Bulletin of Electrical Engineering and Informatics
Vol. 13, No. 5, October 2024, pp. 3552~3567
ISSN: 2302-9285, DOI: 10.11591/eei.v13i5.8004

Comparative analysis of machine learning approaches in Kazakh banknote classification

Ualikhan Sadyk, Makhambet Yerzhan, Rashid Baimukashev, Cemil Turan
Department of Computer Science, Faculty of Engineering and Natural Science, SDU University, Kaskelen, Kazakhstan

Article Info

Article history:

Received Dec 9, 2023
Revised Mar 14, 2024
Accepted Mar 28, 2024

Keywords:

Banknote classification
Hybrid approach
Kazakh banknotes
Machine learning
Model stacking
Traditional approach
Visual impairments

ABSTRACT

Nowadays, smartphones seamlessly blend into every aspect of our lives, including as handheld assistants for individuals with disabilities. Therefore, this research addresses the need for a robust system that can classify Kazakh banknotes. By capitalizing on the availability of smartphones and the ability to integrate detectors with classifiers this study introduces classifiers of Kazakh banknote images specifically designed for banknotes ranging from 500 KZT to 20,000 KZT. It compares traditional and hybrid machine learning (ML) approaches, utilizing a dataset of diverse banknote images, aiming for both lightweight and high accuracy. Competitive performance is demonstrated by the traditional approach, enhanced by thoughtful feature engineering. The hybrid approach, utilizing features from a pre-trained ResNet-18 model, showcases remarkable accuracy and robustness. Evaluation metrics reveal significant achievements, with the traditional approach attaining 94.00% accuracy and the hybrid approach excelling at 99.11%. Model stacking, combining classifiers from both approaches, outperforms individual classifiers, achieving 95.00% and 99.55% accuracy for the traditional and hybrid ML approaches, respectively. Our methodology's comparable outcome in classifying Thai banknotes and coffee beans roasting levels demonstrates their versatility in image classification tasks that rely on color differentiation, showcasing the potential beyond banknote recognition.

Corresponding Author:

Ualikhan Sadyk
Department of Computer Science, Faculty of Engineering and Natural Science
SDU University
Abylaikhan St. 1/1, Kaskelen, Kazakhstan
Email: ualikhan.sadyk@sdu.edu.kz

1. INTRODUCTION

In an era characterized by technological advancements and a heightened emphasis on inclusivity, the creation of efficient and accurate systems to assist individuals with visual impairments is deemed crucial. With the prevalence of smartphones equipped with high-quality cameras in today's society being recognized, a mission is undertaken to address the specific need of classifying Kazakh banknotes [1]. Leveraging the ubiquity of smartphones, which serve as indispensable tools in daily lives, the aim of this study is to pioneer a classifier of Kazakh banknote images.

The primary objective is twofold: the creation of a lightweight and highly accurate Kazakh banknote classifier, and the comparison of the efficacy of traditional and hybrid machine learning (ML) approaches in achieving this goal. Traditional ML approaches, rooted in manual feature extraction and well-established ML techniques, offer interpretability and control [2]. In contrast, hybrid ML approaches, integrating automated feature extraction from pre-trained deep learning models, bring forth the potential to capture intricate patterns and nuances within banknote images [3], [4].

The employed dataset, comprising 4,200 high-quality images of Kazakhstani banknotes under diverse conditions–varying lighting environments, cluttered backgrounds, and folded banknotes–mimics the challenges encountered in everyday life [5]. By narrowing the scope to the denominations commonly utilized, ranging from 500 KZT to 20000 KZT, the aim is to create a classifier that aligns with practical, real-world scenarios.

The potential impact of this research extends beyond technical innovation. By empowering individuals with visual impairments to independently recognize and distinguish Kazakh banknotes through a simple smartphone application, a contribution is made to their financial autonomy. To accomplish this goal, the classifiers developed by this study can be integrated in combination with existing object detection methods in future research [6]–[10]. This aligns with the global efforts towards inclusive technology, recognizing the role of advancements in improving the quality of life for diverse communities.

The subsequent sections of this paper will delve into the methodology employed, datasets utilized, and the detailed processes of traditional and hybrid ML approaches. The evaluation metrics, classifiers, optimization strategies, and model stacking techniques applied will be presented. The potential transferability of the proposed approaches to other currencies and image classification tasks beyond banknote recognition will be discussed. Results and discussions will offer insights into the performance of classifiers, and limitations will be transparently acknowledged. The paper will conclude with a summary of findings, emphasizing the potential societal impact of a lightweight, highly accurate Kazakh banknote classifier.

In today’s world, where technology is advancing rapidly and there is a growing emphasis, on inclusivity it becomes crucial to develop accurate systems that can assist people with impairments. We understand that smartphones with high-quality cameras have become ubiquitous in our society. With this in mind, our goal is to address the need for classifying Kazakh banknotes. By taking advantage of the use of smartphones as tools in our daily lives we aim to lead the way in developing a Kazakh banknote classifier that utilizes camera technology. In this review of existing literature, we explore studies that have examined the application of ML algorithms for banknote classification.

The related works in the field of banknote recognition have witnessed a continual evolution, with researchers exploring various techniques to enhance accuracy in automated systems. A significant contribution in this domain is the work by Hussein *et al.* [11] introducing a method for banknote recognition utilizing a combination of histogram of oriented gradients (HOG) and scale-invariant feature transform (SIFT). The authors addressed the challenges faced by the visually impaired in different countries and proposed a novel banknote recognition system. In their research paper authors used a database containing 12 denominations from three different countries Egypt, Saudi Arabia, and the United States of America. The study employs the speeded-up robust features (SURF) algorithm for keypoint detection, leveraging its speed and robustness to geometry and lighting variations. The subsequent extraction of features through the SIFT and HOG algorithms, both scale-invariant, aims to overcome limitations and reduce computational costs. The research achieves an impressive accuracy of 99.2% in banknote recognition, surpassing the performance of previous methods. The comparison with the Kaggle Egyptian dataset further validates the efficacy of the proposed approach, yielding an accuracy of 98.9%.

Building on this progress, Youn *et al.* [12] have conducted noteworthy research in the domain of efficient multi-currency classification of contact image sensor (CIS) banknotes, establishing a relevant benchmark. Their approach introduces a rapid and effective algorithm that leverages size information and multi-template correlation matching recognizing the distinctive characteristics of different banknote sizes, the researchers generated a size map for grouping banknotes. The algorithm then identified discriminant areas for each banknote, demonstrating high correlations within the same denomination and low correlations with others. To address potential degradations such as writing or aging, post-processing techniques were applied. The evaluation, encompassing 55 banknotes from five countries (Korean Won (KRW), US Dollar (USD), Euro (EUR), Tiongkok Yuan (CNY), and Jordanian Dinar (JOD)), showcased an impressive 100% classification accuracy for unsoiled banknotes and a notable 99.8% accuracy for soiled ones. The algorithm’s efficiency is further highlighted by an average processing time of approximately 4.83 ms per banknote. This research provides valuable insights into methodologies that balance speed, accuracy, and adaptability to real-world conditions.

Another notable work is done by Park and Park [13]. The authors address the growing need for a comprehensive solution to assist visually impaired individuals in recognizing banknotes of various nationalities. The multinational banknote detecting model (MBDM) introduces a novel approach by proposing a multinational banknote detection model capable of identifying currencies such as the KRW, USD, EUR, and RUB. The model’s effectiveness is demonstrated through evaluations of databases captured using smartphone cameras achieving levels of accuracy, recall, and F1 score values. To enhance the models robustness the authors employ mosaic data augmentation techniques to address challenges such as angles, folds, and contrasts commonly found in real-world banknote images. The presented research not only expands the scope of banknote detection but also contributes valuable insights into the development of ML models for aiding visually impaired individuals in diverse currency recognition scenarios.

The research conducted by Sufri *et al*. [14] addresses the challenges faced by visually impaired individuals in identifying and recognizing various banknotes. The study focuses on the impact of region and orientation on the performance of ML and deep learning, employing Malaysian ringgit banknotes as the subject of analysis. The vision-based system involves two experiments, one exploring different regions and the other examining various orientations captured by a smartphone camera in a controlled environment. The first experiment utilizes ML classification algorithms, such as k-nearest neighbors (kNN), decision tree classifier (DTC), support vector machine (SVM), and bayesian classifier (BC), based on red, green, and blue (RGB) values for feature extraction. The second experiment employs AlexNet, a pre-trained convolutional neural network (CNN) model. Notably, SVM and BC outperform kNN and DTC, achieving 100% accuracy. The second stage of the experiment reveals that AlexNet's performance is influenced by the orientation of the training data. Despite this limitation, the overall conclusion suggests that both ML and deep learning models excel in recognizing and classifying Malaysian ringgit banknotes, with a recommendation favoring deep learning for its elimination of manual feature extraction. The developed algorithm holds the potential to enhance the quality of life for visually impaired individuals.

