APTISI Transactions on Management (ATM)
Vol. 7, No. 3, 2023, pp. 194~202
E-ISSN: 2622-6804 P-ISSN: 2622-6812, DOI: 10.33050

Enhancing Machine Learning with Low-Cost $PM_{2.5}$ Air Quality Sensor Calibration using Image Processing

Untung Rahardja[1], Qurotul Aini[2], Danny Manongga[3], Irwan Sembiring[4], Yulia Putri Ayu Sanjaya[5]
[1,2]Master of Information Technology, University of Raharja, Indonesia
[3,4]Faculty of Information Technology, Satya Wacana Christian University, Indonesia
[5]Faculty of Management, University of Raharja, Indonesia.
[1]untung@raharja.info, [2]aini@raharja.info, [3]dmanongga@gmail.com, [4]irwan@staff.uksw.edu, [5]yulia.putri@raharja.info

Article Info

Article history:

Received 04-06-2023
Revised 04-28-2023
Accepted 05-05-2023

Keywords:

Machine Learning
Air Quality
Sensor
AIKU

ABSTRACT

Low-cost particulate matter sensors, due to their increased mobility compared to reference monitors, are transforming air quality monitoring. Calibrating these sensors requires training data from reference monitors, which is traditionally done through conventional procedures or by using machine learning techniques. The latter outperforms traditional methods, but still requires deployment of a reference monitor and significant amounts of training data from the target sensor. In this study, we present a cutting-edge machine learning-based transfer learning technique for rapid sensor calibration with Co-deployment with reference monitors is kept to a minimum. This approach integrates data from a small number of sensors, including the target sensor, reducing the dependence on a reference monitor. Our studies reveal that in recent research, a transfer learning method using a meta-agnostic model has been proposed, and the results proved to be much more effective than the previous method. In trials, calibration errors were successfully reduced by up to 32% and 15% compared to the best raw and base- line observations. This shows the great potential of transfer learning methods to increase the effectiveness of learning in the long term. These results highlight the potential of this innovative transfer learning technique for rapidly and ac- curately calibrating low-cost particulate matter sensors using machine learning.

Corresponding Author:

Untung Rahardja
Master of Information Technology, University of Raharja, Indonesia
Email: untung@raharja.info

1. INTRODUCTION

Exposure to fine particulate matter ($PM_{2.5}$) is the primary environmental risk factor for death globally[1]. Administration normally only deploys a few traditional air quality monitoring stations at key sites in a city due to their high setup costs and labor-intensive maintenance requirements. Better resolution data are urgently needed Accurately assessing people's exposure to air pollution can be challenging due to the multitude of factors affecting the levels of air pollutants. In fact, even locations just a few kilometers apart can exhibit substantial variation in air pollution levels. This highlights the complexity involved in measuring and evaluating the extent of air pollution exposure in individuals[2]. Because of their reduced prices, smaller size, and higher mobility, low-cost sensors fortunately provide an economical solution. As sensors for air pollution monitoring-

are typically less precise than their reference grade counterparts, it is necessary to calibrate them in the field to obtain accurate data. This is commonly done by comparing the sensor data with that of a reference monitor deployed in the same location. The result-ing spatiotemporal measurements are then aggregated in a sensor network, providing a more comprehensive understanding of the air pollution levels in a given area. We emphasize the first setting in this message[3]. Cal- ibrating low-cost sensor data requires capturing the complex and nonlinear relationship between the true values obtained from reference monitors and the raw sensor data[4]. To achieve this, various statistical techniques, such as ARIMA, and supervised machine learning techniques, such as linear regression, closest neighbors, and Gaussian process regression, have been extensively researched. These techniques enable the developmentof accurate calibration models that help to improve the reliability of low-cost sensor readings, thereby facili- tating more accurate monitoring of air pollution levels. Convolutional and recurrent neural networks (CNN) and fully connected neural networks (FCNN) have both been very successful neural network learning tech- niques. efficient in attaining the most advanced sensor calibration performance. Deep neural network-based approaches are generally acknowledged to perform better than traditional machine learning approaches in a va-riety of learning challenges involving structured data. But when compared to the latter, the former often needsa lot more training data. Transfer learning, which makes use of information from copious data from related activities (referred to as the source) and improves the model over the target task, is one method for addressingthis data scarcity issue. Model-agnostic meta learning is a cutting-edge transfer learning method (AIKU). In order for the model to swiftly adapt to the target data, AIKU learns the ideal initialization values for the model parameters[5]. These sophisticated neural network learning techniques primarily handle classification issues, but regression presents unique difficulties[6].

