p-ISSN:2527-5321
e-ISSN:2527-5941

NOAA IMAGE CLASSIFICATION USING GOOGLE'S TEACHABLE MACHINE

Abdul Azzam Ajhari[1], Krisna Jonathan Sitorus[2], Dayang Putri Nur Lyrawati[3], Edward Felix Satria Sridana[4], Ria Lestari[5], Wahyuni Vivi Farida[6]

[1,3,4,5,6]*Pusat Pengkajian dan Pengembangan Teknologi Keamanan Siber dan Sandi,*
Badan Siber dan Siber dan Sandi Negara Republik Indonesia
Jakarta, Indonesia
[2]*Universitas Pembangunan Nasional Veteran Jakarta*
Jakarta, Indonesia
[1]abdul.azzam@bssn.go.id, [2]krisnajs@upnvj.ac.id, [3]dayang.putri@bssn.go.id, [4]edward.felix@bssn.go.id, [5]ria.lestari@bssn.go.id, [6]wahyuni.vivi@bssn.go.id

ABSTRAK

Para peneliti telah mencoba membantu kehidupan manusia dengan memanfaatkan citra satelit. Salah satu satelit yang dapat menyediakan data citra satelit adalah satelit *National Oceanic and Atmospheric Administration* (NOAA). Satelit ini secara khusus menangkap kondisi fisik laut dan atmosfer bumi dengan sisi negatif bahwa tidak semua data yang diperoleh berkualitas baik karena adanya intervensi frekuensi radio. Dalam tulisan ini, penggunaan kecerdasan buatan/*Artificial Intelligence* (AI) diterapkan untuk mengatasi masalah tersebut dengan membuat model pembelajaran mesin/*Machine Learning* (ML) menggunakan mesin yang bisa diajar dari google/*Google's Teachable Machine* untuk mengklasifikasikan gambar dari perangkat frekuensi radio *Weather Satellite Communication* (WeSaCom) yang menangkap gambar menjadi mampu menyaring gambar. Data mining digunakan sebagai metodologi untuk mengumpulkan data citra dari satelit NOAA. Dataset citra dari satelit dianalisis dan dikelompokkan menjadi dua kelas, baik dan buruk. Kelas-kelas ini digunakan untuk membangun model dengan tujuan mengklasifikasikan data gambar yang diperoleh dari satelit NOAA melalui perangkat frekuensi radio WeSaCom. Hasilnya, ditemukan bahwa model Google Teachable Engine yang diterapkan pada perangkat WeSaCom mampu mengklasifikasikan data gambar dari NOAA dengan akurasi 96,72%.

Kata Kunci: kecerdasan buatan, pembelajaran mesin, *google's teachable machine*, satelit administrasi kelautan dan atmosfer nasional, klasifikasi gambar.

ABSTRACT

Researchers have tried to help human life by utilizing satellite imagery. One of the satellites that can provide satellite image data is the National Oceanic and Atmospheric Administration satellite (NOAA). This satellite specifically captures the physical conditions of the sea and the earth's atmosphere with the negative side that not all data obtained is of good quality due to radio frequency intervention. In this paper, the use of Artificial Intelligence (AI) is applied to overcome these problems by creating a Machine Learning (ML) model using Google's Teachable Machine to classify images from a Weather Satellite Communication (WeSaCom) radio frequency device that captures images to be able to filter images. Data mining is used as a methodology to collect image data from the NOAA satellites. Image datasets from satellites are analyzed and grouped into two classes, good and bad. We use these classes to build a model with the aim of classifying image data obtained from the NOAA satellites via the WeSaCom radio frequency device. As a result, we found that the Google's Teachable Machine model applied to the WeSaCom device was able to classify image data from the NOAA with an accuracy of 96.72%.

Key Word: *artificial intelligence, machine learning, google's teachable machine, national oceanic and atmospheric administration satellites, image classification.*

PENDAHULUAN

"National Oceanic and Atmospheric Administration (NOAA) satellite is a satellite owned by United States of America (USA) that can be used to monitor the Earth's weather with one of the parameter used is temperature". (Rasyidi et al, 2020)

One of the sensors in NOAA is Advanced-Very-High-Resolution Radiometer (AVHRR). "With AVHRR sensor, NOAA can get information about physical condition of the sea and the atmosphere". (Akhbar et al, 2018)

p-ISSN:2527-5321
e-ISSN:2527-5941

"Data obtained from NOAA can be processed to study about meteorology parameter, such as to predict temperature and weather forecast". (Wicaksono et al, 2010)