In recent research Sadyk *et al*. [15] compare types of learning models including CNNs, recurrent neural networks (RNNs), and deep belief networks (DBNs) in the context of banknote recognition. According to their findings the CNN model, specifically when using the ResNet 50 architecture demonstrates performance in terms of accuracy, precision, recall, and F1 score when compared to models. The paper presents insights and practical considerations for implementing methods that can advance banknote recognition systems to assist researchers and practitioners in selecting approaches for improving banknote recognition technology.

Existing works showcase promising outcomes, such as the impressive accuracy achieved by Hussein *et al*. [11] using HOG and SIFT, Youn *et al*. [12] efficient multi-currency classification, and the multinational banknote detection model proposed by Park and Park [13]. However, there exist notable gaps in the application of traditional and hybrid ML approaches, especially within the context of Kazakh banknotes. Therefore, the objective of this research is to train a Kazakh banknote classifier that is both lightweight and highly accurate and to assess the effectiveness of traditional and hybrid ML approaches through comparison.

2. METHOD

2.1. Dataset

In this research paper, we utilized a comprehensive dataset titled "dataset of Kazakhstan banknotes with annotations'' from mendeley data for banknote classification [16]. The dataset comprises 4200 high-quality images of Kazakhstani banknotes with a size of 1024×1024 pixels, covering 14 classes that correspond to different denominations ranging from 1 KZT to 20000 KZT. Each class signifies a distinct denomination of Kazakhstani currency. The research primarily centers around banknotes so it only considers the 500 KZT, 1000 KZT, 2000 KZT, 5000 KZT, 10000 KZT, and 20000 KZT categories since the other denominations are in coin form. Consequently, the number of classes was narrowed down to 6.

The banknote images in the dataset of various denominations were captured using the rear cameras of mobile phones under diverse conditions and presented in Figure 1. These images visually demonstrate the challenges faced during the classification process. They showcase scenarios like a rotated 500 KZT banknote against a random background (Figure 1(a)), a flipped 1,000 KZT banknote with some shadows in the background (Figure 1(b)), a crumpled 2,000 KZT banknote being held in hand (Figure 1(c)), a folded 5,000 KZT banknote also being held in hand (Figure 1(d)), a clear view of a 10,000 KZT banknote against a noisy background (Figure 1(e)) and finally a darkened 20,000 KZT banknote against a patterned background (Figure 1(f)). While these images cover real-world scenarios well enough it's important to note that they may not include all possibilities, like moving or partially folded banknotes.

Notably, all images in the dataset are meticulously annotated in the you only look once (YOLO) format [17], [18]. For the research experiments, an equal number of images, specifically 150 for the hybrid approach due to (12 GB) RAM limitations and 200 for the traditional approach, were randomly selected from each class subfolder of the original dataset as shown in Figure 2. Given the context provided above, it is also worth mentioning that given the ample size of the dataset and the random selection of samples, the invocation of dataset augmentation techniques, including rotation, scaling, or modifications in lighting, was deemed unnecessary within the context of model testing and evaluation for this study.

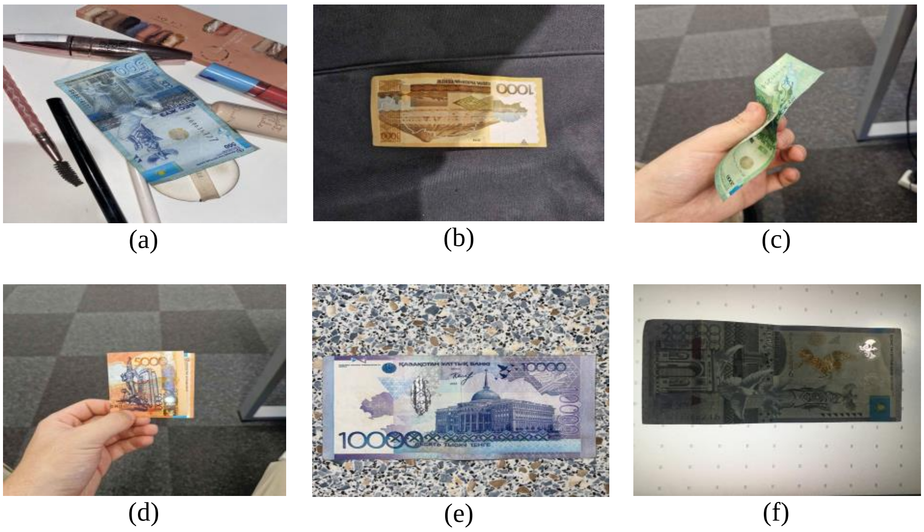


Figure 1. Examples of Kazakh banknote images of each denomination under diverse conditions: (a) rotated 500 KZT with a cluttered background, (b) flipped 1,000 KZT with a partly shadowed background, (c) crumpled 2000 KZT in the hand, (d) folded 5,000 KZT in the hand, (e) clearly visible 10,000 KZT in a noisy background, and (f) darkened 20,000 KZT in a patterned background

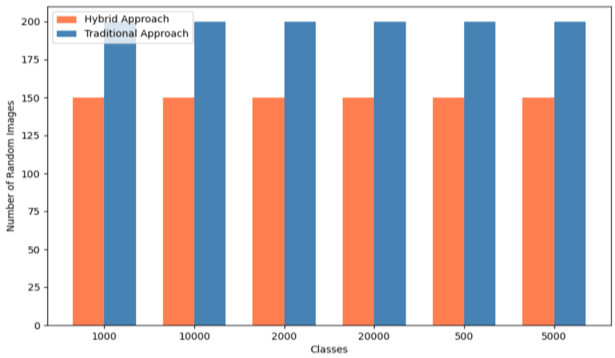


Figure 2. Number of random images of Kazakh banknotes per class

2.2. Data preprocessing and feature extraction

As mentioned earlier, the study covers two approaches which are: a traditional ML approach with manual feature extraction and a hybrid ML approach utilizing automated feature extraction with a pre-trained deep learning model.

2.2.1. Traditional machine learning approach

For the traditional approach, the initial step involved cutting the region of interest (ROI), based on the bounding box parameters specified in the dataset annotations [19], [20]. Notably, to ensure a more precise extraction, the ROIs were cropped with 10% narrowed bounding boxes, as the original bounding boxes covered slightly extra area around the banknotes (see Figure 3). Subsequently, these clippings were resized to a standard shape of 128x128 pixels and normalized dividing the value of each pixel by 255.

Figure 3. The original bounding box (green) and the narrowed bounding box (red)

The manual feature extraction focused on RGB color features, following the method proposed by the related studies [14], [21]. The decision to use RGB color features was motivated by their simplicity and effectiveness in capturing basic color characteristics. The average intensity values for RGB channels were computed, leading to features denoted as RB, RG, and GB, respectively, were calculated using (1) to (3):

$$RB = \underline{r} - \underline{b} \tag{1}$$

$$RB = \underline{r} - \underline{g} \tag{2}$$

$$RB = \underline{g} - \underline{b} \tag{3}$$

Here, $\underline{r}$, $\underline{g}$ and $\underline{b}$ represent the average intensity values for the RGB channels, respectively, within the pixels of the cropped region (see Figure 4). This manual extraction method served as a baseline for comparison against more advanced techniques.

