JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8251

ENHANCING ACCURACY OF WEATHER CLASSIFICATION USING DEEP FEATURES AND SUPPORT VECTOR MACHINE

Raden Sumiharto[1*]; Faisal Dharma Adhinata[2]

Department of Computer Science and Electronics[1]
Universitas Gadjah Mada, Yogyakarta, Indonesia[1]
https://ugm.ac.id/id/[1]
r_sumiharto@ugm.ac.id*

Software Engineering Study Program[2]
Telkom University, Purwokerto, Indonesia[2]
https://purwokerto.telkomuniversity.ac.id/[2]
faisaldharma@telkomuniversity.ac.id

(*) Corresponding Author
(Responsible for the Quality of Paper Content)

Abstract— *Weather is a determinant of farmers' planting calendar. Farmers usually start planting rice in the rainy season because rice requires sufficient water to produce optimal harvests. The weather is almost unpredictable in certain months, so farmers now look at cloud conditions to predict the season. Seasonal predictions based on cloud imagery can be assisted using Artificial Intelligence methods. Previous research used deep learning via transfer learning, but the results were not optimal. This research dataset is sourced from Kaggle and consists of five classes, namely cloudy, foggy, rainy, shine, and sunrise with a total data of 1500 images. This research proposes that a hybrid deep features and machine learning approach be used to increase the accuracy of the results. The MobileNet deep learning method is used at the feature extraction stage, then for classification using the Support Vector Machine (SVM) method. Experimental results with the Radial Basis Function (RBF) kernel on SVM produced an accuracy of 0.9500 for training data. The evaluation results using testing data produced an accuracy of 0.9667. This result also saw an increase of 4.2% in training data compared to previous research. Through these results, MobileNet-SVM is proven to be able to improve classification accuracy when using a small dataset with 1500 images.*

Keywords: *Artificial Intelligence, Deep Features, MobileNet, Support Vector Machine, Weather.*

Intisari— *Cuaca merupakan faktor penentu kalender tanam bagi petani. Petani biasanya mulai menanam padi di musim hujan karena padi membutuhkan air yang cukup untuk menghasilkan panen optimal. Cuaca hampir tidak dapat diprediksi di bulan-bulan tertentu, sehingga petani perlu memperhatikan kondisi awan untuk memprediksi musim tanam. Prediksi musiman berdasarkan citra awan dapat dibantu dengan menggunakan metode Kecerdasan Buatan. Penelitian sebelumnya menggunakan deep learning melalui transfer learning, tetapi hasilnya tidak optimal. Dataset penelitian ini bersumber dari Kaggle terdiri dari lima kelas, yaitu cloudy, foggy, rainy, shine, dan sunrise dengan jumlah total data adalah 1500 citra. Penelitian ini mengusulkan penggunaan pendekatan hybrid deep fitur dan machine learning untuk meningkatkan hasil akurasi. Metode deep learning MobileNet digunakan pada tahap ekstraksi fitur, kemudian untuk klasifikasi menggunakan metode Support Vector Machine (SVM). Hasil eksperimen dengan kernel Radial Basis Function (RBF) pada SVM menghasilkan akurasi 0,9500 pada tahap training. Hasil evaluasi menggunakan data testing menghasilkan akurasi 0,9667. Hasil ini juga menunjukkan peningkatan 4,2% pada data training dibandingkan dengan penelitian sebelumnya. Melalui hasil ini, MobileNet-SVM terbukti dapat meningkatkan akurasi klasifikasi pada penggunaan dataset kecil dengan 1500 citra.*

Kata Kunci: *Cuaca, Deep fitur, Kecerdasan buatan, MobileNet, Support Vector Machine.*

INTRODUCTION

There is an urgent need for highly accurate weather prediction, especially in agriculture. The reason is that unpredictable weather contributes to substantial crop yield losses [1]. Thus, farmers require sound weather information on which to base decisions about when to plant and harvest. Given the unreliability of existing prediction models, farmers need help facing the unprecedented pace of weather shifts. For example, in Indonesia, the peak of rainfall starts in November-December and March-April, while the dry season starts in August until October [2]. However, this month's predictions are uncertain now [3]. Hence, automatic weather decision-making on when and how much to implement plant or harvest and other factors comes with enormous risks to crop cultivation.

Artificial Intelligence (AI) can enhance automatic weather predictions. AI manages large data sets with complex features more effectively than humans [4][5]. Human limitations in complex calculations require AI assistance to solve everyday problems, including weather classification. Previous studies have used deep learning to better understand complex weather patterns [6]. However, the accuracy still needs improvement. The resulting model cannot identify all data classes accurately. This research uses five data classes: cloudy, foggy, rainy, shine, and sunrise. Besides that, using deep learning methods requires many datasets [7]. The dataset used in this research is small, so a combination of appropriate methods is required to increase accuracy. The CNN model used still needs to be optimal for extracting data features [8]. This research aims to increase accuracy in the same case, namely weather classification, using a combination of appropriate feature extraction and classification methods. Therefore, this research proposed a combination of deep features using MobileNet and machine learning using a Support Vector Machine (SVM) to solve this problem. SVM is known for its ability to carry out classification with high accuracy, even with smaller datasets [9]. Thus, combining MobileNet for feature extraction and SVM for classification promises a more optimal solution for improving weather prediction accuracy.

