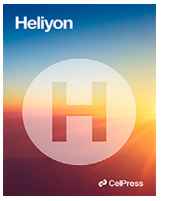
Heliyon 10 (2024) e25714

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon



Research article

Increasing segmentation performance with synthetic agar plate images

Michal Cicatka [a,*], Radim Burget [a], Jan Karasek [b], Jan Lancos [b]

[a] Brno University of Technology, Faculty of Electrical Engineering and Communications, Dept. of Telecommunication, Technicka 12, Brno, 61600, Czech Republic

[b] R&D Automation, Microbiology & Diagnostics, Bruker Daltonics GmbH & Co. KG, Fahrenheitstraße 4, Bremen, 28359, Germany

ARTICLE INFO

Keywords:
Agar plates
Synthetic images generation
Deep learning
Semantic segmentation

ABSTRACT

Background: Agar plate analysis is vital for microbiological testing in industries like food, pharmaceuticals, and biotechnology. Manual inspection is slow, laborious, and error-prone, while existing automated systems struggle with the complexity of real-world agar plates. A shortage of diverse datasets hinders the development and evaluation of robust automated systems.

Methods: In this paper, two new annotated datasets and a novel methodology for synthetic agar plate generation are presented. The datasets comprise 854 images of cultivated agar plates and 1,588 images of empty agar plates, encompassing various agar plate types and microorganisms. These datasets are an extension of the publicly available BRUKERCOLONY dataset, collectively forming one of the largest publicly available annotated datasets for research. The methodology is based on an efficient image generation pipeline that also simulates cultivation-related phenomena such as haemolysis or chromogenic reactions.

Results: The augmentations significantly improved the Dice coefficient of trained U-Net models, increasing it from 0.671 to 0.721. Furthermore, training the U-Net model with a combination of real and 150% synthetic data demonstrated its efficacy, yielding a remarkable Dice coefficient of 0.729, a substantial improvement from the baseline of 0.518. UNet3+ exhibited the highest performance among the U-Net and Attention U-Net architectures, achieving a Dice coefficient of 0.767.

Conclusions: Our experiments showed the methodology's applicability to real-world scenarios, even with highly variable agar plates. Our paper contributes to automating agar plate analysis by presenting a new dataset and effective methodology, potentially enhancing fully automated microbiological testing.

1. Introduction

In the field of biomedical research, where total laboratory automation is a growing trend [1], precise analysis of microbes on agar plates remains a vital task. This is because it is involved in the automation of many routine laboratory procedures, such as disk diffusion tests [2], detection of chromogenic reactions [3], colony counting [4], or colony picking, which is closely related to

* Corresponding author.
E-mail addresses: xcicat02@vut.cz (M. Cicatka), burgetrm@vut.cz (R. Burget), Jan.Karasek@bruker.com (J. Karasek), Jan.Lancos@bruker.com (J. Lancos).

https://doi.org/10.1016/j.heliyon.2024.e25714
Received 15 June 2023; Received in revised form 26 January 2024; Accepted 1 February 2024

Available online 7 February 2024

sample preparation for mass spectrometry analysis [5]. Despite many previous studies on this topic and the availability of automated robots for processing agar plates in the market, a perfect one-size-fits-all solution is still lacking. The proposed solutions often utilize image segmentation, a fundamental computer vision technique employed in various domains ranging from transportation for safety enhancement [6,7], to healthcare for precise tumor detection [8,9]. The lack of robustness of available solutions in real-world scenarios could be attributed to insufficient training data for machine-learning-based segmentation approaches.

The purpose of this work is to present a novel methodology for the generation of synthetic agar plates with microbes and evaluate its performance with a comprehensive publicly available BRUKERDSUT dataset of real-life samples. This dataset helps to ensure that future studies can compare their results to more robust and realistic data, ultimately improving the accuracy and reliability of the field as a whole. The BRUKERDUST dataset is a valuable addition to the existing collection of microbial image datasets. While there are other datasets available, such as the MicrobIA Haemolysis Dataset published in [10], they are limited in their variability and number of samples. The BRUKERDUST dataset, on the other hand, contains unique images of plates with various types of agars and microbes, providing a diverse range of samples for researchers to work with. Additionally, the BRUKERDUST dataset is a natural extension of the previously published BRUKERCOLONY dataset, as it was obtained with the same prototype of a colony-picking robot and under the same lighting conditions, ensuring consistency in the image properties. BRUKERCOLONY dataset along with the BRUKERDUST dataset is to our knowledge the biggest publicly available dataset of agar plates annotated with segmentation maps which also captures the variability of real-life data. We are also introducing a new BRUKERAGAR dataset which contains high-quality images of plates with various agar types, providing a valuable resource for researchers in the field of synthetic data generation and beyond. To our knowledge, this type of dataset has not been previously published, making it a valuable addition to the scientific community. The findings of this work are meant to improve colony localization tasks in the prototype of MBT Pathfinder® colony picking robot.

2. Related work

The first widely known tool for image processing of agar plates is the OpenCFU [11] which was specifically designed to detect circular colonies using traditional image processing techniques from the OpenCV [12] framework such as thresholding and particle filtering. OpenCFU was outperformed by the AutoCellSeg [13] which used techniques such as watershed segmentation. Another tool used is called CellProfiler [14]. It is a tool for analyzing different kinds of cell images including microscopic images. The versatility of this tool stems from the possibility of creating custom processing pipelines. All of these tools require user interaction, and since none of the approaches utilize machine learning, they are likely to be outperformed by more sophisticated solutions that do.

Another approach to microbial colony semantic segmentation was brought by the article [15]. The authors trained a CBDN (convolutional deep belief network) network which acts as a feature extractor and the segmentation itself was performed with SVM (support vector machine). The article faces limitations due to the use of a small custom dataset, making comparisons with other studies challenging. Additionally, the method employed in the study focuses on analyzing small cropped patches of microbial colonies rather than entire images, which may have implications for the accuracy and applicability of the results.

The article [16] targeted the foreground and background contrast improvement by acquiring the image with lighting near the infrared spectrum. The postprocessing of the acquired image was then performed by simple thresholding. The proposed solution requires a custom acquisition unit, making it challenging to integrate into common laboratories. Additionally, the limited number of samples used in their study raises questions about the generalizability of their proposed solution to other types of samples.

The article [17] applies semantic segmentation of agar plate images with a trained U-Net and a U-Net with pretrained backbone models. The authors aim to segment the microbes into two groups (BVG+ and BVG-). The training was performed with a custom dataset with 108 images. Once again the articles use a custom dataset with a limited number of samples.

Analyzing agar plate images poses a significant challenge due to the absence of comprehensive benchmarking data for a unified and reliable evaluation of proposed solutions. This scarcity stems from the vast variability in agar plates, including agar types, microorganisms, and streaking patterns. The extensive diversity in these elements makes the image data highly variable. Moreover, acquiring data requires specialized equipment and is restricted by regulatory constraints, further limiting its availability. Consequently, datasets related to agar plates are often small-scale, akin to numerous instance segmentation datasets in the biomedical technology domain. Notable datasets include those from the MicrobIA project. The MicrobIA Haemolysis dataset contains 235 agar plate images with microbes annotated with pixel-wise segmentation masks. In article [10] the Segments enumeration dataset of approximately 28 500 images of microbial patches was also presented. Each image has a classification label based on the number of colonies in the image (1 to 7 colonies and outliers) which makes them suitable for use in colony counting tasks. Another dataset is the AGAR dataset presented in the article [18] which is annotated with bounding boxes around the colonies. It contains 18,000 plates and it is well suited for use in colony counting applications. For completeness, we also include the dataset DIBas which contains 660 images of microscopic bacterial images (33 species each of 20 images). Another publicly available dataset is the BRUKERCOLONY dataset published in [19]. This dataset is going to be inspected in detail in the methodology section of this article.