This letter discusses resource-constrained calibration methods using neural network learning. In order to prepare, we want to use as little data as possible from a reference monitor that is also present. The calibration method used for a particular sensor. We suggest a brand-new transfer learning technique based on AIKU addressing the current regression issue.

2. METHOD

One low-cost sensor per location is currently deployed across a number of different sites. For a sensor placed at a specific site, the time series of measurements x_p from the inexpensive sensor and the matching readings y_p from the co-deployed reference monitor, respectively, are denoted by ρ,D_p= (x_p,y_p) indicates that the length of both x_p and y_p is the same. We have some target locations and a collection of source sites p_s. A tiny quantity of data is present for each target location, i.e., —D_{pi}| >> |D_{pj}|Ap$_i$ ∈ P^sAp$_j$p^t

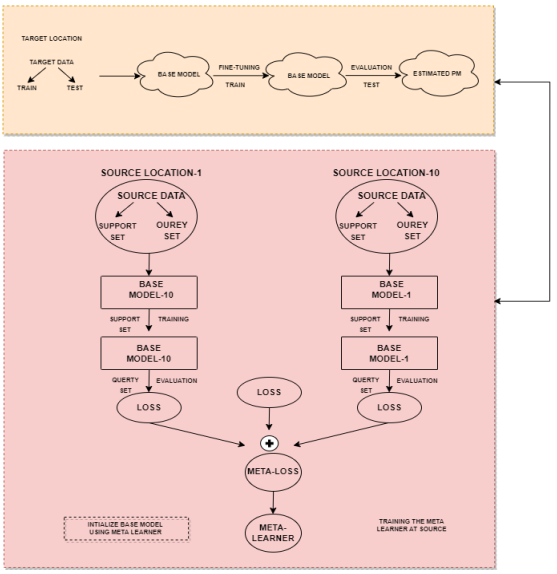

Figure 1. Framework of Brand Image and Green Marketing Towards Purchase Decision

Consider a model with parameters fϑ that maps observations χ to output ˆy, represented by fϑ. To implement fϑ, we employ a neural network. The block diagram of our methodology is shown in Fig 1.

2.1. Using AIKU for Meta Learning

Meta learning is a machine learning technique that involves training a model to adapt to new tasks quickly with minimal data. The goal of meta learning is to develop a more generalizable model that can learn from a diverse set of tasks and apply that knowledge to new tasks[7]. In meta learning, a model is trained across various tasks, each with its unique dataset and loss function, to learn the optimal set of model parameters that can be applied to new tasks. Meta learning involves ”learning to learn,” which means discovering the best way to learn from different tasks and apply that knowledge to new ones. This technique is particularly useful in scenarios where data is scarce, and the model needs to adapt quickly to new situations[8]. By implementing meta learning, we can create more versatile and adaptable algorithms that can be rapidly fine-tuned to suit the needs of different occupations and industries[9].

$$\phi* = argmin_{\phi} E_{D\sim p(D)}[\angle_D(\phi)] \quad (1)$$

The probability distribution of datasets, represented by p(D), plays a critical role in the field of meta- learning. One popular meta-learning approach is the AIKU method, which can be applied to any model that utilizes gradient descent, a powerful optimization-based learning strategy. The main idea behind AIKU-based meta-learning is to efficiently learn the initial parameters of a model so that it can be rapidly adapted for optimal performance on a new task[10]. This approach involves a two-stage optimization process: first, a meta-learner is trained to improve performance across a diverse set of learning tasks, followed by the training of a base model to enhance performance on a specific task. The meta-learner learns to adapt the initial parameters of the base model for each specific task, allowing for efficient and effective adaptation to new environments with limited training data[11]. By rapidly training a base model with minimal data, the meta-learner is then able to identify the most suitable initial parameters for each task-specific model. This allows for the creation of highly versatile and adaptable models that can be efficiently fine-tuned for a wide range of applications[12].