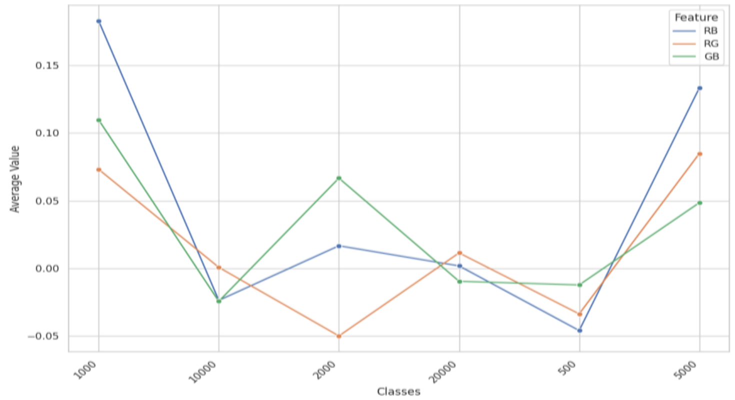


Figure 4. The average value of RB, RG, and GB for each class of Kazakh banknotes

To enhance the traditional approach, the optimization of the color feature extraction process was introduced. The image's colors were transformed into the hue, saturation, and volume (HSV) format, and histograms for hue, saturation, and value were calculated [22], [23]. These histograms were then normalized to ensure comparability across different images. The resulting normalized histograms were combined into a single set of features, providing a more nuanced representation of the color characteristics of the banknotes. Figure 5 shows the average HSV features for each class of Kazakh banknotes in the form of a heatmap.

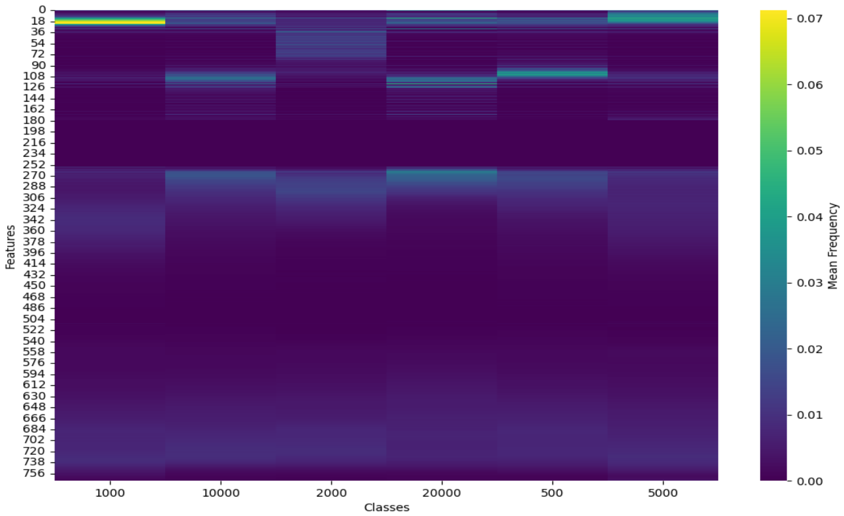


Figure 5. The average HSV features for each class of Kazakh banknotes

2.2.2. Hybrid machine learning approach

For the hybrid approach, we leveraged the power of pre-trained deep learning models. Similar to the traditional approach, the ROI was extracted based on the specified bounding box parameters. However, for the hybrid approach, there was no resizing as the subsequent feature extraction was based on a pre-trained deep learning model. The ROIs were also cropped with 10% narrowed bounding boxes to enhance precision. Feature vectors were obtained using the “avgpool” layer at the end of the ResNet-18 model as shown in Figure 6, which had been pre-trained on the ImageNet dataset [24]. The “avgpool” layer served as a dense representation of the input image, capturing high-level features that could be relevant for banknote classification [25]. This approach allowed us to benefit from the knowledge encoded in the ResNet-18 model, which had learned intricate features from a diverse set of images during its pre-training on ImageNet.

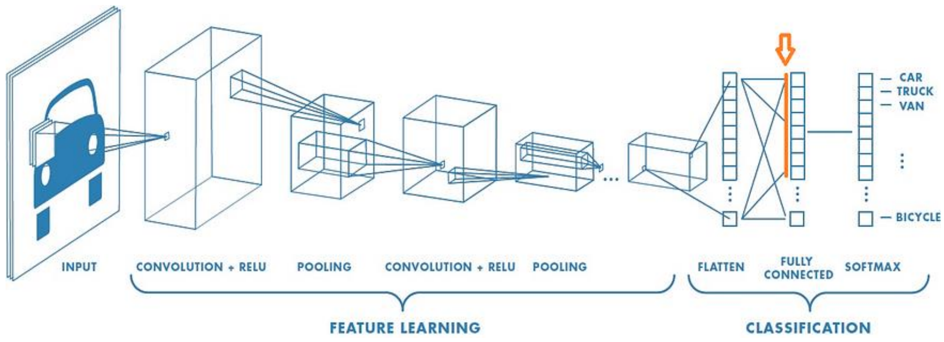


Figure 6. Example of extracting feature vector (orange) from a network [26]

The selection of ResNet-18 was predicated on its optimal balance between the length of the average pooling output vector (512 features) and model accuracy [27]. Furthermore, this model manifests a reduced computational load compared to its counterparts, rendering it an advantageous choice for the objective of constructing a model that is both lightweight and precise [28]. The resulting feature vectors were then utilized for subsequent classification tasks. This hybrid approach offered a more automated and data-driven feature extraction process, potentially capturing complex patterns and nuances in the banknote images.

2.3. Dataset splitting strategy

In this study, the dataset was randomly divided, maintaining a constant random state, to allocate 75% of the data for training purposes and the remaining 25% for testing. This division was designed to provide a comprehensive training dataset that would enable the models to learn the nuances of banknote classifications effectively, while also reserving a substantial portion of the data for an unbiased assessment of the models’ generalization capabilities [29]. The use of a constant random state ensured that the split between training and testing sets was reproducible, facilitating a fair and consistent evaluation of the models performance across different experimental runs.

2.4. Evaluation metrics

To assess the performance of our models, we employed a set of comprehensive evaluation metrics. Accuracy, as a fundamental measure, gauged the overall correctness of our classification. The F1-score, a weighted average of precision and recall, provided insights into the balance between precision and recall [30]. Matthews correlation coefficient (MCC) considered both false positives and false negatives, offering a more robust measure for imbalanced datasets [31]. Receiver operating characteristic area under the curve (ROC-AUC) provided a comprehensive view of the model’s ability to distinguish between classes using the one-versus-rest approach [32], [33].

2.5. Classification algorithms

In this section, the utilization of diverse classifiers for the purpose of banknote classification is explored. A distinct role is fulfilled by each classifier, contributing to the enrichment of the diversity and comprehensiveness of our analysis. kNN: the kNN algorithm was employed as a versatile and intuitive baseline classifier [34]. It relies on the similarity of instances in the feature space, making it particularly suitable for image classification tasks [35]. Utilizing the euclidean distance metric, kNN determined the class of a banknote image based on the classes of its kNN. The choice of k, the number of neighbors, was optimized during the grid search process.

Support vector classifier with radial basis function kernel (SVC-rbf): the SVM with RBF kernel is a powerful algorithm known for its flexibility in capturing complex relationships in data [36]. In our study, we employed the SVC-rbf to account for nonlinear patterns in the banknote images. Random forest classifier: the random forest classifier was chosen for its ability to handle high-dimensional data and capture complex interactions among features. Comprising an ensemble of decision trees, this classifier introduced an element of randomness during training, promoting diversity among the constituent trees [37]. The resulting model offered robustness against overfitting.

Gaussian naive bayes (gaussian NB): gaussian NB is a probabilistic classifier based on bayes theorem, assuming independence among features [38]. Despite its simplicity, gaussian NB has proven effective in various classification tasks [39]. In our study, it served as a probabilistic baseline for comparison with more complex algorithms. Ridge classifier (ridge CV): the ridge CV is a linear model that incorporates L2 regularization [40]. It is particularly suitable for scenarios with multicollinearity among features. In our context, the ridge CV with cross-validation was employed to assess the impact of linear models on banknote classification.

2.6. Optimization

In the pursuit of enhancing model performance, a two-fold optimization strategy was employed. Initially, for the traditional approach, the manual RGB color vector was transitioned to a more robust Color HSV histogram representation. This transformation was aimed at capturing richer color information and improving the ability of the model to distinguish between banknote classes. Simultaneously, a grid search was conducted across the grid of various hyperparameter combinations for each classifier. The grid search focused on optimizing the performance metrics, with a particular emphasis, if possible, on ROC-AUC, given its relevance to our classification task [41], [42]. The grid search process played a crucial role in fine-tuning the models for optimal performance.