Challenges related to biomedical data acquisition can be partially mitigated through the use of data augmentations and synthetic image generation techniques [20]. Augmentations entail controlled modifications applied to existing data, enhancing dataset diversity. Conversely, synthetic image generation enables the creation of artificial data samples that replicate real-world scenarios or anonymize existing data. Furthermore, recent research [21] emphasizes the importance of data-efficient instance segmentation models in computer vision. It introduces the Copy-Paste augmentation technique, where objects are randomly pasted onto images. This method proves effective in enhancing instance segmentation, even without modeling the surrounding visual context for pasting objects.

In the realm of synthetic image generation for agar plate analysis, various approaches have been explored. One notable method, introduced in [22], employs generative adversarial networks to generate binary masks, subsequently used to render colonies into agar plate images. It is worth mentioning that this work is an extension of a previous article [23]. However, the proposed solution primarily utilizes red and white agar from the MicrobIA dataset for both training and evaluation, with the evaluation focused on a limited set of blood agar samples. Another approach to agar plate image generation is presented in [24], where random colony patches from selected dishes in the AGAR dataset are generated and superimposed onto empty agar plates. The study also demonstrates that applying photorealistic style transfer to the resulting agar plate images enhances the performance of benchmarking object detection models. Although the authors employ a more diverse and extensive AGAR dataset for training several object detectors, their performance evaluation focuses solely on bounding box positions.

In biomedical image segmentation, U-Net [25] stands as the gold standard, demonstrating remarkable efficiency in utilizing annotated samples and showcasing superior performance even with a small dataset. Attention U-Net [26] integrates attention gates to focus on target structures, enhancing prediction accuracy and sensitivity without relying on explicit localization modules. Conversely, UNet3+ [27] addresses accurate segmentation by leveraging full-scale skip connections and deep supervisions, notably benefiting organs at varying scales while optimizing computational efficiency.

Additionally, several approaches toward the acquisition of agar plates have been explored. Article [28] discusses an automatic colony counting system that can be implemented with basic image processing knowledge and a low budget, making it accessible to many laboratories. Study [16] proposes a cost-effective system that uses near-infrared light to obtain colony images and an automatic method for detecting and counting colonies with promising precision and time-saving capabilities. These related works collectively highlight the ongoing efforts in automating colony counting processes while considering factors such as cost, ease of implementation, and image quality.

To summarize, state-of-the-art methods for microbial image analysis often use custom datasets that are small, not variable enough, and often not publicly available. This leads to difficulties in comparing proposed solutions and limits their applicability in real-world settings. Moreover, many of these datasets are not annotated, which further impedes progress in the field. While synthetic data generation methods exist, they do not provide enough high-quality data and may not be practical for real-life scenarios. Furthermore, none of the existing solutions address visual changes in agar plates caused by chromogenic agar types or blood haemolysis. To address these challenges, there is a need for a publicly available, complex dataset with labeled images that can serve as a basis for developing and evaluating machine learning algorithms for microbial image analysis. Such a dataset would enable researchers to develop and compare robust, generalized solutions that can be used in a variety of real-world applications.

3. Materials and methods

The main objective of the experiments is to implement and evaluate an image processing pipeline for generating annotated agar plate images with microbes. We put a special emphasis on data variability. The pipeline performance was benchmarked with the gold standard segmentation architecture U-Net and more up-to-date Attention U-Net, and UNet3+ architectures.

3.1. Data acquisition

In our research workflow, we initiated the process by acquiring the essential data for our image generator. The foundation of this generator is an empty agar plate image. This image is subsequently populated with colony patches, and each patch is paired with its corresponding binary mask annotation. As a result, we are able to generate a segmentation annotation mask for each agar plate image.

We leveraged multiple datasets in this study. Our primary dataset, detailed in [19], was the BRUKERCOLONY dataset, serving as our foundational training dataset and the source for extracting microbial colony patches. Additionally, we introduced two novel datasets: the BRUKERAGAR dataset and the BRUKERDUST dataset. These datasets featured images captured using the same colony-picking robot as those in the BRUKERCOLONY dataset, ensuring uniform image sizes and consistent lighting conditions throughout our data sources.

3.1.1. BRUKERCOLONY dataset

The BRUKERCOLONY dataset was partitioned into three segments: training (70%), evaluation (20%), and testing (10%). The training segment played a pivotal role in the extraction of microbial patches, as elaborated in the subsequent subsection. It encompassed 547 meticulously annotated images of agar plates predominantly featuring microbial monocultures on red or white agar. Some of these images capture identical samples acquired under varying lighting conditions and/or rotations.

We present visual representations of the color distribution of the aforementioned datasets as follows: Fig. 1 showcases selected samples from the MicrobIA dataset, and Fig. 2 displays samples from the BRUKERCOLONY dataset. It is apparent that the color diversity within the BRUKERCOLONY dataset surpasses that of the MicrobIA dataset.

For the extraction of colony patch images, we exclusively employed the training portion of the BRUKERCOLONY dataset. This choice aimed to mitigate potential evaluation bias introduced by employing data during the model training phase. In total, we extracted 20,417 colony patch images. The raw colony patches were acquired using a straightforward image processing pipeline same as in [19]. We determined colony positions utilizing watershed segmentation, where ground truth masks served as the markers for the algorithm. The colony center was computed by averaging the coordinates of all points along the contour in the segmented

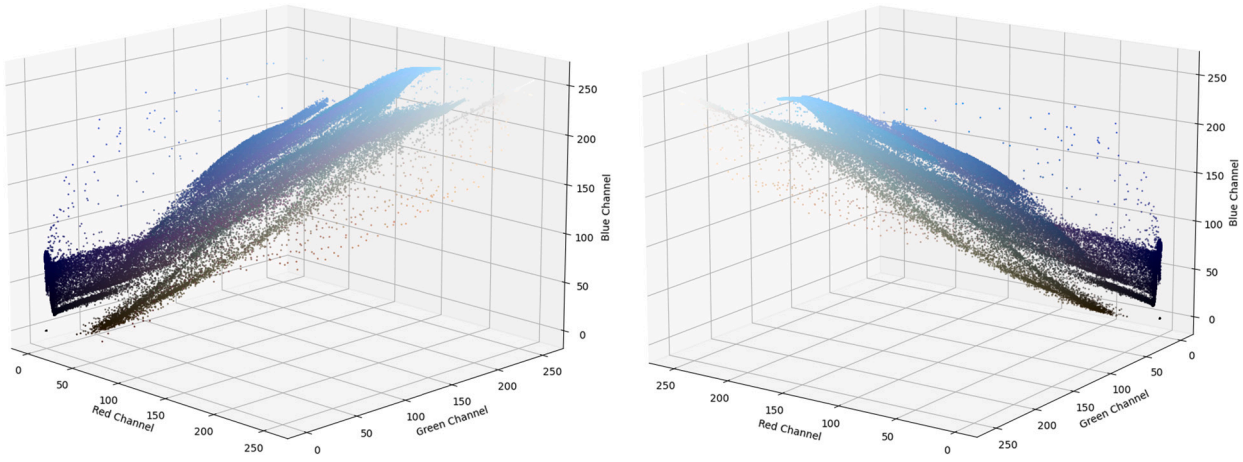


Fig. 1. Color distributions of selected samples from the MICROBIA dataset.

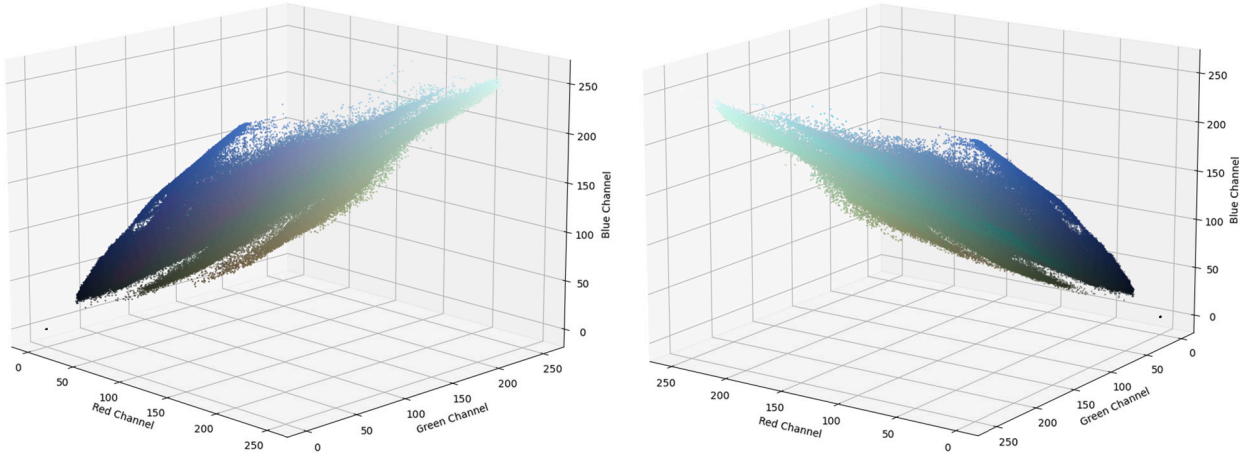


Fig. 2. Color distributions of selected samples from the BRUKERCOLONY dataset.

