การระบุประเภทเหรียญกษาปณ์สำหรับระบบรู้จำเหรียญกษาปณ์ โดยใช้เครือข่ายใยประสาทคอนโวลูชันและการเรียนรู้แบบถ่ายโอน

เติมยศ เสนีวงศ์ ณ อยุธยา*, สิทธิพงศ์ พรอุดมทรัพย์

สาขาวิชาคอมพิวเตอร์แอนิเมชันและมัลติมีเดีย คณะวิทยาศาสตร์และเทคโนโลยี
มหาวิทยาลัยราชภัฏพระนคร กรุงเทพมหานคร
*Corresponding author email: tomeyot@pnru.ac.th

ได้รับบทความ: 7 สิงหาคม 2567
ได้รับบทความแก้ไข: 21 สิงหาคม 2567
ยอมรับตีพิมพ์: 2 กันยายน 2567

บทคัดย่อ

การเก็บสะสมเหรียญเป็นงานอดิเรกที่สามารถดึงดูดผู้คนด้วยเหตุผลหลายประการ ตั้งแต่ความอยากรู้อยากเห็นทางประวัติศาสตร์ โอกาสในการลงทุน ไปจนถึงความเพลิดเพลินและความพึงพอใจส่วนตัว นอกจากนี้ข้อมูลเกี่ยวกับเหรียญประเภทต่างๆ ก็มีความสำคัญและจำเป็นสำหรับนักสะสมเหรียญกษาปณ์รุ่นใหม่ ดังนั้นงานวิจัยนี้จึงนำเสนอการบูรณาการเทคนิคการเรียนรู้เชิงลึกผสมผสานกับการเรียนรู้แบบถ่ายโอนเพื่อการรู้จำเหรียญกษาปณ์ไทย โดยการใช้ประโยชน์จากอัลกอริทึมการเรียนรู้เชิงลึกและการเรียนรู้แบบถ่ายโอนจึงทำให้งานวิจัยนี้จึงสามารถนำไปประยุกต์ใช้ในการพัฒนาระบบค้นหาข้อมูลเหรียญกษาปณ์จากภาพเหรียญสำหรับสกุลเงินไทย โดยประสิทธิภาพของวิธีการที่นำเสนอได้รับการประเมินและเปรียบเทียบประสิทธิภาพกับโมเดลการเรียนรู้แบบถ่ายโอนแบบอื่นๆ และจากผลการทดลองแสดงให้เห็นว่าแบบจำลองฝึกล่วงหน้าที่ใช้วิธีการ MobileNet-v2 สามารถให้ค่า macro average f1-score สูงที่สุดที่ 100% ซึ่งสามารถบ่งบอกได้ว่าวิธีการที่นี้เหมาะสมกับระบบการรู้จำเหรียญกษาปณ์ของไทย

คำสำคัญ: การรู้จำเหรียญ/ การเรียนรู้เชิงลึก/ การเรียนรู้ของเครื่อง/ การเรียนรู้แบบถ่ายโอน

Coin Identification for Coin Recognition Systems Based on Convolutional Neural Network and Transfer Learning

Tomeyot Sanevong Na Ayutaya*, Sittiphong Pornudomthap

[1]Computer Animation and Multimedia Program, Faculty of Science and Technology, Phranakhon Rajabhat University, Bangkok
*Corresponding author email: tomeyot@pnru.ac.th

Received: 7 August 2024
Revised: 21 August 2024
Accepted: 2 September 2024

Abstract

Coin collecting is a multifaceted hobby that appeals to individuals for various reasons, ranging from historical curiosity to investment opportunities to personal enjoyment and satisfaction. Moreover, information about different types of coins is important and necessary for the new generation of coin collectors. This research proposes the integration of deep learning technique incorporating with transfer learning for Thai coin recognition. By leveraging deep learning algorithms and transfer learning, this study can be applied to develop the coin information searching systems from coin images for Thai currency. The performance of the proposed method is evaluated and compared to other transfer learning models. The experimental results signify that the pre-trained MobileNet-v2 method can achieve the highest macro average f1-score of 100%, indicating that this approach is highly suitable for the Thai coin recognition system.