Integrating these two methods is expected to increase accuracy and efficiency in weather data processing. MobileNet is explicitly designed for mobile applications and devices with limited computational capabilities to achieve fast and efficient feature extraction [10][11]. This MobileNet model has relatively few parameters compared to other transfer learning models. Besides that, feature extraction using MobileNet also produces better accuracy when done hybrid with machine learning. The results produced 100% accuracy at the training and validation stages. In turn, SVM with an adjustable kernel can help assess different weather data types, as each has unique and changing features. Using a multiclass SVM makes it easier to differentiate between datasets that are not too large in number [11]. This model is predicted to reduce prediction errors on small datasets [12], which is typical for existing models, and will provide farmers with a more reliable instrument to predict potential unfavourable weather effects.

This approach is expected to significantly help farmers and others in the agriculture and weather sectors. Better prediction models will improve planning and daily work. They will reduce the risk of losing money and make farming more sustainable. Using MobileNet and SVM in this research is expected to change weather prediction technology. It could set a new standard for accurate and reliable weather predictions. So, the contribution to this research is:

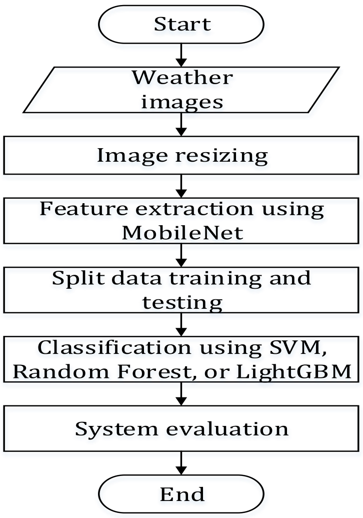
a. Implement a hybrid of MobileNet and SVM to achieve optimal accuracy compared to using all deep learning.
b. Compare the results of using hybrid MobileNet against ensemble machine learning.

In this research, the manuscript's structure is an introduction section that explains the background of the problem and previous research related to weather classification. Then, the methodology section explains the series of methods and evaluation designs used in this research. Then, the results section explores the experimental results and discusses the results obtained. Finally, the final section closes with the conclusion of this research.

MATERIALS AND METHODS

This research uses the MobileNet transfer learning method as the feature extraction stage, and the classifier uses supervised learning. The research stage begins with data acquisition. The raw data is resized to adjust the image size before feature extraction. The pre-processed data was feature extracted using MobileNet and continued with classification. The classification results are five classes, namely cloudy, foggy, rainy, shine, and sunrise. The final stage is to evaluate and achieve the best accuracy. Figure 1 shows the stages of this research.

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8251



Source : (Research Results, 2026)

Figure 1. Proposed Method

Data Acquisition

This research carries out weather classification in agricultural case research, where previous research also used this dataset. This dataset is sourced from the Internet under Creative Commons licenses from platforms like Flickr, Unsplash, and Pexels [13]. The data obtained varies in size, of which there are five classes: cloudy, foggy, rainy, shine, and sunrise. Each image has a category label. Table 1 shows an example of the dataset in this research and its numbers. The total number of datasets is 1500, with two imbalance classes, shine, and sunrise. This dataset also presents testing data of 30 images with labels.

Data Pre-processing

Raw data with various sizes must be resized to prepare for the feature extraction stage. This research uses MobileNet transfer learning, where the input image must be 224 x 224 in size [14]. Therefore, the dataset in this research is resized to adjust the input to the transfer learning. Then, the dataset color remains RGB because the transfer learning input uses three color channels. The resizing process uses nearest interpolation. The color model is RGB, and aspect ratio handling is omitted, so the image is directly resized to 224 x 224. In this research, the program used preprocess_input from MobileNet, so that the pixel values in the image range 0-255 were normalized to -1 to 1. This normalization is necessary to ensure consistent data distribution.

