International Journal of Engineering and Technology Innovation, vol. 14, no. 3, 2024, pp. 321-334

Domain Adaptation for Roasted Coffee Bean Quality Inspection

Cheng-Lung Chang, Shou-Chuan Lai, Ching-Yi Chen*

Department of Information and Telecommunications Engineering, Ming Chuan University, Taoyuan, Taiwan, ROC

Received 27 January 2024; received in revised form 27 March 2024; accepted 28 March 2024

DOI: https://doi.org/10.46604/ijeti.2024.13315

Abstract

Current research in machine learning primarily focuses on raw coffee bean quality, hampered by limited labeled datasets for roasted beans. This study proposes a domain adaptation approach to transfer knowledge acquired from raw coffee beans to the task of inspecting roasted beans. The method maps the source and target data, originating from different distributions, into a shared feature space while minimizing distribution discrepancies with domain adversarial training. Experimental results demonstrate that the proposed approach effectively uses annotated raw bean datasets to achieve a high-performance quality inspection system tailored specifically to roasted coffee beans.

Keywords: machine learning, domain adaptation, domain adversarial training, coffee bean quality inspection

1. Introduction

After undergoing roasting and grinding, coffee beans can be utilized to craft a beverage through hot water extraction. Not only do coffee beverages possess unique flavors, but numerous research studies highlight the benefits of moderate coffee consumption, such as heightened alertness, improved athletic performance, and enhanced metabolism. These advantages have contributed to the widespread global popularity of coffee [1-2].

Coffee beans require immediate entry into the processing pipeline to prevent fermentation and the emergence of undesirable aromas. To facilitate the peeling of coffee beans, a common practice involves their initial immersion to separate floaters, followed by procedures like sun drying and husking to obtain the inner coffee beans. Despite the various merits associated with coffee consumption, the presence of defective beans, if not promptly removed before usage, can negatively impact flavor and potentially compromise human health [3]. While the extraction process for green coffee beans includes an initial screening for immature beans and floaters, the peeling process itself poses a risk of fermentation or mold development. Subsequent post-processing steps, including roasting, may introduce different forms of quality degradation, such as bean fragmentation or charring. Consequently, a final selection of various types of defective beans necessitates manual or machine-based screening.

Common defects found in coffee beans encompass floaters, black beans, sour beans, insect-damaged beans, fungus-damaged beans, broken beans, and foreign matter. Traditionally, the manual visual inspection of defective beans in green coffee beans involves an initial classification using sieves, followed by experiential judgment to assess the quality of coffee beans. This approach is associated with elevated labor costs, time-intensive processes, and potential operational quality instability due to factors such as stress and fatigue, rendering it unsuitable for large-scale processing. Additionally, defective beans may manifest post-roasting, underscoring the necessity for quality testing of roasted coffee beans. Relying solely on manual screening of defective beans in green coffee beans cannot ensure the quality and safety of coffee beverages.

* Corresponding author. E-mail address: chingyi@mail.mcu.edu.tw

English language proofreader: Yen-Chun Hsieh

The traditional method of manually conducting quality inspections on green coffee beans can no longer meet the demands of the vast coffee bean market. Therefore, the industry has begun exploring the use of mechanical equipment for screening. While vibrating mechanisms can facilitate bulk processing by sieving beans of different sizes, their capability to screen various types of defective beans is rather limited. With advancements in industrial technology driving automation in production and quality control, some research institutions have adopted chemical composition analysis to detect varieties or origins [4-5]. Simultaneously, electronic sorting machines are used to inspect and classify defective beans. These devices evaluate the quality of coffee beans through the spectral response characteristics but face challenges due to the similar spectral response patterns among certain defective beans (such as sour beans and immature beans).

Automated Optical Inspection (AOI) is a technology that utilizes visual sensing devices to capture images of the target object's shape and color features. It then employs traditional computer vision (CV) analysis methods to detect abnormalities or defects. AOI possesses characteristics such as speed, accuracy, and repeatability, making it a cost-effective replacement for manual labor in long-term and reliable inspection operations. Given the distinct differences in shape and color between defective beans, branches, pebbles, and regular green coffee beans, integrating AOI into coffee bean quality inspection, coupled with CV and analysis for defect classification, has proven to significantly enhance production efficiency compared to traditional methods. The color sorting machine officially applied in the industry's production lines is an example of an automated mechanical sorting machine based on optical imaging. In addition to industrial color sorting machines, there have been numerous advanced research papers in academia based on AOI [6-7].

Analyzing quality through CV methods demands a profound understanding to construct a classification model. Machine learning (ML) provides a data-driven approach empowering researchers to autonomously learn sample classification from training data. For example, employing a substantial dataset comprising images of defective and high-quality coffee beans enables the development of a classification model for assessing coffee bean quality. This approach mitigates human subjectivity, surpasses the limitations of current electronic sorting machines, and yields more precise results. The integration of ML in CV technology not only enhances the effectiveness of quality inspection in industries but also addresses challenges associated with increasingly complex image quality inspection requirements.

In the realm of grain quality assessment, numerous research papers showcase the efficacy of ML methods. These studies leverage diverse ML algorithms to classify grains based on color and shape features, presenting significant contributions to the field [8-10], this underscores ML's potential and positions it as a valuable asset in the domain of food quality inspection. Noteworthy contributions in the realm of coffee bean quality inspection involve the work of Arboleda et al. [11], who extract crucial features such as area, perimeter, equivalent diameter, and roundness percentage for coffee bean quality assessment. Subsequent classification is achieved using artificial neural networks (ANN) and K-nearest neighbors (KNN) classifiers.

Over the past decade, with the rapid development of graphics processing units (GPUs), their computational capabilities extend beyond traditional matrix operations to encompass a wide range of parallelizable tasks involving multidimensional tensors. This has propelled deep learning (DL) to become the most prominent research topic, its application has been widely applied in machine vision tasks such as image classification, and object detection. Significantly influenced traditional industry technologies. The coffee bean quality detection systems based on DL algorithms are described by Luis et al. [12] and Ruttanadech et al. [13].

Green coffee beans and roasted coffee beans exhibit significant differences in color, appearance, and aroma. Currently, most research on coffee bean quality inspection, realized through CV and ML, primarily focuses on green coffee beans as referenced in [14-15]. This presents a significant gap, as roasted beans, due to their varied conditions post-roasting, and finding a large labeled dataset for roasted coffee beans is not readily available in online resources, introducing unique challenges not addressed by existing methodologies. ML thrives in scenarios with abundant labeled training instances, where training data and test data must share the same distribution.