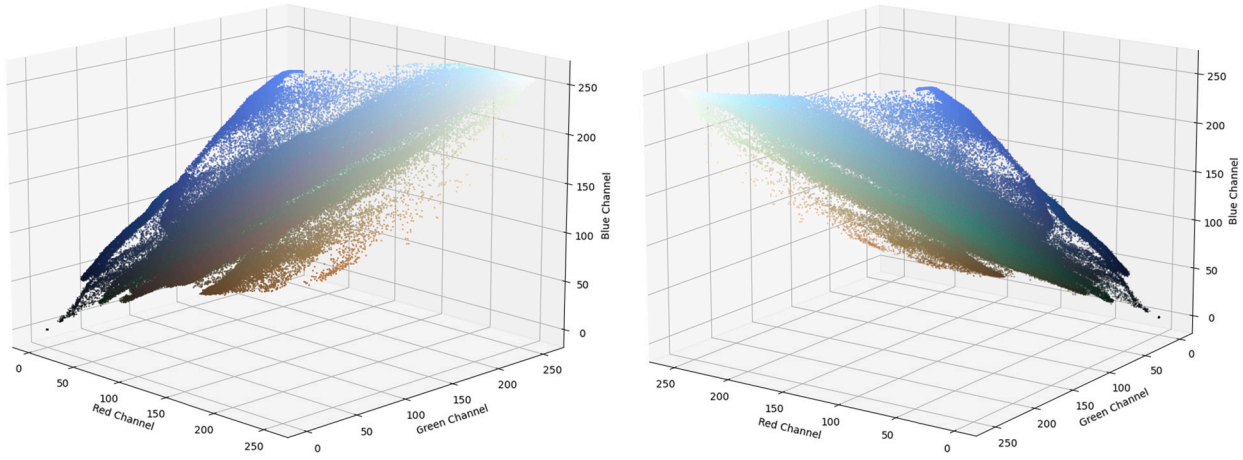


Fig. 3. Color distributions of selected samples from the BRUKERDUST dataset.

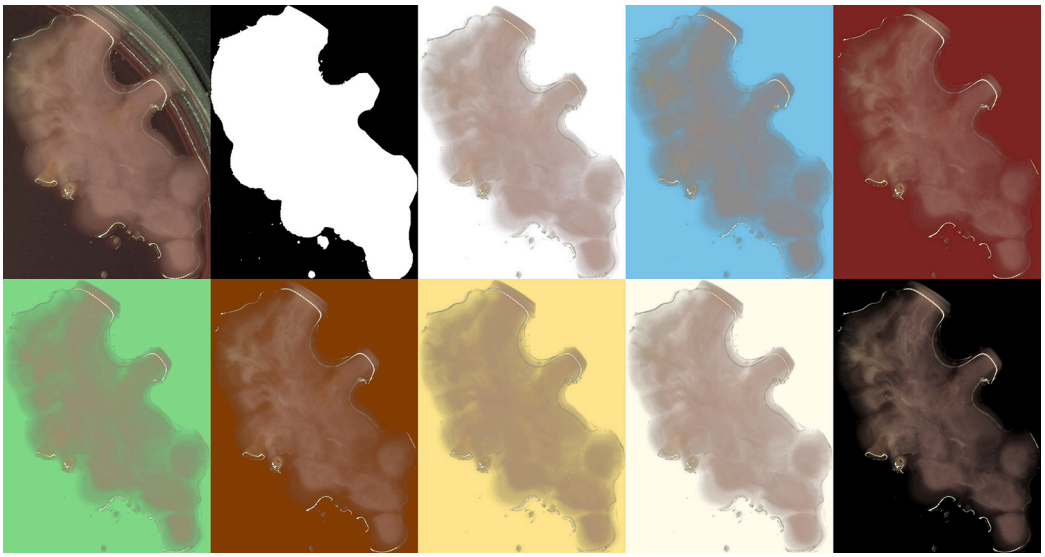


Fig. 4. Colony patch on various backgrounds (the top left image is the original image from which the colony patch was extracted and the image to the right is its mask).

(a) Colony patch (b) Colony patch annotation (c) Adjusted annotation

Fig. 5. Colony patch annotation adjustments.

image. Subsequently, both the mask and the image were cropped and stored using the derived contour. Example of an extracted colony patch can be seen in Fig. 4.

Out of the pool of extracted images, we deliberately selected those depicting intact colonies to ensure the exclusion of images with sharply truncated edges, thus maintaining a real-world appearance in the subsequent image generation process. In cases where a colony patch contained multiple colonies, manual adjustments were applied to the mask to isolate the colony of interest (as exemplified in Fig. 5). Furthermore, we employed morphological dilation on the masks to rectify minor imperfections that may have arisen from potential annotation errors. This step not only contributed to a smoother colony mask but also guaranteed the presence of agar pixels along the colony peripheries.

3.1.2. BRUKERAGAR dataset

As the BRUKERCOLONY dataset lacked images of empty agar plates without cultivated microbes, we took the initiative to create a custom dataset to fill this gap. To ensure the dataset's authenticity and reflect real-world variability, we captured images of empty agar plates, each featuring one of 23 distinct agar types, with ten plates available per type. These plates were photographed from a top-down perspective under six different lighting conditions using the prototype of MBT Pathfinder® colony-picking robot. Additionally, we obtained bottom-view images of each agar plate, although they were not used in this specific work. Consequently, this comprehensive effort resulted in a dataset comprising a total of 1,610 images, which we have made publicly accessible at the following URL: https://github.com/michalcicatka/BRUKERAGAR-DATASET. Exemplary images from this dataset can be found in Fig. 6.

3.1.3. BRUKERDUST dataset

To address the limited variety of agar types and microbes found in the previously mentioned publicly available datasets, we recognized the necessity for a comprehensive evaluation dataset tailored to our image generator. This led to the creation of the bespoke BRUKERDUST dataset. After obtaining images for the BRUKERAGAR dataset, we took Petri dishes and inoculated them with swabs collected from various laboratory surfaces. These inoculated dishes were then left to cultivate at room temperature for a duration of 72 hours. Subsequently, we scanned all the resulting agar plates under six distinct lighting configurations using the MBT