Unfortunately, not all the data acquired from NOAA is in a good quality. This can be caused by intervention from radio frequency signal in urban area, especially Jakarta the place this experiment took place. Because of this very problem, we need an Artificial Intelligence (AI) that could classify between the "good" and the "bad/noise" images. More specifically, we will try to create a Machine Learning (ML) which is a part of AI. ML itself has been increasingly prevalent in today's daily life. "It has powered recommendation systems, translation, speech transcription and many more". (Carney et al, 2020)

"Machine learning consists of feature extraction module that extracts the important features such as edges, textures, etc and a classification module that classify based on the features extracted". (Krisna et al, 2018)

Google's Teachable Machine is a web-based system for teaching and learning the basic principles of ML. "It takes some steps towards exposing the ML training process and combines powerful classification algorithms with intuitive and an easy-to-learn graphical user interface". (Toivonen et al, 2020)

"Google's Teachable Machine can empower accessibility research". (Kacorri, 2017) Based on some of these backgrounds, this study will apply the classification of NOAA satellite imagery using the Google's Teachable Machine. Indonesia, which is an archipelagic country, is very suitable to use satellite imagery with radio frequency. In addition, there are no studies that utilize satellite images obtained from radio frequencies to be trained with ML, images that have been predicted and classified into images with good quality and images with poor quality (bad/noise) can shorten the performance of researchers. Researchers can take advantage of good or bad/noise image quality for the next stage of research.

METODE PENELITIAN

A lot of works regarding either Machine Learning, Artificial Intelligence, Google's Teachable Machine, or NOAA and its use has been published all over the world before this research even took.

In 2016, (Wahyuningsih et al, 2016) conducted research using the NOAA satellite to obtain sea surface temperature conditions as a support for estimating the potential for fish distribution points in the city of Tegal, Central Java.

Meanwhile, in 2018 (Akhbar et al, 2018) did a research using the NOAA satellite to determine the sea surface temperature with the addition of metered float data using the McMillin & Crosby algorithm.

Another application of NOAA satellite can be found in (Wicaksono et al, 2010) research, as stated in his paper published in April 2010. They are using image data from NOAA-17 to determine temperatures on java sea surface. But, during the research, they found data with stripping on it caused by unstable signal when receiving them. Hopefully our AI could easily solve this problem, and by that could encourage other researcher to do more experiment on the application of NOAA based on the data filtered with our AI.

As mentioned before, we would like to conduct a research of NOAA image classification using a tool named Google's Teachable Machine. Google's Teachable Machine is a web-based interface that allows people to train their own Machine Learning classification models, without coding and/or using their webcam, images, or sound. "It uses transfer learning, an ML technique, to find patterns and trends within the images or sound samples and create a simple and easy classification model within seconds". (Carney et al, 2020)

Atmosphere research (Natarajan and Philipoff, 2018) study about earthquake at Sumatra and Nicobar in 2014 using surface and atmospheric parameters from NOAA-18 satellite. The result is they found the anomaly on Nicobar Island earthquake on Feb 27, 2014, and the second anomaly appeared on March 12, 2014 until the earthquake took time on March 21, 2014. For the Sumatra Earthquake, they found three anomalies for three different times. The first anomaly happened on April 16, 2014, the second anomaly happened April 26, 2014, and the last one happened on May 06, 2014. Until the earthquake happened on May 18, 2014. Their paper concludes that an earthquake prediction can be made by using NOAA imagery data.

p-ISSN:2527-5321
e-ISSN:2527-5941

An attempt to do a NOAA classification with an AI was also done by (Chakraborty et al, 1993). Their research in classification of The NOAA Satellite Image Data by Unsupervised Neural Network was successful. However, the algorithm used by the author leads to misclassification when nearby classes have very different variances.

NOAA is not the only earth-monitoring satellite out there that we can find. NASA owns this kind of satellite too, known by AQUA MODIS. In (Hidayat et al, 2015), the researchers use AQUA MODIS imagery to determine the temperature conditions on sea and land surface. Result shows that the distribution of sea surface temperatures in the North Coast of Semarang achieve maximum value degrees in East season, which was 32.3ºC, and minimum value degrees in West Season, which was 24.6º C.