Keywords: Coin Recognition/ Deep Learning/ Machine Learning/ Transfer Learning

Introduction

The Coins have been a fundamental part of human civilization, serving as a medium of exchange, a representation of cultural heritage, and a symbol of national identity. Coin recognition is the process of identifying and categorizing coins based on their specific features. This recognition can be done manually by individuals or automatically using computer vision techniques. In manual coin recognition, individuals visually inspect coins to determine their denomination and value. In addition, the manual recognition relies on the knowledge of different coin types and their physical characteristics. On the other hand, automated coin recognition uses computer vision techniques to identify and classify coins accurately and automatically.

QuanGen Li proposed the coin image recognition based on k-nearest neighbor algorithm [1]. Initially, clear coin images are gathered and processed using the HSV color space model to calculate their edge perimeters. This yields both color and perimeter feature data for each coin. Subsequently, the K-nearest neighbor algorithm is employed to classify these features and train the classifier model. The effectiveness of the classifier is then tested. The system effectively identifies coin denominations while simultaneously decreasing reliance on expensive hardware and reducing costs. Wenze He proposed a novel rotation-invariant CNN called RTN, and applied it to U.S. coin dataset [2]. Comparing to other rotation-invariant methods like SIFT-based approaches and STN, the proposed RTN method demonstrated superior performance in recognition accuracy and achieved state-of-the-art results. Nur Nadirah Zainuddin proposed Malaysian coins recognition using machine learning methods [3]. This research studied many machine learning approaches and applied them to Malaysian coin recognition. Two feature extraction methods were applied to reduce the dimensions by extracting the most meaningful features including the gray-level Co-occurrence matrix (GLCM) and the histogram of oriented gradients (HOG). Six classifiers were trained including Artificial Neural Network (ANN), Linear Discriminant Analysis (LDA), Support Vector Machine (SVM), Adaptive Boosting (AdaBoost), K-Nearest Neighbors (KNN), and Complex Tree (CTREE). For the Grey-Level Co-occurrence

Matrix (GLCM) feature extraction, AdaBoost classifiers were the most accurate, whereas K-Nearest Neighbors (KNN) classifiers were the least accurate. Moreover, the Artificial Neural Networks (ANN) classifier had the highest accuracy in the Histogram of Oriented Gradients (HOG) feature, while the Linear Discriminant Analysis (LDA) classifier had the lowest. A.U. Tajane proposed Deep Learning approach for recognition and detection of Indian coin [4]. The pretrained convolutional neural network, specifically AlexNet, was trained by utilizing features such as textures, colors, and shapes. With a training dataset comprising over 1,600 images, the model was capable of classifying images into four distinct object categories: one, two, five, and ten rupee coins. The trained model was tested on various standard and own recorded datasets consist of rotational, translated and shifted images. The results obtained demonstrate the superior performance of the proposed methodology compared to conventional systems. Debabrata Swain proposed an innovative deep-learning framework designed specifically for the classification of Brazilian coins [5]. This research presented a Repetitive Feature Extractor Convolutional Neural Network (RFE-CNN) model, which significantly enhances the speed and accuracy of currency recognition. The proposed deep learning framework provides a robust and effective solution for the classification of Brazilian coins.

This research aims to explore and adopt a CNN-based approach incorporating with transfer learning for Thai coin recognition. By harnessing the capabilities of deep learning and its capacity to extract intricate features from images, the objective is to create a deep learning model that can precisely identify and classify various Thai coins. The subsequent sections will delve deeper into the methodology employed, experiment details, results obtained, and discussions, culminating in conclusions drawn from the findings and suggestions for future enhancements in coin recognition.