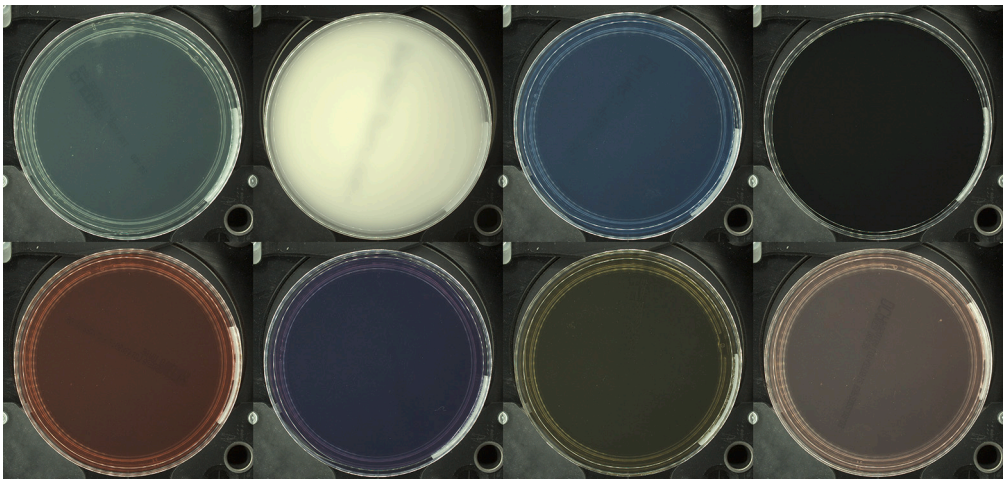
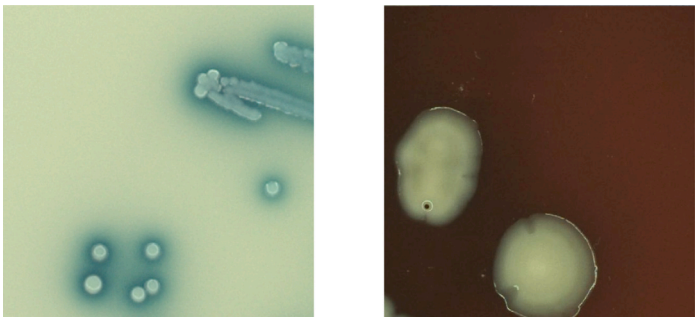


Fig. 6. Examples of BRUKERAGAR dataset.



(a) Chromogenic agar reaction (b) Blood agar haemolysis

Fig. 7. Examples of agar plates visual phenomena.

Pathfinder® colony picking robot. Following visual inspections, we manually annotated the agar plates that were cultivated with microbes. In total, we scanned and processed 854 images for this dataset. It has been made publicly accessible online at the following URL: https://github.com/michalcicatka/BRUKERDUST-DATASET.

Judging from the charts depicted in Fig. 1, 2 and 3, it becomes evident that when considering the color distribution, the BRUKERDUST dataset surpasses both the MicrobIA dataset and the BRUKERCOLONY dataset in terms of diversity.

3.2. Synthetic data generation

Following the data acquisition phase, we proceeded to implement the image generator. Beyond its fundamental image generation capability, our generator incorporates advanced features. Specifically, we integrated functionalities to simulate agar haemolysis and chromogenic reactions. These phenomena are visually represented in Fig. 7 using real agar plate images. These enhancements significantly improve our ability to replicate real-world scenarios in our synthetic images.

The process of generating synthetic agar plate images encompasses several sequential steps. Initially, a random batch of colony patches is selected, and alterations are made to their color, size, and rotation, effectively introducing variability into the dataset. Subsequently, great care is taken in arranging these colonies within the agar plate image to prevent any overlap. Overlaps would lead to an unnatural appearance, which we aim to avoid. Once this meticulous placement is achieved, the synthetic agar plate image is considered complete. Finally, additional augmentations in the form of simulating chromogenic appearance and blood agar haemolysis, can be applied. These augmentations not only further enrich the dataset's diversity, but may also serve future research needs, such as generating specific data for the study of chromogenic samples or the identification of specific microbes that cause haemolysis.

3.2.1. Colony patch batch selection

The selection of colony patches for each synthetic image plays a vital role in achieving the desired microbial coverage on the agar plate. To determine the extent of microbial coverage on the agar plate, we set a specific coverage ratio. Colony patches are then sequentially chosen to populate the agar plate, and this selection process continues until the targeted coverage is achieved. The experiments described in the following subsections used random target coverage levels ranging between 15% and 60%.

In addition to colony patch selection, we introduce further morphological modifications to enrich the dataset. These modifications encompass random flipping and rotation of colony patches. By randomly flipping and rotating colony patches, we enhance dataset diversity, ensuring that colonies can appear in various orientations. To further diversify these synthetic images, we take an additional step. We convert the colony patches into the HSV (Hue, Saturation, Value) color space. Within this color space, we introduce a random hue adjustment ranging from -45 to 45 degrees. This specific hue adjustment was determined empirically, resulting in colony patches displaying a wide range of colors. These combined modifications significantly contribute to the dataset's diversity, enabling the training of segmentation models capable of handling the intricacies of real-world data.

3.2.2. Colony patch positioning

To determine the valid positions of colony patches on the agar plate, we employ a genetic algorithm. This algorithm generates both the resulting agar image and its corresponding segmentation annotation mask. The fitness of the generated image is calculated based on a binary mask, taking into account two crucial factors. The first factor considers the total number of pixels on the binary mask that adheres to a specified maximum distance from the image center, ensuring that colony patches are placed within the Petri dish boundaries in the image. The second factor accounts for the cumulative pixel count of all colony patches in the image, reflecting whether any colony patches overlap. Consequently, agar plates containing colony patches outside the plate's boundaries or overlapping colony patches yield a lower fitness score.

To optimize the algorithm's performance, in our experiments, we used a downsized image mask and colony patches to 512 to 512 pixels. This approach strikes a balance between computational efficiency and the possibility of neighboring colony patches touching. It's important to note that the minimal colony patch size is influenced by the downsampling factor. However, empirical testing has confirmed that this factor has an insignificant impact on the results.

This fitness evaluation is expressed through Eq. (1).

$$\begin{aligned} Current &= \sum_{n=1}^{N} M_{xy} : \ (M_{xy} \neq 0 \wedge |M_{xy}C| < r) \\ Total &= \sum_{n=1}^{N} PM_{n_{xy}} : \ PM_{n_{xy}} \neq 0 \\ Fitness &= \frac{Current}{Total} \end{aligned} \tag{1}$$

Here, M represents the currently assessed mask, PM stands for the mask of a colony patch, C denotes the center of the image, and r represents the radius of the Petri dish.

3.2.3. Colony patch superimposition

The superimposition of colony patches onto the agar plate image is a crucial step in the image generation process. To achieve this, we have developed a straightforward yet effective approach. Initially, we compute the average background (agar) color of the colony patch's neighboring area. This is done by calculating the mean value of all pixels in the source image of the colony patch that are marked as background in the colony patch mask. Mathematically, this process is represented by Eq. (2).

$$PB = Mean(PI_{xy}) : \ PM_{xy} \neq 0 \tag{2}$$

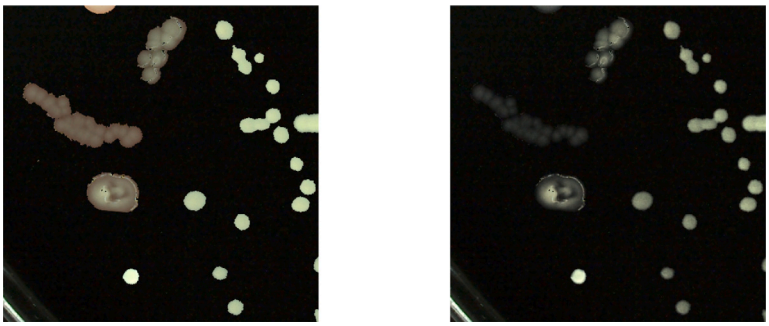
Here, PB represents the mean colony patch background color, PI stands for the colony patch image, and PM is the colony patch mask. With this information, we can calculate the alpha channel of the colony patch as the distance of each pixel channel from the average background color (Eq. (3)). Here, α represents the alpha channel of the colony patch.

$$\begin{aligned} \alpha_R &= |I_{xy_R} - PB_R| \\ \alpha_G &= |I_{xy_G} - PB_G| \\ \alpha_B &= |I_{xy_B} - PB_B| \\ PI_{xy_\alpha} &= 255 - (\alpha_R + \alpha_G + \alpha_B) \end{aligned} \tag{3}$$

Due to the translucent nature of many colonies, the background color can influence the overall color of the colony, which needs to be compensated for. To address this, we apply a color adjustment with Eq. (4). In Eq. (4), CS denotes a color shift, PIM represents the mean color of the colony patch, PB is the mean colony batch background color, API is the adjusted colony patch image, and PI is the colony patch image.