However, a study to compare the accuracy of both satellites was done before by Saputra, 2019). His study uses R-squared (R^2) and Normalize Mean Absolute Error (NMAE) to test the validation of sea surface temperature from both satellites compared to the actual temperature (study was done in Madura Strait). The value of R-squared & NMAE for NOAA and AQUA MODIS respectively are 0.58 & 0.57 and 2.06% & 21.45%. This concludes that R-squared & NMAE has better accuracy than AQUA MODIS.

Several studies using Google Teachable Machine are image classification used to classify diseases. Google’s Teachable Machine was used (Jeong, 2020) for Tooth Marked Tongue Diagnosis. Then, (Forchammer et al, 2022) developed a prognostic score based on image analysis on melanoma disease.

From many studies, Google's Teachable Engine has not been used to predict and classify images from Weather Satellite Communications radio frequency devices. The results of the data captured by the device are NOAA satellite weather image data in the Indonesian archipelago. The data cannot be separated between data with good image quality (good/clean) and poor image quality (noise/bad). This research utilizes an automated model from Google’s Teachable Machine to predict and classify NOAA satellite weather images and separate them between good quality images and poor-quality images.

Data retrieval uses a satellite radio frequency signal capture device, namely weather satellite communication (WeSaCom). In identifying patterns, structures, and insights from the data collected, the data mining methodology has 5 stages, namely business understanding, data acquisition & understanding, data preparation, modeling, and evaluation. The process of these stages is applied in this study with the flow chart in Figure 1 as follows.

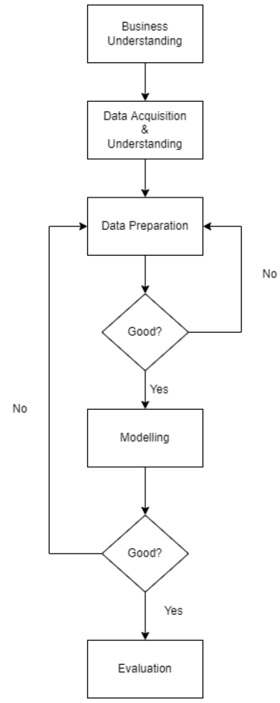

Figure 1. Data Mining process

A. *Business Understanding*

NOAA is a United States of America’s satellite that can predict weather and climate for various purposes in helping human life. In this business process, NOAA's image is used for prediction and classified into images with good quality and poor quality (bad/noise). Poor quality images are images that cannot be seen clearly in parts of the Indonesian archipelago. After finding the business process pattern, we can proceed to the next process.

B. *Data Acquisition and Understanding*

The processed image data comes from the NOAA-15, NOAA-18, and NOAA-19 satellites which passes through Indonesian territory. Ground station receiver device that is used is the WeSaCom APT-06 with a

 p-ISSN:2527-5321
e-ISSN:2527-5941

turnstile antenna and working on frequency 137.10 MHz to 137.9125 MHz (Luo, 2021).

The images obtained from the WeSacom Ground Station System then classified manually with a good image quality percentage which has a percentage of 80% of the Indonesian archipelago can look good.

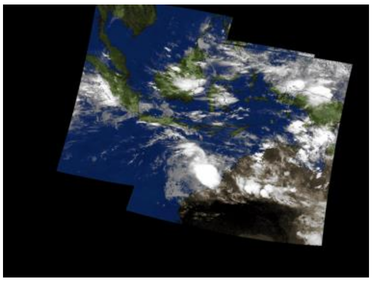

Figure 2. Good image quality

In Figure 2, the image is clean, and more than 80 percent of the Indonesian archipelago is clearly visible.

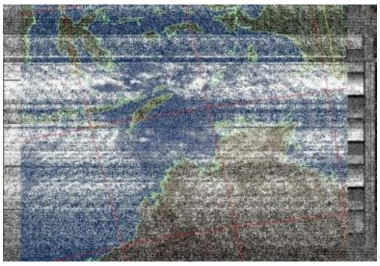

Figure 3. Bad/poor image quality

As for images with poor quality can be seein in Figure 3, the percentage of images is above 30% with the Indonesian archipelago that cannot be seen properly.

C. Data Preparation

The data is divided into two sets, namely training data and test data with each data has noise and clean class and sub-directories. In the clean sub-directory contains NOAA satellites image data that are of good quality. Besides that, noise sub-directory contains poor image data quality. Sub-directories of each class are labeled with tag 0 for noise and tag 1 for clean dataset. The image that is considered suitable can be continued to the modeling stage, but if it is not suitable then it needs to be readjusted. Adjusted images will be reordered manually based on image quality.