Materials and methods

Data Collection

The dataset comprises images of various coins collected from images of Thai coins with 4 different currency values including 1 Baht, 2 Baht, 5 Baht, and 10 Baht coins. We captured two images of each coin, one of the obverse and one of the reverse. With a total of 400 coins, this resulted in 800 images. The researcher meticulously controlled the lighting by using lamps to illuminate the coins from both the left and right sides and positioned the mobile camera directly above the coins. The image size for each coin was 3024 x 4032 pixels. The dataset is divided into the training set, the validation set, and the test set with ratio of 50:20:30 respectively. Since there are totally 800 images in the dataset, the number of coin images in the training set is 400, the number of coin images in the validation set is 160 and the number of coin images in the test set is 240. The number of training images, validation images, and test images in each coin type are illustrated in Table 1.

Table 1. The dataset of coins

Label	Train	Validation	Test	Total
1 Baht	100	40	60	200
2 Baht	100	40	60	200
5 Baht	100	40	60	200
10 Baht	100	40	60	200
Total	**400**	**160**	**240**	**800**

Convolutional Neural Network (CNN)

A Convolutional Neural Network (CNN) is a type of deep learning models used in tasks involving visual imagery such as image classification, object detection, and image segmentation. CNN architectures often include two main parts which are feature learning (or feature extraction) and classification. There are three basic layers used to build CNN architectures including the convolutional layers, pooling layers, and fully connected layers.

The Convolutional layers play a crucial role in extracting meaningful features from the input images using convolution operations that apply filters to input images to detect various features and resulting the output in from of feature maps. Then, the activation function such as ReLU (Rectified Linear Unit) function is applied to these feature maps to give the final outputs of convolution layers. The Pooling layers play a crucial role in reducing the spatial dimensions of the feature maps produced by convolutional layers while retaining the most important information. This reduction helps in controlling the number of parameters and computational complexity of the network, as well as making the learned features more invariant to small translations and distortions in the input images. The fully connected layers are used to perform classification or regression tasks based on the extracted features.

Transfer Learning

Transfer learning is a powerful technique in deep learning, particularly in the domain of image classification, that leverages pre-trained network models to solve new tasks or datasets with limited labeled data [6, 7, 8, 9, 10, 11, 12]. It involves using knowledge gained from training a model on a large and diverse dataset and transferring that knowledge to a new, possibly smaller, dataset for a different but related task.

In this research, the CNN model with transfer learning from MobileNet-v2 is proposed. The MobileNet-v2 model is the convolutional neural network design crafted for optimal performance on mobile devices. It is based on an inverted residual structure where the residual connections are between the bottleneck layers. The intermediate expansion layer utilizes lightweight depthwise convolutions to enhance feature filtering, providing non-linearity. The full architecture of MobileNetV2 comprises the initial fully convolutional layer with 32 filters, succeeded by 19 residual bottleneck layers. The performance of the proposed method is compared to the CNN model with pre-trained EfficientNet-v2, pre-trained ResNet-v2, and pre-trained Inception-v2.

One of the primary advantages of transfer learning in image classification is its ability to expedite model training and improve convergence. By initializing the model with pre-trained weights, the model starts with a more optimal set of parameters, thereby requiring fewer epochs or iterations to achieve good performance on the new dataset. This advantage becomes especially relevant when working with smaller datasets where training from scratch might lead to overfitting due to limited data.

Evaluation Metrics

The Model evaluation in image classification is a critical step in assessing the performance and efficacy of machine learning or deep learning models before their deployment or application to real-world scenarios. It involves assessing various metrics and techniques to measure the performance of the models. The test set is used to evaluate the performance of the model after completing the training. The confusion matrices are also commonly used in model evaluation for image classification. These matrices illustrate the true positive, true negative, false positive, and false negative predictions made by the model across different classes, providing insights into the specific classification errors and performance on individual classes. The true positive (TP) is the number of positive examples correctly predicted by the classification model. The false negative (FN) is the number of positive examples wrongly predicted as negative by the classification model. The false positive (FP) is the number of negative examples wrongly predicted as positive by the classification model. The true negative (TN) is the number of negative examples correctly predicted as negative by the classification model.