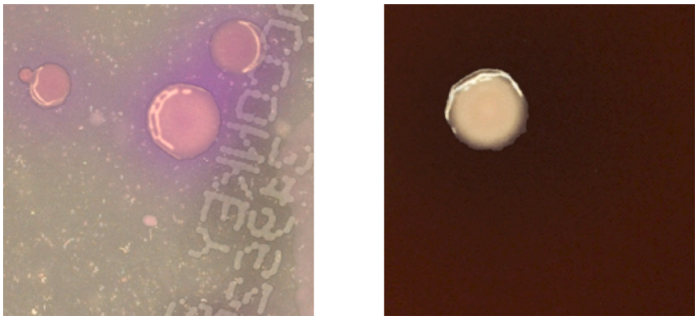
$$\begin{aligned} CS &= PIM - PB \\ API_{xy} &= PI_{xy} - CS \cdot (1 - \frac{PI_{xy_\alpha}}{255}) \end{aligned} \tag{4}$$

Since a simple copy-paste augmentation alone does not produce a realistic appearance, we have taken steps to overcome this limitation. To achieve a more convincing result, we have implemented a blending algorithm. During the colony superimposition process, this algorithm takes into account the color of each colony patch and weights it by its alpha channel. This results in a



(a) Superimposition without blending algorithm (b) Superimposition with blending algorithm

Fig. 8. Comparing superimposition with and without blending algorithm.



(a) Chromogenic agar reaction (b) Blood agar haemolysis

Fig. 9. Examples of simulated agar plate visual phenomena.

seamless blending of the patch's edges into the empty agar plate image. The precise description of this blending process is captured by Eq. (5). In this equation, SI represents the synthetic image, BI is the baseline empty agar image, and PI is the patch image.

$$SIxy = (1 - PI_{ij_\alpha}) \cdot BI_{xy} + PI_{ij_\alpha} \cdot PI_{ij} \tag{5}$$

For a visual representation of the resulting colony patch with an adjusted alpha channel placed on various backgrounds, please refer to Fig. 4. Furthermore, Fig. 8 offers a compelling comparison of agar plate image details acquired without the blending algorithm, underlining the significance of this enhancement in achieving a realistic appearance.

3.2.4. Agar plate image augmentation

During microbial cultivation on the agar plate, some visual phenomena may occur. One of the most common would be haemolysis (breakdown of red blood cells) of the blood agars which causes a change of agar color around the colony's neighboring area. Another common phenomenon would be visual changes in the chromogenic agars. We developed algorithms to simulate these phenomena in synthetic dishes.

The haemolysis simulation uses an adjusted binary mask as a weighting factor for the alteration. The adjustment consists of morphological dilation and Gaussian blur. The visual phenomena are then simulated by decreasing the Value and Saturation of the affected pixels in the HSV model weighted by the adjusted mask. This is done before the superimposition of the colonies, so only the agar is affected.

The principle is very similar when simulating the chromogenic changes. The color of the chromogenic area is selected first. The binary mask adjusted in a very similar way as with the haemolysis is used to weigh the degree of alteration of each processed pixel. The alterations consist of shifting hue, value, and saturation when a pixel is represented in the HSV color space. The agar plate augmentations are going to be further explored in the experiment setup subsection (Sec. 3.3). Examples of the agar augmentations are in Fig. 9

3.3. Experiment setup

The evaluation of the synthetic data generation pipeline is assessed through the performance of trained segmentation models. To optimize the training process, each image sample was divided into 16 images, each sized at 512 by 512 pixels. This division allowed for the inclusion of more images in each training mini-batch. Dice loss [29] was used in all experiments.

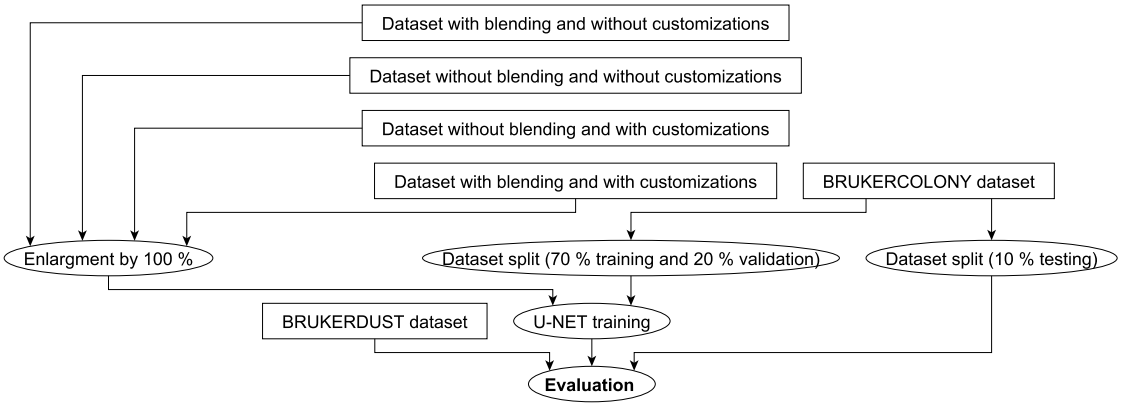


Fig. 10. Diagram of the agar augmentation experiment.

For evaluating model performance, we employed two widely used segmentation metrics: the Jaccard score, also called IoU (Eq. (6)), and the Dice coefficient, also called the F1 score (Eq. (7)). In the mentioned equations, $Dice$ represents the Dice coefficient, $Jaccard$ signifies the Jaccard score, TP stands for true positive, FP represents false positive, and FN denotes false negative.

$$Jaccard = \frac{TP}{TP + FP + FN} \tag{6}$$

$$Dice = \frac{2TP}{2TP + FP + FN} \tag{7}$$

All experiments were evaluated using the same testing dataset, comprising exclusively of real-life images. This dataset was derived from a combination of 10% (51 images) of the BRUKERCOLONY dataset (with the remainder allocated for training and validation experiments) and 43 randomly selected samples from the BRUKERDUST dataset. It is worth noting that during the process of subdividing the sample images into smaller units, some images might not contain any microbes. To prevent division by zero when calculating metrics, these particular images were excluded from the testing dataset.

The first experiment aimed to evaluate how the inclusion of synthetic data, both with and without agar augmentations, affected the performance of the model. In the second experiment, we investigated how altering the ratio of real to synthetic data influenced the model's performance. The first two experiments utilized trained U-Net models for evaluation. Apart from the established U-Net architecture, we conducted an architecture experiment involving Attention U-Net and UNet3+.

3.3.1. Agar augmentation experiment

The agar augmentation experiment, following the diagram in Fig. 10, was designed to understand whether our proposed enhancements such as the blending algorithm and chromogenicity and blood haemolysis simulation had any significant impact on the model's performance. It involved the generation of a dataset using our image generator. Each sample in this dataset was saved in four different configurations of superimpositions:

1. Without blending the agar plate image and colony patch images together.
2. With blending the agar plate image and colony patch images together.
3. With blending the agar plate image and colony patch images together, along with additional augmentations.
4. Without blending the agar plate image and colony patch images together, but with additional augmentations.

These additional augmentations included a 50% probability of haemolysis simulation and a 50% probability of chromogenic agar simulation.

Subsequently, the resulting datasets were utilized to augment the original training and validation datasets by 100%. Following this augmentation, a U-Net model was trained using these datasets, and its performance was evaluated using the testing dataset. The trained U-Net models consisted of 5 encoder/decoder blocks and initiateed with 32 channels in the first convolutional layer. During the training phase, we implemented early stopping with a patience of 10 epochs and utilized exponential learning rate decay, reducing the learning rate by a factor of 0.1, starting from the 15th epoch, from the initial learning rate of 1e-5. The training process was executed in a Python environment using the TensorFlow Framework, leveraging the computational capabilities of an NVIDIA RTX 4090 GPU.

3.3.2. Training data volume experiment

The training data volume experiment, which follows the diagram in Fig. 11, provided insights into whether increasing the proportion of synthetic data in the training dataset had any significant impact on the model's segmentation capabilities.