However, in many real-world ML applications, collecting a sufficient amount of labeled training data is often time-consuming, expensive, or even impractical. Additionally, the assumption that the source domain and target domain share the same distribution does not always hold in practical applications. Discrepancies in data distribution hinder traditional ML algorithms from achieving expected results in similar but new domains, leading to insufficient generalization when the model encounters unseen domains. Transfer learning and domain adaptation (DA) [16- 17] are techniques to enhance the performance of ML models in cross-domain tasks. When the target domain lacks an abundance of labeled data, considering different but relevant labeled datasets as auxiliary data for model pre-training becomes a viable option. Subsequently, adjusting the pre-trained model and applying it to the target domain overcome the challenges of acquiring labeled data in practical applications.

In summary, DA techniques aim to learn a model that enables the knowledge acquired from the source domain to generalize effectively in the target domain. Introducing DA techniques helps mitigate differences in data distribution between the source and target domains, facilitating cross-domain transfer and reuse of domain-invariant knowledge. In this study, a DA-based domain-generalization model is introduced for roasted bean quality inspection. This model utilizes a large labeled dataset of raw beans to train an auxiliary model. Through adversarial learning, the distribution differences between raw and roasted bean features are minimized, achieving the generalization of the raw bean quality detection model to successfully handle roasted bean datasets. Thus, the scope of this system extends beyond traditional methods, offering a novel solution to a largely unaddressed challenge within the field of coffee bean quality inspection. By incorporating both raw and roasted beans into the analysis and utilizing DA techniques to address data scarcity for roasted beans, this study marks a significant advancement in the domain, setting a new precedent for future research in agricultural product quality inspection.

2. Related Approaches

The method proposed in this paper aims to train a DA-based generalized deep model. Therefore, before elaborating on the system architecture, this study briefly introduces the convolutional neural network (CNN) and provides a focused review of methods such as adversarial training, deep domain confusion, and domain-adversarial training of neural networks. These methodologies contribute significantly to achieving the objective of domain generalization.

2.1. Convolutional neural network

CNN is currently one of the mainstream deep neural networks (DNNs), known for its outstanding feature extraction capabilities with local connectivity and parameter sharing. CNN applies convolutional layers to generate feature maps, followed by a pooling layer that facilitates the downsampling of input feature maps. As illustrated in Fig. 1, a typical CNN architecture involves a forward propagation of input data. At the output stage, the network calculates the error between the expected output and the predicted output values. Subsequently, the backpropagation algorithm is employed to update the network's weights. Depending on the type, CNNs have several classic architectures, and a residual neural network (ResNet) is one of them. ResNet utilizes identity mapping to increase the network's depth, addressing the degradation problem in deep networks.

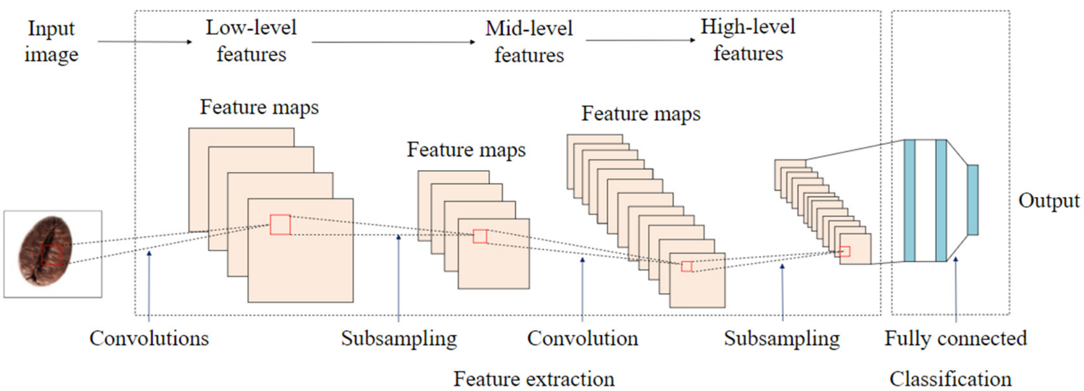


Fig. 1 Typical CNN architecture

2.2. Adversarial training

Autoencoders utilize a combination of an encoder and a decoder to reconstruct their features. The encoder compresses input features from a high-dimensional space to a low-dimensional latent space, while the decoder reconstructs the low-dimensional vector back to the original features, as shown in Fig. 2. The loss function is defined as:

$$loss = \left\| x - Dec\left(Enc(x)\right) \right\|_2 \tag{1}$$

where *Enc* represents the encoder, and *Dec* represents the decoder. Autoencoders are commonly used to learn latent representations of data.

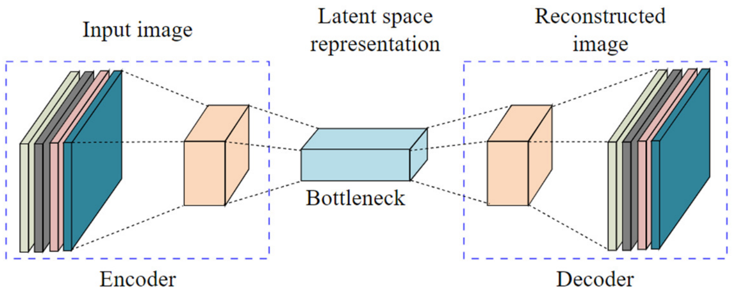


Fig. 2 Autoencoder architecture

A generative adversarial network (GAN) [18] is another common type of deep generative model consisting of a generator (G) and a discriminator (D). The generator's task is to transform samples from a prior distribution into data points, while the discriminator's task is to distinguish whether a data point is real or generated by the generator. The training involves a minimax game between these two components, with the conclusive aim of training the generator to produce data points that closely resemble real data. This adversarial process leads to improvements and enhancements in both the generator and discriminator networks. The loss functions of GAN, G, and D are represented as follows:

$$\min_G \max_D V(D,G) = E_{x\sim p_{data}}\left(\log D(x)\right) + E_{z\sim p_G}\left\{\log\left[1 - D\left(G(z)\right)\right]\right\} \tag{2}$$

$$\max_D V(D) = E_{x\sim p_{data}}\left(\log D(x)\right) + E_{z\sim p_G}\left\{\log\left[1 - D\left(G(z)\right)\right]\right\} \tag{3}$$

$$\min_G V(G) = E_{z\sim p_G}\left\{\log\left[1 - D\left(G(z)\right)\right]\right\} \tag{4}$$

where z is the noise vector, p_{data} represents the distribution of real images, and p_G represents the distribution of generated images.

