image analysis, satellite imagery, and low-contrast object detection tasks [12] –[14]. These capabilities are particularly relevant for crescent moon detection, where the target object appears extremely thin and difficult to distinguish from atmospheric backgrounds. The primary advantage of Convolutional Neural Network (CNN) lies in its ability to learn hierarchical feature representations, enabling it to recognize complex patterns in images with a high degree of accuracy [9], [15], [16]. A number of earlier investigations have sought to utilize computer vision methods in order to spot the crescent moon. Past projects created video-focused systems for identifying the crescent moon by pulling out individual frames and using image modification tools, including Gaussian blur and adaptive thresholding, which help make the crescent moon much clearer to see [17]–[20]. The outcomes of this work show that computer vision strategies are capable of successfully identifying specific frames that show the crescent moon, reaching quite strong accuracy levels when looking at space observation data. Furthermore, alternative research has turned to artificial neural network systems to forecast whether the lunar crescent can be seen based on specific space-related factors, which proves that machine learning options can definitely boost the precision of these visibility forecasts [5]. Alongside this, a handful of papers have used image-enhancement methods, like Contrast Limited Adaptive Histogram Equalization (CLAHE), to raise the overall quality of crescent moon photographs before the actual spotting process begins [20]–[24].

Over the last few years, advanced deep learning strategies have become much more common for spotting and isolating specific items within space-related imagery [25], [26]. Frameworks like U-Net, SegNet, and DeepLab are frequently chosen for a wide variety of picture isolation tasks because they are highly capable of performing classification at the exact pixel level [27], [28]. These specific designs allow software systems to spot tiny items inside pictures with a very high degree of accuracy. On top of that, separate studies have also shown that utilizing standard convolutional neural networks for satellite pictures and medical scans can noticeably boost how accurately the system spots items that have tricky outlines or very low contrast levels [29]. Alongside these highly intricate isolation models, the creation of lightweight convolutional neural network designs has gained a lot of interest for running instant, real-time programs, focusing especially on the MobileNetV2 setup [30]. MobileNetV2 uses special design ideas known as inverted residuals and linear bottlenecks, which allow the overall system to keep its total number of parameters quite small while still keeping its powerful ability to understand and represent important visual features [31]–[34]. Because of this setup, this specific architecture is now widely used across a broad range of computer vision tasks, such as sorting pictures, spotting specific items, and isolating parts of images, especially when working on smaller gadgets that have strict limits on their processing power [34], [35].

MobileNetV2 is primarily designed for image classification tasks and does not inherently generate detailed pixel-level segmentation maps. Therefore, incorporating additional processing layers is essential to reconstruct pixel-wise class predictions from the extracted feature representations, thereby enabling accurate segmentation results. In this architecture, MobileNetV2 serves as a feature extraction backbone, while supplementary segmentation components process the extracted features to generate the final segmentation maps [36]–[39]. Several studies have demonstrated that the efficiency and performance of neural networks can be significantly improved through architectural modifications to convolutional neural networks (CNNs), particularly by integrating additional layers or specialized modules rather than relying solely on segmentation components [40]. The integration of additional layers enables neural networks to capture more complex and discriminative feature representations, thereby enhancing their ability to detect small objects and targets characterized by low contrast. This approach has been widely adopted in modern convolutional neural network (CNN) architectures to improve image segmentation accuracy across various applications. In the context of crescent moon segmentation, a significant challenge arises from the substantial class imbalance between the relatively small area occupied by the crescent moon and the predominantly large sky background, which can adversely affect the performance of conventional semantic segmentation models. Although established architectures such as U-Net, SegNet, and DeepLabV3 have demonstrated promising performance in segmenting relatively large

objects, accurately identifying extremely thin and faint crescent moon structures under challenging atmospheric conditions remains difficult, requiring effective spatial feature reconstruction and enhancement. To address these challenges, this study proposes a modified MobileNetV2 architecture incorporating additional segmentation-specific layers to preserve the fine structural characteristics of the crescent moon during pixel-level classification. Nevertheless, existing crescent moon detection approaches and conventional semantic segmentation architectures continue to exhibit limitations in preserving thin object boundaries and effectively handling low-contrast atmospheric conditions, particularly in datasets characterized by severe class imbalance.

Looking closely at these prior publications, one can deduce that upgrading CNN setups by altering network arrangements and inserting specialized layers offers great promise for boosting a model's skill in isolating objects hidden within highly intricate pictures [41], [42]. This capability becomes incredibly vital when applied to crescent moon tracking, where the target displays a remarkably thin silhouette that frequently blends directly into the surrounding sky. Consequently, the core objective of this research centers on creating an upgraded, lightweight semantic segmentation system built upon MobileNetV2 to detect the crescent moon within real astronomical photography. Instead of inventing a completely unprecedented CNN design from scratch, this project upgrades the conventional MobileNetV2 system by blending in extra segmentation-focused layers, such as transposed convolution, feature polishing, and pixel categorization units, to maximize the detection of slim, faint crescent moons. The most significant value of this research stems from tailoring a compact CNN framework to handle the demanding chore of crescent moon isolation, utilizing authentic astronomical images supplied by BMKG. Furthermore, this study delivers a complete workflow for preparing crescent moon data, which encompasses extracting individual frames, labeling ground truth data based on regions of interest, and testing the system under severe weather disturbances. The newly introduced MobileNetV2-191H system is explicitly engineered to achieve higher segmentation precision while preserving low computational demands for live, real-time crescent moon monitoring operations.

Driven by these recognized difficulties and unanswered research questions, this investigation intends to explore multiple crucial factors connected to the segmentation and tracking of the crescent moon. To begin with, this project analyzes the exact ways a lightweight, CNN-driven semantic segmentation system can be customized to improve the identification of thin, low-contrast crescent moons within astronomical photographs. In the second place, the research measures the precise degree to which reinforcing the MobileNetV2 design with extra segmentation-specific layers boosts overall performance when measured against the baseline MobileNetV2 model. Thirdly, the newly developed MobileNetV2-191H framework undergoes a direct comparison with several established semantic segmentation models, including U-Net, SegNet, DeepLabV3, and EfficientNet, using standard metrics like accuracy, precision, recall, and the F1-score. Lastly, this paper looks into whether the created lightweight framework can sustain top-tier tracking performance while successfully enabling real-time crescent moon observation tasks under heavy atmospheric distortion. Ultimately, the standout offering of this scientific study is the careful transformation of the compact MobileNetV2 structure into the specialized MobileNetV2-191H, achieving highly resource-efficient crescent moon semantic segmentation under taxing astronomical viewing conditions.

2. Method

This study develops an advanced deep learning framework designed to detect the crescent moon by utilizing an optimized Convolutional Neural Network (CNN) model that automatically isolates and emphasizes the moon within various images. The entire development process was executed in a structured, sequential manner, beginning with the collection of a comprehensive image dataset and concluding with the creation of a practical, live-action application that identifies the crescent moon instantly. To provide a clear overview of how this system was structured. The foundational workflow initiates by gathering real-time crescent moon surveillance videos from the BMKG monitoring platform, which are subsequently processed through video enhancement techniques like length modification and individual frame extraction to generate the primary image datasets. After obtaining these frames, ground

truth annotation is conducted by outlining specific Regions of Interest (ROI) on the visuals that display the crescent moon, thereby producing the essential training data for the system. Following this, further image preparation is carried out by adjusting the dimensions of the pictures to match the exact input criteria of the CNN model. The core system itself is built upon an altered MobileNetV2 structure, uniquely designated as MobileNetV2-191H, which serves to enhance the precision of crescent moon image segmentation. Ultimately, this newly developed model undergoes rigorous training and assessment using key performance indicators such as accuracy, precision, recall, and F1-score, before being fully integrated into a live application meant for detecting the crescent moon during active space observations. Fig. 1 illustrates the precise methodology and chronological plan implemented to conceptualize and construct this operational crescent moon tracking system.