The initial step involved training a U-Net model exclusively on real-world data sourced from the BRUKERCOLONY dataset. This model served as our baseline. Subsequently, synthetic datasets were generated using the image generator. Within these datasets,

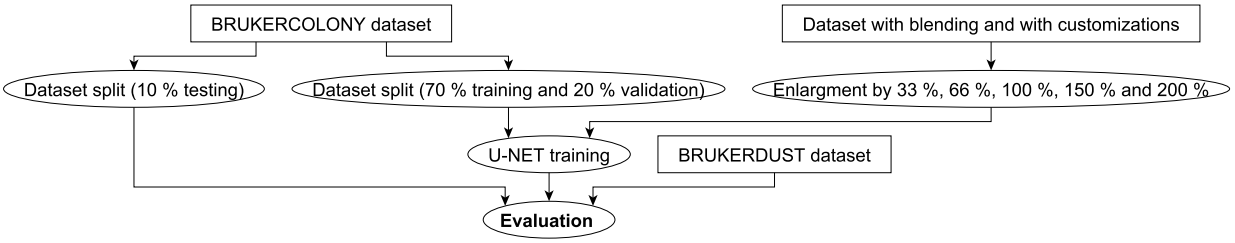


Fig. 11. Diagram of the training data volume experiment.

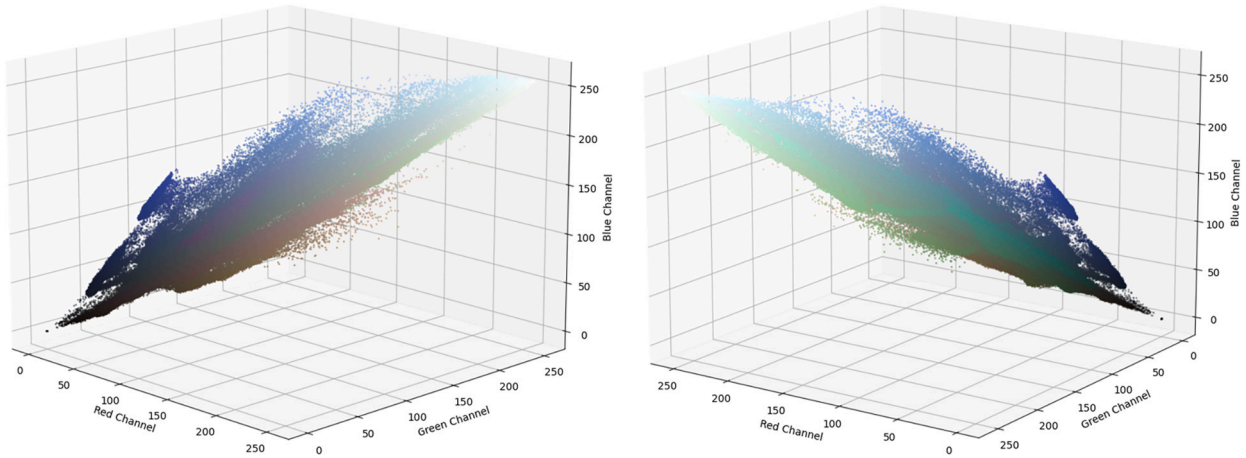


Fig. 12. Color distributions of selected synthetic images.

colony patches were superimposed using our blending algorithm, and each image had a 50% probability of haemolysis simulation and a 50% probability of chromogenic agar simulation. This specific configuration was selected based on the observed performance improvement in the prior experiment, as elaborated upon in the following sections.

For comprehensiveness, we also included a comparison of the color distribution between selected samples from the synthetic images (Fig. 12) and other datasets. These charts illustrate that the synthetic images displayed a more extensive color distribution when compared to the MicrobIA and BRUKERCOLONY datasets. Moreover, they were found to be on par with the color distribution of the BRUKERDUST dataset. This observation underscores the effectiveness of our image generation methodology in capturing the diversity and complexity of real-world agar plates.

Following this, the synthetic datasets were used to augment the original training and validation datasets by various percentages, ranging from 33% to 200%. Each of these augmented datasets was utilized to train separate U-Net models, and the performance of each model was evaluated using the testing dataset. All models were trained with the same configurations as in the agar agumentation experiment.

3.3.3. Architecture experiments

The objective of the architecture experiment was to evaluate the compatibility of our image generation approach with modern deep learning frameworks by assessing its performance across these contemporary architectures, namely Attention U-Net and UNet3+. These architectures were trained using the optimal configuration we had identified, involving a training dataset augmented by 150% with our proposed augmentations. To enhance the model's performance, we employed a VGG16 [30] backbone pretrained on ImageNet [31] for both architectures. The batch size was set to 32 and 20 image samples for Attention U-Net and UNet3+, respectively. The learning rate was initialized at $1e-4$, with an exponential decay of $5e-2$ after 20 epochs. Implementing an early stopping mechanism with a patience of 20 epochs ensured efficient model convergence. For a straightforward comparison, we also trained an additional U-Net model using a pretrained VGG16 backbone and the same training parameters. These architectures were trained in a Python environment utilizing the TensorFlow Framework, harnessing the computational capabilities of two NVIDIA RTX 4090 GPUs. This strategy allowed us to evaluate the adaptability and performance of our proposed image generation technique across a spectrum of modern architectures.

4. Results

We conducted three distinct experiments. Initially, we examined the impact of proposed augmentation techniques on U-Net model performance. Next, we explored how different proportions of synthetic data affected segmentation performance. Lastly, we compared our augmentation approach across contemporary architectures, providing insights into its adaptability and effectiveness.

Table 1
Agar augmentation experiments result.

Experiment	Dice coefficient	Jaccard score
Without blending and without augmentations	0.671	0.601
Without blending and with augmentations	0.693	0.628
With blending and without augmentations	0.697	0.628
With blending and augmentations	**0.721**	**0.648**

Table 2
Volume experiments result.

Experiment	Dice coefficient	Jaccard score
Original BRUKERCOLONY dataset	0.518	0.453
Original BRUKERCOLONY dataset augmented with 33%	0.677	0.606
Original BRUKERCOLONY dataset augmented with 66%	0.719	0.647
Original BRUKERCOLONY dataset augmented with 100%	0.726	0.656
Original BRUKERCOLONY dataset augmented with 150%	**0.729**	**0.656**
Original BRUKERCOLONY dataset augmented with 200%	0.717	0.645

Table 3
Architecture experiment result.

Experiment	Dice coefficient	Jaccard score
U-Net with backbone	0.736	0.660
Attention U-Net with backbone	0.760	0.681
Unet3+ with backbone	**0.767**	**0.690**

4.1. Agar augmentation experiment

The agar augmentation experiment was designed to evaluate the impact of various post-processing algorithms, including blending superimposition and agar augmentations, on the performance of U-Net models. The results of these experiments are summarized in Table 1. The significant improvement in the Dice coefficient from 0.671 to 0.721 and the Jaccard score from 0.601 to 0.648 when applying all augmentations highlights the efficacy of our methodology.

4.2. Volume experiment

Moving on to the volume experiments (Table 2), our objective was to understand how varying the proportion of synthetic data in the training dataset affects model performance. The baseline dataset achieved a Dice coefficient of 0.518 and a Jaccard score of 0.453. The model trained with a 150% enlargement of the original dataset achieved the highest Dice coefficient of 0.729 and a Jaccard score of 0.656. Augmenting the dataset with 200% of the original dataset resulted in a Dice coefficient of 0.717 and a Jaccard score of 0.645.

4.3. Architecture experiment

The results of the architecture experiment are presented in Table 3. Three distinct architectures, namely U-Net with backbone, Attention U-Net with backbone, and Unet3+ with backbone, were evaluated based on their Dice coefficient and Jaccard score. The U-Net model, serving as a baseline, achieved a Dice coefficient of 0.736 and a Jaccard score of 0.660. Introducing attention mechanisms in the Attention U-Net yielded a notable improvement, reaching a Dice coefficient of 0.760 and a Jaccard score of 0.681. The UNet3+ architecture showcased the highest performance, with a Dice coefficient of 0.767 and a Jaccard score of 0.690. These results highlight the efficacy of our proposed image generation methodology in conjunction with different modern architectures, demonstrating the versatility and potential of our approach for enhancing segmentation models.