2.3. Domain adaption

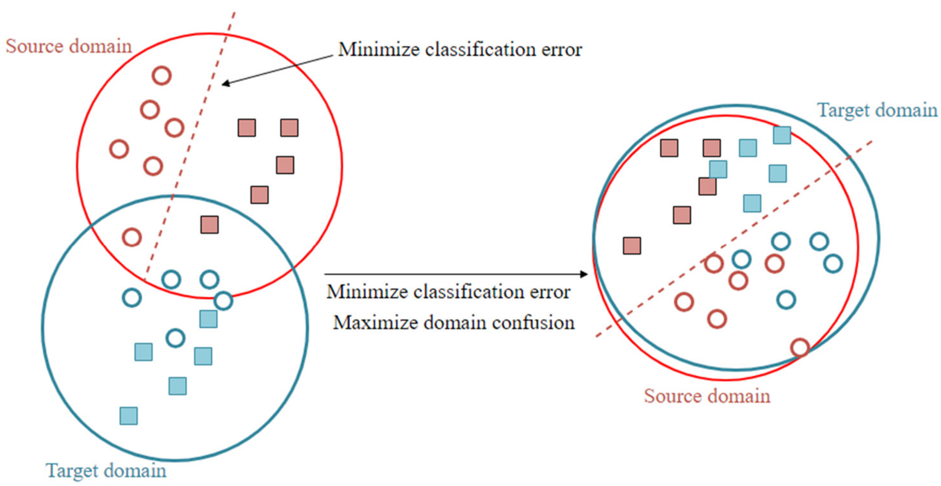


Fig. 3 DA schematic diagram [19]

For supervised learning, DNNs with abundant labeled data often achieve superior performance. However, labeling data typically requires human involvement and relevant expertise, making the process time-consuming, labor-intensive, and expensive. Moreover, such models exhibit weaker adaptability to new environments and tasks. When faced with environments different from the training scenarios, new data must be labeled, and the learning model needs to be retrained. DA techniques address the scarcity of labeled data and enhance model generalization by leveraging related but different source domains to assist learning in the target domain. The DA approach maps source and target domains with different distributions into a common feature space, aiming to bring data from the same category as close as possible in this space while separating data from different categories, as shown in Fig. 3. In practice, there is often a large labeled source domain dataset, while the target domain data may have few or even no labeled instances.

2.4. *Deep domain confusion*

The network architecture employed by deep domain confusion is illustrated in Fig. 4. The first network takes the labeled source dataset as input, while the other network takes the target dataset, which includes a small amount of labeled data or unlabeled data. Both CNNs share weight values. The final loss function is a weighted combination of classification loss and domain loss. The domain loss utilizes the maximum mean discrepancy (MMD) as a metric method to calculate, aiming to obtain domain-invariant features and reduce the distribution gap between the source dataset and the target dataset.

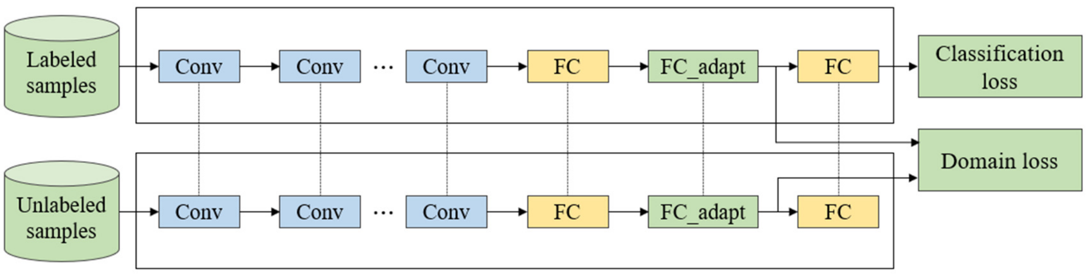


Fig. 4 Deep domain confusion architecture [20]

2.5. *Domain-adversarial training of neural networks* (*DANN*)

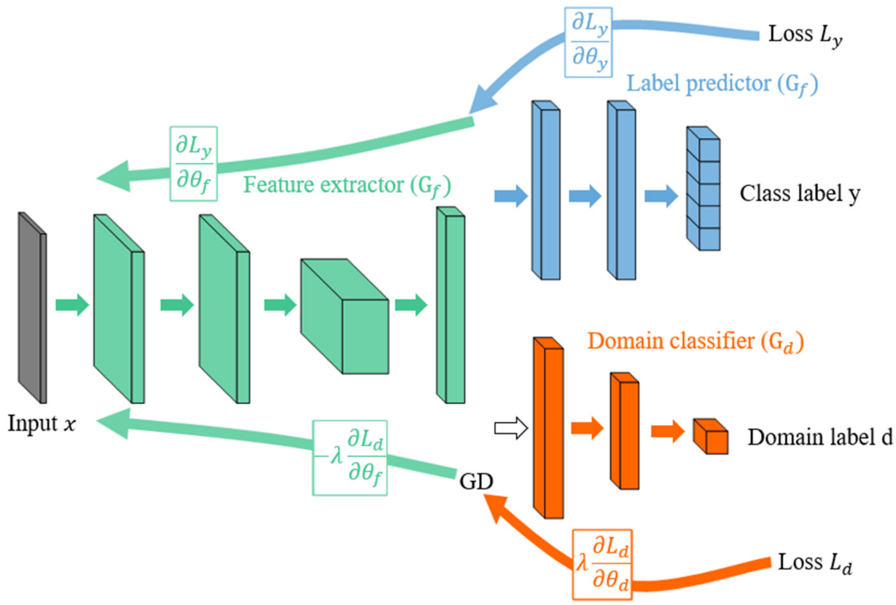


Fig. 5 Domain-adversarial training of neural networks [17]

For many practical applications, the data distributions of the source domain and target domain differ, and obtaining labeled data for the target domain can be challenging or scarce. The fundamental idea of DA is to map both the source and target domain data into the same feature space and find a metric criterion in this space to make the feature distributions of the source and target domain data as close as possible. Consequently, a classifier trained based on the features of the source domain data can be applied to the target domain data. DANN introduces adversarial learning to DA problems, consisting of three network components: feature extractor (G_f), label predictor (G_y), and domain classifier (G_d), as illustrated in Fig. 5. The optimization objective for G_d is to minimize the classification errors of the domain classifier, while the optimization objective for G_f is to maximize the classification errors of the domain classifier. To simultaneously satisfy these opposing optimization objectives during training, DANN utilizes a gradient reversal layer, ensuring that the overall network guarantees the minimization of domain confusion loss by the domain classifier and the maximization of domain confusion loss by the feature extractor.