2.2. Model for Sensor Calibration Proposal

This approach allows the base model to be efficiently and effectively trained on the target task, even when faced with limited data. The result is a highly adaptable and versatile model that can be fine-tuned for a wide range of applications, making it a powerful tool for addressing complex and diverse real-world challenges[13]. Additionally, the meta learner is able to learn from the experiences of multiple source locations, resulting in a more robust and generalizable model. This enables efficient and rapid training of the base model on the target task, even with limited data. As a result, the approach yields a highly adaptable and versatile model that can be fine-tuned for a wide range of applications, making it a powerful tool for tackling complex and diverse real-world challenges[14]. The meta learner is global, while the core model is exclusive to the site. A fully connected neural network is used to build the basic model, represented by the function fϑp with parameters ϑp. To train the base model quickly and efficiently at each location ϑp, a meta learner is utilized, implemented as another fully connected neural network. The architectures of fϑ and fϑp are identical. The proposed calibration scheme is illustrated in Figure 1.

We initiate the process by training our meta learners with the source data[15]. During the training phase, we randomly select samples from the query set and support set of the Dp dataset from the source location p. The meta learner is trained using data from multiple source locations, allowing it to learn a more general model[16]. This model is then used to initialize the base model at the target location, which is subsequently fine-tuned using the limited available data. This approach allows for efficient and effective training of the base model on a new task, even when faced with limited data[17]. The resulting model is highly adaptable and versatile, making it a powerful tool for addressing a wide range of real-world challenges. Fig 1 illustrates the utilization of the meta learner fϑ for initializing the base model fϑp[18].

1. Source Training for the Meta Learner:

The meta learner fϑ is generated at random. When a specific source location pi is chosen, the data Dpi is separated into two mutually exclusive sets: a query set DQ pi and a support set DS pi. Query sets are used to evaluate meta learner performance and train them[19]. On the other hand, support pools are used to train the base model. As shown in the fig 1, during training, we select samples from both data sets from the Dp data source at the p source location to train the base model. After that, we use a trained meta learner to train our simple model at a specific target location with minimal training data from the target site[20]. Gradient descent over DSis used to train the base model, which is initially setup as.

$$\vartheta_{pi} \leftarrow \vartheta_{pi} - \alpha \nabla_{\vartheta pi} \angle \mathrm{D}^{s}_{pi}(\vartheta_{pi}) \tag{2}$$

Here, $\alpha \in$R is the learning rate. The mean absolute error, which is used as the loss function, is defined as

$$\measuredangle_{\mathcal{D}}(\vartheta) \sum_{(x,y)\epsilon d} |\mathrm{f}_{\vartheta}(x) - y| \tag{3}$$

After that, the loss LDQ pi (ϑpi) defined in the base model is computed and evaluated on the query set DQ pi [21]. As illustrated in Fig 1, we repeat the entire process for each source site. The meta learner is trained by minimizing the meta loss, which is the sum of all individual losses LDQ pi (ϑpi) determined on the query set. The goal of the meta learner is to identify optimal initial parameters for thebase model, so that with just a few gradient steps at the new location, the performance on the evaluationset is optimized[22].

$$\phi \leftarrow \phi - \beta \sum_{i} \nabla_{\phi} \angle \mathrm{D}^{Q}_{Pi}(\phi) \tag{4}$$

$\beta \in$ R is the learning rate. In the end, only the meta learners will be kept, while the base models from various sources will be discarded[23].

2. In the Desired Location, Train the Basic Model:

We re-initialize the underlying model with ϕ and train ϑpj utilizing to improve prediction for a target lo- cation pj $\in$ Pt with minimum training data[24]. In this scenario, the test set is used to evaluate calibration performance, while the backing set is used as training data (see blue block in Fig 1). In Algorithm 1, the full calibration model training algorithm is described as a pseudo code[25].

3. EXPERIMENTAL RESULTS

3.1. Dataset Description

During our tests, the low-cost sensors are placed adjacent to the reference screens in 15 different places throughout Mumbai. Instead, the low-cost sensors provide four measurements of the reference monitor'ssingle measurement of $PM_{2.5}$, including $PM_{2.5}$, PM_{10}, temperature, and humidity. Nova fitness sensors, which work by scattering laser light, are used in the low-cost sensors to measure $PM_{2.5}$. Throughout the data collecting period, which lasted from November 1, 2022, to March 5, 2023, one sample was obtained every hour. We deploy five different target sites and 10 different source locations in order to conduct our transfer learning experiments. We chose a two day window for the support pool, and another two day window for the query set from the data set at each source location.