5. Discussion

This article introduces two new datasets of agar plates and a novel image processing pipeline for generating annotated synthetic agar plate images. The performance of the generator was assessed through trained segmentation models.

In the agar augmentation experiment (Table 1), the results clearly demonstrate that the proposed augmentations significantly enhance the performance of the trained models. Specifically, applying blending superimposition and agar augmentations led to a notable improvement in both the Dice coefficient (from 0.671 to 0.721) and the Jaccard score (from 0.601 to 0.648). This suggests that these augmentations effectively enhance the realism of the data and contribute to better model performance in accurately segmenting microbial colonies on the agar plate images. Since a simple copy-paste augmentation already improves the performance

of the trained models, as evident from our experiments and the study [21], our proposed blending technique, which blends colony patch images with the agar plate background, proves to be a crucial component for achieving an even more realistic appearance and improved segmentation performance. Our study extends the findings at [24], which primarily focused on object detection rather than segmentation. Our datasets contain a greater variety of microbes and agar types. Instead of relying on style transfer for enhancing diversity through image transfer, we improved the source data and applied our custom augmentations. Comparing the results presented in this article to the study [22], where a notably higher Jaccard score of 0.936 was achieved, it's important to note that the testing in their study was performed on the MicrobIA dataset with a single type of agar, which typically exhibits high contrast of colonies. This fact underlines the need for a reliable testing dataset such as BRUKERDUST, which facilitates easier comparisons across experiments focused on agar plates.

In the volume experiment (Table 2), we aimed to understand how varying the proportion of synthetic data in the training dataset affects model performance. Interestingly, the best-performing model was achieved by augmenting the dataset with 150% of synthetic data, resulting in the highest Dice coefficient of 0.729 and a Jaccard score of 0.656. However, augmenting the dataset with 200% of the original dataset resulted in a decrease in performance, suggesting a potential issue of overfitting, likely due to the limited number of colony patches, especially those with larger microbial areas. Intriguingly, even a modest proportion of synthetic data (33%) led to significant improvements in the Dice coefficient (0.677) and Jaccard score (0.606). Moreover, models trained with over 66% of synthetic data exhibited consistent and notable improvements in performance, underlining the robustness and effectiveness of incorporating synthetic data in varying proportions.

Moreover, the architecture experiment (Table 3) revealed significant advancements in model performance when incorporating modern architectures such as Attention U-Net and UNet3+, both utilizing a VGG backbone. Notably, these architectures surpassed the baseline U-Net model, emphasizing the adaptability and effectiveness of our augmentation approach across various network architectures. The best UNet3+ model demonstrated exceptional performance, achieving a Dice coefficient of 0.767 and a Jaccard Score of 0.690. Additionally, the U-Net with the backbone gained a Dice coefficient of 0.736 and a Jaccard Score of 0.660, surpassing the U-Net model without the backbone. The Attention U-Net also showcased respectable performance, achieving a Dice coefficient of 0.760 and a Jaccard Score of 0.681. These results underscore not only the efficacy of our augmentation methodology but also the added benefit of integrating a VGG backbone, collectively enhancing segmentation models for more accurate and robust agar plate analysis.

Overall, the integration of our synthetically generated agar plate images has proven to be a valuable technique for enhancing the performance of segmentation models on real-life data. This advancement, coupled with the datasets presented, holds promise for automating laboratory processes in agar plate analysis, traditionally reliant on manual parameter adjustments by experts.

6. Conclusion

This paper has demonstrated the significant impact of synthetic agar plate generation on enhancing automated agar plate analysis by providing a large and diverse dataset for training and testing segmentation models. The paper has presented two new annotated datasets and a new methodology for image generation that simulates various cultivation phenomena.

Through the incorporation of such phenomena in synthetic data generation, the resulting Dice coefficient of trained U-Net models significantly improved from 0.671 to 0.721. This improvement signifies the effectiveness of incorporating simulated cultivation phenomena in enhancing model performance.

The paper has also demonstrated that U-Net models trained with a combination of real and synthetic data can achieve high accuracy and efficiency in segmenting various agar plate types and microorganism species. Specifically, the U-Net model trained using a combination of real data augmented with 150% synthetic data exhibited superior performance compared to the model trained without synthetic data. It achieved a notable improvement, with a Dice coefficient of 0.729, surpassing the baseline result of 0.518.

In addition to the advancements mentioned, this study introduced a comprehensive architecture experiment that further strengthens the contributions of synthetic agar plate generation. By incorporating more up-to-date architectures, namely Attention U-Net and UNet3+, and employing our proposed augmentation techniques, we showcased substantial improvements over the baseline U-Net model. Notably, the UNet3+ architecture exhibited the highest performance, achieving a Dice coefficient of 0.767 and a Jaccard score of 0.690. This experiment highlighted the adaptability and potency of our augmentation approach across different architectures, emphasizing the potential for broader application in various biomedical image segmentation tasks.

Moreover, the paper has introduced two new annotated datasets, offering valuable resources for benchmarking segmentation algorithms. The datasets cover a wide range of agar plate types, microorganism species, and colony characteristics, contributing to the capture of real-world complexity and variability.

The findings presented in this paper contribute to the advancement of automated agar plate analysis, which holds significant implications for microbiological testing in industries like food and beverage, pharmaceuticals, and biotechnology. Looking ahead, future research could explore alternative approaches to augmenting real data with synthetic data, such as leveraging generative adversarial networks or style transfer techniques. Additionally, evaluating the performance of other segmentation models on the new datasets could further enrich the understanding of segmentation methods for agar plate analysis.

In conclusion, this study underscores the value of synthetic data generation and deep learning techniques in enhancing automated agar plate analysis. The combination of real and synthetic data has demonstrated promising results in tackling the challenges posed by diverse agar plates and microorganisms. The paper hopes to inspire more research and innovation at the intersection of computer vision and microbiology, advancing the field of automated microbiological testing.

CRediT authorship contribution statement

Michal Cicatka: Methodology. **Radim Burget:** Supervision. **Jan Karasek:** Resources. **Jan Lancos:** Methodology.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Michal Cicatka reports equipment, drugs, or supplies was provided by Bruker Daltonics GmbH & Co. KG. Jan Karasek reports a relationship with Bruker Daltonics GmbH & Co. KG that includes: employment. Jan Lancos reports a relationship with Bruker Daltonics GmbH & Co. KG that includes: employment. Michal Cicatka reports a relationship with Bruker Daltonics GmbH & Co. KG that includes: employment. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

The authors would like to thank the Bruker Daltonics GmbH & Co. KG for providing the prototype of MBT Pathfinder® colony picking robot and computational resources used in the methodology of this paper.

This work was supported by the Ministry of the Interior of the Czech Republic, (projects no. VK01010153 and VJ02010019).

References

[1] G. Lippi, G.D. Rin, Advantages and limitations of total laboratory automation: a personal overview, Clin. Chem. Lab. Med. 57 (6) (2019) 802–811, https://doi.org/10.1515/cclm-2018-1323 [cited 2023-09-07].

[2] C. Alonso, C. Domínguez, J. Heras, E. Mata, V. Pascual, C. Torres, M. Zarazaga, Antibiogramj: a tool for analysing images from disk diffusion tests, Comput. Methods Programs Biomed. 143 (2017) 159–169, https://doi.org/10.1016/j.cmpb.2017.03.010.

[3] O. Dauwalder, A. Michel, C. Eymard, K. Santos, L. Chanel, A. Luzzati, P. Roy-Azcora, J.F. Sauzon, M. Guillaumont, P. Girardo, C. Fuhrmann, G. Lina, F. Laurent, F. Vandenesch, C. Sobas, Use of artificial intelligence for tailored routine urine analyses, Clin. Microbiol. Infect. 27 (8) (2021) 1168.e1–1168.e6, https://doi.org/10.1016/j.cmi.2020.09.056.