3. DA-Based Roasted Coffee Bean Quality Detector

In the realm of coffee quality inspection, it is common to use a large amount of labeled data from green coffee beans for training, while roasted coffee beans suffer from an insufficiency of adequately labeled data and efficient classifiers. Furthermore, although roasted coffee beans share certain characteristics with green coffee beans, their color and appearance differ significantly, preventing the direct application of classifiers trained on green coffee beans to quality inspection in roasted coffee beans. Enter DA, a conceptual framework aimed at harmonizing data from both the source and target domains within a shared feature space. The key lies in defining a metric criterion within this feature space to align the feature distributions of source and target domain data as closely as possible. By achieving this harmonization, a classifier initially trained on source domain data becomes versatile enough to extend its application to the classification of target domain data.

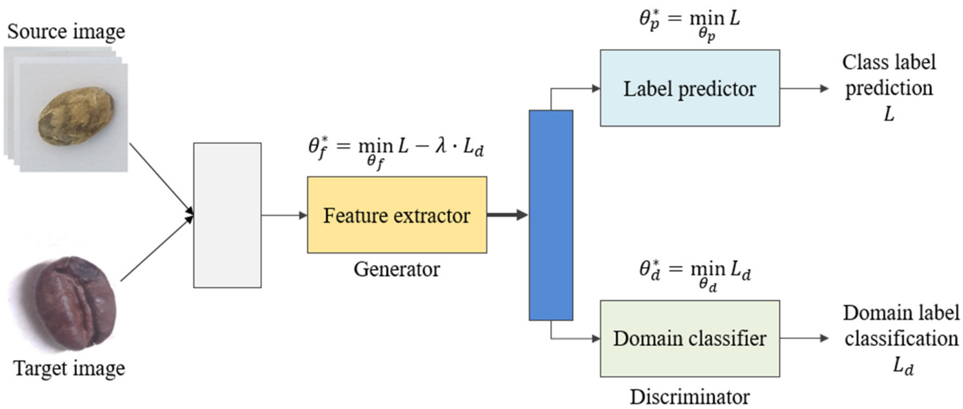


Fig. 6 DA system flowchart

In pursuit of enabling effective quality inspection in roasted coffee beans, the study introduces a pioneering coffee bean quality inspection system model based on DA. The holistic architecture comprises pivotal modules, including a feature extractor, label predictor, and domain classifier. The feature extractor plays a pivotal role in distilling common features from images of both domains. Subsequently, the label predictor adeptly classifies these extracted common features, while the domain classifier discerns the origin domain of the extracted features. The ingenious integration of adversarial learning into the system's architecture serves the dual purpose of maximizing label prediction accuracy and minimizing accuracy loss in domain classification.

This strategic amalgamation empowers the system to transcend the traditional boundaries, enabling classifiers initially trained on green coffee beans to seamlessly adapt to the nuanced task of quality inspection in roasted coffee beans. The conceptual underpinning and operational dynamics of this innovative model are vividly illustrated in Fig. 6.

3.1. Feature extractor & label predictor

The original spatial information in the image contains a large amount of redundant data, which can be transformed into a low-dimensional feature vector. The feature extractor composed of a DNN is responsible for transforming the raw input image into a lower-dimensional feature representation. These representations capture fundamental patterns and features relevant to the ongoing task, enhancing key characteristics while preserving the accuracy and integrity of the original data description. Extracting good representations not only contributes to subsequent learning steps and generalization capabilities but also, in some cases, aids in the interpretability of artificial intelligence (AI) models.

The key role of the feature extractor in Fig. 6 is to make the feature points of the source domain and target domain indistinguishable in distribution. The optimization objective of the parameter set θ_f^* of the feature extractor is to minimize the total loss function derived from both the weighted L and L_d, as shown below:

$$\theta_f^* = \min_{\theta_f}\left(L - \lambda \cdot L_d\right) \tag{5}$$

The label predictor, also known as a task classifier or label classifier, primarily functions to predict the labels or categories corresponding to input data. In contrast to traditional supervised learning where labels are only used for training and evaluation, in DA training, the label predictor is an integral part of the model architecture. Its predictions are used not only for the main task but also in the DA process. The label predictor is typically trained using annotated data from the source domain, and its learning process maps the feature representations generated by the feature extractor to the correct class labels [21].

In the training process of DA, the label predictor's training is crucial not only for the main task to be addressed but also indirectly contributes to the adversarial training objective. θ_p^* represents the optimal parameter set of the label predictor. L represents the classification cross-entropy for the source domain in the label predictor. The goal of this optimization process is to find the optimal parameter settings for the label predictor, achieving accurate predictions in label prediction.

$$\theta_p^* = \min_{\theta_p} L \tag{6}$$

3.2. Domain classifier

The domain classifier is a binary classifier with the primary function of assessing the domain from which the feature representation generated by the feature extractor originates [21]. Collaborating with the feature extractor, it aims to minimize specific domain confusion loss (L_d). Through adversarial learning between the domain classifier and the feature extractor, domain-invariant cross-domain representation learning can be achieved, thereby fulfilling the goal of DA.

During the training process, the parameter set θ_d of the domain classifier gradually adjusts to enhance its ability to correctly classify the source domain of the features generated by the feature extractor [21]. Simultaneously, the feature extractor is adjusted to ensure the generation of feature representations that are difficult to distinguish between two different domains. θ_d^* represents the optimal parameter set of the domain classifier. The optimization goal is to minimize the domain confusion loss (L_d) by adjusting the parameter set of the domain classifier, aiming to enhance the domain classifier's ability to correctly identify the domain from which the feature representation generated by the feature extractor originates.

$$\theta_d^* = \min_{\theta_d} L_d \tag{7}$$

3.3. Domain-invariant representation learning

The representation extracted by the feature extractor is passed to the domain classifier, which then determines whether the incoming representation originates from the source domain or the target domain, subsequently computing the loss. The training objective of the domain classifier is to classify the incoming representation into the correct domain category, while the training objective of the feature extractor is precisely the opposite. It aims to extract features that are sufficient to confuse the domain classifier, making it unable to identify from which domain the output representation comes. This creates an adversarial relationship, as illustrated in Fig. 7, outlining the distinct objectives of each module. In contrast to the traditional approach seen in GANs, DA problems deviate from the process of generating samples using random noise, as illustrated in Fig. 8. Instead, DA directly considers the data within the target domain as the generated samples. This departure signifies a shift in the role of the generator, no longer focused on generating samples but now serving as a crucial feature extraction component in the adaptation process.