The query set takes place later in the timeline, and the twosets are incompatible. For the training and validation sets, we choose a window of three days for each target city, and the next 15 days of data serve as the testing set. We feed the calibration model data from the low-costsensors that monitor temperature, humidity, $PM_{2.5}$, and PM_{10} levels.

Table 1. 15 days of data testing set

Loc	B1	B2	B3	MAML
	MAE (Standard deviation)			
1	36.4 (29.1)	13.7 (15.2)	9.0 (18.2)	8.0 (6.5)
2	21.8 (31.2)	14.5 (27.9)	12.8 (25.3)	12.1 (17.5)
3	18.9 (17.1)	12.8 (14.9)	10.0 (17.2)	8.7 (8.2)
4	21.5 (22.6)	17.1 (19.2)	19.2 (18.3)	15.9 (18.7)
5	18.8 (20.9)	14.6 (19.4)	12.6 (21.0)	10.1 (8.2)
	RMSE			
1	46.6	16.5	11.2	10.3
2	32.0	24.3	21.7	20.2
3	24.7	15.8	14.1	12.0
4	32.3	25.2	26.4	24.5
5	24.1	18.8	16.3	12.9
	R2			
1	-32.4	46.6	75.3	79.9
2	0.0	41.4	53.0	55.7
3	-56.7	46.9	56.3	68.9
4	-90.0	-10.4	-12.1	2.5
5	22.1	51.4	64.5	77.0

3.2. Implementation Details

The proposed technique employs a standard feed-forward neural network consisting of two hidden layers, each containing 128 neurons activated with the ReLU activation function, to model fθ. The output layer comprises only one linear neuron[26]. The meta learner is trained using the Adam optimizer with a learning rate of 104, while the base model is trained using the stochastic gradient descent optimizer with a learning rate of 103. Hyperparameter tuning is performed using the validation set[27].

3.3. Implementation Details

The proposed technique employs a standard feed-forward neural network consisting of two hidden layers, each containing 128 neurons activated with the ReLU activation function, to model fϑ. The output layer comprises only one linear neuron[26]. The meta learner is trained using the Adam optimizer with a learning rate of 104, while the base model is trained using the stochastic gradient descent optimizer with a learning rate of 103. Hyperparameter tuning is performed using the validation set[27].

3.4. Performance Metrics

Once the $PM_{2.5}$ was calibrated, we assessed the performance of the calibration model by comparing its readings with the corresponding measurements obtained from colocated reference monitors[28]. We used three metrics for the comparison: mean absolute error (MAE), root mean square error (RMSE), and coefficient of determination (R2). To further analyze the accuracy of the calibration model, we also calculated the standard deviation of the mean absolute error[29].

3.5. Results

We implemented our strategy and compared calibrated $PM_{2.5}$ readings for five target cities using various methodologies.

1. The baseline B_1 performs the worst, illustrating that neural network learning fails when training instances are scarce.
2. The baseline B_2 outperforms B_1 by a wide margin, demonstrating that sensor calibration is successful when knowledge is transferred across two locations. In our earlier work, we explore the baseline B_2 in great depth.
3. The majority of the time, as evidenced by the better results with B_3, the transfer learning performance increases with the quantity of source data. Location 4's baseline B_2 performs better than B_3 there, nevertheless. The findings effectively show that just combining data from different source locations may worsen performance if a source site can be optimally transferred to a specific destination location.
4. The suggested approach beats every other approach, demonstrating the superiority of AIKU-based transfer learning over fine-tuning-based transfer learning. For an independent two-sample t test, the improvement has a p-value of less than 0.01 and is statistically significant.
5. While the MAE, RMSE, and R2 values of the suggested technique all indicate improvements, In addition, the standard deviation of the absolute error also decreases, indicating that the calibrated $PM_{2.5}$ reading is very close to that of the reference monitor.

Examples of corrected $PM_{2.5}$ the values are compared to raw values from low-cost sensors and readings from reference monitors. As can be observed, the calibrated values closely mirror the reference monitor when compared to the raw sensor data.