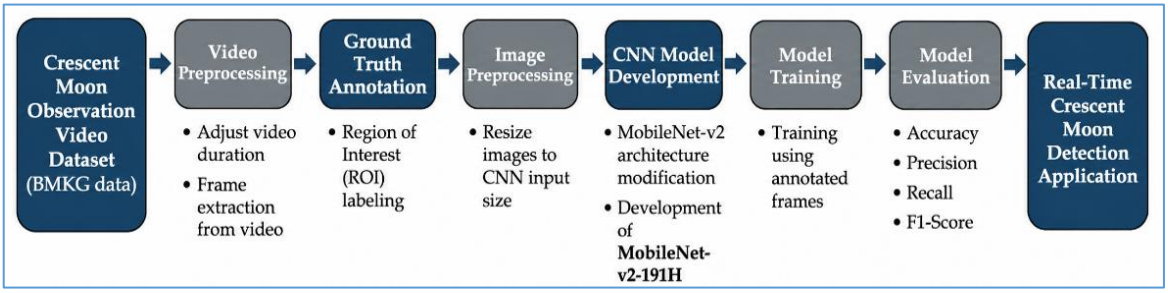


Fig. 1. Research Framework

The model development process, as presented in Fig. 1, was conducted through a series of systematic modifications to the Convolutional Neural Network (CNN) architecture to enhance the detection and segmentation of crescent moon objects in astronomical images. Fig. 2 illustrates the successive stages of architectural development, beginning with MobileNetV2 as the baseline model and progressing toward the proposed MobileNetV2-191H architecture, specifically optimized for crescent moon detection and segmentation. As illustrated in Fig. 2, the initial stage involved adopting the original MobileNetV2 architecture as the baseline for subsequent architectural modifications and performance improvements.

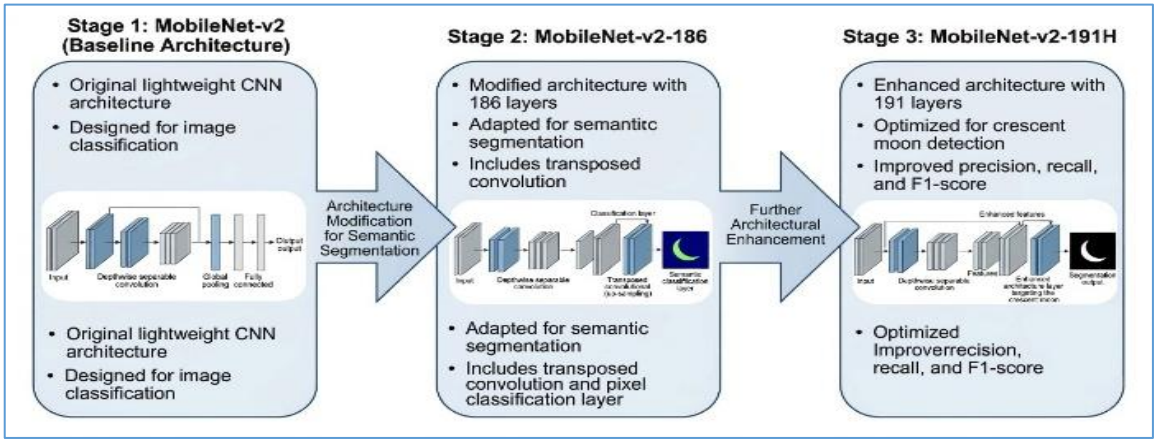


Fig. 2. Architectural Development of CNN for Crescent Moon Detection

This lightweight CNN was originally designed for image classification tasks. In the second stage, the architecture was modified into MobileNetV2-186 by adding several additional layers and adjusting the network structure to support semantic segmentation. These modifications included the incorporation of transposed convolution layers and a pixel classification layer, enabling the model to generate segmentation maps of crescent moon objects in the images. The MobileNetV2-186 architecture was previously tested by the authors and achieved an accuracy of 99.68% in crescent moon detection [43]. In the third stage, further development was carried out to create MobileNetV2-191H by refining the architecture and adding several layers to enhance the feature extraction capabilities for crescent moon objects, which are characterized by their extremely thin shape and low contrast. This architectural enhancement aims to improve the model's performance in crescent moon detection, as reflected by increased precision, recall, and F1-score compared to the previous model. The proposed model

architecture in this study was developed from MobileNetV2 by modifying the network structure to enhance the segmentation capability for crescent moon objects in astronomical observation images. Fig. 3 illustrates the MobileNetV2-191H architecture used in this study as the CNN model for crescent moon detection and segmentation. Fig. 3 illustrates the detailed flow of feature extraction, spatial downsampling, feature refinement, transposed convolution upsampling, and pixel-level classification within the proposed MobileNetV2-191H framework for crescent moon semantic segmentation. The MobileNetV2-191H architecture processes input images of size 240 × 424 × 3 through initial convolution and batch normalization layers, followed by depthwise separable convolutions using 3 × 3 and 1 × 1 kernels to efficiently extract features with increasing channel depth. The original MobileNetV2 is then enhanced by adding convolution, max pooling, and transposed convolution layers to improve segmentation performance, where transposed convolution enables upsampling for reconstructing feature maps. In the final stage, a softmax and pixel classification layer classify each pixel as crescent moon or background. With a total of 191 layers, this modified architecture improves the detection of thin, low-contrast crescent moon objects, leading to better accuracy, precision, recall, and F1-score. The notation 191H refers to the total number of layers incorporated in the enhanced architecture, while the letter H represents the hybrid enhancement strategy that combines lightweight feature extraction with segmentation-oriented reconstruction modules for crescent moon segmentation. The architectural refinement was designed to balance segmentation performance and computational efficiency, where the additional layers were gradually introduced to improve spatial reconstruction capability without significantly increasing model complexity.

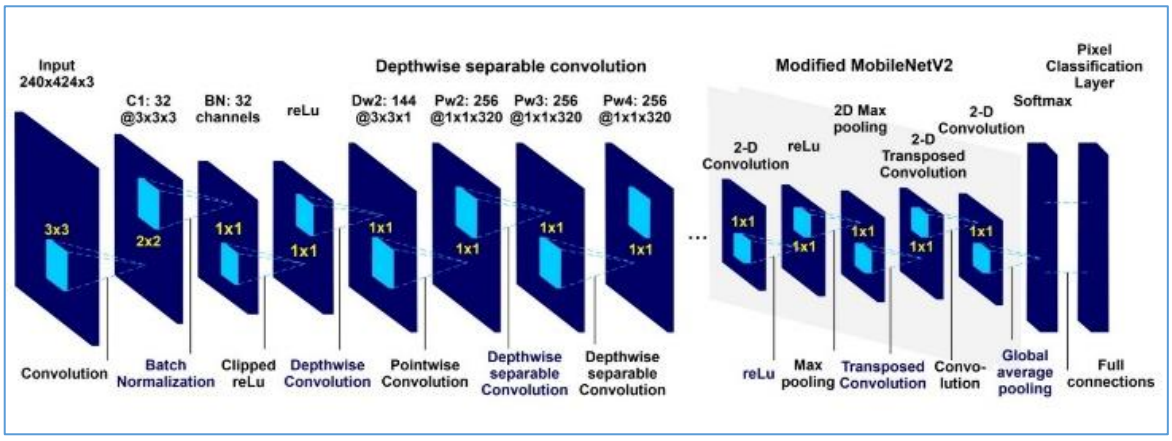


Fig. 3. MobileNetV2-191H Architectural Model

The proposed MobileNetV2-191H model is not intended as a completely new backbone architecture, but rather as an enhanced adaptation of MobileNetV2 for lightweight semantic segmentation of crescent moon objects. The structural enhancement mainly focuses on improving the capability of MobileNetV2 to preserve spatial information and reconstruct thin object boundaries during segmentation. Several modifications were introduced compared to the standard MobileNetV2 architecture. First, additional transposed convolution layers were incorporated to perform progressive upsampling and restore spatial resolution lost during feature extraction. Second, feature refinement layers were added to strengthen the representation of extremely thin crescent moon structures with low contrast against atmospheric backgrounds. Third, a pixel-level classification layer was integrated to transform the original classification-oriented MobileNetV2 into a semantic segmentation framework. Unlike standard MobileNetV2-based segmentation approaches such as MobileNetV2 combined with DeepLabV3, the proposed framework emphasizes a lightweight decoder configuration with lower computational complexity while maintaining high segmentation accuracy for crescent moon detection. The proposed enhancement is specifically optimized for astronomical observation images that contain noise, illumination variation, cloud interference, and extremely small target objects. Table 1 presents a comparison of architectural components, operation types, and the function of each layer between MobileNetV2 and MobileNetV2-191H used in this research.