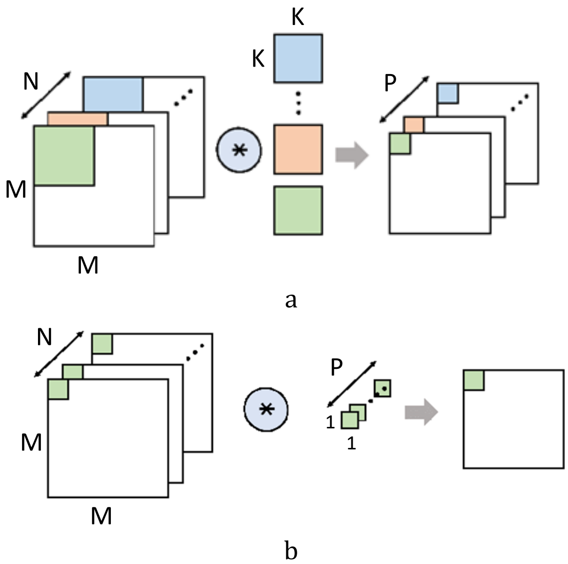
Feature extraction

This research uses MobileNet transfer learning at the feature extraction stage. Two layers are often repeated in the MobileNet architecture: depthwise convolution (conv_dw) and pointwise convolution (conv_pw). In the MobileNet architecture, conv_dw and conv_pw achieve higher computational efficiency. Figure 2 shows the depthwise and pointwise convolution process [15].

Table 1. Sample Weather Classification Dataset

Sample of dataset	Class	Amount of data
	Cloudy	300
	Foggy	300
	Rainy	300
	Shine	250
	Sunrise	350

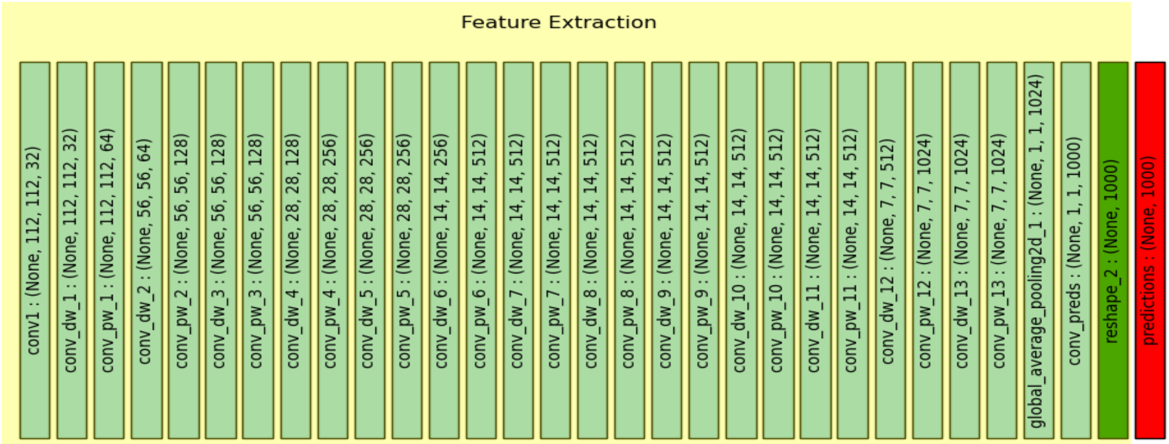
Source : (Gupta, 2020 [13])



Source : (Tsaif, 2023 [15])

Figure 2. Convolution Filters, a) Depthwise Convolution, b) Pointwise Convolution

Conv_dw involves using a separate filter for each input channel. If a layer has 32 channels, 32



Source : (Research Results, 2026)

Figure 3. Feature Extraction on MobileNet Architecture

filters will be applied to each channel. This approach aims to reduce the number of operations and parameters because each filter only needs to learn from the data on one channel. This step's result is an output with the same number of channels as the input, where each channel has been filtered separately. Next, conv_pw is used to combine the output of depthwise convolution. This convolution uses a 1x1 filter to combine information from all available channels, allowing the network to create more prosperous and complex feature representations. Using MobileNet makes the training stage faster [16]. Figure 3 shows the MobileNet architecture used in this research. In this research, the MobileNet layers used in the feature extraction stage are from conv2D (conv1) to reshape. The number of vectors produced is 1000 vectors per image, which are deep features. The prediction layer is not used because it will be replaced using supervised learning.

Classification

In this research, the prediction layer in MobileNet is not used. We propose using supervised learning for weather classification. We also propose traditional machine learning methods, namely SVM and ensemble learning, namely Random Forest and LightGBM.

Super Vector Machine

SVM is a powerful and efficient algorithm primarily developed for binary classification [17], which implies distinguishing and separating between two classes. However, it is possible to apply SVM for multiclass classification [18]. For instance, the algorithm can categorize meteorological conditions into five classes: cloudy, foggy, rainy, shine, or sunrise. A multiclass problem can be solved by two basic approaches instead of abandoning SVM: one-vs-one and one-vs-all. This research applied the one-vs-all technique. The one-vs-all, or one-vs-rest method, implies training a single instance of SVM per class, where a model separates one class from the others. For example, one model will distinguish cloudy from foggy, rainy, shine, and sunrise classes. Then, each data sample is tested against all these models, and the model that provides the highest decision value (distance from the margin) determines the class of that sample.