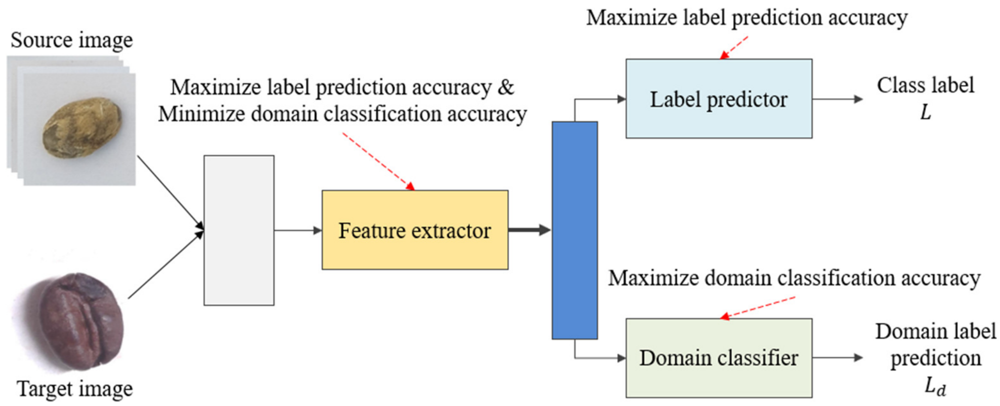


Fig. 7 Leveraging distinct module objectives for domain-invariant representation learning

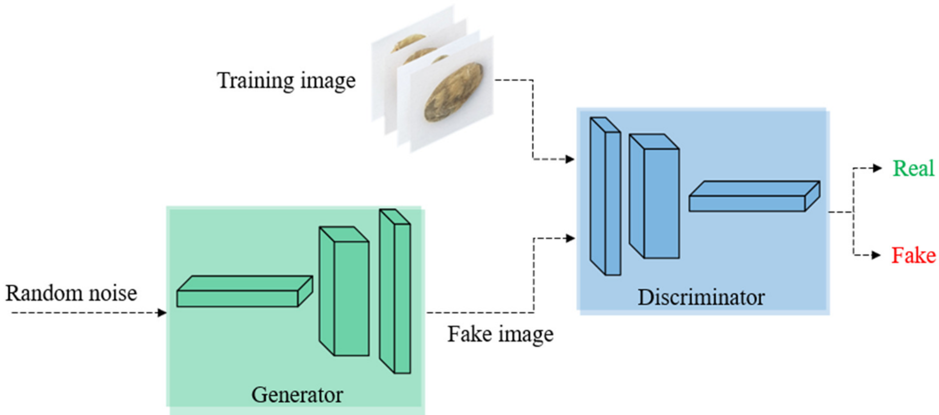


Fig. 8 The process of generating samples by GAN

4. Experimental Results and Discussion

This study utilized PyTorch 1.12.0 for the development of DL models. The processor used in the computer is a Core(TM) i5-12400F CPU, 16GB DDR4 RAM, and is complemented by an NVIDIA GeForce RTX 3060 graphics card to enhance computational capabilities.

4.1. Coffee bean dataset

In the development of the coffee bean quality inspection system, the dataset was meticulously curated to ensure robust model training and evaluation. The dataset comprised 4626 green coffee bean images, obtained from coffee-cobra [22], designated as the source domain, and 886 roasted coffee bean images, obtained from aistudio-Baidu [23], serving as the target domain. Both datasets featured images with a resolution of 400 × 400 pixels. To facilitate a comprehensive evaluation of the model's performance, the initial dataset was divided into distinct subsets for training and testing. The training subset, crucial for model learning, constituted approximately 80% of the green coffee bean images and 80% of the roasted coffee bean images. The testing subset comprised the remaining 20% of images, enabling an unbiased assessment of the model's generalization capability on unseen data. The division of the dataset was performed to ensure a balanced representation of various bean qualities within each subset, thereby mitigating potential biases and facilitating a comprehensive evaluation across different stages of model development. The composition of both datasets is visualized in Fig. 9.

(a) Good green bean

(b) Bad green bean

(c) Good roasted bean

(d) Bad roasted bean

Fig. 9 Partial image in the dataset

4.2. *Coffee bean quality inspection based on domain adaptation*

This section delves into a comprehensive examination of domain adaptation techniques applied to coffee bean quality inspection. The analytical endeavors are bifurcated into two principal subsections. The first assesses the classification accuracy of the proposed method across varying values of λ within different domains. The second evaluates the distribution of features between the source and target domains to ascertain the effectiveness of the domain adaptation strategy.

4.2.1. *Analysis of classification accuracy*

Table 1 Efficiency comparison of each level to control the balance of DA and classification

λ	Target data accuracy	Source data accuracy
0.0	65.16%	98.52%
0.1	73.33%	97.52%
0.2	80.60%	96.45%
0.3	77.74%	95.25%
0.4	73.55%	94.54%
0.5	83.15%	93.23%
0.6	79.45%	93.60%
0.7	85.16%	94.55%
0.8	**92.24%**	**98.96%**

In this endeavor, a coffee bean quality inspection system was developed utilizing EfficientNetV2-S [24], and the λ value in Eq. (5) was fine-tuned to dynamically adjust the balance between L and L_d, facilitating adversarial learning between the domain classifier and feature extractor. Without accounting for domain confusion loss ($\lambda = 0$), the system exhibited a 98.52% true positive rate for raw beans after training on the labeled raw bean dataset. However, when encountering a domain shift with roasted beans, the accuracy decreased to 65.16%. Incrementally adjusting the λ value, indicating an increased focus on L_d, led to improvements in the system's accuracy for target domain data. Notably, with λ set to 0.8, the system achieved a 98.96% true positive rate for raw beans, and the correct classification rate for roasted beans improved from 65.16% to 92.24%, as meticulously detailed in Table 1.

