Enhancing AI microscopy for foodborne bacterial classification using adversarial domain adaptation to address optical and biological variability

Siddhartha Bhattacharya[1,2], Aarham Wasit[1,2], J Mason Earles[3,4], Nitin Nitin[3,5] and Jiyoon Yi[1]*

[1]Department of Biosystems and Agricultural Engineering, Michigan State University, East Lansing, MI, United States, [2]Department of Computer Science and Engineering, Michigan State University, East Lansing, MI, United States, [3]Department of Biological and Agricultural Engineering, University of California, Davis, Davis, CA, United States, [4]Department of Viticulture and Enology, University of California, Davis, Davis, CA, United States, [5]Department of Food Science and Technology, University of California, Davis, Davis, CA, United States

AI-enabled microscopy is emerging for rapid bacterial classification, yet its utility remains limited in dynamic or resource-limited settings due to imaging variability. This study aims to enhance the generalizability of AI microscopy using domain adaptation techniques. Six bacterial species, including three Gram-positive (*Bacillus coagulans*, *Bacillus subtilis*, *Listeria innocua*) and three Gram-negative (*Escherichia coli*, *Salmonella* Enteritidis, *Salmonella* Typhimurium), were grown into microcolonies on soft tryptic soy agar plates at 37°C for 3–5 h. Images were acquired under varying microscopy modalities and magnifications. Domain-adversarial neural networks (DANNs) addressed single-target domain variations and multi-DANNs (MDANNs) handled multiple domains simultaneously. EfficientNetV2 backbone provided fine-grained feature extraction suitable for small targets, with few-shot learning enhancing scalability in data-limited domains. The source domain contained 105 images per species (*n* = 630) collected under optimal conditions (phase contrast, 60× magnification, 3-h incubation). Target domains introduced variations in modality (brightfield, BF), lower magnification (20×), and extended incubation (20x-5h), each with <5 labeled training images per species (*n* ≤ 30) and test datasets of 60–90 images. DANNs improved target domain classification accuracy by up to 54.5% for 20× (34.4% to 88.9%), 43.3% for 20x-5h (40.0% to 83.3%), and 31.7% for BF (43.4% to 73.3%), with minimal accuracy loss in the source domain. MDANNs further improved accuracy in the BF domain from 73.3% to 76.7%. Feature visualizations by Grad-CAM and t-SNE validated the model's ability to learn domain-invariant features across conditions. This study presents a scalable and adaptable framework for bacterial classification, extending the utility of microscopy to decentralized and resource-limited settings where imaging variability often challenges performance.

1 Introduction

Rapid detection and identification of foodborne bacteria, including both pathogenic and spoilage species, are essential for ensuring food safety and quality. However, traditional methods rely on time-intensive culture-based approaches, requiring prolonged incubation to form visible bacterial colonies (Ferone et al., 2020). These methods are further constrained by the need for selective media tailored to specific pathogens, limiting their scope and generalizability (Ferone et al., 2020; Qiu et al., 2021). Such challenges not only delay the implementation of corrective actions but also increase the risk of outbreaks, product recalls, and economic losses across the food supply chain (Qiu et al., 2021; Hoffmann et al., 2024). Addressing these limitations requires innovative approaches that combine speed, accuracy, scalability, and minimized resource demands in bacterial classification.

Artificial intelligence (AI)-enabled microscopy has emerged as a promising solution, integrating deep learning for rapid analysis of microscopic patterns captured through quick imaging snapshots. Our previous work demonstrated that convolutional neural networks (CNNs) can classify bacteria at the microcolony stage, significantly reducing testing time compared to traditional methods that require lengthy enrichment and full incubation (Ma et al., 2023). However, this approach relied on an existing CNN architecture optimized for computational performance on generic datasets rather than addressing the unique challenges of microscopic imaging. Moreover, model training was conducted on datasets collected under controlled laboratory conditions, limiting its generalizability to real-world scenarios characterized by optical and biological variability. Similarly, other early efforts in AI-enabled microscopy relied on small datasets and generic architectures, increasing the risk of overfitting and restricted their applicability to diverse imaging conditions (Melanthota et al., 2022; Wu and Gadsden, 2023). These constraints highlight the critical need for tailored architectural advancements and methods that enhance robustness to variability.

Recent advancements in CNN architecture design and data-centric techniques have shown potential to address these challenges. EfficientNet, for example, employs a compound scaling approach that systematically balances depth, width, and resolution, enabling the capture of fine-grained morphological features in small targets while minimizing computational cost (Tan and Le, 2019). EfficientNet variants have been applied to identify cellular defects and detect blood cells in hematology, showing their adaptability to small targets within cell imaging contexts (Otamendi et al., 2021; Xu et al., 2022). The newer EfficientNetV2 incorporates faster training and enhanced regularization, making it even more suitable for tasks requiring subtle or fine-grained feature extraction on limited datasets (Tan and Le, 2021). Additionally, image augmentation techniques have proven effective in improving bacterial classification and generalization across growth stages in clinical microscopy by simulating variability inherent to imaging conditions (Chin et al., 2024; Jeckel and Drescher, 2020). Together, these architectural innovations and data augmentation strategies address some limitations of previous approaches, enabling more effective analysis of small bacterial targets.

Beyond architectural advancements, domain adaptation techniques have been increasingly explored to address variability in optical setups and biological conditions (Tomczak et al., 2021). Traditional deep learning models often assume that training and testing data follow the same probability distribution. However, real-world applications like bacterial image classification encounter significant distribution shifts due to variations in microscopes, imaging conditions, or bacterial growth conditions across laboratories and testing sites. Collecting and annotating large datasets for every new environment is both impractical and expensive for biological images. Domain adaptation addresses this challenge by enhancing model robustness, enabling generalization to diverse setups without extensive data annotation. Previous studies applying domain adaptation in biomedical imaging have shown enhanced generalization across imaging modalities and biological conditions for tasks such as image classification and cell segmentation (Tomczak et al., 2021; Xing et al., 2021). Such techniques have shown promise for addressing variability while reducing reliance on annotated datasets.

Thus, this study aims to develop a domain-adaptive image classification model that generalizes from controlled laboratory conditions (*source domain*) to variable laboratory conditions (*target domains*), with only a few labeled examples. The objectives are to i) enhance the generalizability of AI-enabled microscopy for bacterial classification using adversarial domain adaptation and image augmentation, and ii) compare the performance of classical single-target domain adaptation with multi-domain adaptation for simultaneous generalization across multiple domains. To achieve these goals, domain-adversarial neural networks (DANNs) and multi-domain adversarial neural networks (MDANNs) were employed to learn shared, domain-invariant feature representations optimized through adversarial training. By implementing these advancements with as few as 1–5 labeled samples per bacterial species, this study effectively addresses challenges posed by natural biological variability and domain shifts in microscopic imaging.

2 Materials and methods

2.1 Source and target domain conditions

In domain adaptation, the *source domain* is the dataset used for training, collected under controlled conditions, while *target domains* are datasets with variations used to test the model's generalizability. The source domain ("PC") dataset, primarily derived from our previous work (Ma et al., 2023), was collected under controlled laboratory conditions using phase contrast microscopy at 60× magnification with microcolonies incubated for 3 h. These conditions were identified as optimal for bacterial microcolony imaging in our prior study. To evaluate generalizability in this study, additional datasets were collected under varying conditions simulating optical setups and microbial variability. These target domains included: i) brightfield microscopy at 60× magnification ("BF" domain), representing lower-contrast imaging often used in resource-limited setups; ii) phase contrast microscopy at a lower magnification of 20× (i.e., "20×" domain), using more accessible, less specialized equipment that trades resolution for broader applicability; and iii) phase contrast microscopy at a 20× magnification with an extended

TABLE 1 Source and target domains with laboratory conditions.

Abbreviation	Domain type	Microscopy modality	Magnification	Incubation time
PC	Source	Phase contrast	60×	3 h
BF	Target	Brightfield	60×	3 h
20 ×	Target	Phase contrast	20×	3 h
20x-5h	Target	Phase contrast	20×	5 h

incubation time of 5 h (i.e., "20x-5h" domain), capturing the additional biological variability introduced by longer growth periods, which might arise from deviations in protocol or the need to enhance detectability with lower magnification. A summary of laboratory conditions for all domains is presented in Table 1.