In this study, the Scikit-Learn Python library was utilized extensively for the training of selected models and the optimization of their hyperparameters [43]. Highlighting these key hyperparameters and their optimized values in the context of this study is crucial for ensuring reproducibility and facilitating a deeper understanding of the models' sensitivity. For the kNN algorithm, the critical hyperparameter was the number of neighbors (neighbors=integer). In the case of SVC-rbf, the essential hyperparameters included the Kernel coefficient for 'rbf' (gamma=non-negative float) and the regularization parameter (C=positive float). For the random forest classifier, the significant hyperparameter was the number of trees in the forest (n_estimators=integer). Regarding the gaussian NB model, the key parameter was the portion of the largest variance of all features that is added to variances for calculation stability (var_smoothing=float). Lastly, for ridge CV, the important hyperparameters were the weights associated with classes (class_weight=dict or "balance"'), and the flag indicating whether the cross-validation values corresponding to each alpha should be stored (store_cv_values=true or false).

The careful adjustment of these hyperparameters has not only fine-tuned our models but also significantly boosted their efficiency and accuracy, ensuring they are perfectly suited for the task at hand. This meticulous process has led to a more robust and reliable classifier, capable of delivering superior results under a variety of conditions [44]. Therefore, these optimized hyperparameter values were instrumental in tailoring the models to achieve optimal performance on the dataset in question.

2.7. Model stacking ensemble

After optimizing individual classifiers for both the traditional and hybrid approaches, we pursued a model stacking ensemble. The intuition behind model stacking lies in combining the strengths of diverse classifiers to achieve a more robust and accurate final prediction [45]. In our study, the optimized classifiers were used as base estimators, and SVC-rbf served as the final estimator as shown in Figure 7.

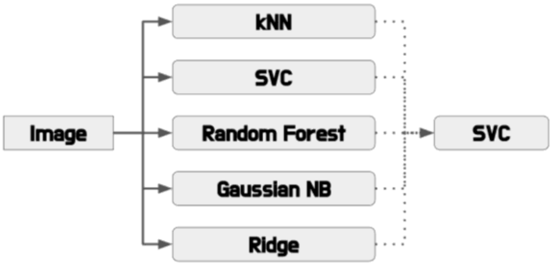


Figure 7. The simplified demonstration of model stacking ensemble

The stacking ensemble aimed to leverage the complementary strengths of the individual classifiers to enhance overall performance. By integrating diverse predictive models, it sought to mitigate the weaknesses inherent in any single model, aiming for a more robust and accurate prediction system [46]. This approach capitalizes on the varied perspectives offered by each classifier, fostering a synergistic effect that potentially elevates the ensemble's ability to tackle complex pattern recognition tasks with improved reliability [47].

2.8. Potential transferability of proposed approaches

To evaluate the potential transferability of the approaches, proposed in this study, to other image classification tasks, we conducted experiments on additional datasets beyond Kazakh banknotes. These tests aimed to assess the adaptability and effectiveness of our methodologies across different domains and challenges. Our first extension involved testing the proposed approaches on the "dataset of Indian and Thai banknotes with annotations" [48]. This exploration aimed to ascertain whether our methods could effectively classify Thai banknotes, which exhibit notable differences in color schemes, textures, and security features compared to Kazakh banknotes. Some examples of Thai banknote images can be seen in Figure 8, which showcases a diverse set across various denominations, ranging from a 20 Baht note with a partly shaded background (Figure 8(a)), a rotated 50 Baht note with a wooden background (Figure 8(b)), a folded 100 Baht note in the hand (Figure 8(c)), a folded 500 Baht note (Figure 8(d)), a clearly visible 1,000 Baht note in a cluttered background (Figure 8(e)), and a crumpled 1,000 Baht note in the hand (Figure 8(f)), thereby presenting a comprehensive test scenario for evaluating the robustness and adaptability of image classification techniques. The goal was to assess the capability of our classification techniques to adapt to and accurately recognize banknotes from different countries, each with its unique design specifications and printing technologies.

Figure 8. Examples of Thai banknote images of each denomination under diverse conditions: (a) 20 Baht with a partly shaded background, (b) rotated 50 Baht with a wooden background, (c) folded 100 Baht in the hand, (d) folded 500 Baht, (e) clearly visible 1,000 Baht in a cluttered background, (f) crumpled 1,000 Baht in the hand

Further exploring the versatility of our approaches, we applied our methodologies to the "coffee roast intelligence" dataset, a departure from banknote recognition to a domain that also heavily relies on color for classification [49]. The challenge here focused on categorizing coffee beans into dark, green, light, and medium classes based on their color characteristics, a task fundamentally different from banknote classification but similarly dependent on subtle color variations as illustrated in Figure 9, which consists clear images of the four roasting levels of coffee beans: (Figure 9 (a)) dark coffee bean, (Figure 9 (b)) green coffee bean, (Figure 9 (c)) light coffee bean, and (Figure 9 (d)) medium coffee bean. This extension aimed to test the applicability of our approaches to broader image classification tasks where color plays a crucial role in distinguishing between classes, such as in quality control in the food industry or agricultural product analysis. These experimental applications were designed to closely follow the methodological framework established for Kazakh banknote classification, including similar preprocessing steps, feature extraction methods encompassing both traditional and deep learning techniques, and the utilization of model stacking strategies. The results from these experiments are intended to illuminate the generalizability and flexibility of our classification approaches, demonstrating their potential utility across a diverse array of image classification scenarios.

(a) (b) (c) (d)

Figure 9. Well-lit images of the four roasting levels of coffee beans: (a) dark coffee bean, (b) green coffee bean, (c) light coffee bean, and (d) medium coffee bean

3. RESULTS AND DISCUSSION

The evaluation of the traditional and hybrid ML approaches reveals distinct performances of Kazakh banknote classifiers, both before and after optimization. The analysis includes traditional RGB color vectors, proposed HSV color features in the traditional approach, and features extracted from the pre-trained ResNet-18 model in the hybrid ML approach. Model stacking, combining classifiers from each approach, is also explored to harness the synergies of both methodologies.

3.1. Traditional machine learning approach

Table 1 showcases the performance of classifiers based on traditional RGB color vectors. The results indicate varying degrees of accuracy, F1-score, MCC, and receiver ROC-AUC. The SVC-rbf stands out with the highest accuracy of 71% and MCC of 65.52%. However, other classifiers like kNN and Random Forest also demonstrate competitive performance.

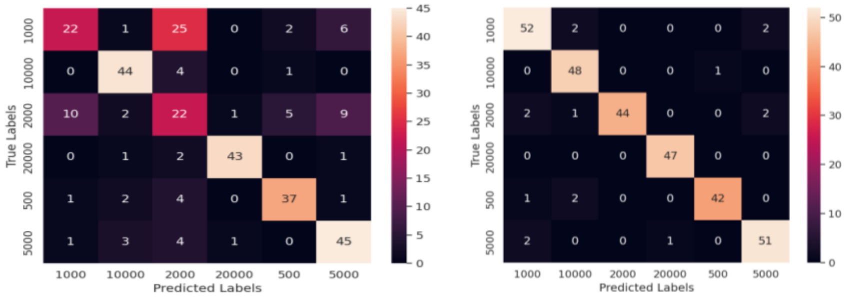
Table 1. Traditional approach (color RGB vectors-3 features)

Classifier	Accuracy (%)	F1-score (%)	MCC (%)	ROC-AUC (%)
kNN	70.48	70.74	64.61	88.99
SVC-RBF	71.00	70.69	65.52	93.07
Random forest	69.33	69.05	63.19	91.90
Gaussian NB	56.19	57.73	47.89	88.87
Ridge CV	67.94	63.66	62.48	None

Table 2 presents the outcomes after optimizing the traditional approach by transitioning to color HSV histograms. Notably, the optimization substantially improves the performance across all classifiers. random forest achieves an impressive accuracy of 94.67%, demonstrating the effectiveness of the optimized features in enhancing classification accuracy. The detailed parameter configurations for each classifier are included for transparency. The confusion matrix presented in Figure 10 provides a visual representation of the classification performance, comparing the pre-optimization results of SVC-pol (Figure 10(a)) and the post-optimization outcomes of random forest (Figure 10(b)) for the traditional approach. Each cell in the matrix represents the true labels and predicted labels for each image giving us a clearer understanding of how well the classifiers can identify different banknote classes [50].