D. Modelling

After the image data passes through the preparation stage, the next step is modelling using Google's Teachable Machine. By default, the data is dividing into 85% for training data and 15% for test data.
Standard image processing used with TensorFlow at 224x224 pixel color images. There are 127 images for each noise and good class. So that the total images owned in this research amounted to 254 images.
The total training data used is 214 data and the remaining 40 data is used for test data in the evaluation process. Some of the hyperparameters carried out in this research model can be seen in Table 1.

TABLE I
HYPERPARAMETER SUMMARY

No	Hyperparameter	Alternative Value
1	Epochs	50
2	Batch Size	[32, 64]
3	Learning Rate	[0.001, 0.0001]

Model parameters that are suitable and considered good for comparison can then proceed to the evaluation stage, but if the model is not optimal on the parameters that have been set, the data will be re-examined in the preparation of the data.

E. Evaluation

The performance of the model is measured based on the classification of the testing data compared to its ground truth. Comparison of classification results and ground truth data testing is tabulated into a confusion matrix can be seen in Figure 6. The confusion matrix score reflected the model's ability to predict accurately all the positive predictions it generates (Soni et al, 2021).

		Actual	
		Positive	Negative
Prediction	Positive	TP	FP
	Negative	FN	TN

Figure 4. Confusion matrix evaluation

Figure 4 regarding the confusion matrix can be explained as follows:

- True Positive (TP) is data that is predicted to be positive and has a positive value.
- True Negative (TN) is data that is predicted to be negative and has a negative value.
- False Positive (FP) is data that is predicted to be positive and has a negative value.
- False Negative (FN) is data that is predicted to be negative and has a positive

 p-ISSN:2527-5321
e-ISSN:2527-5941

value.

To measure the performance results, the confusion matrix classification results have four terms, namely overall accuracy, precision, recall, and f1-score, as explained below.

Performance measurement was predicted using the overall accuracy as the following equation (1). Model accuracy score reflected the model's capacity to predict both accurately positive and negative of all results.

$$accuracy = \frac{TP+TN}{TP+FP+FN+TN} \quad (1)$$

Then, to measure the level of positive observation ratio that was correctly predicted, precision equation (2) was used.

$$precision = \frac{TP}{TP+FP} \quad (2)$$

The ratio of positive observations was correctly predicted to all statements in the actual class to measure the recall (3) value. The precision score was model accuracy.

The score reflected the model's ability to predict accurately all the positive predictions it generates.

$$recall = \frac{TP}{TP+FN} \quad (3)$$

Finally, to see the balance of performance, the f1-score calculation (4) was carried out from the average precision and recall value. The recall evaluated the effectiveness of the proposed model in identifying positive samples.

$$f1 - score = 2x\frac{precision\ x\ recall}{precision + recall} \quad (4)$$

HASIL DAN PEMBAHASAN

The four experiments that have been carried out at the evaluation stage are ranked where the largest f1-score is the best hyperparameter model.
The model was trained four times with the evaluation results can be seen in Table 2.

In the first experiment in Table 2, training was conducted with hyperparameters with 50 epochs, 32 batch size, and a learning rate with a value of 0.001. The results of the evaluation obtained training performance accuracy of 95.32% and f1-score of 86.48%.

TABLE II
CONFUSION MATRIX

No	Overall Accuracy	Precision	Recall	F1-score
1	87.50%	80.00%	94.12%	86.48%
2	87.50%	85.00%	89.47%	87.18%
3	87.50%	90.00%	85.71%	87.80%
4	87.50%	85.00%	89.47%	87.18%

Second experiment, training was conducted with hyperparameters with same epochs and learning rate, but batch size modified to 64. The results of the second evaluation obtained training performance accuracy of 94.85% and f1-score of 87.18%.
The third experiment still uses the same hyperparameters as the first experiment, but by changing the learning rate to a value of 0.0001. The results of the third evaluation experiment obtained training performance accuracy of 96.72% and f1-score 87.80%.
The last experiment is the fourth, the hyperparameter used is the same as the second experiment but is carried out by modifying the learning rate to 0.0001. The results of the last experiment obtained training performance accuracy of 96.26% and f1-score 87.18%.
The best evaluation performance of the whole experiment is the third trial, the results are shown in the visualization of accuracy per epoch shown in Figure 5.