2.2 Data collection

2.2.1 Bacterial strains

To enable controlled evaluation of domain adaptation across biologically meaningful variation, six foodborne bacterial species were selected to span three Gram-positive (*Bacillus coagulans*, *Bacillus subtilis*, *Listeria innocua*) and three Gram-negative strains (*Escherichia coli*, *Salmonella enterica* serovar Enteritidis, *Salmonella enterica* serovar Typhimurium), reflecting distinct surface structures and microcolony morphologies relevant to microscopy-based classification. The selection included both pathogenic and spoilage organisms commonly encountered in food safety contexts. Species were drawn from multiple genera, with two *Bacillus* species and two *Salmonella* serovars included to capture within-genus variability at the species and serovar levels. All bacterial strains were stored in tryptic soy broth (TSB, Sigma-Aldrich, St. Louis, MO, USA) supplemented with 15% v/v glycerol at -80°C. Prior to experimentation, a bacterial glycerol stock was streaked onto a tryptic soy agar (TSA, Sigma-Aldrich) plate and incubated for 24 h. Subsequently, a single bacterial colony was transferred from the TSA plate to 10 mL of TSB, followed by overnight shaking at 175 rpm. With the exception of *Pseudomonas fluorescens*, which was incubated at 30°C, all strains were incubated at 37°C. The fresh overnight culture was then diluted with sterile phosphate-buffered saline (PBS, Fisher Scientific, Pittsburg, CA, USA) to obtain specific concentrations for microcolony cultivation.

2.2.2 Microcolony cultivation and microscopy

Bacterial microcolonies were formed following our previously published method (Ma et al., 2023). Briefly, 1 mL of bacterial suspension was deposited onto soft TSA plates (0.7% w/v agarose) and incubated at 37°C for 3 h. To ensure consistency, the thickness of the soft TSA plates was maintained at 1 mm by adding 2 mL of growth media into a 60-mm petri dish. Microscopic images were acquired using an Olympus IX71 inverted microscope, equipped with 20× and 60× objective lenses. For phase contrast imaging, Ph2 objective lenses were used in conjunction with a phase turret to match the phase ring corresponding to the selected objective. For brightfield imaging, the phase turret was adjusted to the open aperture position to allow standard brightfield illumination with

a standard objective lens. Raw images were captured using a CCD camera (Model C4742-80-12AG, Hamamatsu, Tokyo, Japan) and the Metamorph imaging software (version 7.7.2.0, Universal Imaging Corporation). All images were acquired as TIF files and subsequently converted into JPG format using the image processing software ImageJ (Schneider et al., 2012). Eeah image had a resolution of 672 × 512 pixels with a pixel size of 107.5 nm.

2.3 Data preparation and augmentation

To prepare the datasets for model training, image files were structured to work seamlessly with the PyTorch deep learning framework, leveraging its built-in tools for data loading and augmentation. Images were organized into a hierarchical directory structure, with bacterial species (serving as the true class labels) represented as folder names within parent directories that encoded metadata for laboratory conditions. Image files were loaded using the `torchvision.datasets.ImageFolder`, and pixel values were normalized to the range of 0–1 by dividing by 255. To maximize data efficiency, augmentation techniques such as flips, random rotations, and random brightness contrast adjustments were applied using the `albumentations` library (Buslaev et al., 2020). During model training, each image had a 50% probability of undergoing transformation at each step of the augmentation pipeline. Consequently, the specific combination of augmentations applied to a given image varied across epochs. For example, an image could undergo random rotation and Gaussian blur in one epoch, and to random rotation and stretching in another epoch. Augmentations were applied dynamically rather than generating a static, expanded dataset, in order to reduce memory overhead and allow for greater variability. This stochastic augmentation process was repeated for each training epoch, exposing the model to a unique variant of each image throughout training. These augmentation strategies were designed to mimic physical variability inherent in microscopic imaging and to mitigate overfitting by increasing dataset diversity (Simard et al., 2003; Olenskyj et al., 2022).

Datasets were split into training, validation, and test sets based on the specific requirements of source and target domains. Domain adaptation in this study involved training on a larger dataset from a single source domain and testing on smaller datasets from multiple target domains. For the source domain, 15% of the images for each of the 6 bacterial species were held out as a test dataset, while the remaining images were randomly split into training and validation sets in a 70/30 ratio. This resulted in 377 training images, 162 validation images, and 90 test images. Target domains contained fewer labeled samples compared to the source domain, so

a maximum of 5 images per bacterial species was used for training (30 images total). The test sets for the target domains consisted of 90 images for 20×, and 60 images each for BF and 20x-5h domains.

2.4 Model architecture and training

2.4.1 Feature extractor

EfficientNetV2 was selected as the backbone of the proposed architecture due to its superior accuracy and computational efficiency in feature extraction. EfficientNets are a family of CNN architectures with state-of-the-art performance on image classification tasks, offering better accuracy and efficiency compared to other CNN models (Tan and Le, 2019). EfficientNetV2, introduced in 2021, further improves training speed and parameter efficiency through a combination of training-aware neural architecture search and progressive scaling (Tan and Le, 2021). This progressive learning strategy involves training the model on smaller image size with weak regularization initially (i.e., weak constraints to the learning process), then transitioning to larger image sizes with stronger regularization. This approach ensures efficient model training while improving generalization by capturing both low-level and high-level features. In this study, the EfficientNetV2 backbone served as the shared feature extractor for both classification and domain adaptation tasks, as detailed in the following sections.

2.4.2 Adversarial domain adaptation

To address domain shifts across optical and biological variability, DANNs were implemented. These networks were designed to achieve domain invariance and accurate classification across multiple target domains while requiring a minimal number of labeled samples. The model was trained on 539 bacterial images from the source (PC) domain (including 377 training and 162 validation) under the controlled laboratory conditions and evaluated on target domains with variations in optical setups and microbial sample incubation times, as detailed in Table 1. Each target domain had fewer than 5 labeled samples available per bacterial species.

A variant of DANNs, referred to as MDANN, was also employed to extend domain adaptation to multiple domains simultaneously (Ganin et al., 2016). As shown in Figure 1, the model was designed to learn a shared feature extractor capable of capturing domain-invariant features from microscopic images. These extracted features were then passed to a task-specific classification head to predict class labels (i.e., bacterial species), optimized using a task loss calculated as the cross-entropy between predicted and true class labels. To enforce domain invariance, a multi-domain discriminator network was trained concurrently to predict the domain labels of input images based on their feature representations. The domain loss, calculated as the cross-entropy across predicted domain probabilities (with 0 representing the source domain and $1, \ldots, m$ representing the m target domains), was backpropagated through the discriminator network. A gradient reversal layer was applied to multiply the gradient of weights of the discriminator network with respect to

the training loss by $-\lambda$, a tunable hyperparameter that controls the strength of domain alignment. The adjusted gradients were then propagated through the backbone network, promoting the extraction of domain-invariant feature representations. As the discriminator network improved in distinguishing domain labels, the backbone network simultaneously adapted to generate features that minimized domain-specific biases.

2.4.3 Model training

The model training process was structured to achieve simultaneous classification and domain adaptation using a multi-task learning framework. Let S denote the source domain and $T = \{T_1, \ldots, T_{m-1}\}$ denote the target domains. The combined dataset is denoted as $(S \cup T) = \{S, T_1 \ldots, T_{m-1}\}$, where the class labels for any $x \sim (S \cup T)$ are defined as $\{0, \ldots, 5\}$. The shared EfficientNetV2 feature extractor, f_e, is parameterized by θ_e. The forward pass of the extractor is represented as $f_e(x, \theta_e) = e$, where $e \in \mathbb{R}^{2152}$ is the feature embedding of the input x and θ_e are the learnable parameters of the feature extractor. The extracted features are processed by two separate heads: a classification head and a domain regressor head. The classification head, $f_c(f_e(x, \theta_e), \theta_c) : \mathbb{R}^{2152} \to \mathbb{R}^6$, maps the embeddings to probabilities for each class label, while the domain regressor head, $f_d(f_e(x, \theta_e), \theta_d) : \mathbb{R}^{2152} \to \mathbb{R}^m$, maps the embeddings to domain labels, where m is the total number of domains (including the source domain).

The multi-task loss function comprises the classification loss, $\mathcal{L}_C$, and the domain adaptation loss, $\mathcal{L}_D$, and is defined as follows:

$$\mathcal{L} = \mathcal{L}_C + \mathcal{L}_D \tag{1}$$

$$\mathcal{L}_C = -\sum_{i=1}^{m} y_i \cdot log f_c(f_e(x_i, \theta_e), \theta_c) \tag{2}$$

$$\mathcal{L}_D = -\lambda \sum_{i=1}^{m} d_i \cdot log f_d(f_e(x_i, \theta_e), \theta_d) \tag{3}$$

where $x_i \sim (S \cup T)$ is an input sample with class label $y_i \in \{0, \ldots, 5\}$ and domain label $d_i \in \{0, \ldots, m-1\}$, where $d_i = 0$ corresponds to S, $d_i = 1$ corresponds to T_1, and so on. The hyperparameter λ controls the relative importance of the domain loss. A higher value of λ increases the alignment of features representations across domains.