Visualization of these metrics aid in understanding where the model excels and where it struggles, thereby guiding improvements or fine-tuning efforts.

In this research, the confusion matrix is used to calculate the performance of the models using several evaluation metrices including accuracy, precision, recall, and F1-score. The accuracy measures the percentage of correctly classified images from the total number of images in the test dataset. The accuracy metric can be calculated using the following formula.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

The precision measures the ratio of correctly predicted positive observations to the total predicted positive observations. The precision metric can be calculated using the following formula.

$$Precision = \frac{TP}{TP + FP}$$

The recall measures the ratio of correctly predicted positive observations to the actual positives in the dataset. The precision metric can be calculated using the following formula.

$$Recall = \frac{TP}{TP + FN}$$

The F1-score is the harmonic mean of precision and recall that can be calculated using the following formula.

$$F = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

The macro average precision, macro average recall, and macro average F1-score are the metrices used to evaluate the performance of a classification model, particularly in scenarios where there are multiple classes involved. The macro average precision is computed by taking the average of the precision scores calculated for each class separately. The macro average recall is computed by taking the average of the recall scores calculated for each class separately. Similarly, the macro average F1-score is computed by taking the average of the F1-scores calculated for each class separately.

Results

In this research, all pre-trained CNN models are imported from Kaggle and their classification heads are replaced by the new classification heads. The new classification heads of all CNN pre-trained models are fully connected layers. The fully connected layer consists of a hidden layer with 128 nodes and an output layer with 4 nodes. All CNN models with pre-trained CNN networks are trained on the training set and evaluated on the test set. For each pre-trained model, the experiments are conducted five times and then the best result is selected. The predicted results from the best model are used to compute the confusion matrix, used to evaluate the performances of all CNN models based on evaluation matrices including precision, recall, F1-score, and their macro average. The process mentioned above is shown in Figure1.

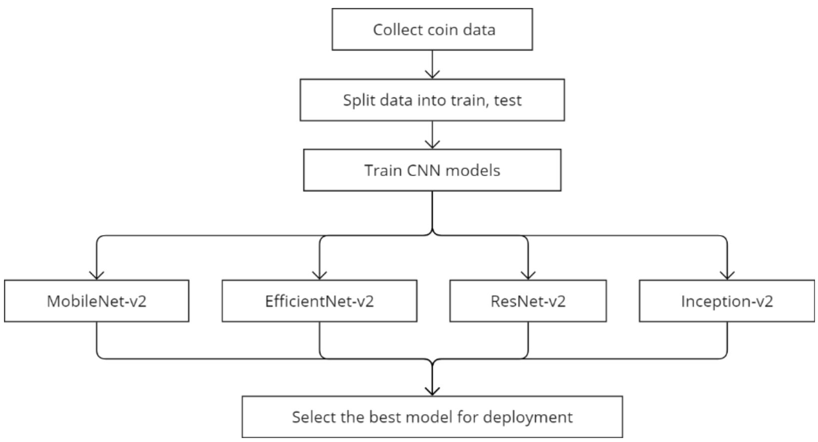

Figure 1. The workflow diagram of training and testing models.

The CNN model with pre-trained MobileNet-v2 is trained on the training set and tested on the test set. Then, the confusion matrix is computed as illustrated in Figure 2.