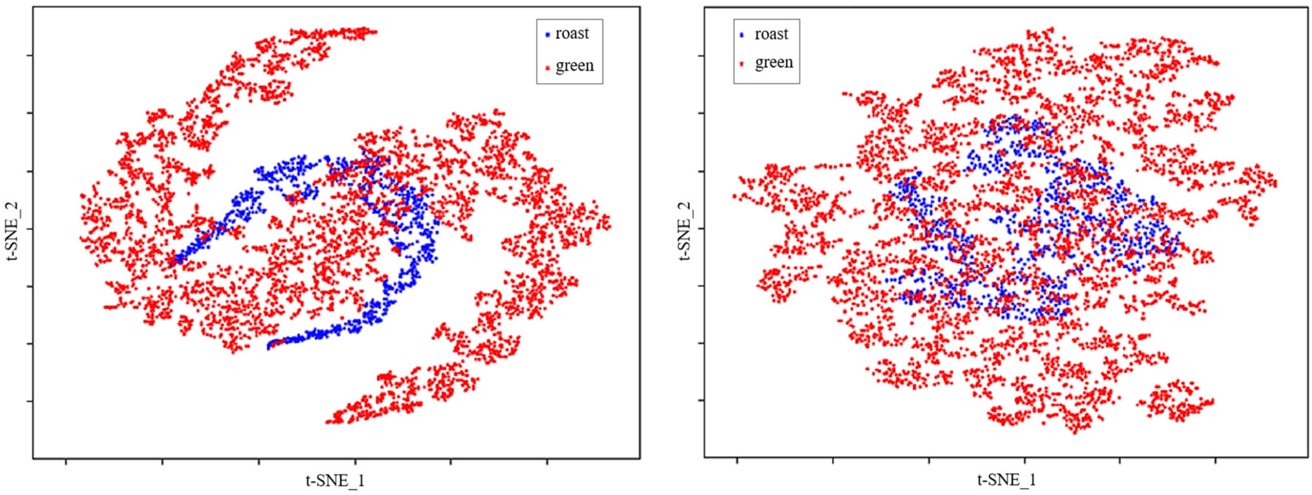
In line with the experiments outlined in Table 1, the evaluation was extended to various deep-learning models as backbone networks for the coffee bean quality detector. Among these, EfficientNetV2-S distinguished itself as the top performer, as detailed in Table 2. This superiority is attributed to its innovative architecture that optimally balances depth, width, and resolution through compound scaling, allowing it to efficiently process the complexity of coffee bean images. Additionally, EfficientNetV2-S's advanced regularization techniques and efficient convolutional blocks significantly enhance its capability to discern intricate details, further validating its selection for accurately assessing coffee bean quality among the diverse models considered.

Table 2 Efficiency comparison of different deep learning models

Models	Accuracy	Precision	Recall	F1 Score
VGG	84.99%	84.01%	91.15%	86.74%
Xception	85.71%	86.01%	85.75%	86.31%
ResNet18	87.52%	87.07%	94.13%	89.47%
DenseNet121	88.51%	89.03%	88.61%	89.21%
MobileNetV3small	90.72%	90.24%	90.83%	90.40%
EfficientNetV2-S	**92.24%**	**92.34%**	**92.24%**	**91.43%**

4.2.2. Analysis of feature distribution results between source and target domains

To validate the effectiveness of DA in roasted coffee bean quality inspection, the t-distributed stochastic neighbor embedding (t-SNE) method was employed to reduce the dimensionality of feature vectors output by the feature extractor. This projection transformed the high-dimensional data into a two-dimensional feature space, facilitating the visualization of DA experimental results. Fig. 10 illustrates the visualization results after applying t-SNE for data dimensionality reduction. Fig. 10(a) displays the distribution of feature points for green and roasted coffee beans before DA. Significant differences in the distribution of green and roasted coffee beans were observed. In Fig. 10(b), the system's feature points for raw and roasted beans overlap in the two-dimensional space after DA. This figure presents the distribution of coffee beans of different categories in the two-dimensional space. In this scenario, knowledge acquired from the source domain can be generalized better in the target domain, achieving the cross-domain transfer and reuse of domain-invariant knowledge.



(a) Before DA (b) After applying DA

Fig. 10 The distribution of feature points for raw and roasted beans in a two-dimensional space

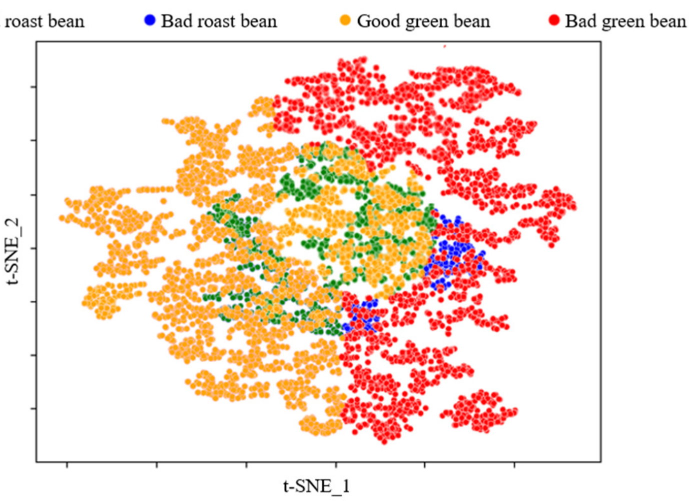
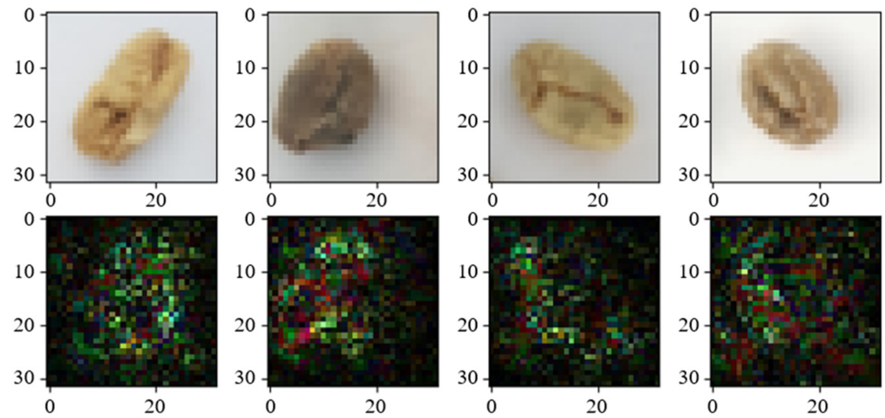


Fig. 11 Distribution analysis of features for different categories

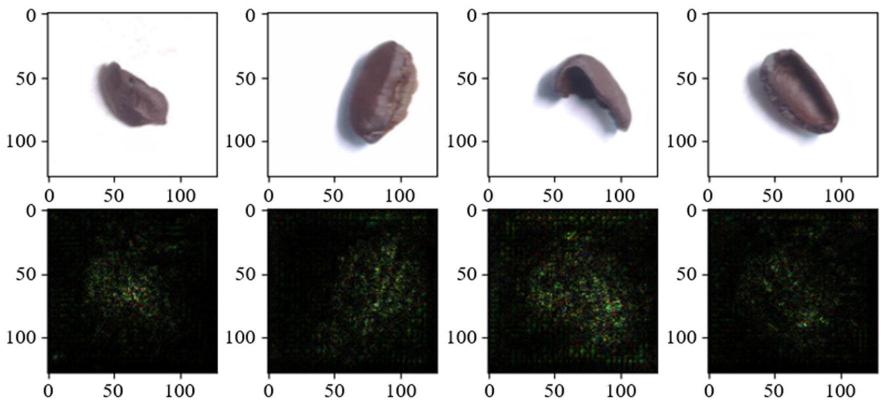
As shown in Fig. 11, the feature distribution chart illustrates the characteristics of good beans and defective beans. The orange dots represent the good bean category for green coffee beans, while the red dots represent the defective bean category for green coffee beans. These two clusters are distinctly separated in the feature space. As for the green dots in the figure, they represent the good bean category for roasted coffee beans, and the blue dots represent the defective bean category for roasted coffee beans. Similarly, these two clusters in the feature space also exhibit clear separation. It is noteworthy that the red dots and blue dots belonging to the defective bean category overlap in a large cluster, and the orange dots and green dots belonging to the good bean category similarly overlap in another large cluster. There is a clear boundary between these two large clusters, indicating that the feature extractor of the proposed architecture possesses excellent feature extraction capabilities. It can extract well-distinguished features, demonstrating a clear distinction between the two categories.