Finally, the model parameter are updated during training using the following equations:

$$\theta_f \leftarrow \theta_f - \alpha \frac{\partial L_C}{\partial \theta_f} + \alpha \cdot \tau \cdot \lambda \frac{\partial L_D}{\partial \theta_f} \tag{4}$$

$$\theta_c \leftarrow \theta_c - \alpha \frac{\partial L_C}{\partial \theta_c} \tag{5}$$

$$\theta_d \leftarrow \theta_d - \alpha \cdot \lambda \frac{\partial L_D}{\partial \theta_d} \tag{6}$$

where α is the learning rate and p is the scaling factor. Gradient reversal is applied in Equation 4 to the domain-regressor loss in

FIGURE 1
Schematic of the model training process using a domain-adversarial neural network (DANN) for domain adaptation for *M* domains.
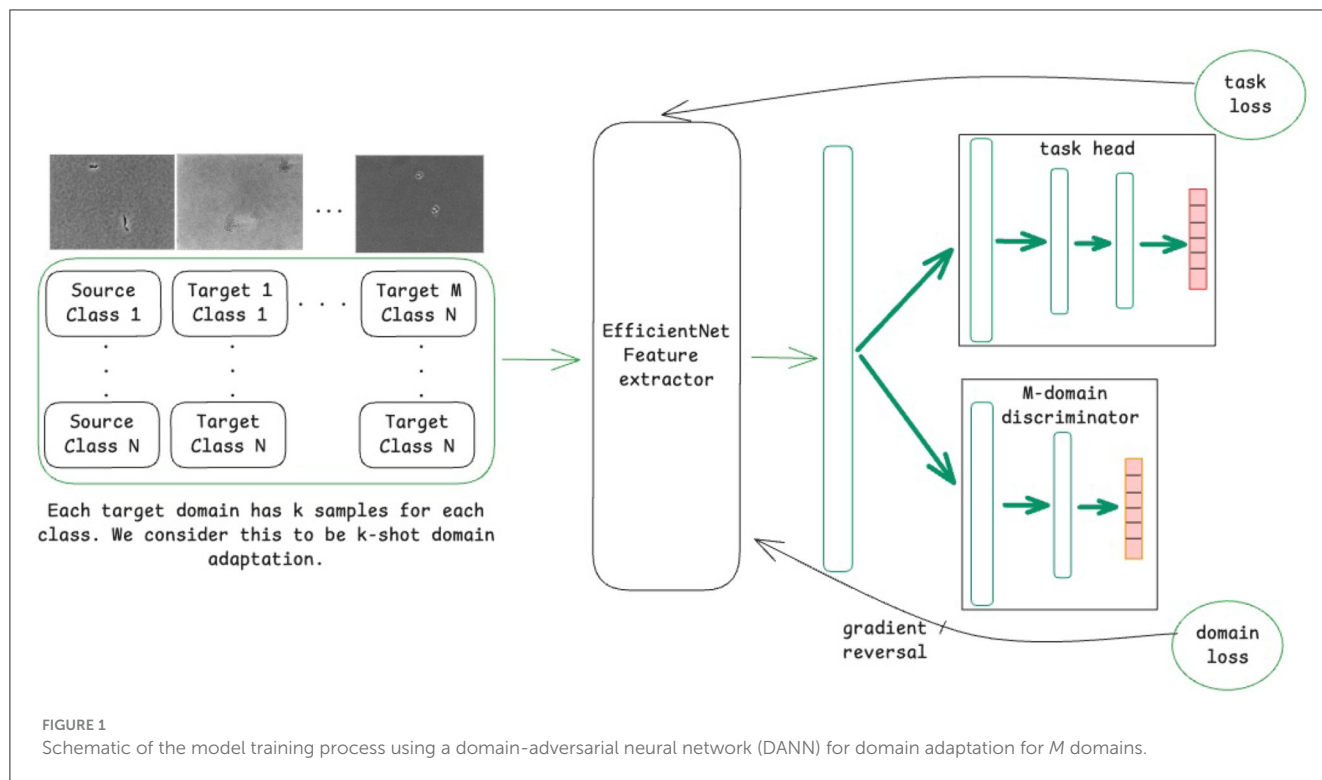
the update of θ_f by adding rather than subtracting its gradient. The scaling factor τ is applied to the domain-regressor loss during the gradient reversal step, a monotonically increasing sigmoid function as follows:

$$\tau = \frac{2}{1 + e^{\frac{-10t_i}{t}}} - 1 \qquad (7)$$

where t_i denotes the current epoch and t is the maximum number of epochs set before training. This adaptive weighting mechanism allows the feature extractor to learn discriminative features in early epochs and transition to domain-invariant features in later epochs, typically after 30–40 epochs (Chen et al., 2019).

Training was conducted for a maximum of 90 epochs with a batch size of 6. The AdamW optimizer was used, with a learning rate of 0.001 and a weight decay of 0.001. For each experiment, the model checkpoint corresponding to the lowest validation loss was selected for evaluation on the held-out test set.

2.5 Model evaluation and visualization

The performance of bacterial image classification was evaluated using classification accuracy, precision, recall, and F1-score. These metrics were calculated for each bacterial species, and the final reported values reflect the average across all classes to ensure balanced assessment regardless of class size. Models were evaluated on both the source and target domains under different training conditions, including source-only and domain-adversarial training with 1-shot, 3-shot, and 5-shot labeled samples per class.
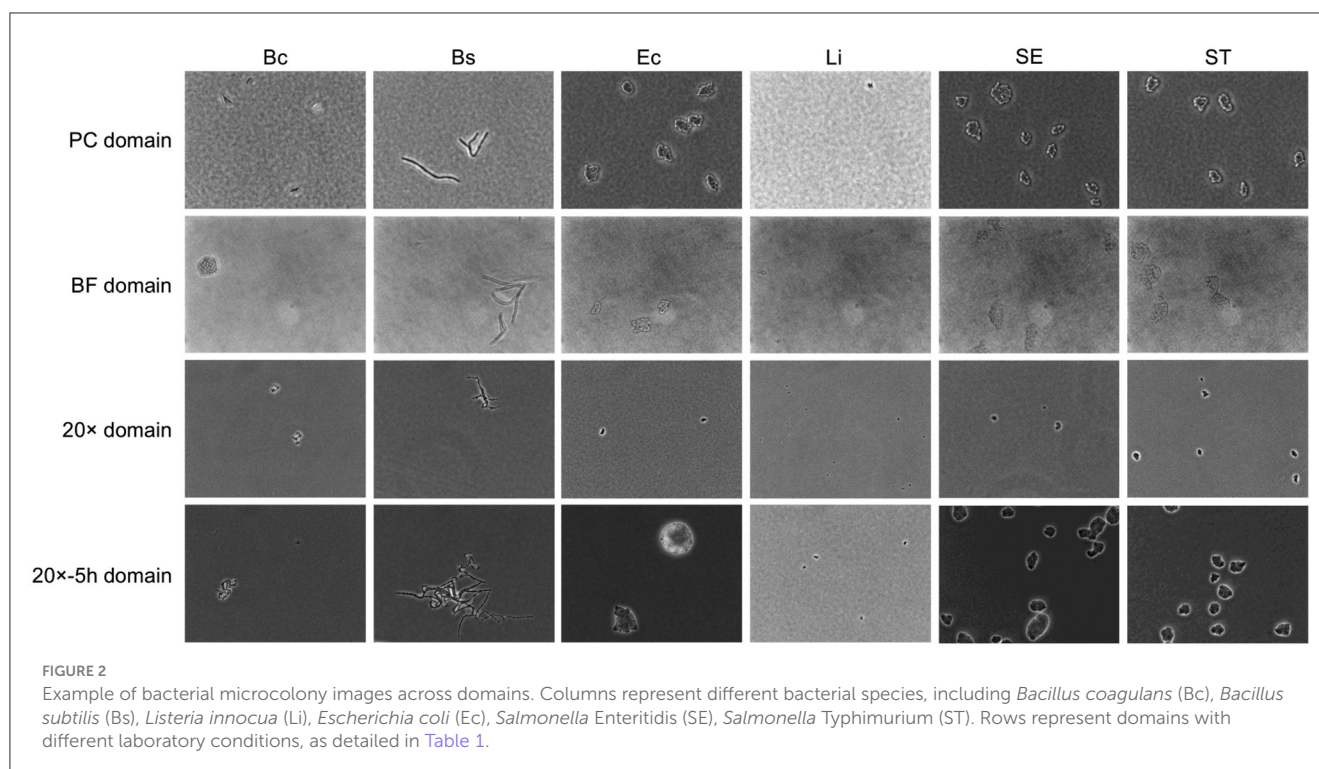
Additionally, Gradient-based class activation mapping (Grad-CAM) (Selvaraju et al., 2017) was employed to gain qualitative insights into the specific image features utilized by the model to predict bacterial species. Grad-CAM is widely used visualization technique in image classification tasks that highlights regions of an input image receiving significant attention from the model during prediction. While the model outputs a single value corresponding to the predicted bacterial species, the Grad-CAM visualizations indicate the regions in the input image that most influence the prediction. In this study, it was hypothesized that these regions would correspond to areas of the input image where individual microcolony is visible, although other features of the input, such as the spatial arrangement of multiple microcolonies, might also contribute to the predictions.