Random Forest

Random Forest is a very effective and versatile machine learning algorithm for classification and regression [19]. In the context of multiclass classification, such as differentiating between weather classes such as cloudy, foggy, rainy, shine, and sunrise, Random Forest combines many decision trees built at training time. Each tree in Random Forest provides class predictions based on the features provided. The final prediction for a data sample is obtained by aggregating the results from all trees, usually through majority voting. It means that the class most frequently predicted by the trees in the forest will be taken as the final prediction result. The advantages of Random Forest in such cases include its ability to handle large databases with many features [20], its tolerance for missing data [21], and its capacity to assess the importance of different features for classification due to its ensemble nature [22], where it combines many simple models to make robust predictions.

LightGBM

LightGBM (Light Gradient Boosting Machine) is a machine learning algorithm based on gradient boosting techniques that is efficient and effective for various tasks, including classification [23]. In

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8251

Table 2. Experimental Results on The Effect of The Percentage of Training Data and Data Validation

Split Data		SVM			Random Forest			LightGBM		
Train	Validation	Precision	Recall	Accuracy	Precision	Recall	Accuracy	Precision	Recall	Accuracy
50	50	0.9377	0.9396	0.9400	0.9021	0.8979	0.9013	0.9152	0.9118	0.9147
60	40	0.9380	0.9391	0.9400	0.9092	0.9048	0.9083	0.9216	0.9196	0.9217
70	30	0.9325	0.9345	0.9333	0.9045	0.8977	0.9022	0.9279	0.9250	0.9267
80	**20**	**0.9462**	**0.9498**	**0.9467**	**0.9174**	**0.9133**	**0.9133**	**0.9465**	**0.9414**	**0.9433**

Source : (Research Results, 2026)

weather classification with five different classes, cloudy, foggy, rainy, shine, and sunrise, LightGBM processes this classification task by optimizing a predictive model consisting of a small number of decision trees which are added iteratively. LightGBM uses histograms to build decision trees [24].

Each newly added tree attempts to correct errors made by models previously in the chain, so that with each iteration, the model becomes increasingly accurate. In the training process, LightGBM treats multiclass classification as a series of binary classification tasks.

System Evaluation

Weather classification with multiclass cloudy, foggy, rainy, shine, and sunrise in this research was evaluated using macro recall, macro precision, and accuracy matrices. Macro recall measures the average success of the model in identifying all positive cases for each class separately.

Ensuring that each weather condition is identified correctly without ignoring other conditions is important. Macro precision calculates the average proportion of correct predictions among those predicted for each class, indicating how accurate the model is in predicting each type of weather. T

here's also the macro F1 matrix, which combines macro recall and macro precision. In this research, the macro F1 was used to evaluate the testing data. The final matrix is accuracy, which measures the proportion of correct predictions for all classes. So, accuracy provides an overall picture of the general effectiveness of the model. The equations of macro recall, macro precision, and macro F1 are shown in equations (1), (2), and (3).

$$macro\ Recall = \frac{1}{n}\sum_i \frac{TP_i}{TP_i + FN_i} \quad (1)$$

$$macro\ Precision = \frac{1}{n}\sum_i \frac{TP_i}{TP_i + FP_i} \quad (2)$$

$$macro\ F1 = \frac{1}{n}\sum_i \frac{2P_iR_i}{P_i + R_i} \quad (3)$$

RESULTS AND DISCUSSION

Weather classification research uses a Core i5 computer configuration with 8GB RAM. The Python programming configuration used is Python 3 and CPU-based. The research experiment began with an experimental percentage of training data and data validation. The best results were used to test parameters in supervised learning. The results of this research will be compared with previous research using the same dataset.

Experimental percentage of training data and data validation

In this experiment, the percentage of training data and data validation will use four scenarios, namely 50 : 50, 60 : 40, 70 : 30, and 80 : 20. Then the parameters in the SVM method used are linear kernels, and in the Random Forest and LightGBM methods using the default parameter, namely n_estimator 100. Table 2 shows the experimental results of the training data and data validation percentage in this research.

The results of this experiment show that increasing the percentage of training data can increase the precision, recall, and accuracy values of all supervised learning models. In the SVM model, the highest precision, recall, and accuracy were 0.9462, 0.9498, and 0.9467, respectively, where these results were achieved with an 80 : 20 data split configuration. The Random Forest method increased performance with the highest precision, recall, and accuracy at the 80 : 20 split data. The result using this configuration is 0.9174, 0.9133, and 0.9133, respectively, likewise with LightGBM, which shows excellent performance with the highest precision, recall, and accuracy at an 80:20 split with results of 0.9465, 0.9414, and 0.9433, respectively. SVM and LightGBM show significant performance improvements compared to Random Forest. From these results, a more significant percentage of training data results in better model performance. The best configuration is 80 : 20 for data training and validation data. Therefore, in the next experiment, we will use this configuration. The SVM method uses various kernels for its experiments, while Random Forest and LightGBM use n_estimator as their experimental variable. This parameter is used because it regulates the number of trees, which provides predictive stability and affects classification performance. The other parameters are the default parameters for each supervised learning method.