Table 1. Comparison of the MobileNetV2 and MobileNetV2-191H Architectures

MobileNetV2		MobileNetV2-191H	
Convolution	Input 224×224×3	Convolution	Input 240×424×3
Batch Normalization	C1:32 @3×3×3	Batch Normalization	C1:32 @3×3×3
Clipped ReLU	BN:32 channels	Clipped ReLU	BN:32 channels
Depthwise Convolution	ReLU	Depthwise Convolution	ReLU
Pointwise Convolution	Dw2:144 @3×3×3 Pw2:256 @1×1×320	Pointwise Convolution	Dw2:144 @3×3×3 Pw2:256 @1×1×320
Depthwise Separable Convolution	Pw3:256 @1×1×320	Depthwise Separable Convolution	Pw3:256 @1×1×320
Depthwise Separable Convolution	Pw4:256 @1×1×320	Depthwise Separable Convolution	Pw4:256 @1×1×320
		2-D Convolution	Feature refinement layer
		Max Pooling	Spatial down-sampling
		Transposed Convolution	Feature up-sampling
		2-D Convolution	Feature reconstruction
Global Average Pooling	Softmax	Global Average Pooling	Softmax
Full Connections	Classification Layer	Full Connections	Pixel Classification Layer

According to the data presented in Table 1, both the MobileNetV2 and MobileNetV2-191H structures have many identical foundational elements in common, such as convolution, batch normalization, ReLU activation, alongside depthwise and pointwise convolution segments that work together to effectively pull out important visual characteristics from images. The main distinction between them is found in the specific adjustments made to the MobileNetV2-191H setup, specifically the inclusion of extra components like 2D convolution, max pooling, transposed convolution, and a dedicated feature reconstruction step, all of which are designed to significantly improve how well the system separates and identifies different parts of an image. Furthermore, MobileNetV2 produces an output in the form of a classification layer for overall image classification, whereas MobileNetV2-191H employs a pixel classification layer that enables the model to perform segmentation at the pixel level. These modifications allow MobileNetV2-191H to achieve better performance in detecting crescent moon objects, which are characterized by thin shapes and low contrast, compared to the standard MobileNetV2 architecture. The training process of the MobileNetV2-191H architecture consists of several main stages, including convolution operations, ReLU activation, max pooling, transposed convolution (upsampling), softmax-based pixel classification, and parameter optimization using the Stochastic Gradient Descent with Momentum (SGDM) algorithm. The following mathematical formulations describe the main operations employed in the proposed MobileNetV2-191H framework for crescent moon semantic segmentation and optimization during the training process. Compared with the original MobileNetV2 architecture consisting of 155 layers, the proposed MobileNetV2-191H architecture incorporates 36 additional layers, resulting in a total of 191 layers. The architectural enhancement includes additional max pooling layers for improved spatial feature abstraction, transposed convolution layers for spatial reconstruction during upsampling, feature refinement layers for preserving thin crescent moon boundaries, and a pixel classification layer for accurate semantic segmentation. These modifications were specifically designed to improve segmentation performance for small, low-contrast crescent moon objects while maintaining the lightweight characteristics of MobileNetV2.

2.1. Transposed Convolution

The transposed convolution layer is used to perform upsampling on the feature map, allowing the spatial resolution to be restored during the semantic segmentation process.

$$Y_{i,j} = \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} X_{i-m,j-n} \cdot W_{m,n} \tag{1}$$

This process allows the network to generate more detailed spatial representations, facilitating the detection of crescent moon objects that exhibit extremely thin shapes in images. The transposed convolution process plays an important role in reconstructing the spatial structure of extremely thin crescent moon regions during pixel-level segmentation.

2.2. Loss Function (Cross-Entropy Loss)

To calculate the prediction error, the Cross-Entropy Loss function is used.

$$L = -\sum_{i=1}^{N}\sum_{c=1}^{K} y_{i,c} \log\left(p_{i,c}\right) \tag{2}$$

where $y(i,c)$ is the ground truth label, p(i,c) is the predicted probability, and N is the total number of pixels in the image. This loss function measures the difference between the predicted probability distribution and the true labels. The optimization process aims to improve segmentation consistency between predicted crescent moon regions and ROI-based ground truth annotations under challenging atmospheric conditions.

2.3. Parameter Optimization (SGDM)

The network parameters are updated using the Stochastic Gradient Descent with Momentum (SGDM) algorithm.

$$v_{t+1} = \mu v_t - \eta \nabla L(\theta_t) \tag{3}$$

$$\theta_{t+1} = \theta_t + v_{t+1} \tag{4}$$

where θ represents the model parameters, η is the learning rate, μ is the momentum, and $\nabla L(\theta)$ is the gradient of the loss. The SGDM algorithm is used to accelerate the convergence of the training process and help the model reach optimal parameters.

2.4. Objective Function Training

In general, the objective of the training process is to find the optimal parameters θ that minimize the loss value on the training dataset.

$$\theta^* = \arg\min_{\theta} \frac{1}{N}\sum_{i=1}^{N} L(y_i, f(x_i;\theta)) \tag{5}$$

where $f(xi;\theta)$ represents the network's prediction function for the input image xi.

2.5. Training Configuration

To train the recommended MobileNetV2-191H framework, the researchers utilized the SGDM optimization method alongside a starting learning rate set at 0.001, a momentum value of 0.9, and a mini-batch capacity of 16. This learning procedure lasted for 50 full epochs, employing input photographs that were adjusted to a resolution of 240 × 424 × 3 pixels. For the pixel-level categorization goal, cross-entropy loss served as the primary optimization target. Throughout the training phase, the system randomly mixed the order of the images during each epoch, which helped boost its overall generalization power and minimized the chances of overfitting. Furthermore, they put early stopping rules into action so that training would automatically halt once the validation loss leveled off, thereby avoiding needless processing cycles. The actual testing and experiments were carried out via MATLAB R2024a on a computer workstation that featured an NVIDIA RTX-series graphics card, an Intel Core i7 central processing unit, and 32 GB of random-access memory. Additionally, simple data enhancement methods, specifically flipping the pictures horizontally and tweaking their brightness levels, were introduced to make sure the model stays highly durable and dependable when facing different atmospheric viewing situations.

3. Results and Discussion

In this research, the identification of the crescent moon within images was achieved by building and adjusting a Convolutional Neural Network (CNN) architecture, utilizing a collection of data gathered from BMKG's own crescent moon observation videos. This specific dataset comprised 27 distinct observation videos that went through a frame extraction phase to ultimately yield 6,283 individual images, which were then split into a training group of 5,241 images and a testing group of 1,042 images.