For qualitative understanding of domain alignment within the model, t-distributed stochastic neighbor embedding (t-SNE) (Hinton and Roweis, 2002) was employed to project the feature vectors of test images from the source and target domains into two-dimensional space. This technique enabled visualization of whether domain-adversarial training facilitated domain alignment between the input images of different domains.

3 Results

3.1 Domain variability in microcolony imaging

This study explores domain adaptation across one *source domain* and three *target domains*, each characterized by distinct laboratory conditions (Table 1). The source domain (PC) dataset consisted of images collected under controlled laboratory

FIGURE 2
Example of bacterial microcolony images across domains. Columns represent different bacterial species, including *Bacillus coagulans* (Bc), *Bacillus subtilis* (Bs), *Listeria innocua* (Li), *Escherichia coli* (Ec), *Salmonella* Enteritidis (SE), *Salmonella* Typhimurium (ST). Rows represent domains with different laboratory conditions, as detailed in Table 1.

conditions using phase contrast microscopy at 60× magnification from microcolonies incubated for 3 h. Target domains were chosen to simulate variations encountered in real-world applications, including brightfield microscopy at 60× magnification (BF), phase contrast microscopy at a lower magnification (20×), and phase contrast microscopy at 20× magnification with an extended incubation time to capture additional biological variability (20x-5h).

Substantial variability was observed in microcolony imaging across domains, as illustrated in Figure 2. Source domain images displayed granular backgrounds and fine distinctions between individual bacterial cells, in addition to well-defined microcolonies. These features were attributable to the higher magnification and phase contrast microscopy under controlled cell growth conditions. In contrast, images from the BF target domain, exhibited lower contrast, rendering microcolonies less discernible against the background. Similarly, images from the 20× domain presented smaller targets with reduced cellular features resolution, attributable to the lower magnification. Extended incubation in the 20x-5h domain resulted in larger microcolony sizes but introduced variability in spatial distributions and subtle focal shifts, further complicating feature discrimination.
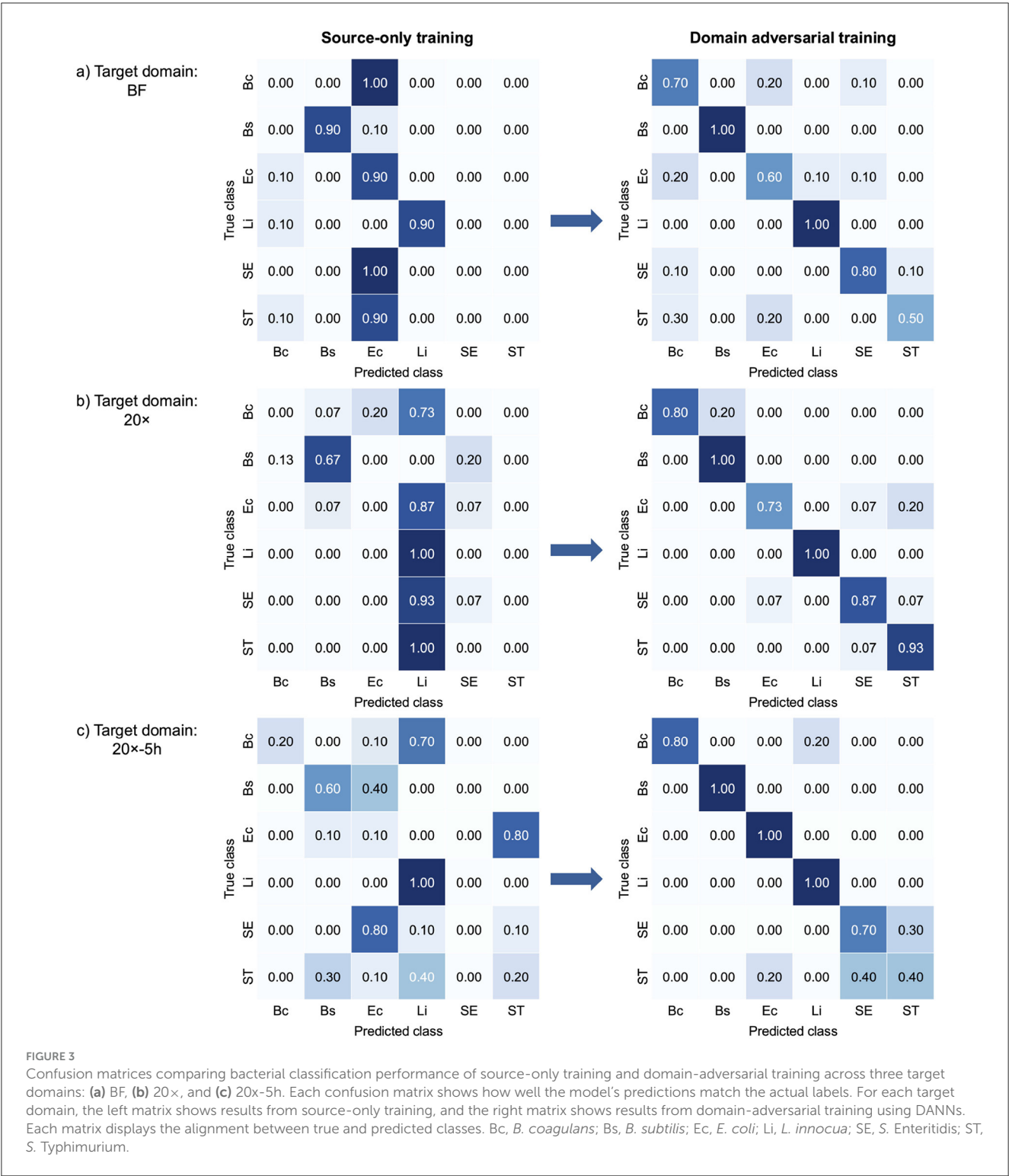
This variability in optical and biological factors diminishes the discriminative quality of image features, presenting challenges for bacterial detection and classification. Addressing this variability necessitated models capable of effectively transferring performance from the high-quality source domain (PC) to variable target domains (BF, 20×, 20x-5h). Given that the bacterial species remained consistent across domains, the model needed to learn domain-invariant features that were unaffected by optical and biological variability, while simultaneously maintaining the ability to classify species-specific traits. To achieve this, DANNs and MDANNs were further implemented, as detailed in Section 2.4. These architectures and training processes were specifically designed to align feature representations across domains, ensuring both domain invariance and accurate classification across all target domains.

3.2 Classification performance improvements with domain adaptation

The ability of DANNs for single-target domain adaptation and MDANNs for multi-domain adaptation was evaluated to address the classification performance gaps between the source domain (PC) and target domains (BF, 20×, 20x-5h). Single-target domain adaptation aligned feature representations for individual target domains, while multi-domain adaptation generalized across multiple target domains simultaneously. Both approaches leveraged limited labeled samples (1-shot, 3-shot, or 5-shot) from the target domains, enabling significant performance improvements despite substantial variability in optical setups and microbial sample incubation times.