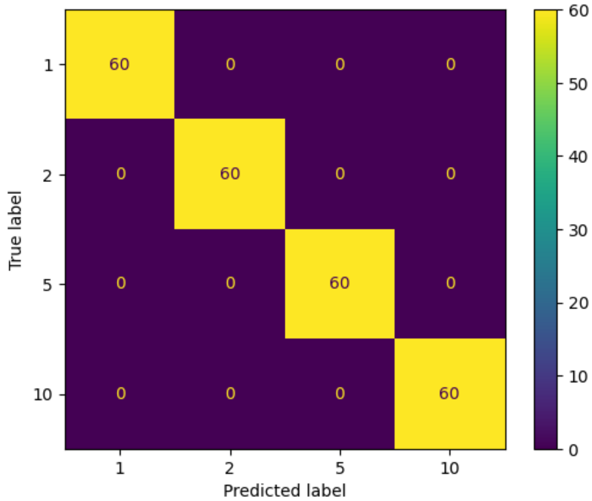

Figure 2. The confusion matrix of the CNN model with pre-trained MobileNet-v2.

The confusion matrix of the CNN model with pre-trained MobileNet-v2 is used to calculate precision, recall, F1-score, their macro average for each individual coin class as shown in Table 2.

Table 2. The performance of the CNN model with pre-trained MobileNet-v2.

Coin Type	Precision	Recall	F1-score
1 Baht	1.00	1.00	1.00
2 Baht	1.00	1.00	1.00
5 Baht	1.00	1.00	1.00
10 Baht	1.00	1.00	1.00
macro average	1.00	1.00	1.00

From Table 2., it can be seen that the CNN model with transfer learning from pre-trained MobileNet-v2 model can achieve the macro average precision of 100%, the macro average recall of 100%, and the macro average F1-score of 100%.

The CNN model with pre-trained EfficientNet-v2 is trained on the training set and tested on the test set. Then, the confusion matrix is computed as illustrated in Figure 3.

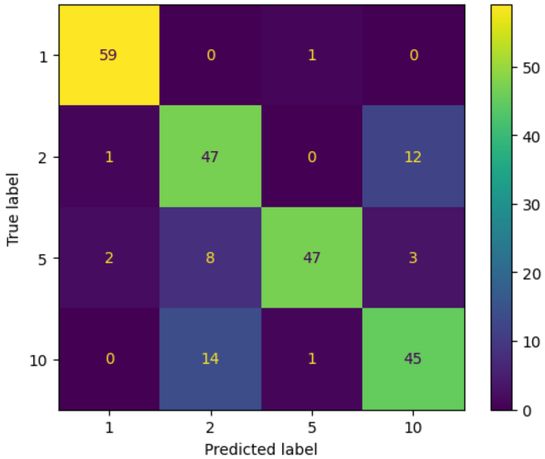

Figure 3. The confusion matrix of the CNN model with pre-trained EfficientNet-v2

The confusion matrix of the CNN model with pre-trained EfficientNet-v2 is used to calculate precision, recall, F1-score, their macro average for each individual coin class as shown in Table 3.

Table 3. The performance of the CNN model with pre-trained EfficientNet-v2.

Coin Type	Precision	Recall	F1-score
1 Baht	0.95	0.98	0.97
2 Baht	0.68	0.78	0.73
5 Baht	0.96	0.78	0.86
10 Baht	0.75	0.75	0.75
macro average	0.84	0.82	0.83

From Table 3., it can be seen that the CNN model with transfer learning from pre-trained EfficientNet-v2 model can achieve the macro average precision of 84%, the macro average recall of 82%, and the macro average F1-score of 83%.

The CNN model with pre-trained ResNet-v2 is trained on the training set and tested on the test set. Then, the confusion matrix is computed as illustrated in Figure 4.

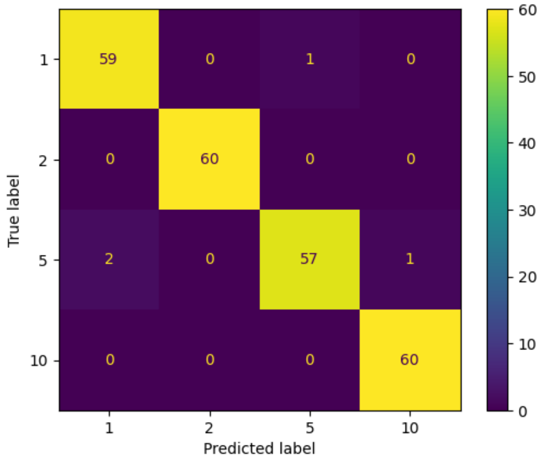

Figure 4. The confusion matrix of the CNN model with pre-trained ResNet-v2.