Experiment of hybrid MobileNet-SVM

The hybrid MobileNet-SVM uses an 80:20 training data and data validation. In this experiment, three kernels in SVM will be tested, namely polynomial, RBF, and linear. Table 3 shows the experimental results of the influence of the kernel in SVM on precision, recall, and classification accuracy results.

The RBF kernel provides the best performance with precision, recall, and accuracy of 0.9525, 0.9538, and 0.9500, respectively. The RBF kernel has the advantage of using a flexible Gaussian function to adapt to the data vector's complexity resulting from feature extraction via the gamma parameter. Precise gamma creates a hyperplane to form a more precise boundary around the data, capturing subtle patterns that linear and polynomial kernels might overlook. Therefore, the RBF kernel can produce more effective hyperplanes in separating data classes, thereby providing better precision, recall, and accuracy.

Table 3. Experimental Results on The Effect of The Kernel on SVM

Kernel	Precision	Recall	Accuracy
polynomial	0.9429	0.9447	0.9433
RBF	**0.9525**	**0.9538**	**0.9500**
linear	0.9462	0.9498	0.9467

Source : (Research Results, 2026)

Experiment of hybrid MobileNet-Random Forest

The MobileNet-Random Forest hybrid experiment also uses an 80 : 20 training data and data validation. This experiment will evaluate the n_estimator, namely the number of trees in the Random Forest. Table 4 shows the experimental results of the influence of n_estimator in Random Forest on precision, recall, and classification accuracy results.

Table 4. Experimental Results on The Effect of The n_estimator on Random Forest

n_estimator	Precision	Recall	Accuracy
80	0.9070	0.8991	0.9000
100	0.9174	0.9133	0.9133
120	0.9182	0.9120	0.9133
140	**0.9226**	**0.9160**	**0.9167**
160	0.9212	0.9149	0.9167
180	0.9177	0.9138	0.9133

Source : (Research Results, 2026)

The results of weather classification using the hybrid MobileNet-Random Forest method show variations in performance based on the n_estimator parameter. This parameter of n_estimator represents the number of trees used in the Random Forest method. Table 6 shows that performance values increase as the n_estimator value increases from 80 to 180. The highest value was achieved using n_estimator 140, namely precision of 0.9226, recall of 0.9160, and accuracy of 0.9167. These results show that increasing the number of trees in Random Forest to a certain point increases the model's ability to capture patterns and variations in weather image data. However, when the n_estimator value was increased to 180, the model performance decreased slightly, with precision of 0.9177, recall of 0.9138, and accuracy of 0.9133. This decrease is caused by overfitting, where many trees cause the model to be too tied to the training data and less able to generalize to the testing data. It shows that although increasing the number of trees can improve performance to a certain extent, there is a point where further increase no longer improves data classification performance.

Experiment of hybrid MobileNet-LightGBM

The final hybrid experiment is the MobileNet-LightGBM combination. This experiment also uses an 80:20 configuration of training data and data validation. The parameter that will be evaluated is n_estimator, which refers to the number of trees the model will build, the same as the Random Forest method. Table 5 shows the experimental results of the influence of the n_estimator in LightGBM on the precision, recall, and classification accuracy results.

Table 5. Experimental Results on The Effect of The n_estimator on LightGBM

n_estimator	Precision	Recall	Accuracy
80	0.9437	0.9381	0.9400
100	0.9465	0.9414	0.9433
120	0.9465	0.9414	0.9433
140	**0.9524**	**0.9481**	**0.9500**
160	0.9524	0.9481	0.9500
180	0.9524	0.9481	0.9500

Source : (Research Results, 2026)

The weather classification using the MobileNet-LightGBM hybrid method produce performance variations based on the n_estimators parameter value. Based on Table 5, precision, recall, and accuracy increase with increasing number of trees up to a certain point. At n_estimators 100, the model achieves a precision of 0.9465, recall of 0.9414, and accuracy of 0.9433.

Increasing the n_estimator parameter value from 100 to 140 in LightGBM can improve classification performance. In this experiment, the

Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)
VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8251

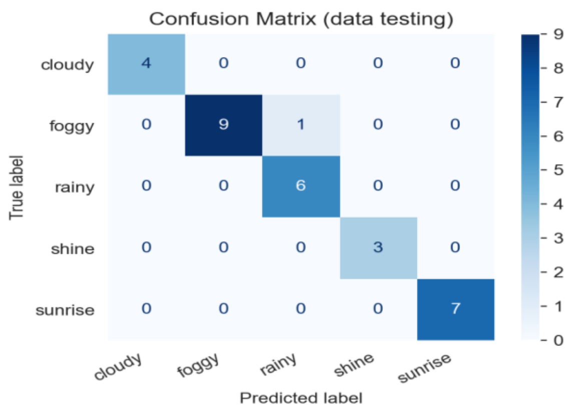
best performance of LightGBM was achieved using n_estimator value of 140. The precision, recall, and accuracy results obtained are 0.9524, 0.9481, and 0.9500. Then, increasing the n_estimator value further to 180 does not improve classification performance. These results indicate that increasing the number of trees to more than 140 does not significantly increase model performance.