Figure 3 presents confusion matrices that illustrate class-wise changes in bacterial classification performance across target domains, highlighting the effect of domain-adversarial training. The left panels illustrate the limited performance of source-only training, which relied solely on data from the source domain. This approach struggled to generalize to target domains due to domain shifts. In contrast, domain-adversarial training effectively aligned feature representations across domains, enhancing classification accuracy in the target domains, as shown in the right panels. In the BF target domain (Figure 3a), *B. coagulans*, *S.* Enteritidis, and

FIGURE 3
Confusion matrices comparing bacterial classification performance of source-only training and domain-adversarial training across three target domains: **(a)** BF, **(b)** 20×, and **(c)** 20x-5h. Each confusion matrix shows how well the model's predictions match the actual labels. For each target domain, the left matrix shows results from source-only training, and the right matrix shows results from domain-adversarial training using DANNs. Each matrix displays the alignment between true and predicted classes. Bc, *B. coagulans*; Bs, *B. subtilis*; Ec, *E. coli*; Li, *L. innocua*; SE, *S.* Enteritidis; ST, *S.* Typhimurium.

S. Typhimurium exhibited near-zero or zero classification accuracy under source-only training, but all showed marked improvement with domain-adversarial training. Similarly, in the 20× domain (Figure 3b), *B. coagulans*, *E. coli*, and *S.* Typhimurium were largely misclassified in the absence of adaptation and benefited from adversarial training. In the 20x-5h domain (Figure 3c), *S.* Enteritidis, which had zero classification accuracy under

source-only training, also demonstrated substantial gains. Overall, the results show that domain-adversarial training led to substantial improvements in classification accuracy across all target domains and all bacterial species.

The classification performance of single-target domain adaptation using DANNs is quantified in Table 2, with target domain accuracy increasing by up to 54.5%, while source domain

TABLE 2 Classification performance for domain-adversarial neural networks (DANNs) with EfficientNetV2 backbone.

Domain	Metric	Source-only	1-shot DANN	3-shot DANN	5-shot DANN
BF source	Accuracy (%)	94.4	91.1	**94.4**	93.3
	Precision (%)	94.6	91.1	94.5	93.6
	Recall (%)	94.4	91.1	94.4	93.3
	F1-score (%)	94.4	91.0	94.4	93.3
BF target	Accuracy (%)	41.7	56.7	**75.0**	73.3
	Precision (%)	37.6	56.2	76.1	73.4
	Recall (%)	41.7	56.7	75.0	73.3
	F1-score (%)	38.2	53.4	73.1	73.3
20× source	Accuracy (%)	**94.4**	90.0	90.0	93.3
	Precision (%)	94.6	90.4	90.1	93.5
	Recall (%)	94.4	90.0	90.0	93.3
	F1-score (%)	94.4	90.0	89.9	93.4
20× target	Accuracy (%)	33.3	54.4	82.2	**88.9**
	Precision (%)	40.3	43.3	83.3	89.9
	Recall (%)	33.3	54.4	82.2	88.9
	F1-score (%)	28.1	46.8	82.0	88.8
20x-5h source	Accuracy (%)	94.4	94.4	94.4	**94.4**
	Precision (%)	94.6	94.6	94.7	94.8
	Recall (%)	94.4	94.4	94.4	94.4
	F1-score (%)	94.4	94.4	94.5	94.5
20x-5h target	Accuracy (%)	40.0	71.7	83.3	**83.3**
	Precision (%)	43.4	80.1	84.1	84.1
	Recall (%)	40.0	71.7	83.3	83.3
	F1-score (%)	34.1	70.9	81.3	81.3

Results are presented for source-only training and domain-adversarial training with 1-shot, 3-shot, and 5-shot labeled samples. Bold values indicate the highest accuracy for each domain and setting.

accuracy experienced minimal degradation, within 1.1% of source-only training. The model performed better on the 20× and 20x-5h domains compared to BF. Specifically, accuracy improvements reached up to 54.5% and 43.3% for 20× and 20x-5h domains, respectively, while BF domain showed a more modest improvements of 31.7%. In addition to accuracy, other performance metrics also improved across target domains with DANN. In the 20× domain, F1-score increased from 28.1% (source-only) to 88.8% with 5-shot DANN, with similar gains observed in precision (from 40.3% to 89.9%) and recall (from 33.3% to 88.9%). The 20x-5h domain showed a rise in F1-score from 34.1% to 81.3%, while precision and recall improved to 84.1% and 83.3%, respectively. Improvements in the BF domain were more modest, with F1-score rising from 38.2% to 73.3% under 5-shot DANN. These consistent gains across metrics support the effectiveness of adversarial training for enhancing performance in label-limited, variable imaging conditions.

Multi-domain adaptation using MDANNs further showed strong generalization across multiple target domains, as shown in Table 3. MDANNs achieved accuracy comparable or higher than single-target domain adaptation for BF domain, with accuracy improvement of 33.3% (from 43.3% to 76.7%), while preserving source domain accuracy to within 4.4% of source-only training. However, for the 20× domain, MDANNs achieved slightly lower performance compared single-target DANNs, despite showing substantial improvement of 47% (from 34.4% to 82.2%). Precision, recall, and F1-score also improved with MDANN. In the BF target domain, F1-score increased from 57.8% (1-shot) to 76.2% (5-shot), with corresponding gains in precision (from 57.4% to 77.3%) and recall (from 60.0% to 76.7%). In the 20× domain, F1-score reached 74.8% with 5-shot MDANN, up from 48.4% with 1-shot training, alongside increases in precision (from 60.6% to 75.3%) and recall (from 54.4% to 75.6%). These improvements further demonstrate that MDANNs effectively generalize across domains even under few-shot constraints.

Overall, these results highlight the effectiveness of domain-adversarial training in both single-target and multi-domain scenarios. DANNs and MDANNs demonstrated the ability to learn domain-invariant features and align feature representations across domains, enabling accurate bacterial classification despite the variations in optical setups and microbial sample incubation times.

TABLE 3 Classification performance for multi-domain adversarial neural networks (MDANNs) with EfficientNetV2 backbone.

Domain	Metric	1-shot MDANN	3-shot MDANN	5-shot MDANN
Source	Accuracy (%)	87.8	**92.2**	90.0
	Precision (%)	88.0	92.4	90.5
	Recall (%)	87.8	92.2	90.0
	F1-score (%)	87.8	92.1	89.8
Target (BF)	Accuracy (%)	60.0	68.3	**76.7**
	Precision (%)	57.4	67.7	77.3
	Recall (%)	60	68.3	76.7
	F1-score (%)	57.8	65.1	76.2
Target (20×)	Accuracy (%)	54.4	**82.2**	75.6
	Precision (%)	60.6	84.0	75.3
	Recall (%)	54.4	82.2	75.6
	F1-score (%)	48.4	82.1	74.8

The source domain was PC and the target domains were BF and 20×. Results are shown for domain-adversarial training with 1-shot, 3-shot, and 5-shot labeled samples. Bold values indicate the highest accuracy for each domain and setting.

3.3 Visualizing feature representations and domain alignment

Grad-CAM visualizations (Figure 4) provided insights into how the model identified important features across different imaging domains, highlighting the regions of microcolony images that were most influential during classification. Warmer colors (e.g., red and yellow) represented areas with stronger activation, indicating the model's focus during prediction. In the PC domain, activations were sharply localized along the well-defined edges of individual microcolonies, reflecting the high-contrast details available in phase contrast microscopy under controlled laboratory conditions. In the BF domain, activations were broader and slightly less concentrated. While the overall activation structure resembled that of the PC domain, the model adapted to the lower contrast of brightfield microscopy by expanding its focus to include central regions of microcolonies. In the 20× domain, activations were noticeably more diffuse and less sharply defined. The lower magnification reduced the resolution of individual cells, making edge-specific features less informative. As a result, the model shifted its attention to the spatial arrangement and overall morphology of microcolonies, which became the most informative distinguishing features.

The alignment of feature representations was further validated using t-SNE visualization (Figure 5). This visualization projected high-dimensional feature embeddings into a two-dimensional space, allowing the formation of visually distinct clusters for each bacterial species. In the visualization, the source domain (PC) embeddings are represented by circles and the target domain (e.g., 20x-5h) embeddings are represented by crosses, with colors indicating different bacterial species. Overlap or close alignment between source and target embeddings within the same bacterial species cluster indicates that domain-adversarial training successfully aligned features across domains. For most species, such as *B. coagulans*, *B. subtilis*, *E. coli*, and *L. innocua*, the source and target clusters overlap significantly, demonstrating effective alignment of features across domains. However, for closely related species such as *S. Enteritidis* and *S. Typhimurium*, the visualization shows overlapping clusters with minimal separability between these two species, revealing challenges in distinguishing them.