The confusion matrix of the CNN model with pre-trained ResNet-v2 is used to calculate precision, recall, F1-score, their macro average for each individual coin class as shown in Table 4.

Table 4. The performance of the CNN model with pre-trained ResNet-v2.

Coin Type	Precision	Recall	F1-score
1 Baht	0.97	0.98	0.98
2 Baht	1.00	1.00	1.00
5 Baht	0.98	0.95	0.97
10 Baht	0.98	1.00	0.99
macro average	0.98	0.98	0.98

From Table 4., it can be seen that the CNN model with transfer learning from pre-trained ResNet-v2 model can achieve the macro average precision of 98%, the macro average recall of 98%, and the macro average F1-score of 98%.

The CNN model with pre-trained Inception-v2 is trained on the training set and tested on the test set. Then, the confusion matrix is computed as illustrated in Figure 5.

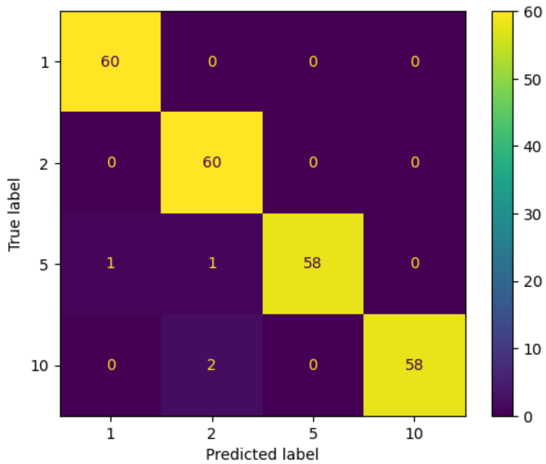

Figure 5. The confusion matrix of the CNN model with pre-trained Inception-v2.

The confusion matrix of the CNN model with pre-trained Inception-v2 is used to calculate precision, recall, F1-score, their macro average for each individual coin class as shown in Table 5.

Table 5. The performance of the CNN model with pre-trained Inception-v2.

Coin Type	Precision	Recall	F1-score
1 Baht	0.98	1.00	0.99
2 Baht	0.95	1.00	0.98
5 Baht	1.00	0.97	0.98
10 Baht	1.00	0.97	0.98
macro average	0.98	0.98	0.98

From Table 5., it can be seen that the CNN model with transfer learning from pre-trained Inception-v2 model can achieve the macro average precision of 98%, the macro average recall of 98%, and the macro average F1-score of 98%.

From all experimental results, it can signify that CNN model with transfer learning from pre-trained MobileNet-v2 demonstrates robustness to variations in coin orientation. Furthermore, the CNN model with pre-trained MobileNet-v2 can achieve the highest precision, recall, F1-score, their macro average comparing to others. The highest performance achieved by the CNN model with pre-trained MobileNet-v2 significant promises for real-world applications, such as Thai coin recognition systems.

Discussion

Numerous studies [1, 2, 3, 4, 5] have utilized deep learning models to recognize coins within their respective countries, aiming to develop systems [13] for identifying various coin types. The materials and images used for coins differ across countries, resulting in distinctive coin characteristics. Consequently, models trained on one country's dataset are ineffective for predicting coin types from other countries. This research focuses on applying deep learning principles to create a model for identifying coin types in Thailand. In this study, the researcher aims to select the best result of each model instead of using the average to identify the model with the highest performance. Averaging would obscure the maximum potential of each model. Furthermore, this research seeks to compare the best outcomes among models to select the most suitable one for practical application. Since models in real world applications are often fine-tuned for optimal performance, selecting the best result is more appropriate in this context. Experimental results demonstrate that the CNN model with pre-trained MobileNet-v2 surpasses all other compared models, making it the most suitable choice for predicting coin types in Thailand.