4. CONCLUSION

The issue of quick and accurate calibration of inexpensive sensors with the least amount of co-deployment with reference monitors is addressed in this letter. It employs transfer learning for low co-deployment and neural networks for efficiency. An efficient strategy for calibrating inexpensive sensors and a model-independent transfer learning technique based on meta-learning for regression are the two key contri- butions of this letter. It is shown that the suggested technique outperforms alternative learning systems in the transfer learning environment. Both stationary sensors and mobile sensors with moving reference monitors can be calibrated using the described technique. We wish to address issues with maintenance and sensor drift with future extended deployments. Additionally, we intend to expand our model to include calibration for a range of gas sensors, including those for NO_x, O_3, and CO.

ACKNOWLEDGMENT

We thank the Higher Education Excellence Basic Research (PDUPT), which funded this research from the Government Scientific Research Grant Program number 0357/E5/AK.04/2022. And thanks also to the University of Raharja, particularly Alphabet Incubator, who helped complete this research

REFERENCES

[1] Q. Aini, W. Febriani, C. Lukita, S. Kosasi, and U. Rahardja, "New normal regulation with face recognition technology using attendx for student attendance algorithm," in *2022 International Conference on Science and Technology (ICOSTECH)*. IEEE, 2022, pp. 1–7.

[2] Q. Aini, D. Manongga, U. Rahardja, I. Sembiring, V. Elmanda, A. Faturahman, and N. P. L. Santoso, "Security level significance in dapps blockchain-based document authentication," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 4, no. 3, pp. 292–305, 2022.

[3] Q. Aini, D. Manongga, I. Sembiring, D. Apriliasari *et al.*, "Transformation of payment in education use bitcoin with reduced confirmation times," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 5, no. 1, pp. 1–8, 2023.

[4] M. R. Anwar, F. P. Oganda, N. P. L. Santoso, and M. Fabio, "Artificial intelligence that exists in the human mind," *International Transactions on Artificial Intelligence*, vol. 1, no. 1, pp. 28–42, 2022.

[5] P. A. Sunarya, "Machine learning and artificial intelligence as educational games," *International Transactions on Artificial Intelligence*, vol. 1, no. 1, pp. 129–138, 2022.

[6] B. K. Bintaro, P. Sokibi, I. Amsyar, and Y. P. A. Sanjaya, "Utilizing digital marketing as a business strategy: Utilizing digital marketing as a business strategy," *Startupreneur Bisnis Digital (SABDA Journal)*, vol. 1, no. 1, pp. 63–71, 2022.

[7] B. K. Bintoro, N. Lutfiani, D. Julianingsih *et al.*, "Analysis of the effect of service quality on company reputation on purchase decisions for professional recruitment services," *APTISI Transactions on Management (ATM)*, vol. 7, no. 1, pp. 35–41, 2023.

[8] C. S. Bangun, S. Purnama, and A. S. Panjaitan, "Analysis of new business opportunities from online informal education mediamorphosis through digital platforms," *International Transactions on Education Technology*, vol. 1, no. 1, pp. 42–52, 2022.

[9] A. Fernanda, A. R. F. Geovanni, and M. Huda, "Application of artificial intelligence to the development of playing ability in the valorant game," *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, vol. 4, no. 1, pp. 22–31, 2022.

[10] S. Kosasi, U. Rahardja, N. Lutfiani, E. P. Harahap, and S. N. Sari, "Blockchain technology-emerging research themes opportunities in higher education," in *2022 International Conference on Science and Technology (ICOSTECH)*. IEEE, 2022, pp. 1–8.

[11] S. Kosasi, I. D. A. E. Yuliani, U. Rahardja *et al.*, "Boosting e-service quality of online product businesses through it leadership," in *2022 International Conference on Science and Technology (ICOSTECH)*. IEEE, 2022, pp. 1–10.

[12] N. Lutfiani, S. Wijono, U. Rahardja, A. Iriani, Q. Aini, and R. A. D. Septian, "A bibliometric study: Recommendation based on artificial intelligence for ilearning education," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 5, no. 2, pp. 112–119, 2023.

[13] L. Nirmalasari, A. Alwiyah, P. A. Sunarya, and A. S. Panjaitan, "A digital marketing strategy based on applications to rise customer satisfaction," *International Journal of Cyber and IT Service Management*, vol. 2, no. 2, pp. 139–145, 2022.

[14] A. G. Pamungkas, A. Suharko, D. Apriani, E. A. Nabila *et al.*, "Analysis of the effect of quality, service price and satisfaction on patients and their impact on visits to exclusive dental clinics in south jakarta," *APTISI Transactions on Management (ATM)*, vol. 7, no. 1, pp. 9–14, 2023.