Table 2. Traditional approach after the optimization (color HSV histogram-768 features)

Classifier	Accuracy (%)	F1-score (%)	MCC (%)	ROC-AUC (%)	Best parameters
kNN	88.00	88.12	85.65	96.24	3 neighbors
SVC-RBF	94.00	93.99	92.81	98.99	gamma=4, C=1000
Random forest	94.67	94.67	93.61	98.99	n_estimators=100
Gaussian NB	79.67	79.94	75.61	96.55	var smoothing=0.01
Ridge CV	86.67	86.49	84.16	None	class_weight=‘balanced’, store_cv_values=True



(a) (b)

Figure 10. Confusion matrix of: (a) SVC-rbf pre-optimization and (b) random forest post-optimization

3.2. Hybrid machine learning approach

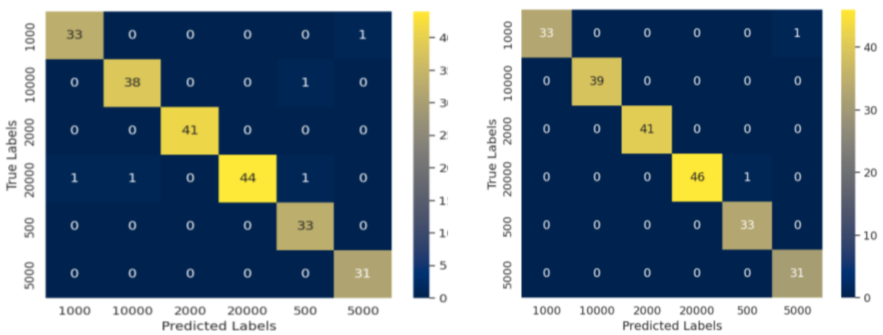
In Table 3, the hybrid approach, utilizing features extracted from the pre-trained ResNet-18 model, showcases remarkable performance. All classifiers exhibit high accuracy, F1-score, MCC, and ROC-AUC values, reflecting the power of leveraging deep learning features. SVC-rbf achieves outstanding accuracy and MCC scores of 97.78% and 97.34%, respectively. After optimization, presented in Table 4, the hybrid approach continues to deliver exceptional results. SVC-rbf now attains a near-perfect accuracy of 99.11% and MCC of 99.93%, further emphasizing the efficacy of combining deep learning features with optimized classifiers. The confusion matrix presented in Figure 11 provides a visual representation of the classification performance, comparing the pre-optimization results (Figure 11(a)) and the post-optimization outcomes (Figure 11(b)) of SVC-pol for the hybrid approach. In those matrices, every cell represents the true labels and predicted labels for each image, giving us a better grasp of how accurately the classifiers can recognize various banknote categories.

Table 3. Hybrid approach (ResNet-18-512 features)

Classifier	Accuracy (%)	F1-score (%)	MCC (%)	ROC-AUC (%)
kNN	91.55	91.61	89.97	99.25
SVC-rbf	97.78	97.77	97.34	99.93
Random forest	94.22	94.25	93.11	99.98
Gaussian NB	92.89	92.98	91.51	99.16
Ridge CV	96.44	96.44	95.73	None

Table 4. Hybrid approach after the optimization (ResNet-18-512 features)

Classifier	Accuracy (%)	F1-score (%)	MCC (%)	ROC-AUC (%)	Best parameters
kNN	93.33	93.39	92.11	99.22	3 neighbors
SVC-rbf	99.11	99.11	98.93	99.98	gamma=0.0001, C=1000
Random forest	95.11	95.12	94.17	99.98	n_estimators=100
Gaussian NB	93.78	93.80	92.55	99.45	var smoothing=0.1
Ridge CV	98.22	98.22	97.86	None	class_weight=‘balanced’, store_cv_values=True



(a) (b)

Figure 11. Confusion matrix of: (a) SVC-rbf pre-optimization and (b) SVC-rbf post-optimization

3.3. Model stacking ensemble

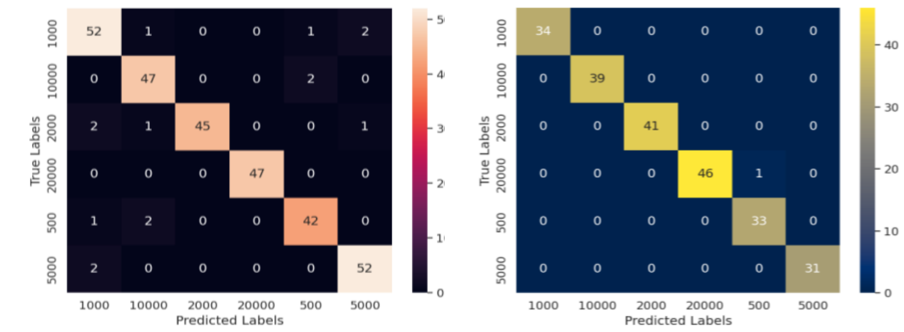
The trained models of both ML approaches were saved in a ".p" format. Since the hybrid approach utilizes the pre-trained deep neural network, the weights for the hybrid model were calculated by combining the weights of ResNet-18 (44.7 MB, the model is available at [28]) with the model stacking ensemble. Furthermore, beyond the classification of Kazakhstani banknotes, the efficacy of the proposed approaches in categorizing Thai banknotes and roasted coffee beans was rigorously examined. This investigation aimed to ascertain the adaptability and versatility of the proposed methodologies across varied image classification tasks, including other currencies and domains extending beyond banknote recognition.

3.3.1. Kazakh banknotes

Table 5 explores the performance of the model stacking ensemble, combining Kazakh banknote classifiers from both approaches for Kazakh banknote classification. This ensemble, utilizing kNN, SVC-rbf, random forest, gaussian NB, and ridge CV, achieved notable accuracy and F1-score values. The traditional approach shows an accuracy of 95.00%, and MCC of 95.01%, with as model weight of 10.9 MB. The hybrid approach, however, significantly outperforms the traditional method, achieving 99.55% accuracy, and 99.56% MCC, but at an increased model size of 52.97 MB, illustrating the superior performance of the hybrid method at the cost of larger model complexity. Figure 12 displays confusion matrices for the model stacking ensemble using optimized models from the traditional (Figure 12(a)) and hybrid (Figure 12(b)) ML approaches for Kazakh banknote classification. These matrices highlight the classification accuracy and instances of misclassification for different banknote denominations, illustrating the comparative effectiveness and areas for enhancement in each approach.

Table 5. Model stacking after the optimization SVC-rbf (gamma=0.0001, C=1000) for Kazakh banknotes

Approach	Accuracy (%)	F1-score (%)	MCC (%)	Weights (MB)
Traditional	95.00	94.01	95.01	10.9
Hybrid	99.55	99.47	99.56	52.97



(a) (b)

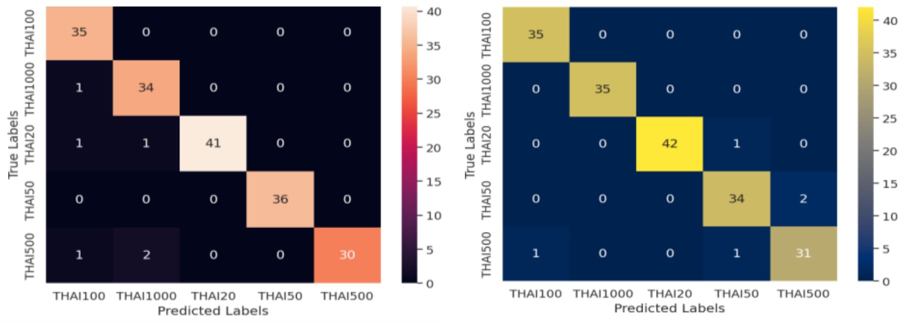
Figure 12. Confusion matrix of the model stacking ensemble in which optimized models of: (a) traditional ML approach and (b) hybrid ML approach for Kazakh banknote classification were used

3.3.2. Thai banknotes

Table 6 presents the outcomes of the model stacking ensemble, which integrates classifiers including kNN, SVC-rbf, random forest, gaussian NB, and ridge CV from both traditional and hybrid approaches. The table details the post-optimization performance of SVC-rbf across various metrics such as accuracy, F1-score, MCC, and model size in megabytes (MB). The traditional approach demonstrates an accuracy of 96.70%, with a model size of 7.24 MB. The hybrid approach shows slightly higher performance with an accuracy of 97.25% and a considerably larger model size of 52.28 MB. This comparison underscores the efficacy and efficiency of model stacking in enhancing classification performance, while also highlighting the trade-offs in model size between traditional and hybrid methods. Figure 13 presents confusion matrices for the model stacking ensemble with optimized models from the traditional (Figure 13(a)) and hybrid (Figure 13(b)) ML approaches, specifically applied to Thai banknote classification. These matrices reveal the accuracy and misclassification rates for various Thai banknote denominations, demonstrating the effectiveness and potential areas of improvement for each method.