To completely avoid the issue of data leakage and guarantee a genuinely unbiased assessment, this division of data was intentionally executed based on the original videos instead of individual image frames. Consequently, 21 observation videos were designated solely for creating the training data, whereas the 6 leftover videos were set aside strictly for testing purposes to create a completely autonomous evaluation environment. This meant that every single frame pulled from a specific crescent moon video went entirely into either the training collection or the testing collection, thereby preventing any frames from a single video from being shared between the two groups. Ultimately, the visual images produced by this extraction procedure served as the core material for training and evaluating the crescent moon identification model, with a clear visual sample of this extracted image dataset presented directly in Fig. 4.

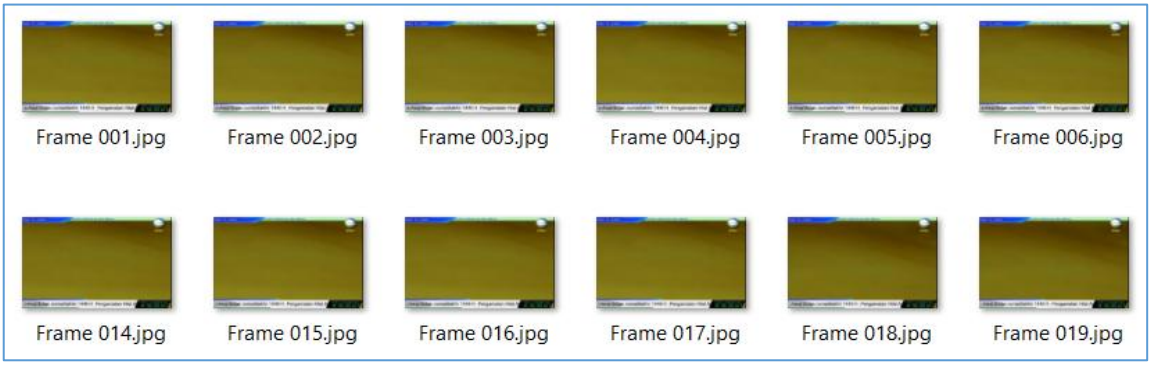


Fig. 4. Result of Frame Extraction from the Crescent Moon Observation Videos

Based on Fig. 4, after extracting frames from crescent moon observation videos, the next stage involves creating Region of Interest (ROI) ground truth by manually labeling the crescent moon areas in each image. This process generates annotated images that serve as reference targets during training, helping the CNN model learn the distinct visual characteristics of crescent moon objects, which are typically thin and low in contrast against the sky background. The resulting ROI ground truth improves the models ability to perform accurate segmentation on observational images. Furthermore, to evaluate the training performance, loss values are analyzed across epochs, where the training loss curve of the MobileNetV2-191H model illustrates how the model learns and converges during the training process. Based on Fig. 5, each image extracted from crescent moon observation videos is manually labeled using the Region of Interest (ROI) method to mark the crescent moon area and generate ground truth images for CNN training.

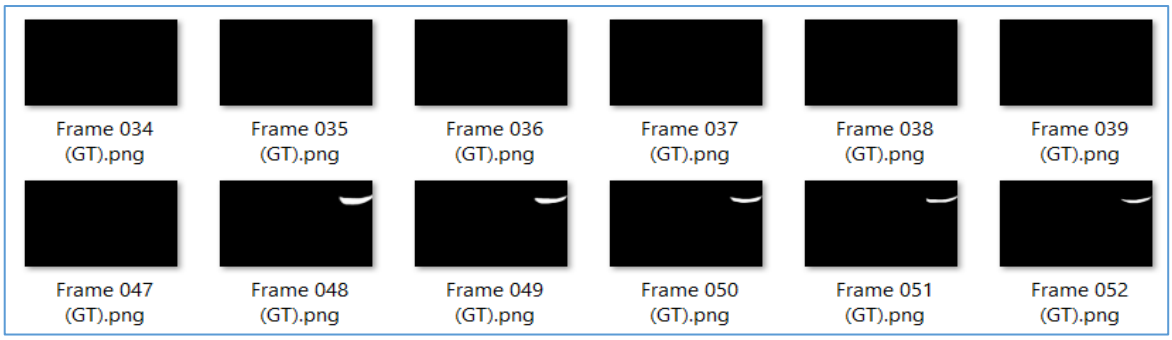


Fig. 5. ROI Ground Truth Results on the Images

This process helps the model learn the unique visual characteristics of the crescent moon, which is thin and has low contrast against the sky background, thereby improving segmentation accuracy. Furthermore, to evaluate the training performance, loss analysis is conducted, where Fig. 6 illustrates the training loss curve of the MobileNetV2-191H model, showing the changes in loss values at each epoch throughout the training process.

Based on Fig. 6, the training loss of the MobileNetV2-191H model decreased sharply from around 0.68 in the first epoch to approximately 0.01 within the early epochs, indicating rapid learning of crescent

moon feature patterns. After the 10th epoch, the loss stabilized at a very low level until the 50th epoch, showing that the model had converged and learned consistently without significant fluctuations. Similarly, training accuracy increased significantly from about 58% in the first epoch to nearly 100% within a few epochs, then remained stable throughout the training process. The model achieved a final accuracy of 99.7138%, demonstrating its strong capability in learning and accurately classifying crescent moon object features, with the stable accuracy curve further confirming successful convergence. To evaluate the models ability to detect crescent moon objects in observational images, testing was conducted using frames extracted from the observation videos. Fig. 7 shows the results of testing the MobileNetV2-191H model on these image frames to assess the detection of crescent moon objects.

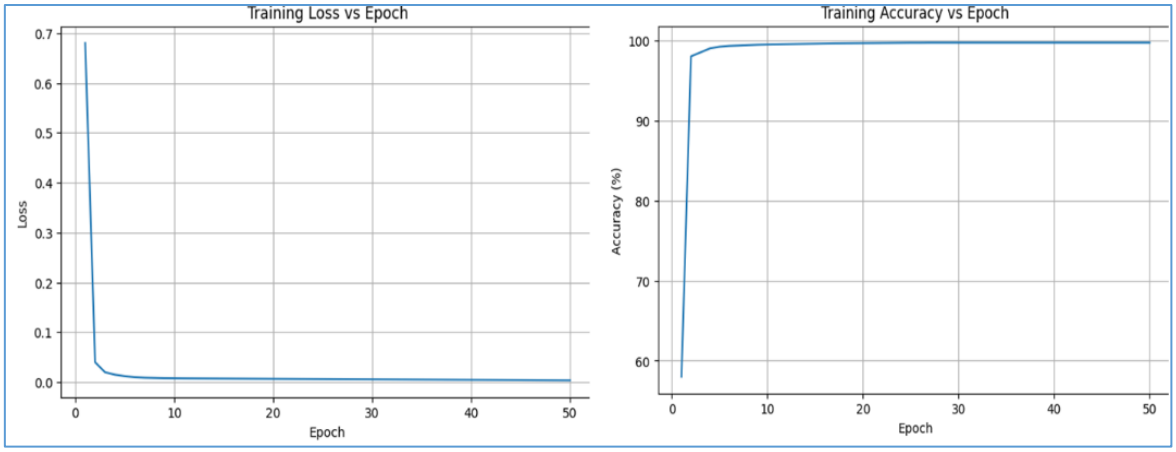


Fig. 6. Training Accuracy and Loss Curve of the Proposed MobileNetV2-191H Model



Fig. 7. Crescent Moon Segmentation Results under Various Atmospheric Observation Conditions, Including Low Contrast, Cloud Interference, and Noise

As shown in Fig. 7, the evaluation procedure involves a detailed comparison across various phases of detection, which encompass the initial unaltered photograph, the identification outcomes highlighted by red bounding boxes, the segmented binary graphics, and the actual ground truth benchmarks. The striking resemblance between the final segmented data and the ground truth references serves as clear evidence that the specialized MobileNetV2-191H framework possesses a high level of precision when identifying and isolating crescent moon shapes. After successfully completing the initial training phase and subsequent performance reviews, this developed framework was deployed into a functional, live detection application built entirely within the MATLAB programming environment. To thoroughly check how durable and dependable this suggested segmentation setup remains when facing practical, everyday stargazing and astronomical conditions, it was put to the test against a variety of difficult

situations, such as overcast and cloudy skies, interference from atmospheric noise, poor lighting setups, and situations where the crescent moon itself showed very low contrast against its background.