[15] Y. Y. R. Rachmawati, Y. P. A. Sanjaya, and S. Edilia, "Web-based temperature, oxygen saturation, and heart rate monitoring system," *IAIC Transactions on Sustainable Digital Innovation (ITSDI)*, vol. 4, no. 1, pp. 38–45, 2022.

[16] A. A. A. R. Pudyanti, A. N. A. Redioka, and V. T. Devana, "Analyses based on theory of capital based approach on indonesian graduate employability," *ADI Journal on Recent Innovation*, vol. 4, no. 1, pp. 25–33, 2022.

[17] U. Rahardja, "Camera trap approaches using artificial intelligence and citizen science," *International Transactions on Artificial Intelligence*, vol. 1, no. 1, pp. 71–83, 2022.

[18] U. Rahardja, Q. Aini, P. A. Sunarya, D. Manongga, and D. Julianingsih, "The use of tensorflow in analyzing air quality artificial intelligence predictions pm2. 5," *Aptisi Transactions on Technopreneurship (ATT)*, vol. 4, no. 3, pp. 313–324, 2022.

[19] U. Rahardja, "The economic impact of cryptocurrencies in indonesia," *ADI Journal on Recent Innovation*, vol. 4, no. 2, pp. 194–200, 2023.

[20] U. Rahardja, Q. Aini, D. Manongga, I. Sembiring, and I. D. Girinzio, "Implementation of tensor flowin air quality monitoring based on artificial intelligence," *International Journal of Artificial Intelligence Research*, vol. 6, no. 1, 2023.

[21] B. Rawat, A. S. Bist, U. Rahardja, Q. Aini, and Y. P. A. Sanjaya, "Recent deep learning based nlp tech- niques for chatbot development: An exhaustive survey," in *2022 10th International Conference on Cyber and IT Service Management (CITSM)*. IEEE, 2022, pp. 1–4.

[22] D. Rustiana, J. D. Pratama, T. Mudabbir, M. A. Fahmi, and G. A. Rofei, "Adoption computerized certifi- cate transparency and confidentiality," *Int. J. Cyber IT Serv. Manag*, vol. 2, no. 1, pp. 1–10, 2022.

[23] Y. P. A. Sanjaya and M. A. Akhyar, "Blockchain and smart contract applications can be a support for msme supply chain finance based on sharia crowdfunding," *Blockchain Frontier Technology*, vol. 2, no. 1, pp. 44–49, 2022.

[24] S. Santoso, P. Harsani, E. P. Harahap, P. A. Sunarya, and Y. P. A. Sanjaya, "Enrichment program using promethee for decision support systems of prospective assistance funds disabilities," in *2022 1st Interna- tional Conference on Technology Innovation and Its Applications (ICTIIA)*. IEEE, 2022, pp. 1–5.

[25] F. Septiyana, M. S. Shihab, H. Kusumah, D. Apriliasari *et al.*, "Analysis of the effect of product quality, price perception and social value on purchase decisions for lampung tapis fabrics," *APTISI Transactions on Management (ATM)*, vol. 7, no. 1, pp. 54–59, 2023.

[26] Q. Aini, U. Rahardja, and T. Hariguna, "The antecedent of perceived value to determine of student con- tinuance intention and student participate adoption of ilearning," *Procedia Computer Science*, vol. 161, pp. 242–249, 2019.

[27] U. Rahardja, Q. Aini, M. A. Ngadi, M. Hardini, and F. P. Oganda, "The blockchain manifesto," in *2020 2nd International Conference on Cybernetics and Intelligent System (ICORIS)*. IEEE, 2020, pp. 1–5.

[28] U. Rahardja, C. T. Sigalingging, P. O. H. Putra, A. Nizar Hidayanto, and K. Phusavat, "The impact of mobile payment application design and performance attributes on consumer emotions and continuance intention," *SAGE Open*, vol. 13, no. 1, p. 21582440231151919, 2023.

[29] C. Lukita, M. Hatta, E. Harahap, and U. Rahardja, "Crowd funding management platform based on block chain technology using smart contracts," *J. Adv. Res. Dyn. Control Syst*, vol. 12, no. 2, pp. 1928–1933,2020