4 Discussion

4.1 Generalizing bacterial classification beyond sample preparation and long incubation constraints

This study demonstrates the application of adversarial domain adaptation to enhance the generalizability of AI-enabled microscopy for bacterial classification under diverse laboratory conditions. By integrating a pretrained EfficientNetV2 backbone with domain-adversarial training, our approach achieved up to a 54.5% improvement in classification accuracy across target domains, effectively addressing domain shifts such as variations in optical setups and microbial sample incubation times. Previous studies in AI-enabled bacterial classification predominantly focused on optimizing model performance within controlled laboratory settings, often relying on sample preparation techniques such as staining, selective enrichment, or the use of specialized reagents, which extended processing times and required skilled personnel (Chen et al., 2024; Ramesh et al., 2024; Wakabayashi et al., 2024). For example, Chen et al. (2024) utilized selective media and Gram staining to isolate and chemically label bacterial samples, enabling their model to classify pre-processed colonies but constraining the scope to well-prepared datasets, with incubation times of 18–24 h for enrichment and an additional 18–48 h for further processing. Similarly, many studies, including Ramesh et al. (2024), utilized the publicly available Digital Images of Bacteria Species (DIBaS) dataset, which consisted of Gram-stained samples imaged at high magnification (100×), providing high-quality data but without addressing generalizability across broader sample preparation and imaging conditions. Wakabayashi et al. (2024) introduced variability by working with inoculated
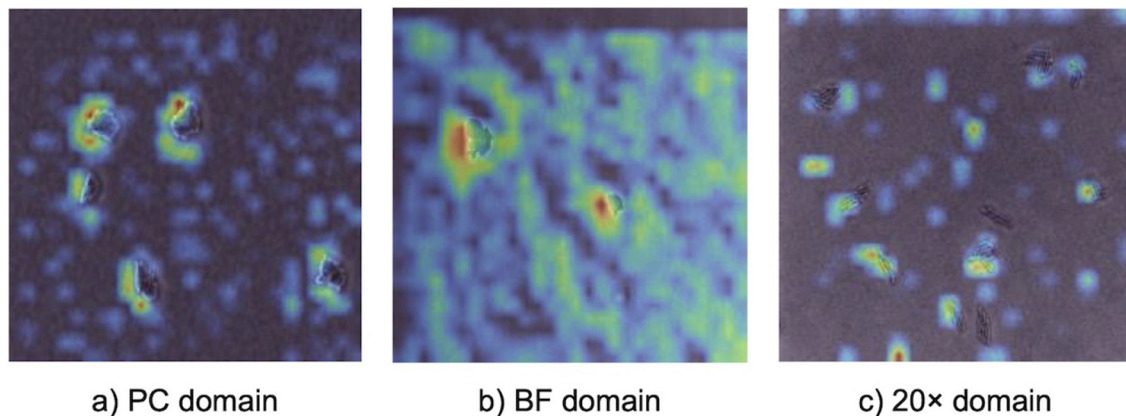
FIGURE 4
Gradient-based class activation mapping (Grad-CAM) visualizations illustrating model attention during prediction with a 5-shot MDANN. Panels show activation heatmaps for correctly classified bacterial microcolonies in **(a)** PC (source domain), **(b)** BF (target domain), and **(c)** 20× (target domain). Warmer colors indicate regions with stronger contribution to the model's prediction, providing model interpretability.
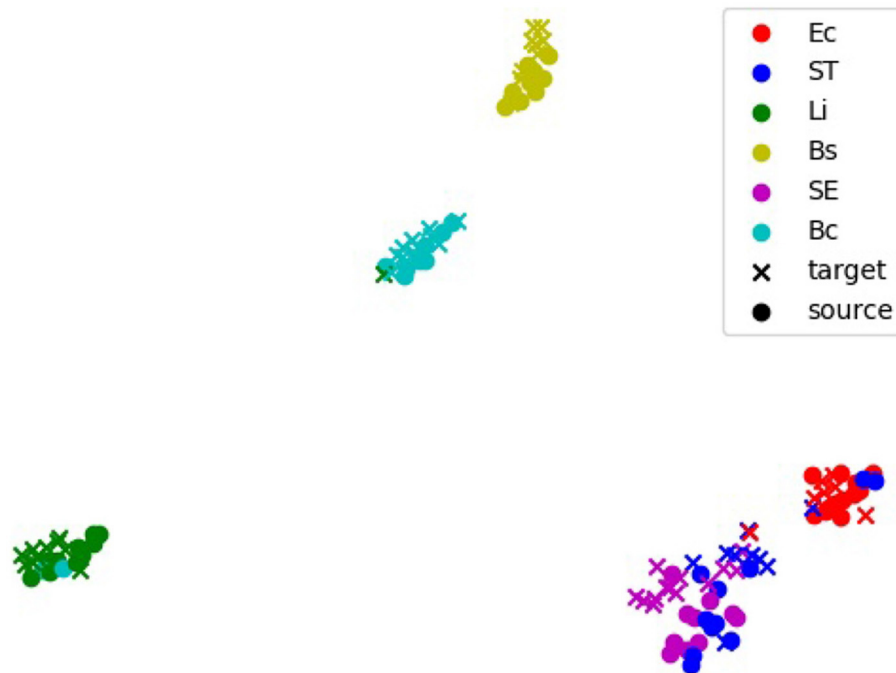

FIGURE 5
t-distributed stochastic neighbors embedding (t-SNE) visualization of feature embeddings extracted by the 5-shot DANN. The source domain is PC, and the target domain is 20x-5h. Clusters represent feature embeddings for specific bacterial species, with source samples (black dots) and target samples (colored crosses). Overlapping clusters indicate effective alignment of features across domains. Bc, *B. coagulans*; Bs, *B. subtilis*; Ec, *E. coli*; Li, *L. innocua*; SE, *S.* Enteritidis; ST, *S.* Typhimurium.

meat samples, yet relied on extended incubation times (20 h) and selective media, which may not be feasible in resource-limited settings. In these studies, AI models were employed to analyze highly processed or enriched datasets, often focusing on tasks like counting fully grown colonies or identifying pre-processed features rather than generalizing to diverse conditions. In contrast, our study focused on minimizing the reliance on extensive sample preparation, long incubation times, and precise skillsets required for high-quality image acquisition. By using general-purpose media with a reduced incubation time of 3 h, we demonstrated the potential to shift bacterial classification from labor-intensive preparation to automated abstract feature recognition. To the best of our knowledge, this is the first study to explicitly evaluate the generalizability of a model trained on high-quality, controlled datasets to diverse image domains, encompassing optical and biological variability. These advancements highlight the model's potential for resource-limited and variable laboratory environments.

4.2 Adversarial domain adaptation for cross-modality imaging in bacterial classification

The results of this study highlight the effectiveness of adversarial domain adaptation for bacterial classification in microscopy, showcasing its ability to address optical variability and reduce dependence on specific microscopy modalities and resource-intensive sample preparation. Cross-modality microscopy has been extensively explored in biomedical imaging, particularly for tasks like cell and nucleus detection, where large, curated datasets are often used to achieve robust performance. For example, previous work utilized k-means clustering and modality-specific deep learning models for cell segmentation across diverse microscopy modalities, such as brightfield, phase contrast, differential interference contrast (DIC), and fluorescence microscopy, emphasizing the need for domain adaptation to improve generalizability beyond individual models (Wang et al., 2023). Similarly, adversarial domain adaptation was applied to translate between stained and unstained images for cell and nucleus detection, improving detection performance compared to baseline "source-only" training (Xing et al., 2021). However, these studies primarily focused on instance detection and often relied on staining information for classification. Another study applied domain adaptation to digitally stain white blood cells, reducing the need for staining procedures and preserving structural information critical for classification, though it did not address optical variability, such as changes in magnification (Tomczak et al., 2021). In contrast, this study emphasizes optical variability for bacterial classification, exploring the practicality of adversarial domain adaptation with small datasets containing as few as 1–5 labeled samples per species in target domains. It demonstrated robust performance across variations in optical setups, including transitions between microscopy modalities (phase contrast to brightfield) and magnifications ($60\times$ to $20\times$). This few-shot approach can be particularly advantageous in scenarios where collecting and annotating large datasets is impractical. By reducing dependency on extensive labeled data, the proposed framework ensures scalability for real-world applications. A key advantage of this approach lies in its ability to learn abstract features that are domain-invariant, enabling robust classification of both Gram-positive and Gram-negative bacteria across diverse imaging conditions without relying on staining or contrast-enhancing techniques. Unlike earlier bacterial classification methods that required Gram-staining to enhance image contrast (Chen et al., 2024; Ramesh et al., 2024) or phase contrast microscopy for microcolony visualization (Ma et al., 2023), this method shifts the classification burden to the model, bypassing optical and chemical dependencies inherent to traditional workflows. Additionally, transitioning from $60\times$ to $20\times$ magnification facilitates the use of more accessible, less specialized equipment, balancing resolution with broader applicability. By leveraging a few-shot domain adaptation approach to address variability in optical setups, this domain-adversarial framework ensures scalability while maintaining robust performance. This makes it particularly suitable for decentralized and resource-limited environments such as onsite food safety testing.