Conclusion

This research concentrates on revolutionizing coin recognition in Thai coin recognition systems through the application of deep learning and transfer learning. The proposed method adopts the convolutional neural network with the transfer learning for Thai coin recognition. This can be achieved by adopting the pre-trained MobileNet-v2 model whose classification head is replaced by the new fully connected layer consisting of a hidden layer and an output layer. This proposed model is trained and tested on Thai coin images. In addition, the performance of the proposed method is evaluated and compared to other CNN models with pre-trained EfficientNet-v2, pre-trained ResNet-v2, and pre-trained Inception-v2 models. All CNN models with pre-trained models are trained and tested on the same dataset. The experimental results illustrate that the proposed method can overcome all comparative models in term of precision, recall, and f1-score.

Acknowledgments

The research ethics in human of this research have been approved by the academic and ethics committees of the Institute of Research and Development at Phranakhon Rajabhat University. The research project number is 03.007/67.

References

1. Li Q, Ji L, Wu J. Coin image recognition based on K-nearest neighbor algorithm. In: ECITech 2022; The 2022 International Conference on Electrical, Control and Information Technology; 2022; Kunming, China.
2. He W, Wu H. Ancient and modern coin recognition using a novel rotation-invariant convolutional neural network. In: 2021 6th International Symposium on Computer and Information Processing Technology (ISCIPT); 2021; Changsha, China. p. 395-8.
3. Zainuddin NN, Azhari MSNBN, Hashim W, Alkahtani AA, Mustafa AS, Alkawsi G, et al. Malaysian coins recognition using machine learning methods. In: 2021 2nd International Conference on Artificial Intelligence and Data Sciences (AiDAS); 2021; Ipoh, Malaysia. p. 1-5.
4. Tajane AU, Patil JM, Shahane AS, Dhulekar PA, Gandhe ST, Phade GM. Deep learning based Indian currency coin recognition. In: 2018 International Conference on Advances in Communication and Computing Technology (ICACCT); 2018; Sangamner, India. p. 130-4.

5. Swain D, Rupapara V, Nour A, Satapathy S, Acharya B, Mishra S, et al. A deep learning framework for the classification of Brazilian coins. IEEE Access. 2023;11:109448-61.
6. Zhou J, Zhuang J, Li B, Zhou L. Research on underwater image recognition based on transfer learning. In: OCEANS 2022; 2022; Hampton Roads, VA, USA. p. 1-7.
7. Kondo K, Hasegawa T. Deep transfer learning using class augmentation for sensor-based human activity recognition. IEEE Sens Lett. 2022 Oct;6(10):1-4. Art no. 6003604.
8. Tai TT, Thanh DNH, Hung NQ. A dish recognition framework using transfer learning. IEEE Access. 2022;10:7793-9.
9. Goel P, Ganatra A. Handwritten Gujarati numerals classification based on deep convolution neural networks using transfer learning scenarios. IEEE Access. 2023;11:20202-15.
10. Li J, Luo X, Ma H, Zhao W. A hybrid deep transfer learning model with kernel metric for COVID-19 pneumonia classification using chest CT images. IEEE/ACM Trans Comput Biol Bioinform. 2023 Jul-Aug;20(4):2506-17.
11. Huang X. High resolution remote sensing image classification based on deep transfer learning and multi feature network. IEEE Access. 2023;11:110075-85.
12. Raza A, Munir K, Almutairi MS, Sehar R. Novel transfer learning based deep features for diagnosis of Down syndrome in children using facial images. IEEE Access. 2024;12:16386-96.
13. Capece N, Erra U, Ciliberto AV. Implementation of a coin recognition system for mobile devices with deep learning. In: 2016 12th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS); 2016; Naples, Italy. p. 186-92.