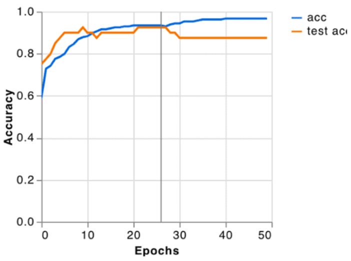

Figure 5. Performance evaluation accuracy per epoch

The best accuracy evaluation performance is shown in Figure 5 of the third experiments carried out, obtained an accuracy of 96.72% with an f1-score of 87.80%. In addition to performance evaluation, which is assessed from accuracy per epoch, performance is also

p-ISSN:2527-5321
e-ISSN:2527-5941

measured by performance evaluation of loss per epoch which can be seen in Figure 6.

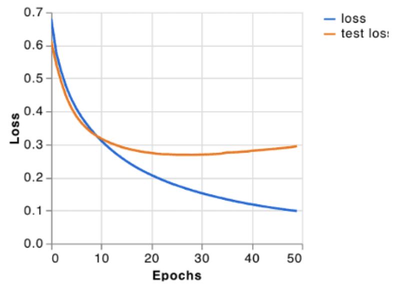

Figure 6. Performance evaluation loss per epoch

While in Figure 6 the performance evaluation of the third experiment for loss accuracy has a large enough gap with the test loss. training loss produces an optimal value of 9.83% and a test loss of 29.41% on the parameters of batch size 32 and learning rate of 0.0001. So, it can be said that this model has an underfit character.

Although Google's Teachable Machine automated ML model has obtained a good accuracy score, this model needs to be compared to using Deep Learning (DL) models such as Convolution Neural Network. Thus, precise parameters and better comparison models are needed to improve model performance.

For more insight, there are other ways to classify image data, for example is to use a subset of deep learning, famously known as Convolution Neural Network (CNN). CNN is a deep learning approach that is widely used for solving complex problems (Indolia et al, 2018). CNN algorithm has advantages in image classification tasks. Preliminary study (Luo, 2021) on the deep learning image classification method based on CNN, which can achieve high recognition accuracy. The research for NOAA image classification using CNN can be a recommendation for future works.

SIMPULAN DAN SARAN

The machine learning model used is Google's Teachable Machine, the classification between good and noise data collected from NOAA could work with the the most optimal model is the third experimental model with an overall accuracy of 96.72% and an f1-score of 87.80%. But the model has underfit characteristics.

After getting the performance results, this study tries to predict the image that has never been tested using the optimal Google's Teachable Machine model.

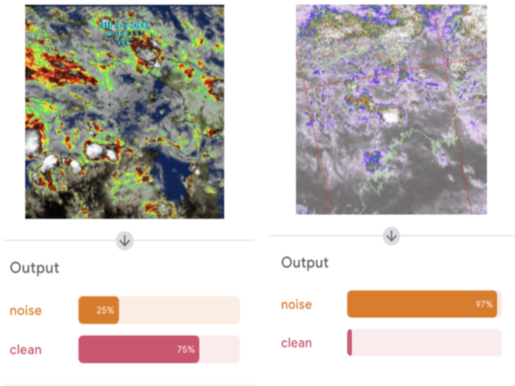

Figure 7. Model image test prediction results

It can be seen in Figure 7, the image checked is accurate even though the model is underfit. The underfit model in this study can occur due to the lack of image data. Image captured by WeSaCom with different sensors. So, it is necessary to balance the data between different sensors.

The downside of this experiment is that we must manually labelled for each good and noise class so it could be used for the training set. This can affect our set to be incorrect as there might be a human-error in labelling them manually. Appropriate and evaluated Google's Teachable Machine Models can be downloaded and implemented to IoT devices and smartphones making it a multi-device system.

The current research results can classify images into clean and noise classes. So that the model can be used for more in-depth research such as weather prediction or natural disasters. In addition, another model is needed in further research to compare the performance of a better evaluation in classifying weather images from NOAA.

UCAPAN TERIMAKASIH

We would like to thanks to Pusat Pengkajian dan Pengembangan Teknologi Siber dan Sandi, Badan Siber dan Sandi Negara for providing us with the information that is obtained from their NOAA Weather Satellite Communication (WeSaCom) receiver.

p-ISSN:2527-5321
e-ISSN:2527-5941

DAFTAR PUSTAKA

Rasyidi, E. S., Sandi, R. and Buraerah, M. F. (2020). Monitoring Perubahan Suhu Ibu Kota Negara Tahun 1993-2019 Menggunakan Citra Satelit Landsat (Studi Kasus: Jakarta, Singapura, Kuala Lumpur Dan Bangkok). *J. Ilm. Ecosyst.*, 50–58.