The user-facing software interface is designed to present both the original incoming images and the final processed results simultaneously, drawing clear attention to any detected crescent moon targets using helpful bounding rectangles while offering a collection of practical tools like immediate real-time video stream processing, historical offline data examination, and a structured system for logging daily observations. Furthermore, the software interface keeps the operator informed by constantly broadcasting the current tracking status, displaying clear indicators like "Visible" to verify when a crescent moon has been successfully found, which ultimately proves that the newly introduced framework works exceptionally well not just in controlled training scenarios or standard test setups, but also when put to work in actual, live crescent moon tracking and monitoring deployments. Fig. 8 presents the real-time testing process for crescent moon detection using the proposed MobileNetV2-191H framework. Although the proposed MobileNetV2-191H model achieved superior segmentation performance, several false positive and false negative cases were still observed during testing. False positive detections generally occurred when bright atmospheric artifacts, cloud boundaries, or image noise exhibited visual patterns similar to thin crescent moon structures. Conversely, false negative cases were primarily observed under extremely low-contrast observation conditions where the crescent moon boundary became difficult to distinguish from the surrounding sky background. These observations indicate that atmospheric interference remains one of the main challenges in crescent moon segmentation and may affect detection performance under unfavorable observation conditions.



Fig. 8. Real-Time Testing for Crescent Moon Detection

Based on Fig. 8, it can be observed that the proposed MobileNetV2-191H model successfully performs real-time crescent moon detection despite the small object size and low contrast against the sky background. To measure how much better the newly suggested model performs, the original MobileNetV2 structure was directly compared against the updated MobileNetV2-191H version. In Table 2, the detailed breakdown of the accuracy, precision, recall, and F1-score results for both MobileNetV2 and MobileNetV2-191H is provided, which are entirely determined by the specific evaluation datasets analyzed throughout this research project.

Table 2. Performance Comparison between MobileNetV2 and MobileNetV2-191H

Indicator	MobileNetV2	MobileNetV2-191H	Description
Accuracy	97.8421%	99.7138%	Improvement of 1.8717%
Precision	80.2145%	84.805%	Improvement of 4.5905%
Recall	81.0032%	85.3231%	Improvement of 4.3199%
F1-Score	79.1456%	84.0673%	Improvement of 4.9217%

The MobileNetV2-191H model demonstrates a significant performance improvement compared to the MobileNetV2 model (Table 2). For the accuracy indicator, MobileNetV2 achieved 97.8421%, whereas MobileNetV2-191H reached 99.7138%, showing an improvement of 1.8717%. For precision, the value increased from 80.2145% to 84.805%, an improvement of 4.5905%. Similarly, recall improved from 81.0032% to 85.3231%, an increase of 4.3199%. The largest improvement is observed in the F1-score, which rose from 79.1456% in MobileNetV2 to 84.0673% in MobileNetV2-191H, representing

an increase of 4.9217%. The performance improvement achieved by MobileNetV2-191H is mainly influenced by the integration of additional segmentation-oriented layers that improve spatial reconstruction capability and feature refinement for extremely thin crescent moon objects. The incorporation of transposed convolution layers enables better recovery of spatial information lost during feature extraction, while the additional refinement layers help preserve low-contrast crescent boundaries under atmospheric interference conditions. These architectural enhancements contribute to improved pixel-level segmentation consistency compared to the standard MobileNetV2 architecture. Based on these findings, it is clear that the structural changes made, which involved putting in a few extra layers to create the MobileNet-V2-191H version, successfully improve the model's overall capability to recognize and isolate crescent moon shapes with much higher precision than the regular MobileNetV2 setup. Furthermore, Fig. 9 displays a direct visual contrast of how both the training loss and accuracy levels progress for both MobileNetV2 and MobileNetV2-191H throughout every single epoch of the entire training phase.

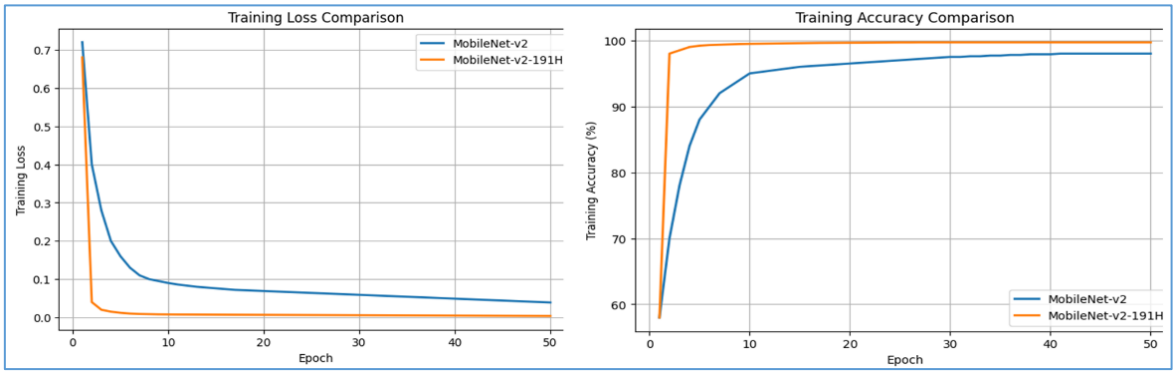


Fig. 9. Training Loss and Accuracy Comparison between MobileNetV2 and MobileNetV2-191H

Based on Fig. 9, both models show decreasing training loss as epochs increase, but the MobileNetV2-191H model achieves a faster and more significant reduction, reaching around 0.01 compared to approximately 0.04 in MobileNetV2. This indicates a more effective feature learning capability and improved convergence due to the architectural enhancements. Similarly, training accuracy in both models increases over time, but MobileNetV2-191H improves more rapidly, achieving about 99.71% within early epochs and maintaining stable performance, while MobileNetV2 gradually reaches around 98% by the end of training. These results demonstrate that the additional layers in MobileNetV2-191H enable more efficient learning, faster convergence, and more optimal overall performance in detecting crescent moon objects. To assess how well the newly designed architecture performs, a detailed comparison was carried out using three distinct models: the standard MobileNetV2 serving as the baseline, MobileNetV2-186 representing the initial modification stage, and MobileNetV2-191H acting as the final recommended model. In Table 3, you can find the comprehensive evaluation of these three specific models, which highlights their exact values for accuracy, precision, recall, and F1-score as determined by the final testing dataset outcomes. Table 3 shows that every phase of the architectural modification led to a noticeable boost in the system's capability to detect the crescent moon. Initially, the standard MobileNetV2 framework reached an accuracy level of 97.8421%, alongside a precision of 80.2145%, a recall of 81.0032%, and an F1-score of 79.1456%. Once the structure was upgraded to the MobileNetV2-186 version, the overall efficiency of the model climbed higher, successfully hitting an accuracy of 99.6835%, a precision of 82.1152%, a recall of 82.2122%, and an F1-score of 80.5725%.