Experiment using Data Testing

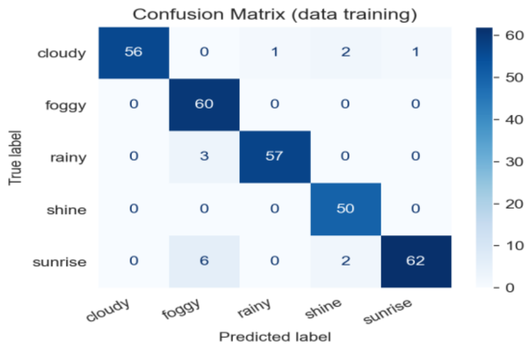
In this experiment, we used three models with configurations based on previous experiments and produced the best accuracy. The kernel used by the SVM method is RBF. The Random Forest and LightGBM methods use n_estimator 140. Table 6 shows experiments using testing data. Weather classification test results show that the MobileNet-SVM hybrid method produced the highest accuracy, namely 0.9667. Figure 4 shows the confusion matrix using the training and testing data.

The confusion matrix results from the training process showed excellent performance, with almost all classes located on the main diagonal. The foggy and shine classes performed best in predicting result. The class with the most prediction errors was sunrise, with eight of the 70 data. Furthermore, in the 30 testing data sets, 29 were correctly classified in all evaluation methods. Only one image was detected incorrectly.

The actual label was foggy, but the predicted result was rainy. Based on the training results, the combination of the MobileNet-SVM and MobileNet-LightGBM models produced the same accuracy, namely 0.9500. However, the recall value of MobileNet-SVM was higher than that of MobileNet-LightGBM, so MobileNet-SVM produced the highest testing accuracy. Statistically, the difference is not very significant. However, this research focused on achieving optimal accuracy, so the model that produced the highest accuracy was selected.



(a)



(b)

Source : (Research Results, 2026)

Figure 4.(a)(b) Confusion matrix from data training and data testing using MobileNet-SVM

Table 6. Experimental results using data testing

Method	Accuracy	F1
MobileNet-SVM	0.9667	0.9741
MobileNet-Random Forest	0.9000	0.9223
MobileNet-LightGBM	0.9333	0.9467

Source : (Research Results, 2026)

Discussion

Experimental results show that the combination of MobileNet-SVM and MobileNet-LightGBM is superior to MobileNet-Random Forest. Using MobileNet for feature extraction makes the resulting features more discriminatory between classes in high-dimensional spaces. SVM can perform better by maximizing margins with a non-linear kernel. LightGBM can also handle non-linearities and relevant feature interactions. However, the use of bagging-based Random Forest is less optimal for exploring subtle discriminatory patterns. Therefore, the hybrid results with Random Forest have the lowest accuracy.

Hybrid transfer learning, combining feature extraction using a deep learning model such as MobileNet with an SVM-supervised learning classifier, provides better results than just using transfer learning. Table 7 shows the classification report results using the MobileNet-SVM hybrid. The macro F1 and accuracy scores are not significantly different because the model performance in each class is fairly even and the data used is relatively balanced. Classification using flatten or global average pooling effectively reduces feature dimensions and prepares them for dense layers. However, they could be more optimal in handling the complexity and variability of the extracted feature data. Supervised learning methods like SVM have more sophisticated and special mechanisms for handling structured data. For example, SVM can effectively maximize the margin between classes, providing more apparent separation between different classes. In other words, supervised learning classifiers better capture complex and non-linear patterns in the extracted feature data. Therefore, combining powerful feature extraction from a transfer learning model with a supervised learning classifier can result in superior classification performance. Table 8 shows a comparison of the results of this research with previous research using the same dataset. The configuration for dividing training and data validation is also made the same, namely 80% training data and 20% data validation. The number of training data for cloudy, foggy, rainy, shine, and sunrise classes is 240, 240, 240, 200, and 280, while the number of validation data is 60, 60, 60, 50, and 70.