Akhbar, A. I., Jaya, Y. V. and Febrianto, T. (2018). Kajian Suhu Permukaan Laut Berdasarkan Data Citra Satelit NOAA-AVHRR Dan Data Argo Float Di Perairan Selatan Jawa. *Din. Marit.*, 27–32.

Wicaksono, A., Muhsoni, F. F. and Fahrudin, A. (2010). Aplikasi Data Citra Satelit NOAA-17 Untuk Mengukur Variasi Suhu Permukaan Laut Jawa. *J. Kelaut.*, 70-74, doi: 10.21107/jk.v3i1.849.

Carney, M., Webster, B., Alvarado, I., Phillips, K., Howell, N., Griffith, J., Jongejan, J., Pitaru, A. and Chen, A. (2020). Teachable Machine : Approachable Web-Based Tool for Exploring Machine Learning Classification. *2020 CHI Conference on Human Factors in Computing Systems*, doi: 10.1145/3334480.3382839.

Krisna, M. M., Neelima, M., Harshali, M. and Rao, M. V. G. (2018). Image Classification Using Deep Learning. *International Journal of Engineering & Technology*, 614-617, doi: 10.14419/ijet.v7i2.7.10892.

Toivonen, T., Jormanainen, I., Kahila, J., Tedre, M., Valtonen, T. and Vartiainen, H. (2020). Co-Designing Machine Learning Apps in K-12 with Primary School Children. *2020 IEEE 20th International Conference on Advanced Learning Technologies (ICALT)*, doi: 10.1109/ICALT49669.2020.00099.

Kacorri, H. (2017). Teachable Machines for Accessibility. *ACM SIGACCESS Accessibility and Computing*, doi: 10.1145/3167902.3167904.

Wahyuningsih, D. S., Wulan, T. R., Ibrahim, F. and Maulana, E. (2016). Application of NOAA Imagery to Predict the Fish Distribution Point Case Study: Pelabuhan Perikanan Pantai (PPP) Tegalsari, Tegal, Center of Java. *Seminar Nasional Penginderaan Jauh*.

Natarajan, V. and Philipoff, P. (2018). Observation of Surface and Atmospheric Parameters Using “NOAA 18” Satellite : A Study on Earthquakes of Sumatra and Nicobar Is Regions for the Year 2014. *Nat. Hazards*, doi: 10.1007/s11069-018-3242-y.

Chakraborty, G., Kudoh, J., Shiratori, N. and Noguchi, S. (1993). Classification of the NOAA Satellite Image Data by Unsupervised Neural Network. *Trans. of the Society of Instrument and Control Engineers*, 281-287.

Hidayat, R., Subardjo, P. and Ismanto, A. (2015). Variabilitas Suhu Permukaan Laut Di Pantai Utara Semarang Menggunakan Citra Satelit Aqua Modis. *Journal of Oceanography*, 166-170.

Saputra, M. R. (2019). Analisis Perbandingan Citra Satelit Aqua Modis Dan NOAA AVHRR Untuk Pemetaan Suhu Permukaan Laut Dengan Menggunakan Acuan Data In Situ (Studi Kasus : Perairan Pesisir Selat Madura).

Jeong, H. (2020). Feasibility Study of Google's Teachable Machine in Diagnosis of Tooth-Marked Tongue. *Journal of Dental Hygiene Science. The Korean Society of Dental Hygiene Science*, 206-212, doi: 10.17135/jdhs.2020.20.4.206.

Forchhammer, S., Abu-Ghazaleh, A., Metzler, G., Garbe, C. and Eigentler, T. (2022). Development of an Image Analysis-Based Prognosis Score Using Google's Teachable Machine in Melanoma. *Cancers 2022*, doi: 10.3390/cancers14092243.

Indolia, S., Goswami, A. K., Mishra, S. P. and Asopa, P. (2018). Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach. *Procedia Comput. Sci.*, 679-688, doi: 10.1016/j.procs.2018.05.069.

Luo, L. (2021). Research on Image Classification Algorithm Based on Convolutional Neural Network. *J. Phys.: Conf. Ser.*, doi: 10.1088/1742-6596/2083/3/032054.

Soni, K. B., Chopade, M. and Vaghela, R. (2021). Credit Card Fraud Detection Using Machine Learning Approach. *Appl. Inf. Syst. Manag.*, 71-76, doi: 10.15408/aism.v4i2.20570.