4.3. Analysis of explainable models

While DNN models have been successfully applied across diverse applications, such as intelligent detection and medical diagnosis services, the inherent opacity of AI poses a challenge. The machine's inability to explain the decision-making process and the factors behind its actions affects users' trust and confidence in its capabilities. To address this issue, saliency maps, as illustrated by Brocki and Chung [25], offer a means of visualizing the significance score assigned to each pixel in an image, presenting it as a heatmap. In this representation, pixels with higher scores manifest in brighter colors, while lower-scored pixels appear darker. Leveraging saliency maps allows for the interpretability of the model, providing valuable insights into the critical information within an image.



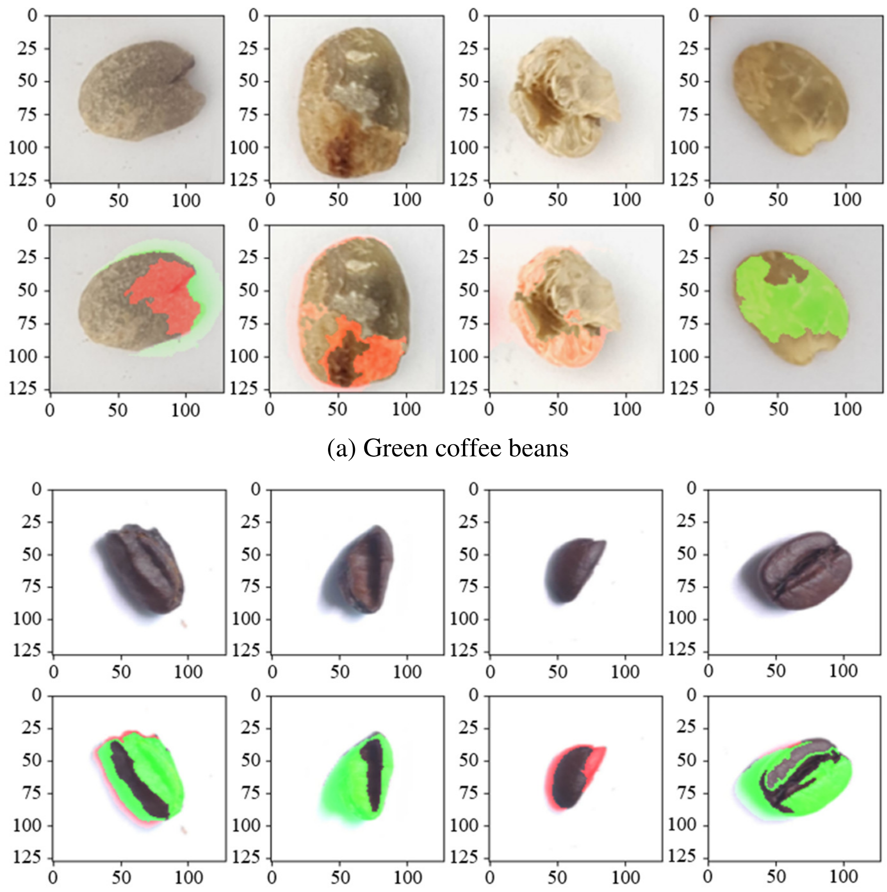
(a) Green coffee beans



(b) Roasted coffee beans

Fig. 12 Saliency maps

Fig. 12(a) and Fig. 12(b), display the original images of green and roasted coffee beans, along with their corresponding saliency maps. These saliency maps highlight potential defective areas in the form of green dots, offering a visual way to emphasize these problematic regions. In the Saliency map, green areas represent regions where the model has high confidence in its predictions, implying accuracy in these predicted areas. Conversely, red areas indicate regions where the model considers there might be poorer predictions, potentially related to lower-quality areas of the coffee beans. These explanatory visualizations provide a clearer insight into the model's confidence in predictions for various regions, allowing us to focus on areas where potential issues may arise.



(a) Green coffee beans

(b) Roasted coffee beans

Fig. 13 Generating visual explanation maps using LIME

Local interpretable model-agnostic explanation (LIME) [26], categorized under global explanation, utilizes simple regression or linear models to locally approximate the prediction results of a target black-box complex model, aiming to achieve interpretable models. LIME is a rapid method for analyzing the contribution of each feature, providing human-interpretable explanations, thus aiding in understanding the decision behavior of ML models. As shown in Fig. 13(a) and Fig. 13(b), LIME was employed to visualize explanations for the approach, revealing that the model emphasizes identifying distinguishing clues from the internal regions of coffee beans, rather than relying on the background of specific areas. The added clarification explains that green areas in the figures indicate positive influences, or features supporting the model's decision, while red areas signify negative influences, pointing towards alternate classifications. This color-coded approach offers deeper insights into the model's decision process by pinpointing crucial influencing features.

5. Conclusions

In this study, DA techniques are utilized to formulate models with domain-generalization capabilities tailored for the inspection of roasted coffee bean quality. The model leverages an extensive labeled dataset of green coffee beans to train an auxiliary model, and through adversarial learning, systematically reduces the distributional disparities in features between green and roasted coffee beans. Experimental results illustrate a significant improvement in the classification accuracy for roasted coffee beans, escalating from an initial 65.16% to 92.24%. This achievement underscores the proficiency of the approach in broadening the scope of green coffee bean quality inspection models, effectively accommodating datasets that include roasted coffee beans. Future research is set to delve into sophisticated object detection and instance segmentation methodologies, such as those within the YOLO series or employing mask R-CNN techniques. Such methods show potential for advancing the system's ability to assess multiple beans simultaneously and accurately evaluate partially occluded beans. The integration of these complex techniques is expected to address current limitations, enhancing the robustness and practical utility of the quality inspection system across a range of scenarios.