Table 3. Performance Comparison among MobileNetV2, MobileNetV2-186, and MobileNetV2-191H

Indicator	MobileNetV2	MobileNetV2-186	MobileNetV2-191H
Accuracy	97.8421%	99.6835%	99.7138%
Precision	80.2145%	82.1152%	84.805%
Recall	81.0032%	82.2122%	85.3231%
F1-Score	79.1456%	80.5725%	84.0673%

Subsequent architectural enhancements led to the creation of the MobileNetV2-191H variant, which delivered the most outstanding outcomes by securing an accuracy of 99.7138%, a precision of 84.805%, a recall of 85.3231%, and an F1-score of 84.0673%. These findings clearly demonstrate that progressively adding more layers to the original MobileNetV2 setup successfully maximized the network's capacity to pull out key features and identify crescent moon structures characterized by their slender outlines and weak contrast, thereby generating a far more precise identification output than the older versions. Additionally, Fig. 10 illustrates a visual comparison of the training loss and accuracy progressions for the MobileNetV2, MobileNetV2-186, and MobileNetV2-191H setups throughout each training epoch.

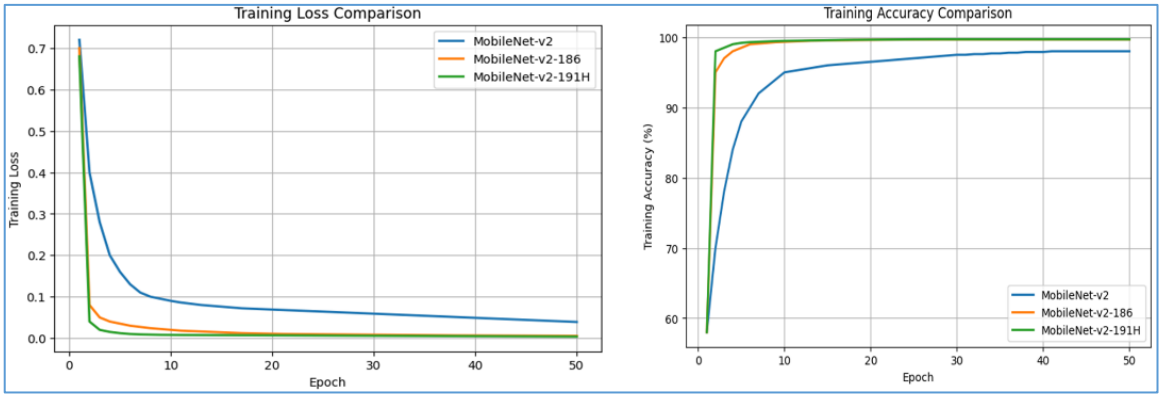


Fig. 10. Training Loss and Accuracy Comparison between MobileNetV2, MobileNetV2-186, and MobileNetV2-191H

All models show decreasing training loss as epochs increase (Fig. 10), but the proposed MobileNetV2-191H achieves the fastest reduction and lowest final loss (≈0.005), outperforming both the base MobileNetV2 (≈0.04) and MobileNetV2-186 (≈0.01). This indicates that the architectural enhancements in MobileNetV2-191H improve feature learning and accelerate convergence. Similarly, as shown in Fig. 10, all models exhibit increasing training accuracy, with MobileNetV2 gradually reaching around 98%, MobileNetV2-186 achieving approximately 99.68%, and MobileNetV2-191H delivering the best performance with faster, more stable convergence and a final accuracy of about 99.71%. These results confirm that the refined architecture of MobileNetV2-191H enhances the extraction of crescent moon features and provides superior overall training performance compared to previous models. To assess how well the newly introduced model works in contrast to alternative segmentation techniques, an analysis of its performance was carried out by testing various deep learning architectures on an identical dataset. Table 4 shows a detailed breakdown that contrasts the performance of different segmentation frameworks, specifically U-Net, DeepLabv3, SegNet, EfficientNet, and MobileNetV2-191H, by analyzing crucial evaluation metrics such as overall accuracy, precision rates, recall scores, and F1-score values.

Table 4. Segmentation Performance Comparison of Deep Learning Models

Model	Accuracy (%)	Precision (%)	Recall (%)	F1-Score (%)	IoU (%)	Dice (%)
U-Net	98.30	78.12	79.45	78.78	65.11	78.85
DeepLabv3	98.74	80.36	81.12	80.74	67.52	80.60
SegNet	98.55	79.21	80.08	79.64	66.38	79.78
EfficientNet	99.12	82.45	83.01	82.73	70.15	82.52
MobileNetV2-191H	99.71	84.81	85.32	84.07	72.64	84.13

The MobileNetV2-191H model outperforms all other segmentation models evaluated on the identical dataset (Table 4). Looking at the specific metrics, the U-Net model registered an accuracy of 98.30%, alongside a precision of 78.12%, a recall of 79.45%, and an F1-score of 78.78%. A higher level of effectiveness was observed in the DeepLabv3 model, which reached an accuracy of 98.74%, precision of 80.36%, recall of 81.12%, and an F1-score of 80.74%. Meanwhile, the SegNet model managed to achieve an accuracy of 98.55%, a precision score of 79.21%, a recall rate of 80.08%, and an F1-score of 79.64%. The EfficientNet model displayed an even stronger capability, securing an accuracy of 99.12%,

precision of 82.45%, recall of 83.01%, and an F1-score of 82.73%. Ultimately, the newly proposed MobileNetV2-191H model surpassed them all by delivering the highest metrics, which include an outstanding 99.71% accuracy, 84.81% precision, 85.32% recall, and an F1-score of 84.07%. These outcomes clearly show that modifying and adding specific layers to the standard MobileNetV2 structure successfully boosts its capacity to capture spatial characteristics of crescent moon shapes, thereby producing far more precise segmentation outcomes than the competing models on this dataset. Furthermore, to get a clearer picture of the segmentation quality during severe class imbalance, where the tiny crescent moon takes up very little space compared to the massive sky backdrop, the Intersection over Union (IoU) was also carefully measured

The newly introduced MobileNetV2-191H successfully secured the top IoU score out of all the tested models, proving its superior ability to protect and accurately isolate the delicate structures of the thin crescent moon during the pixel-by-pixel segmentation process. Even though the system achieved a massive overall accuracy of 99.71%, this figure needs to be analyzed with caution because the crescent moon targets cover only a minor fraction of the picture while the sky background takes up the majority of the pixels, and standard accuracy can be misleading. Because of this heavily unbalanced pixel layout, specialized evaluation metrics like the IoU, Dice coefficient, precision, recall, and F1-score offer a much fairer and more reliable assessment of how well the crescent moon is actually being segmented. It is important to note that these published outcomes were derived strictly from the specific testing setup established throughout this research project. To build on these findings, upcoming research projects could focus on conducting multiple experimental runs as well as incorporating deep statistical validation, such as calculating standard deviations or confidence intervals, to thoroughly verify how reliable and steady the model remains across different scenarios. Lastly, Table 5 provides a detailed breakdown of the computational performance comparison among the various models that were tested.

Table 5. Computational Performance Comparison of Segmentation Models

Model	Parameters (M)	Inference Time (ms)	FPS
U-Net	31.0	85	11.7
DeepLabv3	39.6	112	8.9
SegNet	29.4	96	10.4
EfficientNet	20.1	74	13.5
MobileNetV2-191H	8.7	41	24.3

The computational performance evaluation indicates that the proposed MobileNetV2-191H maintains lower computational complexity compared to several baseline segmentation models (Table 5). The lightweight architecture enables faster inference speed and higher FPS performance while maintaining competitive segmentation accuracy, demonstrating its suitability for computationally efficient crescent moon observation applications. Despite achieving strong segmentation performance, several limitations were observed during testing. Some false negative cases occurred when the crescent moon appeared extremely faint due to cloud interference, atmospheric turbulence, or low illumination conditions, causing the model to miss very thin crescent structures. In addition, several false positive detections were observed in frames containing bright atmospheric noise or light reflections near the horizon that visually resembled crescent moon patterns. These challenging conditions explain why the precision and recall values remained around 84–85%, indicating that distinguishing extremely thin crescent moon objects from complex sky backgrounds remains a difficult semantic segmentation problem. Future work may focus on integrating attention mechanisms, adaptive contrast enhancement, or temporal information from sequential video frames to further improve segmentation robustness under challenging atmospheric conditions. These limitations indicate the presence of several false positive and false negative segmentation cases under highly challenging atmospheric observation conditions.