Table 6. Model stacking after the optimization SVC-rbf (gamma=0.0001, C=1000) for Thai banknotes

Approach	Accuracy (%)	F1-score (%)	MCC (%)	Weights (MB)
Traditional	96.70	96.71	95.92	7.24
Hybrid	97.25	97.25	96.56	52.28



(a) (b)

Figure 13. Confusion matrix of the model stacking ensemble in which optimized models of: (a) traditional ML approach and (b) hybrid ML approach for Thai banknote classification were used

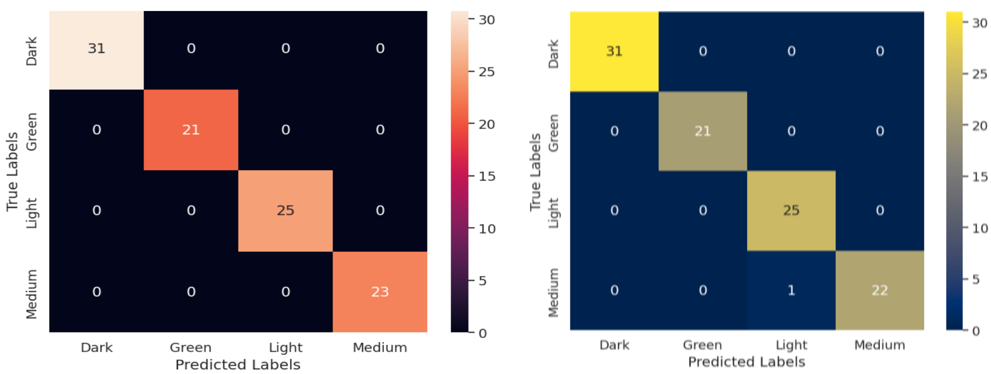
3.3.3. Roasted coffee beans

Table 7 summarizes the model stacking ensemble's performance for coffee bean roasting level classification post-optimization with SVC-rbf. The traditional approach achieves perfect scores across accuracy, F1-score, and MCC, with a model size of 3.58 MB. In contrast, the hybrid approach scores 99.00% accuracy, 98.99% F1-score, and 98.67% MCC, requiring a larger model size of 48.4 MB, highlighting the exceptional accuracy of traditional methods in specific contexts while noting the hybrid method's slight decrease in performance but significant increase in model size.

Table 7. Model stacking after the optimization SVC-rbf (gamma=0.0001, C=1000) for coffee beans

Approach	Accuracy (%)	F1-score (%)	MCC (%)	Weights (MB)
Traditional	100	100	100	3.58
Hybrid	99.00	98.99	98.67	48.4

Figure 14 showcases confusion matrices for the model stacking ensemble utilizing optimized models from both the traditional (Figure 14(a)) and hybrid (Figure 14(b)) ML approaches, aimed at classifying coffee bean roasting levels. The first matrix highlights a perfect classification outcome using the traditional ML approach, whereas the hybrid approach also shows a high level of accuracy with only one false prediction. This comparison underlines the efficacy of both approaches in distinguishing between different roasting levels of coffee beans, with the traditional approach achieving remarkable precision.



(a) (b)

Figure 14. Confusion matrix of the model stacking ensemble in which optimized models of: (a) traditional ML approach and (b) hybrid ML approach for coffee bean classification were used

3.4. Discussion

Discussing the performance of classifiers based on RGB color vectors, a diverse range of results emerges. SVC-rbf ranks first in terms of accuracy and MCC, demonstrating its ability to capture non-linear relationships in the dataset. However, recognizing the potential for improvement, a shift to Color HSV histograms significantly enhances classification accuracy, with random forest achieving a reasonable 94.67%. Moreover, the classifier’s primary reliance on color features minimizes the impact of environmental factors, such as reflections and shadows, on its performance.

Moving to a hybrid approach that utilizes features from the pre-trained ResNet-18 model results in a paradigm shift. This approach consistently outperforms the traditional one across all classifiers. SVC-RBF and random forest, in particular, exhibit exceptional accuracy and ROC-AUC, which emphasizes the superiority of deep learning features. After the optimization, the hybrid approach maintains its superior performance, emphasizing the sustained effectiveness of leveraging deep learning in feature extraction.

These findings highlight the importance of using a combination of automated methods in ML to classify banknotes. Even though the traditional approach relies on designed features it still benefits greatly from optimization and lightweight emphasizing the significance of carefully selecting relevant features. On the other hand, by utilizing deep learning features the hybrid method excels at capturing intricate patterns within banknote images.

Lastly, model stacking emerges as a technique that effectively combines different classifiers strengths. The stacked model performs better than separate classifiers in both traditional and hybrid approaches, suggesting that incorporating diverse sources of information enhances predictive capabilities.

The exploration of our proposed approaches applicability to Thai banknote classification and the classification of roasted coffee beans underscores the methods remarkable versatility and potential for transferability across diverse image classification tasks. Successfully adapting these methodologies to the distinct challenges presented by Thai banknotes-with their unique color schemes, textures, and security features-demonstrates the approaches robustness in handling different currency types. Similarly, the effective application to the roasted coffee bean classification task, which hinges on subtle color variations to distinguish between roasting levels, further validates the utility of our methods beyond the realm of banknote recognition. These results underscore the flexibility of the proposed approaches, indicating their wide-ranging applicability across diverse domains where color distinctions between classes play a critical role.

However, this study unveils several limitations impacting the generalizability of the classifiers. Relying heavily on color features assumes that these features alone can accurately represent the characteristics of banknotes without considering other factors like text or materials. The assumption of feature homogeneity across denominations may overlook variations between them. This limited scope and the hybrid approach’s dependence on image quality introduce potential inconsistencies in real-world applications. Additionally, the hybrid approach, which utilizes pre-trained deep learning architectures, may compromise on interpretability relative to conventional ML methodologies. Comprehending the decision-making mechanism of the model, particularly in intricate situations, could prove to be formidable. Acknowledging these limitations is crucial for interpreting the findings and guiding future research in refining practical banknote classification systems.

4. CONCLUSION

In the pursuit of creating a lightweight, highly accurate Kazakh banknote classifier to address challenges posed by visual impairments, a comprehensive study was conducted to compare traditional and hybrid ML approaches. Utilizing a meticulously annotated dataset of 4200 high-quality images of Kazakhstani banknotes, the focus was placed on 6 denominations crucial for practical applications. Initially, the traditional approach, employing RGB color vectors and manual feature extraction, demonstrated commendable performance. The adoption of color HSV histograms further boosted the accuracy of this method, underscoring the pivotal role of meticulous feature engineering in enhancing classification outcomes. On the other hand, remarkable accuracy and robustness were showcased by the hybrid approach, utilizing features from the pre-trained ResNet-18 model.

Through a detailed evaluation process that included metrics such as accuracy, F1-score, MCC, and ROC-AUC, the research unveiled the high efficiency of both approaches. Following optimization, the traditional approach attained a significant accuracy of 94.00%, while the hybrid approach excelled with a near-perfect accuracy of 99.11%. Transitioning to model stacking, an ensemble technique that synergizes classifiers from both the traditional and hybrid models, surpassed individual classifier performance. Featuring kNN, SVC-rbf, random forest, gaussian NB, and ridge CV as base estimators, the stacked model demonstrated accuracies of 95.00% and 99.55% for the traditional and hybrid approaches, respectively.

The effectiveness of our methods in classifying Thai banknotes and roasted coffee beans highlights their broad utility and adaptability across various image classification tasks. This ability, especially in tasks where color plays a crucial role, emphasizes the potential applications beyond banknote recognition. However, acknowledgment of several limitations is crucial in interpreting the study's findings. The hard reliance on color features, exclusion of non-color features, and potential deviations from real-world scenarios due to the focus on color-centric analysis are acknowledged. The assumed homogeneity of features and the limited consideration of cultural elements further impact the generalizability of findings.

In conclusion, the research advances the field of Kazakh banknote classification, providing a roadmap for the creation of practical, accurate, and lightweight classifiers. The hybrid approach, leveraging the strengths of both manual and automated feature extraction, emerges as a promising solution. Model stacking, combining the best of both worlds, showcases the potential for enhanced predictive capabilities. As achievements and limitations are navigated, the study lays the foundation for future research, encouraging the exploration of additional features, cultural considerations, advanced deep learning architectures, and combinations of banknote detection techniques with the banknote classifiers. The ultimate goal is to contribute to the development of inclusive and efficient banknote classification systems, addressing challenges faced by individuals with visual impairments.