Conflicts of Interest

The authors declare no conflict of interest.

References

[1] J. Yu, D. Liang, J. Li, Z. Liu, F. Zhou, T. Wang, et al., "Coffee, Green Tea Intake, and the Risk of Hepatocellular Carcinoma: A Systematic Review and Meta-Analysis of Observational Studies," Nutrition and Cancer, vol. 75, no. 5, pp. 1295-1308, 2023.

[2] S. Shin, J. E. Lee, E. Loftfield, X. O. Shu, S. K. Abe, M. S. Rahman, et al., "Coffee and Tea Consumption and Mortality from All Causes, Cardiovascular Disease and Cancer: A Pooled Analysis of Prospective Studies from the Asia Cohort Consortium," International Journal of Epidemiology, vol. 51, no. 2, pp. 626-640, April 2022.

[3] K. N. Papadopoulos, Food Chemistry Research Developments, New York: Nova Science Publishers, pp. 105-138, 2008.

[4] L. Servillo, A. Giovane, R. Casale, D. Cautela, N. D'Onofrio, M. L. Balestrieri, et al., "Glucosylated Forms of Serotonin and Tryptophan in Green Coffee Beans," LWT, vol. 73, pp. 117-122, November 2016.

[5] B. Mehari, M. Redi-Abshiro, B. S. Chandravanshi, S. Combrinck, M. Atlabachew, and R. McCrindle, "Profiling of Phenolic Compounds Using UPLC–MS for Determining the Geographical Origin of Green Coffee Beans from Ethiopia," Journal of Food Composition and Analysis, vol. 45, pp. 16-25, February 2016.

[6] E. R. Arboleda, A. C. Fajardo, and R. P. Medina, "An Image Processing Technique for Coffee Black Beans Identification," IEEE International Conference on Innovative Research and Development, pp. 1-5, May 2018.

[7] Radi, M. Rivai, and M. H. Purnomo, "Combination of First and Second Order Statistical Features of Bulk Grain Image for Quality Grade Estimation of Green Coffee Bean," ARPN Journal of Engineering and Applied Sciences, vol. 10, no. 18, pp. 8165-8174, October 2015.

[8] D. Mandal, "Adaptive Neuro-Fuzzy Inference System Based Grading of Basmati Rice Grains Using Image Processing Technique," Applied System Innovation, vol. 1, no. 2, article no. 19, June 2018.

[9] E. Ordukaya and B. Karlik, "Quality Control of Olive Oils Using Machine Learning and Electronic Nose," Journal of Food Quality, vol. 2017, article no. 9272404, 2017.

[10] H. Zareiforoush, S. Minaei, M. R. Alizadeh, and A. Banakar, "Qualitative Classification of Milled Rice Grains Using Computer Vision and Metaheuristic Techniques," Journal of Food Science and Technology, vol. 53, no. 1, pp. 118-131, January 2016.

[11] E. R. Arboleda, A. C. Fajardo, and R. P. Medina, "Classification of Coffee Bean Species Using Image Processing, Artificial Neural Network and K Nearest Neighbors," IEEE International Conference on Innovative Research and Development, pp. 1-5, May 2018.

[12] V. A. M. Luis, M. V. T. Quiñones, and A. N. Yumang, "Classification of Defects in Robusta Green Coffee Beans Using YOLO," IEEE International Conference on Artificial Intelligence in Engineering and Technology, pp. 1-6, September 2022.

[13] N. Ruttanadech, K. Phetpan, N. Srisang, S. Srisang, T. Chungcharoen, W. Limmun, et al., "Rapid and Accurate Classification of Aspergillus Ochraceous Contamination in Robusta Green Coffee Bean Through Near-Infrared Spectral Analysis Using Machine Learning," Food Control, vol. 145, article no. 109446, March 2023.

[14] M. García, J. E. Candelo-Becerra, and F. E. Hoyos, "Quality and Defect Inspection of Green Coffee Beans Using a Computer Vision System," Applied Sciences, vol. 9, no. 19, article no. 4195, October 2019.

[15] Y. Chen, B. Gao, and W. Lu, "Recent Research Advancements of Coffee Quality Detection: Targeted Analyses vs. Nontargeted Fingerprinting and Related Issues," Journal of Food Quality, vol. 2023, article no. 6156247, 2023.

[16] Y. Kim, D. Cho, K. Han, P. Panda, and S. Hong, "Domain Adaptation Without Source Data," IEEE Transactions on Artificial Intelligence, vol. 2, no. 6, pp. 508-518, December 2021.

[17] M. HassanPour Zonoozi and V. Seydi, "A Survey on Adversarial Domain Adaptation," Neural Processing Letters, vol. 55, no. 3, pp. 2429-2469, June 2023.

[18] G. Iglesias, E. Talavera, and A. Díaz-Álvarez, “A Survey on GANs for Computer Vision: Recent Research, Analysis and Taxonomy,” Computer Science Review, vol. 48, article no. 100553, May 2023.
[19] E. Tzeng, J. Hoffman, N. Zhang, K. Saenko, and T. Darrell, “Deep Domain Confusion: Maximizing for Domain Invariance,” https://doi.org/10.48550/arXiv.1412.3474, December 10, 2014.
[20] F. Yu, X. Xiu, and Y. Li, “A Survey on Deep Transfer Learning and Beyond,” Mathematics, vol. 10, no. 19, article no. 3619, October 2022.
[21] M. Martini, V. Mazzia, A. Khaliq, and M. Chiaberge, “Domain-Adversarial Training of Self-Attention-Based Networks for Land Cover Classification Using Multi-Temporal Sentinel-2 Satellite Imagery,” Remote Sensing, vol. 13, no. 13, article no. 2564, July 2021.
[22] “Small Optical Sorter,” https://github.com/edgeryders/coffee-cobra, March 26, 2021.
[23] “Coffee Bean Defect Detection,” https://aistudio.baidu.com/datasetdetail/150459, June 04, 2022.
[24] M. Tan and Q. Le, “Efficientnetv2: Smaller Models and Faster Training,” International Conference on Machine Learning, vol. 139, pp. 10096-10106, July 2021.
[25] L. Brocki and N. C. Chung, “Input Bias in Rectified Gradients and Modified Saliency Maps,” IEEE International Conference on Big Data and Smart Computing, pp. 148-151, January 2021.
[26] J. Hu, K. Zhu, S. Cheng, N. M. Kovalchuk, A. Soulsby, M. J. H. Simmons, et al., “Explainable AI Models for Predicting Drop Coalescence in Microfluidics Device,” Chemical Engineering Journal, vol. 481, article no. 148465, February 2024.