4.3 Learning abstract features for small targets and biological variability

The EfficientNetV2 backbone showed remarkable effectiveness in this study, particularly for classifying small bacterial microcolonies under challenging imaging conditions, such as lower magnifications in the $20\times$ domain. Its compound scaling strategy, which balances depth, width, and resolution, enables hierarchical feature learning that is well-suited for capturing fine-grained features of small targets. Unlike previous approaches that employed shallow CNN architectures, such as the five-layer convolutional model used for foodborne pathogen classification (Chen et al., 2024), EfficientNetV2's deeper architecture (13–60 convolutional layers) excelled in extracting information-rich abstract features. While Chen et al. (2024) achieved promising results for rapid detection using brightfield microscopy, their approach relied on Gram-staining, high magnifications with oil immersion objectives, and extended incubation times to grow fully formed colonies. In contrast, our EfficientNetV2 backbone, combined with adversarial domain adaptation, enabled classification of microcolonies cultivated for shorter incubation periods without staining or high-resolution imaging. Additionally, previous approaches in bacterial image classification often excluded low-quality images manually to mitigate errors (Chen et al., 2024), whereas EfficientNetV2's ability to generalize stemmed from its progressive learning strategies, advanced regularization techniques, and integration with an image augmentation pipeline. These capabilities allowed the model to perform robustly even with suboptimal image quality. EfficientNetV2 also addressed limitations of object detection models like YOLOv3, which struggled with small bacterial targets due to specific design constraints (Wakabayashi et al., 2024). Its grid cell mechanism often failed to localize small microcolonies accurately, downsampling erases fine-grained details, and predefined anchor boxes were poorly suited to microcolony sizes and shapes. These limitations, observed in our preliminary experiments with YOLOv4, hindered performance in the $20\times$ domain. In contrast, EfficientNetV2's hierarchical feature learning and dynamic scaling captured subtle spatial and morphological details, making it well-suited for small bacterial target analysis under varying conditions.

The flexibility of EfficientNetV2 is further supported by Grad-CAM and t-SNE analyses, which highlighted its ability to adapt to varying imaging conditions. Grad-CAM visualizations revealed that in the PC domain, the model focused on edge-specific features, leveraging high-contrast details provided by phase contrast microscopy. For the $20\times$ domain, where reduced magnification limited the visibility of fine cellular features, the model shifted its focus to broader microcolony-level characteristics such as shape, size, and spatial arrangement, demonstrating dynamic feature extraction. Similarly, the t-SNE visualization for the 20x-5h domain showed significant alignment between source (PC) and target embeddings for most species, confirming the model's ability to learn domain-invariant features and generalize across optical and biological variability. However, overlapping t-SNE clusters for closely related species, such as *S. Enteritidis* (SE) and *S. Typhimurium* (ST), underscored the challenge of distinguishing morphologically similar species, as also reflected

in the confusion matrix (Figure 3c). These findings emphasized the combined strength of EfficientNetV2 and adversarial domain adaptation in handling variability across domains.

4.4 Future directions for enhancing domain adaptation in microscopy

This study highlights the strengths and challenges of DANNs and MDANNs in addressing domain variability across diverse optical and biological conditions. While DANN improved classification performance across multiple target domains, results from Table 2 and Section 3.2 indicate that low contrast in the BF domain posed greater alignment difficulties than variations in magnification or microbial sample incubation times in other domains. MDANNs demonstrated strong generalization across multiple domains by leveraging shared features across bacterial species, but results from Table 3 and Section 3.2 show that the BF domain consistently exhibited lower improvements compared to others, underscoring the persistent challenges of low-contrast imaging. Building on prior multi-domain adversarial approaches, which have demonstrated success with larger datasets and complex discriminator architectures (Zhao et al., 2018; Pei et al., 2018), this study extended these advancements by achieving domain-adversarial training with as few as 1–5 labeled samples per bacterial species. This few-shot approach reduces the labor-intensive process of manual annotation and aligns with the broader CNN-based strategies prevalent in this field. To support this proof-of-concept evaluation under constrained labeling conditions, we selected six foodborne species that enabled controlled testing of domain generalization across biologically meaningful variation. These taxa span Gram-positive and Gram-negative groups and multiple genera and include two *Bacillus* species and two *Salmonella* serovars, allowing us to assess model robustness across genus-, species-, and serovar-level distinctions. Future work will expand to broader taxa to further enhance generalizability. In particular, including additional Gram-negative genera such as *Shigella*, *Yersinia*, and *Vibrio* will allow us to evaluate model performance under greater phenotypic ambiguity. These organisms can exhibit overlapping growth characteristics in non-selective media like TSB, despite significant genetic and physiological differences (Schiemann, 1983; Dekker and Frank, 2015). Incorporating these genera would provide a more rigorous test of classification robustness across closely related taxa. Moreover, expanding the range of environmental growth conditions, including variations in media composition, pH, and substrate type, may help capture the phenotypic plasticity observed in real-world bacterial populations. Such variability is likely to influence morphological and physiological traits relevant to imaging-based classification. Broader validation efforts could help align model predictions with established taxonomic criteria and support integration with conventional classification frameworks.

Another promising future direction would be to transition toward unsupervised domain adaptation, eliminating the reliance on labeled samples entirely. Leveraging generative adversarial networks (GANs), we could digitally generate pixel-level labels or synthetic training data, enabling an unsupervised workflow that emphasizes pixel-level classification. This shift would address both labor limitations and the need for enhanced alignment at finer spatial resolutions, such as in low-contrast modalities like brightfield microscopy. GAN-based approaches for digital staining or synthetic dataset generation have shown potential in recent studies on biomedical imaging (Mukherjee et al., 2023; Goyal et al., 2024), and their integration could further improve domain alignment and classification performance. Additionally, future efforts could incorporate spectral or biochemical features into the dataset, improving the differentiation of closely related bacterial species, such as *S.* Enteritidis and *S.* Typhimurium, providing additional discriminatory power. Building on this direction, integrating other data modalities may further improve classification resolution beyond morphology alone. Our recent study using AI-enabled hyperspectral microscopy (Papa et al., 2025) demonstrates that spectral features, which are often associated with physiological and biochemical differences, can enhance the differentiation of closely related bacterial strains. Combining such multimodal approaches with domain adaptation could further strengthen performance in complex classification tasks involving subspecies, pathovars, or functional variants. Dynamic phenotypic traits such as microcolony spreading or motility may also offer discriminative features beyond morphology. Incorporating multi-frame or time-resolved imaging could allow models to learn behavioral patterns that vary across subspecies or functional variants. These advancements would enhance the robustness of domain-adversarial frameworks and broaden their scalability and applicability to decentralized or resource-limited laboratory environments.

5 Conclusions

This study demonstrates the potential of adversarial domain adaptation for bacterial classification in microscopy, addressing variability in optical setups and microbial sample incubation times. By integrating EfficientNetV2 image classification backbone with DANNs and MDANNs, the framework achieved robust generalization from controlled laboratory conditions to various setups with optical and biological variability. Single-target adaptation improved performance across individual domains, while multi-domain adaptation achieved strong generalization across multiple domains with slight trade-offs in some cases. Grad-CAM and t-SNE visualizations validated the model's ability to extract domain-invariant features, enabling robust adaptation to challenges such as low-contrast imaging, low magnifications, and extended incubation times. The few-shot learning approach underscores the scalability of this framework for real-world applications, such as onsite food safety testing, where labeled data are limited. Follow-on studies could explore unsupervised approaches to further enhance domain adaptation by reducing reliance on labeled samples, improving contrast in low-quality images, and enabling alignment at finer spatial resolutions. This study establishes a strong foundation for optimizing multi-domain adaptation and enhancing

scalability and accuracy in decentralized and resource-limited settings.

Data availability statement

The code is available at https://github.com/food-ai-engineering-lab/microcolony-domain-adaptation-frai. Future updates will be integrated to this repository.

Author contributions

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Generative AI statement

The author(s) declare that no Gen AI was used in the creation of this manuscript.