ACKNOWLEDGEMENTS

This work is supported by funding from the Ministry of Science and Higher Education of the Republic of Kazakhstan, IRN AP19675614. This work is supported by the internal grant of SDU University for 2023-2024.

REFERENCES

[1] F. Laricchia, "Global smartphone penetration rate as share of population from 2016 to 2022," Statista.com. [Online]. Available: https://www.statista.com/statistics/203734/global-smartphone-penetration-per-capita-since-2005. (Accessed: Feb. 01, 2024).

[2] S. Khan, M. Sajjad, T. Hussain, A. Ullah, and A. S. Imran, "A review on traditional machine learning and deep learning models for WBCs classification in blood smear images," *IEEE Access*, vol. 9, pp. 10657–10673, 2021, doi: 10.1109/ACCESS.2020.3048172.

[3] K. Merkelbach *et al.*, "HybridML: Open source platform for hybrid modeling," *Computers and Chemical Engineering*, vol. 160, p. 107736, Apr. 2022, doi: 10.1016/j.compchemeng.2022.107736.

[4] M. O. Khairandish, M. Sharma, V. Jain, J. M. Chatterjee, and N. Z. Jhanjhi, "A Hybrid CNN-SVM threshold segmentation approach for tumor detection and classification of MRI brain images," *Irbm*, vol. 43, no. 4, pp. 290–299, Aug. 2022, doi: 10.1016/j.irbm.2021.06.003.

[5] U. Sadyk, R. Baimukashev, A. Bozshina, and C. Turan, "KZ-BD: dataset of Kazakhstan banknotes with annotations," *Social Science Research Network*, vol. 53, 2024, doi: 10.1016/j.dib.2024.110076.

[6] Q. Chen, Z. Song, J. Dong, Z. Huang, Y. Hua, and S. Yan, "Contextualizing object detection and classification," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 37, no. 1, pp. 13–27, Jan. 2015, doi: 10.1109/TPAMI.2014.2343217.

[7] A. Khammari, F. Nashashibi, Y. Abramson, and C. Laurgeau, "Vehicle detection combining gradient analysis and AdaBoost classification," in *IEEE Conference on Intelligent Transportation Systems*, IEEE, 2005, pp. 66–71, doi: 10.1109/ITSC.2005.1520202.

[8] X. Li *et al.*, "A comprehensive review of computer-aided whole-slide image analysis: from datasets to feature extraction, segmentation, classification and detection approaches," *Artificial Intelligence Review*, vol. 55, no. 6, pp. 4809–4878, Aug. 2022, doi: 10.1007/s10462-021-10121-0.

[9] H. Harzallah, F. Jurie, and C. Schmid, "Combining efficient object localization and image classification," in *Proceedings of the IEEE International Conference on Computer Vision*, IEEE, Sep. 2009, pp. 237–244, doi: 10.1109/ICCV.2009.5459257.

[10] U. Sadyk, R. Baimukashev, and C. Turan, "State-of-the-art review of deep learning methods in fake banknote recognition problem," *International Journal of Advanced Computer Science and Applications*, vol. 15, no. 1, pp. 848–856, 2024, doi: 10.14569/IJACSA.2024.0150185.

[11] G. S. Hussein, S. Elseuofi, W. H. Dukhan, and A. H. Ali, "A novel method for banknote recognition using a combined histogram of oriented gradients and scale-invariant feature transform," *Information Sciences Letters*, vol. 12, no. 9, pp. 2121–2131, Sep. 2023, doi: 10.18576/isl/120918.

[12] S. Youn, E. Choi, Y. Baek, and C. Lee, "Efficient multi-currency classification of CIS banknotes," *Neurocomputing*, vol. 156, pp. 22–32, May 2015, doi: 10.1016/j.neucom.2015.01.014.

[13] C. Park and K. R. Park, "MBDM: Multinational banknote detecting model for assisting visually impaired people," *Mathematics*, vol. 11, no. 6, p. 1392, Mar. 2023, doi: 10.3390/math11061392.

[14] N. A. J. Sufri, N. A. Rahmad, N. F. Ghazali, N. Shahar, and M. A. As'ari, "Vision based system for banknote recognition using different machine learning and deep learning approach," in *2019 IEEE 10th Control and System Graduate Research Colloquium*, IEEE, Aug. 2019, pp. 5–8, doi: 10.1109/ICSGRC.2019.8837068.

[15] U. Sadyk, C. Turan, and R. Baimukashev, "Overview of deep learning models for banknote recognition," in *Proceedings-2023 17th International Conference on Electronics Computer and Computation*, IEEE, Jun. 2023, pp. 1–5, doi: 10.1109/ICECCO58239.2023.10147142.

[16] U. S. and A. Bozshina, "Dataset Of Kazakhstan banknotes with annotations," *Mendeley Data*, vol. 1, 2023, doi: 10.17632/dny3dgvvw8.1.

[17] N. Mamat, M. F. Othman, R. Abdulghafor, A. A. Alwan, and Y. Gulzar, "Enhancing image annotation technique of fruit classification using a deep learning approach," *Sustainability (Switzerland)*, vol. 15, no. 2, p. 901, Jan. 2023, doi: 10.3390/su15020901.