4. Conclusion

In this research, scientists successfully created an updated version of the MobileNetV2 architecture, named MobileNetV2-191H, to serve as a specialized segmentation model that can more precisely identify and locate crescent moon shapes within astronomical photos. They achieved this by taking the standard

MobileNetV2 design and inserting extra layers, which allowed the system to switch its focus from simply identifying whole images to understanding specific shapes within those images, known as semantic segmentation. When tested, this newly suggested framework performed much better than the original MobileNetV2, the older MobileNetV2-186 version, and several other well-known segmentation tools like U-Net, DeepLabv3, SegNet, and EfficientNet, reaching an impressive accuracy of 99.71%, a precision rate of 84.81%, a recall rate of 85.32%, and an F1-score of 84.07%. Additionally, the charts tracking how well the system learned showed that MobileNetV2-191H managed to learn much quicker and showed a more consistent, stable improvement during its training phase. These clear results prove that adding and fine-tuning these specific layers within the MobileNetV2 setup greatly boosts its power to pick up on the unique details of crescent moons, which are notoriously difficult to spot because they are very thin and easily blend into the dark sky behind them. It is important to note that this project does not claim to invent a completely new base for deep learning from scratch; rather, its main value comes from adjusting and improving the lightweight MobileNetV2 structure so it can perform semantic segmentation on crescent moons during tough, real-world nighttime viewing situations. By adding these new, segmentation-focused layers, the system does a much better job of keeping track of thin, delicate shapes without using up too much computer power, making it perfectly ready for real-time applications. Looking ahead, upcoming projects in this field might explore creating even more flexible models by using next-generation deep learning methods and connecting them directly to live telescope systems to make finding the crescent moon even more precise and dependable.

Acknowledgment

The authors express gratitude to Universitas Putra Indonesia YPTK for the support that enabled the successful completion of this article.

Declarations

Author contribution. Conceptualization and methodology, M.H.Y. and Y.; dataset collection and annotation, M.H.Y. and S.; software implementation and experimentation, M.H.Y.; formal analysis and validation, Y. and S.; visualization and data interpretation, M.H.Y. and S.; manuscript preparation, M.H.Y.; manuscript review and editing, Y. and S. All authors have read and approved the final version of the manuscript.

Funding statement. This research was independently funded.

Conflict of interest. The authors declare no conflict of interest.

Additional information. No additional information is available for this paper.

References

[1] Z. T. Allawi, "A pattern-recognizer artificial neural network for the prediction of new crescent visibility in Iraq," *Computation*, vol. 10, no. 10, p. 186, 2022, doi: 10.3390/computation10100186.

[2] M. S. A. M. Nawawi, M. S. Faid, M. H. M. Saadon, R. A. Wahab, and N. Ahmad, "Hijri month determination in southeast asia: An illustration between religion, science, and cultural background," *Heliyon*, vol. 10, no. 20, 2024, doi: 10.1016/j.heliyon.2024.e38668.

[3] J. A. Utama *et al.*, "Young lunar crescent detection based on video data with computer vision techniques," *Astron. Comput.*, vol. 44, p. 100731, 2023, doi: 10.1016/j.ascom.2023.100731.

[4] M. Al-Rajab, S. Loucif, and Y. Al Risheh, "Predicting new crescent moon visibility applying machine learning algorithms," *Sci. Rep.*, vol. 13, no. 1, p. 6674, 2023, doi: 10.1038/s41598-023-32807-x.

[5] S. Loucif, M. Al-Rajab, R. Abu Zitar, and M. Rezk, "Toward a globally lunar calendar: a machine learning-driven approach for crescent moon visibility prediction," *J. Big Data*, vol. 11, no. 1, p. 114, 2024, doi: 10.1186/s40537-024-00979-6.

[6] S. Sen, S. Agarwal, P. Chakraborty, and K. P. Singh, "Astronomical big data processing using machine learning: A comprehensive review," *Exp. Astron.*, vol. 53, no. 1, pp. 1–43, 2022, doi: 10.1007/s10686-021-09827-4.

[7] J.-V. Rodríguez, I. Rodríguez-Rodríguez, and W. L. Woo, "On the application of machine learning in astronomy and astrophysics: A text-mining-based scientometric analysis," *WIREs Data Min. Knowl. Discov.*, vol. 12, no. 5, p. e1476, Sep. 2022, doi: 10.1002/widm.1476.

[8] S. Madkar, D. Y. Sakhare, K. A. Phutane, A. P. Haral, K. B. Nikam, and S. Tharunyha, "Video Based Forest Fire and Smoke Detection Using YoLo and CNN," *3rd Int. Conf. Power, Energy, Control Transm. Syst. ICPECTS 2022 - Proc.*, pp. 1–5, 2022, doi: 10.1109/ICPECTS56089.2022.10046717.

[9] S. Frizzi, R. Kaabi, M. Bouchouicha, J. M. Ginoux, E. Moreau, and F. Fnaiech, "Convolutional neural network for video fire and smoke detection," *IECON Proc. (Industrial Electron. Conf.*, pp. 877–882, 2016, doi: 10.1109/IECON.2016.7793196.

[10] U. Mittal and P. Chawla, "Acoustic Based Emergency Vehicle Detection Using Ensemble of deep Learning Models," *Procedia Comput. Sci.*, vol. 218, pp. 227–234, 2022, doi: 10.1016/j.procs.2023.01.005.

[11] V. Bacu, C. Nandra, A. Sabou, T. Stefanut, and D. Gorgan, "Assessment of Asteroid Classification Using Deep Convolutional Neural Networks," *Aerospace*, vol. 10, no. 9. p. 752, 2023, doi: 10.3390/aerospace10090752.

[12] A. Febriani, R. Wahyuni, Y. Irawan, and R. Melyanti, "Improved Hybrid Machine and Deep Learning Model for Optimization of Smart Egg Incubator," *J. Appl. Data Sci.*, vol. 5, no. 3, pp. 1052–1068, 2024, doi: 10.47738/jads.v5i3.304.

[13] X. Liu, K. H. Ghazali, F. Han, and I. I. Mohamed, "Review of CNN in aerial image processing," *Imaging Sci. J.*, vol. 71, no. 1, pp. 1–13, Jan. 2023, doi: 10.1080/13682199.2023.2174651.

[14] D. Setiawan, R. N. Putri, S. Herlina, A. N. Hidayanto, Y. Irawan, and N. Hohashi, "A Hybrid YOLO–CNN Model for Automatic Detection and Severity Assessment of Atopic Dermatitis in Infant Images," *J. Appl. Data Sci.*, vol. 7, no. 2, pp. 1079–1091, May 2026, doi: 10.47738/jads.v7i2.1212.

[15] M. M. Taye, "Theoretical Understanding of Convolutional Neural Network: Concepts, Architectures, Applications, Future Directions," *Computation*, vol. 11, no. 3. p. 52, 2023, doi: 10.47738/jads.v6i4.977.

[16] D. Setiawan, R. N. Putri, I. Fitri, A. N. Hidayanto, Y. Irawan, and N. Hohashi, "Improved Deep Learning Model for Prediction of Dermatitis in Infants," *J. Appl. Data Sci.*, vol. 6, no. 2, pp. 871–884, 2025, doi: 10.47738/jads.v6i2.542.