Publisher's note

References

Buslaev, A., Iglovikov, V. I., Khvedchenya, E., Parinov, A., Druzhinin, M., and Kalinin, A. A. (2020). Albumentations: fast and flexible image augmentations. *Information* 11:125. doi: 10.3390/info11020125

Chen, C., Xie, W., Huang, W., Rong, Y., Ding, X., Huang, Y., et al. (2019). "Progressive feature alignment for unsupervised domain adaptation," in *2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, 627–636. doi: 10.1109/CVPR.2019.00072

Chen, Q., Bao, H., Li, H., Wu, T., Qi, X., Zhu, C., et al. (2024). Microscopic identification of foodborne bacterial pathogens based on deep learning method. *Food Control* 161:110413. doi: 10.1016/j.foodcont.2024.110413

Chin, S. Y., Dong, J., Hasikin, K., Ngui, R., Lai, K. W., Yeoh, P. S. Q., et al. (2024). Bacterial image analysis using multi-task deep learning approaches for clinical microscopy. *PeerJ Comput. Sci.* 10:e2180. doi: 10.7717/peerj-cs.2180

Dekker, J. P., and Frank, K. M. (2015). Salmonella, shigella, and yersinia. *Clin. Lab. Med.* 35, 225-246. doi: 10.1016/j.cll.2015.02.002

Ferone, M., Gowen, A., Fanning, S., and Scannell, A. G. M. (2020). Microbial detection and identification methods: bench top assays to omics approaches. *Compr. Rev. Food Sci. Food Saf.* 19, 3106–3129. doi: 10.1111/1541-4337.12618

Ganin, Y., Ustinova, E., Ajakan, H., Germain, P., Larochelle, H., Laviolette, F., et al. (2016). Domain-adversarial training of neural networks. *J. Mach. Learn. Res.* 17, 1–35.

Goyal, V., Liu, M., Bai, A., Lin, N., and Hsieh, C. (2024). "Generalizing microscopy image labeling via layer-matching adversarial domain adaptation," in *ICML'24 Workshop ML for Life and Material Science: From Theory to Industry Applications*.

Hinton, G. E., and Roweis, S. (2002). "Stochastic neighbor embedding," in *Advances in Neural Information Processing Systems, volume 15*, eds. S. Becker, S. Thrun, and K. Obermayer (Cambridge, MA: MIT Press).

Hoffmann, S., White, A. E., McQueen, R. B., Ahn, J.-W., Gunn-Sandell, L. B., and Scallan Walter, E. J. (2024). Economic burden of foodborne illnesses acquired in the United States. *Foodborne Pathog. Dis.* 22, 4–14. doi: 10.1089/fpd.2023.0157

Jeckel, H., and Drescher, K. (2020). Advances and opportunities in image analysis of bacterial cells and communities. *FEMS Microbiol. Rev.* 45:fuaa062. doi: 10.1093/femsre/fuaa062

Ma, L., Yi, J., Wisuthiphaet, N., Earles, M., and Nitin, N. (2023). Accelerating the detection of bacteria in food using artificial intelligence and optical imaging. *Appl. Environ. Microbiol.* 89, e01828–e01822. doi: 10.1128/aem.01828-22

Melanthota, S. K., Gopal, D., Chakrabarti, S., Kashyap, A. A., Radhakrishnan, R., and Mazumder, N. (2022). Deep learning-based image processing in optical microscopy. *Biophys. Rev.* 14, 463–481. doi: 10.1007/s12551-022-00949-3

Mukherjee, S., Sarkar, R., Manich, M., Labruyère, E., and Olivo-Marin, J.-C. (2023). Domain adapted multitask learning for segmenting amoeboid cells in microscopy. *IEEE Trans. Med. Imaging* 42, 42–54. doi: 10.1109/TMI.2022.3203022

Olenskyj, A. G., Donis-González, I. R., Earles, J. M., and Bornhorst, G. M. (2022). End-to-end prediction of uniaxial compression profiles of apples during in vitro digestion using time-series micro-computed tomography and deep learning. *J. Food Eng.* 325:111014. doi: 10.1016/j.jfoodeng.2022.111014

Otamendi, U., Martinez, I., Quartulli, M., Olaizola, I. G., Viles, E., and Cambarau, W. (2021). Segmentation of cell-level anomalies in electroluminescence images of photovoltaic modules. *Solar Energy* 220, 914–926. doi: 10.1016/j.solener.2021.03.058

Papa, M., Wasit, A., Pecora, J., Bergholz, T. M., and Yi, J. (2025). Detection of viable but nonculturable E. coli induced by low-level antimicrobials using AI-enabled hyperspectral microscopy. *J. Food Prot.* 88:100430. doi: 10.1016/j.jfp.2024.100430

Pei, Z., Cao, Z., Long, M., and Wang, J. (2018). Multi-adversarial domain adaptation. *Proc. AAAI Conf. Artif. Intell.* 32, 3934–3941. doi: 10.1609/aaai.v32i1.11767

Qiu, Q., Dewey-Mattia, D., Subramhanya, S., Cui, Z., Griffin, P. M., Lance, S., et al. (2021). Food recalls associated with foodborne disease outbreaks, United States, 2006–2016. *Epidemiol. Infect.* 149:e190. doi: 10.1017/S0950268821001722

Ramesh, H., Elshinawy, A., Ahmed, A., Kassoumeh, M. A., Khan, M., and Mounsef, J. (2024). "Biointel: real-time bacteria identification using microscopy imaging," in *2024*

IEEE International Symposium on Biomedical Imaging (ISBI) (Piscataway, NJ: IEEE), 1–4. doi: 10.1109/ISBI56570.2024.10635473

Schiemann, D. (1983). Comparison of enrichment and plating media for recovery of virulent strains of Yersinia enterocolitica from inoculated beef stew. *J. Food Prot.* 46, 957–964. doi: 10.4315/0362-028X-46.11.957

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2017). "Grad-CAM: Visual explanations from deep networks via gradient-based localization," in *Proceedings of the IEEE International Conference on Computer Vision* (Piscataway, NJ: IEEE), 618–626. doi: 10.1109/ICCV.2017.74

Simard, P. Y., Steinkraus, D., and Platt, J. C. (2003). Best practices for convolutional neural networks applied to visual document analysis. *In Seventh International Conference on Document Analysis and Recognition (ICDAR)* 3, 958–963. doi: 10.1109/ICDAR.2003.1227801

Tan, M., and Le, Q. (2019). "Efficientnet: rethinking model scaling for convolutional neural networks," in *International Conference on Machine Learning* (PMLR), 6105–6114.

Tan, M., and Le, Q. (2021). "Efficientnetv2: smaller models and faster training," in *International Conference on Machine Learning* (PMLR), 10096–10106.

Tomczak, A., Ilic, S., Marquardt, G., Engel, T., Forster, F., Navab, N., et al. (2021). Multi-task multi-domain learning for digital staining and classification of leukocytes. *IEEE Trans. Med. Imaging* 40, 2897–2910. doi: 10.1109/TMI.2020.3046334

Wakabayashi, R., Aoyanagi, A., and Tominaga, T. (2024). Rapid counting of coliforms and Escherichia coli by deep learning-based classifier. *J. Food Saf.* 44:e13158. doi: 10.1111/jfs.13158

Wang, Z., Fang, Z., Chen, Y., Yang, Z., Liu, X., and Zhang, Y. (2023). "Semi-supervised cell instance segmentation for multi-modality microscope images," in *Proceedings of The Cell Segmentation Challenge in Multi-modality High-Resolution Microscopy Images, volume 212 of Proceedings of Machine Learning Research*, eds. J. Ma, R. Xie, A. Gupta, J. Guilherme de Almeida, G. D. Bader, and B. Wang (PMLR), 1–11.

Wu, Y., and Gadsden, S. A. (2023). Machine learning algorithms in microbial classification: a comparative analysis. *Front. Artif. Intell.* 6:1200994. doi: 10.3389/frai.2023.1200994

Xing, F., Cornish, T. C., Bennett, T. D., and Ghosh, D. (2021). Bidirectional mapping-based domain adaptation for nucleus detection in cross-modality microscopy images. *IEEE Trans. Med. Imaging* 40, 2880–2896. doi: 10.1109/TMI.2020.3042789

Xu, F., Li, X., Yang, H., Wang, Y., and Xiang, W. (2022). Te-yolof: Tiny and efficient yolof for blood cell detection. *Biomed. Signal Process. Control* 73:103416. doi: 10.1016/j.bspc.2021.103416

Zhao, H., Zhang, S., Wu, G., Moura, J. M. F., Costeira, J. P., and Gordon, G. J. (2018). "Adversarial multiple source domain adaptation," in *Advances in Neural Information Processing Systems, volume 31*, eds. S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett (Red Hook, NY: Curran Associates, Inc.).