Table. Classification Report Using MobileNet-SVM

Class	Precision	Recall	F1-Score	Support
Cloudy	1.0000	0.9333	0.9655	60
Foggy	0.8696	1.0000	0.9302	60
Rainy	0.9828	0.9500	0.9661	60
Shine	0.9259	1.0000	0.9615	50
Sunrise	0.9841	0.8857	0.9323	70
Accuracy			0.9500	300
Macro Avg	0.9525	0.9538	0.9511	300
Weighted Avg	0.9544	0.9500	0.9502	300

Source : (Research Results, 2026)

Table 8. Comparison With Previous Research

Methods	Accuracy using data training	Accuracy using data testing
DenseNet-201 [6]	0.908	-
VGG-19 [6]	0.882	-
MobileNet-SVM (Ours)	**0.950**	**0.967**

Source : (Research Results, 2026)

Based on Table 8, the research we propose has better results in accuracy using data training. In the experiment using testing data, only one image was mispredicted because the image displayed contained another large object, so the weather was not clear. The MobileNetV2-SVM combination yielded better training accuracy than DenseNet-201 and VGG-19 because MobileNetV2 has fewer parameters. On a smaller dataset, complex models can easily overfit. DenseNet-201 and VGG-19 have very large parameters, requiring more training data. Based on these results, the weakness of this research is that weather classification that is obstructed by other objects in the image can result in classification errors. The strategy needed is to segment by taking only images that describe the weather. Other solutions can also use a combination of other transfer learning for feature extraction methods to be more accurate in obtaining features as input for supervised learning.

CONCLUSION

Using the hybrid transfer learning method as feature extraction and supervised learning as a classifier provides better results than using only transfer learning via fully connected layers. The three hybrid methods were evaluated using testing data, namely MobileNet-SVM, MobileNet-Random Forest, and MobileNet-LightGBM, where the best results were obtained from MobileNet-SVM with an accuracy of 0.9667. However, it not only considers the accuracy of the testing data but also the accuracy of the training data. So, in this research, the best results were achieved by hybrid MobileNet-SVM, where the training accuracy was 0.9500. T

hrough these results, the MobileNet-SVM combination shows to be more optimal when using small datasets, which is proven to produce higher accuracy than pure deep learning models. The advantage of this hybrid method lies in the ability of the transfer learning model to extract rich and meaningful features from images. It is proven that each image produces 1000 feature vectors. These features are then processed further by a supervised learning classifier, effectively handling non-linear feature data.

Future research can clean the data before processing it into feature extraction so that the weather appears clearer than other objects. Data processing can use segmentation methods before becoming input for feature extraction. In addition, fine-tuning other parameters in MobileNet is used to achieve optimal accuracy. These results allow the classification to be implemented in the real world, where farmers can recognize weather through input image data. The use of larger datasets is also needed to improve the generalization of the classification. The industry can also integrate this method with IoT devices so that farmers can more easily implement it for weather classification. The results of this integration also need to be evaluated regarding the method's classifying speed.

REFERENCE

[1] J. Schmitt, F. Offermann, M. Söder, C. Frühauf, and R. Finger, “Extreme weather events cause significant crop yield losses at the farm level in German agriculture,” *Food Policy*, vol. 112, no. August, 2022, doi: 10.1016/j.foodpol.2022.102359.

[2] R. Ramadhan *et al.*, “Trends in rainfall and hydrometeorological disasters in new capital city of Indonesia from long-term satellite-based precipitation products,” *Remote Sensing Applications: Society and*

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8251

Environment, vol. 28, p. 100827, 2022, doi: https://doi.org/10.1016/j.rsase.2022.100827.

[3] A. Kolios, M. Richmond, S. Koukoura, and B. Yeter, “Effect of weather forecast uncertainty on offshore wind farm availability assessment,” *Ocean Engineering*, vol. 285, no. P1, p. 115265, 2023, doi: 10.1016/j.oceaneng.2023.115265.

[4] A. Haleem, M. Javaid, M. Asim Qadri, R. Pratap Singh, and R. Suman, “Artificial intelligence (AI) applications for marketing: A literature-based study,” *International Journal of Intelligent Networks*, vol. 3, pp. 119–132, 2022, doi: https://doi.org/10.1016/j.ijin.2022.08.005.

[5] S. Joksimovic, D. Ifenthaler, R. Marrone, M. De Laat, and G. Siemens, “Opportunities of artificial intelligence for supporting complex problem-solving: Findings from a scoping review,” *Computers and Education: Artificial Intelligence*, vol. 4, p. 100138, 2023, doi: https://doi.org/10.1016/j.caeai.2023.100138.

[6] G. F. Fitriana, A. B. Arifa, A. Prasetiadi, F. D. Adhinata, and N. G. Ramadhan, “Improving Accuracy of Cloud Images Using DenseNet-VGG19,” *International Journal on Advanced Science, Engineering and Information Technology*, vol. 13, no. 2, pp. 688–693, 2023, doi: 10.18517/ijaseit.13.2.18293.

[7] M. M. Taye, “Understanding of Machine Learning with Deep Learning: Architectures, Workflow, Applications and Future Directions,” *Computers*, vol. 12, no. 5, 2023, doi: 10.3390/computers12050091.