[4] H.R. Frost, S.K. Tsoi, C.A. Baker, D. Laho, M.L. Sanderson-Smith, A.C. Steer, P.R. Smeesters, Validation of an automated colony counting system for group a streptococcus, BMC Res. Notes 9 (1) (Feb. 2016), https://doi.org/10.1186/s13104-016-1875-z.

[5] R. Heestermans, P. Herroelen, K. Emmerechts, K. Vandoorslaer, D.D. Geyter, T. Demuyser, I. Wybo, D. Piérard, A. Muyldermans, Validation of the colibrí instrument for automated preparation of MALDI-TOF MS targets for yeast identification, J. Clin. Microbiol. 60 (7) (Jul. 2022), https://doi.org/10.1128/jcm.00237-22.

[6] Z. Chen, J. Yang, L. Chen, Z. Feng, L. Jia, Efficient railway track region segmentation algorithm based on lightweight neural network and cross-fusion decoder, Autom. Constr. 155 (2023) 105069, https://doi.org/10.1016/j.autcon.2023.105069.

[7] Z. Feng, J. Yang, Z. Chen, Z. Kang, Lrseg: an efficient railway region extraction method based on lightweight encoder and self-correcting decoder, Expert Syst. Appl. 238 (2024) 122386, https://doi.org/10.1016/j.eswa.2023.122386.

[8] D. Cheng, X. Gao, Y. Mao, B. Xiao, P. You, J. Gai, M. Zhu, J. Kang, F. Zhao, N. Mao, Brain tumor feature extraction and edge enhancement algorithm based on u-net network, Heliyon 9 (11) (2023) e22536, https://doi.org/10.1016/j.heliyon.2023.e22536.

[9] W.K. Soh, H.Y. Yuen, J.C. Rajapakse, Hut: hybrid unet transformer for brain lesion and tumour segmentation, Heliyon 9 (12) (2023) e22412, https://doi.org/10.1016/j.heliyon.2023.e22412.

[10] A. Ferrari, S. Lombardi, A. Signoroni, Bacterial colony counting with convolutional neural networks in digital microbiology imaging, Pattern Recognit. 61 (2017) 629–640, https://doi.org/10.1016/j.patcog.2016.07.016, https://www.sciencedirect.com/science/article/pii/S0031320316301650.

[11] Q. Geissmann, Opencfu, a new free and open-source software to count cell colonies and other circular objects, PLoS ONE 8 (2013) e54072, https://doi.org/10.1371/journal.pone.0054072.

[12] G. Bradski, The OpenCV library, Dr. Dobb's J. Softw. Tools (2000).

[13] A.u.M. Khan, A. Torelli, I. Wolf, N. Gretz, Autocellseg: robust automatic colony forming unit (cfu)/cell analysis using adaptive image segmentation and easy-to-use post-editing techniques, Sci. Rep. 8 (1) (2018) 7302, https://doi.org/10.1038/s41598-018-24916-9.

[14] D.R. Stirling, M.J. Swain-Bowden, A.M. Lucas, A.E. Carpenter, B.A. Cimini, A. Goodman, Cellprofiler 4: improvements in speed, utility and usability, BMC Bioinform. 22 (1) (2021) 433, https://doi.org/10.1186/s12859-021-04344-9.

[15] D. Nie, E.A. Shank, V. Jojic, A deep framework for bacterial image segmentation and classification, in: Proceedings of the 6th ACM Conference on Bioinformatics, Computational Biology and Health Informatics, BCB '15, Association for Computing Machinery, New York, NY, USA, 2015, pp. 306–314, https://doi.org/10.1145/2808719.2808751.

[16] G. Zhu, B. Yan, M. Xing, C. Tian, Automated counting of bacterial colonies on agar plates based on images captured at near-infrared light, J. Microbiol. Methods 153 (2018) 66–73, https://doi.org/10.1016/j.mimet.2018.09.004, https://www.sciencedirect.com/science/article/pii/S0167701218306146.

[17] T. Naets, M. Huijsmans, P. Smyth, L. Sorber, G. de Lannoy, A mask R-CNN approach to counting bacterial colony forming units in pharmaceutical development, https://doi.org/10.48550/ARXIV.2103.05337, 2021, https://arxiv.org/abs/2103.05337.

[18] S. Majchrowska, J. Pawłowski, G. Guła, T. Bonus, A. Hanas, A. Loch, A. Pawlak, J. Roszkowiak, T. Golan, Z. Drulis-Kawa, Agar a microbial colony dataset for deep learning detection, arXiv:2108.01234, 2021.

[19] M. Cicatka, R. Burget, J. Karasek, Machine-learning approach to microbial colony localisation, in: 2022 45th International Conference on Telecommunications and Signal Processing (TSP), 2022, pp. 206–211, https://doi.org/10.1109/TSP55681.2022.9851236.

[20] M. Xu, S. Yoon, A. Fuentes, D.S. Park, A comprehensive survey of image augmentation techniques for deep learning, Pattern Recognit. 137 (2023) 109347, https://doi.org/10.1016/j.patcog.2023.109347, https://www.sciencedirect.com/science/article/pii/S0031320323000481.

[21] G. Ghiasi, Y. Cui, A. Srinivas, R. Qian, T. Lin, E.D. Cubuk, Q.V. Le, B. Zoph, Simple copy-paste is a strong data augmentation method for instance segmentation, CoRR, arXiv:2012.07177 [abs], 2020, arXiv:2012.07177, https://arxiv.org/abs/2012.07177.

[22] P. Andreini, S. Bonechi, M. Bianchini, A. Mecocci, F. Scarselli, Image generation by gan and style transfer for agar plate image segmentation, Comput. Methods Programs Biomed. 184 (2020) 105268, https://doi.org/10.1016/j.cmpb.2019.105268, https://www.sciencedirect.com/science/article/pii/S0169260719311216.

[23] P. Andreini, S. Bonechi, M. Bianchini, A. Mecocci, F. Scarselli, A deep learning approach to bacterial colony segmentation, in: V. Kůrková, Y. Manolopoulos, B. Hammer, L. Iliadis, I. Maglogiannis (Eds.), Artificial Neural Networks and Machine Learning – ICANN 2018, Springer International Publishing, Cham, 2018, pp. 522–533.
[24] J. Pawłowski, S. Majchrowska, T. Golan, Generation of microbial colonies dataset with deep learning style transfer, Sci. Rep. 12 (1) (2022) 5212, https://doi.org/10.1038/s41598-022-09264-z.
[25] O. Ronneberger, P. Fischer, T. Brox, U-net: convolutional networks for biomedical image segmentation, https://doi.org/10.48550/ARXIV.1505.04597, 2015, https://arxiv.org/abs/1505.04597.
[26] O. Oktay, J. Schlemper, L.L. Folgoc, M. Lee, M. Heinrich, K. Misawa, K. Mori, S. McDonagh, N.Y. Hammerla, B. Kainz, B. Glocker, D. Rueckert, Attention u-net: learning where to look for the pancreas, arXiv:1804.03999, 2018.
[27] H. Huang, L. Lin, R. Tong, H. Hu, Q. Zhang, Y. Iwamoto, X. Han, Y.-W. Chen, J. Wu, Unet 3+: a full-scale connected unet for medical image segmentation, arXiv:2004.08790, 2020.
[28] L. Hogekamp, S.H. Hogekamp, M.R. Stahl, Experimental setup and image processing method for automatic enumeration of bacterial colonies on agar plates, PLoS ONE 15 (6) (2020) 1–17, https://doi.org/10.1371/journal.pone.0232869.
[29] C.H. Sudre, W. Li, T. Vercauteren, S. Ourselin, M.J. Cardoso, Generalised dice overlap as a deep learning loss function for highly unbalanced segmentations, CoRR, arXiv:1707.03237 [abs], 2017, arXiv:1707.03237, http://arxiv.org/abs/1707.03237.
[30] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, arXiv:1409.1556, 2015.
[31] J. Deng, W. Dong, R. Socher, L.-J. Li, K. Li, L. Fei-Fei, Imagenet: a large-scale hierarchical image database, in: 2009 IEEE Conference on Computer Vision and Pattern Recognition, 2009, pp. 248–255, https://doi.org/10.1109/CVPR.2009.5206848.