[18] C. Sachin, N. Manasa, V. Sharma, and N. A. A. Kumaar, “Vegetable classification using you only look once algorithm,” in *2019 International Conference on Cutting-Edge Technologies in Engineering*, IEEE, Nov. 2019, pp. 101–107, doi: 10.1109/ICon-CuTE47290.2019.8991457.
[19] A. Nieto-Castanon, S. S. Ghosh, J. A. Tourville, and F. H. Guenther, “Region of interest based analysis of functional imaging data,” *NeuroImage*, vol. 19, no. 4, pp. 1303–1316, Aug. 2003, doi: 10.1016/S1053-8119(03)00188-5.
[20] J. Askelöf, M. L. Carlander, and C. Christopoulos, “Region of interest coding in JPEG 2000,” *Signal Processing: Image Communication*, vol. 17, no. 1, pp. 105–111, Jan. 2002, doi: 10.1016/S0923-5965(01)00026-1.
[21] R. Khan, A. Hanbury, J. Stöttinger, and A. Bais, “Color based skin classification,” *Pattern Recognition Letters*, vol. 33, no. 2, pp. 157–163, Jan. 2012, doi: 10.1016/j.patrec.2011.09.032.
[22] S. Sural, G. Qian, and S. Pramanik, “Segmentation and histogram generation using the HSV color space for image retrieval,” in *IEEE International Conference on Image Processing*, IEEE, 2002, pp. II-589-II–592, doi: 10.1109/icip.2002.1040019.
[23] I. Kurniastuti, E. N. I. Yuliati, F. Yudianto, and T. D. Wulan, “Determination of hue saturation value (HSV) color feature in kidney histology image,” *Journal of Physics: Conference Series*, vol. 2157, no. 1, p. 012020, Jan. 2022, doi: 10.1088/1742-6596/2157/1/012020.
[24] K. He, X. Zhang, S. Ren, and J. Sun, “Deep residual learning for image recognition,” *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 2016-Decem, pp. 770–778, 2016, doi: 10.1109/CVPR.2016.90.
[25] H. Song, Y. Zhou, Z. Jiang, X. Guo, and Z. Yang, “ResNet with global and local image features, stacked pooling block, for semantic segmentation,” in *2018 IEEE/CIC International Conference on Communications in China*, IEEE, Aug. 2018, pp. 79–83, doi: 10.1109/ICCChina.2018.8641146.
[26] C. Safka, “Extract a feature vector for any image with PyTorch,” Medium.com. [Online]. Available: https://miro.medium.com/v2/resize:fit:1100/format:webp/1*1BEDb6N5T4ZRZVb31IpKsw.png. (Accessed: Jan. 05, 2023).
[27] P. Singh, P. Raj, and V. P. Namboodiri, “EDS pooling layer,” *Image and Vision Computing*, vol. 98, p. 103923, Jun. 2020, doi: 10.1016/j.imavis.2020.103923.
[28] PyTorch, “Resnet models-PyTorch Hub,” PyTorch.org. [Online]. Available: https://pytorch.org/hub/pytorch_vision_resnet. (Accessed: Jan. 06, 2024).
[29] K. K. Dobbin and R. M. Simon, “Optimally splitting cases for training and testing high dimensional classifiers,” *BMC Medical Genomics*, vol. 4, no. 1, p. 31, Dec. 2011, doi: 10.1186/1755-8794-4-31.
[30] Ž. Vujović, “Classification Model Evaluation Metrics,” *International Journal of Advanced Computer Science and Applications*, vol. 12, no. 6, pp. 599–606, 2021, doi: 10.14569/IJACSA.2021.0120670.
[31] D. Chicco and G. Jurman, “The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation,” *BMC Genomics*, vol. 21, no. 1, p. 6, Dec. 2020, doi: 10.1186/s12864-019-6413-7.
[32] A. J. Bowers and X. Zhou, “Receiver operating characteristic (ROC) area under the curve (AUC): a diagnostic measure for evaluating the accuracy of predictors of education outcomes,” *Journal of Education for Students Placed at Risk*, vol. 24, no. 1, pp. 20–46, Jan. 2019, doi: 10.1080/10824669.2018.1523734.
[33] Scikit Learn, “sklearn.metrics.roc_auc_score — scikit-learn 0.23.2 documentation,” scikit-learn.org. [Online]. Available: https://scikit-learn.org/stable/modules/generated/sklearn.metrics.roc_auc_score.html. (Accessed: Jan. 06, 2023).
[34] O. Kramer, “K-Nearest Neighbors,” *Dimensionality Reduction with Unsupervised Nearest Neighbors*, vol. 51, pp. 13–23, 2013, doi: 10.1007/978-3-642-38652-7_2.
[35] Z. Fan, J. K. Xie, Z. Y. Wang, P. C. Liu, S. J. Qu, and L. Huo, “Image classification method based on improved KNN Algorithm,” *Journal of Physics: Conference Series*, vol. 1930, no. 1, p. 012009, May 2021, doi: 10.1088/1742-6596/1930/1/012009.
[36] K. Thurnhofer-Hemsi, E. López-Rubio, M. A. Molina-Cabello, and K. Najarian, “Radial basis function kernel optimization for Support Vector Machine classifiers,” *arXiv*, 2020, doi: 10.48550/arXiv.2007.08233.
[37] Ö. Akar and O. Güngör, “Rastgele orman algoritması kullanılarak çok bantlı görüntülerin sınıflandırılması,” *Journal of Geodesy and Geoinformation*, vol. 1, no. 2, pp. 139–146, 2012, doi: 10.9733/jgg.241212.1t.
[38] S. McCann and D. G. Lowe, "Local Naive Bayes Nearest Neighbor for image classification," *2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence*, RI, USA, 2012, pp. 3650-3656, doi: 10.1109/CVPR.2012.6248111.
[39] H. Kamel, D. Abdulah, and J. M. Al-Tuwaijari, “Cancer classification using Gaussian Naive Bayes Algorithm,” in *Proceedings of the 5th International Engineering Conference*, IEEE, Jun. 2019, pp. 165–170, doi: 10.1109/IEC47844.2019.8950650.
[40] C. Peng and Q. Cheng, “Discriminative Ridge Machine: A classifier for high-dimensional data or imbalanced data,” *IEEE Transactions on Neural Networks and Learning Systems*, vol. 32, no. 6, pp. 2595–2609, Jun. 2021, doi: 10.1109/TNNLS.2020.3006877.
[41] H. Alibrahim and S. A. Ludwig, “Hyperparameter optimization: comparing genetic Algorithm against Grid Search and Bayesian Optimization,” in *2021 IEEE Congress on Evolutionary Computation*, IEEE, Jun. 2021, pp. 1551–1559, doi: 10.1109/CEC45853.2021.9504761.
[42] Q. Huang, J. Mao, and Y. Liu, “An improved grid search algorithm of SVR parameters optimization,” in *International Conference on Communication Technology Proceedings*, IEEE, Nov. 2012, pp. 1022–1026, doi: 10.1109/ICCT.2012.6511415.
[43] Scikit Learn, “API Reference — scikit-learn 1.4.0 documentation,” scikit-learn.org. [Online]. Available: https://scikit-learn.org/stable/modules/classes.html. (Accessed: Jan. 06, 2023).
[44] A. M. Carrington *et al.*, “Deep ROC Analysis and AUC as balanced average accuracy, for improved classifier selection, audit and explanation,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 45, no. 1, pp. 329–341, Jan. 2023, doi: 10.1109/TPAMI.2022.3145392.
[45] A. Chatzimparmpas, R. M. Martins, K. Kucher, and A. Kerren, “StackGenVis: alignment of data, algorithms, and models for stacking ensemble learning using performance metrics,” *IEEE Transactions on Visualization and Computer Graphics*, vol. 27, no. 2, pp. 1547–1557, Feb. 2021, doi: 10.1109/TVCG.2020.3030352.
[46] J. Dou *et al.*, “Improved landslide assessment using support vector machine with bagging, boosting, and stacking ensemble machine learning framework in a mountainous watershed, Japan,” *Landslides*, vol. 17, no. 3, pp. 641–658, Mar. 2020, doi: 10.1007/s10346-019-01286-5.
[47] P. Charoenkwan, W. Chiangjong, C. Nantasenamat, M. M. Hasan, B. Manavalan, and W. Shoombuatong, “StackIL6: A stacking ensemble model for improving the prediction of IL-6 inducing peptides,” *Briefings in Bioinformatics*, vol. 22, no. 6, Nov. 2021, doi: 10.1093/bib/bbab172.
[48] V. Meshram, K. Patil, and P. Chumchu, “Dataset of Indian and Thai banknotes with annotations,” *Data in Brief*, vol. 41, 2022, doi: 10.1016/j.dib.2022.108007.
[49] S. Ontoum, T. Khemanantakul, P. Sroison, T. Triyason, and B. Watanapa, “Coffee roast intelligence,” *arXiv*, 2022, doi:

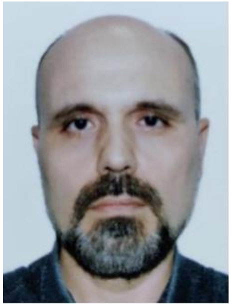
10.48550/arXiv.2206.01841.
[50] N. D. Marom, L. Rokach, and A. Shmilovici, "Using the confusion matrix for improving ensemble classifiers," in *2010 IEEE 26th Convention of Electrical and Electronics Engineers in Israel*, IEEE, Nov. 2010, pp. 555–559, doi: 10.1109/EEEI.2010.5662159.

BIOGRAPHIES OF AUTHORS

Ualikhan Sadyk earned a Bachelor of Technology (Second Class Hons) in Radio Engineering, Electronics and Telecommunications from Karaganda State University, Kazakhstan, and a Master of Science in Computer Science from SDU University, Kazakhstan. Currently in the second year of his Ph.D. program in Computer Science at SDU University. Formerly serving as the Head of the Computer Science Department at SDU University. He currently holds the positions of Educational Program Coordinator of Computer Science and Senior Lecturer at the Department of Computer Science, SDU University, Kaskelen, Kazakhstan. His research interests encompass recommendation systems, image recognition, and computer vision. He can be contacted at email: ualikhan.sadyk@sdu.edu.kz.

Makhambet Yerzhan is currently a 4th-year Information Systems student at SDU University with a year of experience as a Data Scientist at FelenaSoft, a company specializing in multimedia software development. In this role, he has developed various detection technologies, including the abandoned object detection system. His primary research interests lie in computer vision, human-AI interaction-specifically, machine-based human emotion recognition-and the vast field of generative pre-trained transformers (GPT). He can be contacted at email: makhambeterzhan@gmail.com.

Rashid Baimukashev received a Bachelor of Sciences in Radio Engineering, Electronics, and Telecommunications from Almaty Institute of Power Engineering and Telecommunications, Kazakhstan, Master of Sciences in Communications Engineering from RWTH-Aachen University, Germany and currently in a second year Ph.D. program in Computer Science at SDU University, Kazakhstan. His research interests are in image processing and network security. He can be contacted at email: rashid.baimukashev@sdu.edu.kz.



Cemil Turan is an Associated Professor at SDU University, Kazakhstan, where he has been a faculty member since 2008. From 2021-2022, he was also the Head of the Computer Science Department. He graduated bachelor's degree in Electrical Engineering from Yildiz Technical University, Turkey, in 1995, and completed his Ph.D. in Electrical and Computer Engineering from Mevlana Rumi University, Turkey, in 2016. Her research interests are primarily in the area of digital signal and image processing. He can be contacted at email: cemil.turan@sdu.edu.kz.