[17] S. C. Nagavarapu, A. Abraham, S. Li, and J. Dauwels, "Enhancing autonomous vehicle navigation through computer vision: techniques for lane marker detection and rain removal," in *Internet of Vehicles and Computer Vision Solutions for Smart City Transformations*, Springer, 2025, pp. 167–191, doi: 10.1007/978-3-031-72959-1_8.

[18] C. Zhao *et al.*, "A review of optical image enhancement for extreme space environments," *Adv. Astronaut.*, vol. 8, no. 2, pp. 171–199, 2025, doi: 10.1007/s42423-025-00175-5.

[19] Z. Azouz, B. Honarvar Shakibaei Asli, and M. Khan, "Evolution of crack analysis in structures using image processing technique: A review," *Electronics*, vol. 12, no. 18, p. 3862, 2023, doi: 10.3390/electronics12183862.

[20] A. L. A. M. Nasir *et al.*, "New crescent moon detection using Circular Hough Transform (CHT)," *Astron. Comput.*, vol. 51, p. 100902, 2025, doi: 10.1016/j.ascom.2024.100902.

[21] Z. Tu, Q. Wang, Z. Jiang, and J. Jiang, "Wavelet Quadtree Contrast Limited Adaptive Histogram Equalisation Tablet Enhancement Method for Defect Detection in Low-Contrast Tablets," *IET Image Process.*, vol. 20, no. 1, p. e70293, 2026, doi: 10.1049/ipr2.70293.

[22] M. S. Z. Ahmad, N. A. A. Aziz, H. S. Lim, A. K. Ghazali, and 'Afif Abdul Latiff, "Impact of Image Enhancement Using Contrast-Limited Adaptive Histogram Equalization (CLAHE), Anisotropic Diffusion, and Histogram Equalization on Spine X-Ray Segmentation with U-Net, Mask R-CNN, and Transfer Learning," *Algorithms*, vol. 18, no. 12, p. 796, 2025, doi: 10.3390/a18120796.

[23] M. Al-Rajab, S. Loucif, R. A. Zitar, and M. G. Abdu-Aguye, "An AI approach to lunar phase detection: enhancing the identification of the new crescent with astronomical data integration," *Front. Artif. Intell.*, vol. 9, p. 1727824, 2026, doi: 10.3389/frai.2026.1727824.

[24] Y. Zhang, Y. Wang, Y. Zhu, L. Yang, L. Ge, and C. Luo, "Visibility prediction based on machine learning algorithms," *Atmosphere (Basel).*, vol. 13, no. 7, p. 1125, 2022, doi: 10.3390/atmos13071125.

[25] J. He, H. Luo, W. Xiao, S. Liu, and Z. Qi, "ATD-DL: a deep learning framework for faint astronomical target detection," *Front. Astron. Sp. Sci.*, vol. 13, p. 1782465, 2026, doi: 10.3389/fspas.2026.1782465.

[26] F. Dumitrescu, B. Ceachi, C.-O. Truică, M. Trăscău, and A. M. Florea, "A novel deep learning-based relabeling architecture for space objects detection from partially annotated astronomical images," *Aerospace*, vol. 9, no. 9, p. 520, 2022, doi: 10.3390/aerospace9090520.

[27] A. Oubaalla, H. El Moubtahij, and N. El Akkad, "Medical image segmentation using deep learning: A survey," in *International Conference on Digital Technologies and Applications*, 2023, pp. 974–983, doi: 10.1007/978-3-031-29860-8_97.

[28] L. Lei, Q. Yang, L. Yang, T. Shen, R. Wang, and C. Fu, "Deep learning implementation of image segmentation in agricultural applications: a comprehensive review: Lei et al.," *Artif. Intell. Rev.*, vol. 57, no. 6, p. 149, 2024, doi: 10.1007/s10462-024-10775-6.

[29] X. Li *et al.*, "Satellite cloud image segmentation based on lightweight convolutional neural network," *PLoS One*, vol. 18, no. 2, p. e0280408, 2023, doi: 10.1371/journal.pone.0280408.

[30] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "Mobilenetv2: Inverted residuals and linear bottlenecks," in *2018 IEEE/CVF conference on computer vision and pattern recognition*, 2018, pp. 4510–4520, doi: 10.1109/CVPR.2018.00474.

[31] R. GR, S. Rohit, C. G Raghavendra, P. R Shetty, and S. Hegde MP, "Fish Classification with Machine Learning: Enhancing Accuracy and Efficiency," 2024.

[32] E. H. Alkhammash, "Multi-classification using YOLOv11 and hybrid YOLO11n-MobileNet models: A fire classes case study," *Fire*, vol. 8, no. 1, p. 17, 2025, doi: 10.3390/fire8010017.

[33] Q. Zhu, H. Zhuang, M. Zhao, S. Xu, and R. Meng, "A study on expression recognition based on improved mobilenetV2 network," *Sci. Rep.*, vol. 14, no. 1, p. 8121, 2024, doi: 10.1038/s41598-024-58736-x.

[34] S. Khalid, R. H. Ali, and H. Bin Khalid, "Fruit and vegetable recognition using mobilenetv2: An image classification approach," *Eng. Proc.*, vol. 87, no. 1, p. 108, 2025, doi: 10.3390/engproc2025087108.

[35] R. Indraswari, R. Rokhana, and W. Herulambang, "Melanoma image classification based on MobileNetV2 network," *Procedia Comput. Sci.*, vol. 197, pp. 198–207, 2022, doi: 10.1016/j.procs.2021.12.132.

[36] Y. Gulzar, "Fruit image classification model based on MobileNetV2 with deep transfer learning technique," *Sustainability*, vol. 15, no. 3, p. 1906, 2023, doi: 10.3390/su15031906.

[37] T. Soni, S. Gupta, S. Bharany, A. U. Rehman, R. M. Ghoniem, and B. M. Taye, "Retinal vessel segmentation using multi scale feature attention with MobileNetV2 encoder," *Sci. Rep.*, vol. 15, no. 1, p. 43369, 2025, doi: 10.1038/s41598-025-28707-x.

[38] T. Öznacar and N. Varol Kayapunar, "Advanced skin cancer prediction with medical image data using MobileNetV2 deep learning and optimized techniques," *Sci. Rep.*, vol. 15, no. 1, p. 28962, 2025, doi: 10.1038/s41598-025-14963-4.

[39] Y. Zou, L. Wu, C. Zuo, L. Chen, B. Zhou, and H. Zhang, "White blood cell classification network using MobileNetv2 with multiscale feature extraction module and attention mechanism," *Biomed. Signal Process. Control*, vol. 99, p. 106820, 2025, doi: 10.1016/j.bspc.2024.106820.

[40] M. A. Saleem, N. Senan, F. Wahid, M. Aamir, A. Samad, and M. Khan, "Comparative analysis of recent architecture of convolutional neural network," *Math. Probl. Eng.*, vol. 2022, no. 1, p. 7313612, 2022, doi: 10.1155/2022/7313612.

[41] A. Bayegizova *et al.*, "Fire detection using deep learning methods," *Int. J. Electr. Comput. Eng.*, vol. 14, no. 1, pp. 547–555, 2024, doi: 10.11591/ijece.v14i1.pp547-555.

[42] F. M. Talaat, "Real-time facial emotion recognition system among children with autism based on deep learning and IoT," *Neural Comput. Appl.*, vol. 35, no. 17, pp. 12717–12728, 2023, doi: 10.1007/s00521-023-08372-9.

[43] M. H. Yuhandri, Yuhandri, and Sumijan, "Modified Convolutional Neural Networks (CNN) with MobileNet-v2-186 Architecture in Crescent Moon Image Detection," in *2025 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*, 2025, pp. 32–37, doi: 10.1109/ICIMCIS68501.2025.11326934.