[8] Y. Liu, H. Pu, and D. W. Sun, “Efficient extraction of deep image features using convolutional neural network (CNN) for applications in detecting and analysing complex food matrices,” *Trends in Food Science and Technology*, vol. 113, no. May, pp. 193–204, 2021, doi: 10.1016/j.tifs.2021.04.042.

[9] S. A. Salleh *et al.*, “Support Vector Machine (SVM) and Object Based Classification in Earth Linear Features Extraction: A Comparison,” *Revue Internationale de Géomatique*, vol. 33, no. 1, pp. 183–199, 2024, doi: 10.32604/rig.2024.050723.

[10] W. Rahmaniar and A. Hernawan, “Real-time human detection using deep learning on embedded platforms: A review,” *Journal of Robotics and Control (JRC)*, vol. 2, no. 6, pp. 462-468Y, 2021, doi: 10.18196/jrc.26123.

[11] F. D. Adhinata, N. G. Ramadhan, A. Amrulloh, and A. R. Bahtiar, “Comparison of Supervised Learning Methods for COVID-19 Classification on Chest X-Ray Image,” *CommIT Journal*, vol. 16, no. 2, pp. 195–201, 2022, doi: 10.21512/commit.v16i2.7970.

[12] Y. Wu and G. Tao, “Application of a New Loss Function-Based Support Vector Machine Algorithm in Quality Control of Measurement Observation Data,” *Mathematical Problems in Engineering*, vol. 2022, 2022, doi: 10.1155/2022/7266719.

[13] V. Gupta, “Weather Classification,” *Kaggle.com*, 2020. https://www.kaggle.com/datasets/vijaygiitk/multiclass-weather-dataset/ (accessed May 22, 2024).

[14] B. T. Felix and Suharjito, “Face Liveness Classification Using Mobilenet and Support Vector Machines,” *ICIC Express Letters*, vol. 16, no. 7, pp. 779–786, 2022, doi: 10.24507/icicel.16.07.779.

[15] T. H. Tsai, Y. C. Ho, and P. T. Chi, “Hardware Architecture Design for Hand Gesture Recognition System on FPGA,” *IEEE Access*, vol. 11, no. June, pp. 51767–51776, 2023, doi: 10.1109/ACCESS.2023.3277857.

[16] L. A. Latumakulita, F. J. Paat, Saroyo, I. Karim, I. N. G. A. Astawa, and H. Sirait, “Comparison of MobileNet and VGG16 CNN Architectures for Web-based Starfish Species Identification System,” *Journal of Applied Data Sciences*, vol. 5, no. 4, pp. 2117–2130, 2024, doi: 10.47738/jads.v5i4.456.

[17] R. Guido, S. Ferrisi, D. Lofaro, and D. Conforti, “An Overview on the Advancements of Support Vector Machine Models in Healthcare Applications: A Review,” *Information (Switzerland)*, vol. 15, no. 4, 2024, doi: 10.3390/info15040235.

[18] R. Krebs, S. S. Bagui, D. Mink, and S. C. Bagui, “Applying Multi-CLASS Support Vector Machines: One-vs.-One vs. One-vs.-All on the UWF-ZeekDataFall22 Dataset,” *Electronics*, vol. 13, no. 19, 2024, doi: 10.3390/electronics13193916.

[19] Z. Sun, G. Wang, P. Li, H. Wang, M. Zhang, and X. Liang, “An improved random forest based on the classification accuracy and correlation measurement of decision trees,” *Expert Systems with Applications*, vol. 237, no. PB, p. 121549, 2024, doi: 10.1016/j.eswa.2023.121549.

[20] D. Ghosh and J. Cabrera, “Enriched Random Forest for High Dimensional Genomic Data.,” *IEEE/ACM transactions on computational biology and bioinformatics*, vol. 19, no. 5, pp.

2817–2828, 2022, doi: 10.1109/TCBB.2021.3089417.

[21] F. Mostafa, E. Hasan, M. Williamson, and H. Khan, "Statistical Machine Learning Approaches to Liver Disease Prediction," *Livers*, vol. 1, no. 4, pp. 294–312, 2021, doi: 10.3390/livers1040023.

[22] A. Kharwar and D. Thakor, "A random forest algorithm under the ensemble approach for feature selection and classification," *International Journal of Communication Networks and Distributed Systems*, vol. 29, no. 4, pp. 426–447, 2023, doi: 10.1504/IJCNDS.2023.131737.

[23] T. O. Omotehinwa, D. O. Oyewola, and E. G. Moung, "Optimizing the light gradient-boosting machine algorithm for an efficient early detection of coronary heart disease," *Informatics and Health*, vol. 1, no. 2, pp. 70–81, 2024, doi: 10.1016/j.infoh.2024.06.001.

[24] Y. Liu and Z. Chen, "LightGBM-Based Human Action Recognition Using Sensors," *Sensors*, vol. 25, no. 12, pp. 1–17, 2025, doi: 10.3390/s25123704.