Automated image analysis tool to measure microbial growth on solid cultures

Ancin-Murguzur, Fransisco Javier

Elsevier BV

Tieteelliset aikakauslehtiartikkelit
© Elsevier B.V.
CC BY-NC-ND https://creativecommons.org/licenses/by-nc-nd/4.0/
http://dx.doi.org/10.1016/j.compag.2018.06.031

https://erepo.uef.fi/handle/123456789/7361
Downloaded from University of Eastern Finland's eRepository
Automated image analysis tool to measure microbial growth on solid cultures

Francisco Javier Ancin-Murguzur\(^1\,*\), Aitor Barbero-López\(^2\), Sari Kontunen-Soppela\(^3\)
and Antti Haapala\(^2\)

\(^1\)Norwegian Institute for Bioeconomy Research, P.O. Box 115, N-1431 Ås, Norway
\(^2\)School of Forest Sciences, University of Eastern Finland
\(^3\)Department of Environmental and Biological Sciences, University of Eastern Finland

Address correspondence to Francisco Javier Ancin Murguzur, x.ancin@gmail.com.

*Current address: Norwegian Institute for Bioeconomy Research, P.O. Box 115, N-1431 Ås, Norway

Abstract

Microbial growth on culture media is a commonplace technique to estimate the growth rate and virulence of microbes, assess inhibitory effects of compounds and estimate potential damages of plant pathogens in agriculture. Growth area measurement of solid cultures is still commonly performed as a manual process that requires skilled technicians and substantial time, thus warranting an automated system to reduce the workload and increase measurement efficiency. A machine learning approach (Support Vector Machines) was developed to fully automate the area measurement process. We developed a functional model that processes images and returns the microbial area coverage considerably faster than a manual measurement method, with minimal user input and highly comparable results (R\(^2\)=0.97) applicable over large datasets.
Determination of microbial growth on cell cultures is a common practice in natural sciences, medicine and bioprocess engineering to provide relevant information on growth rates of different microbial strains or species (Jo, Kim, and Jung 2009; Masso and Vaisman 2007; Raz et al. 2012). Microbial growth measurements are applied in agriculture to identify plant pathogens, and assess the effectiveness of antimicrobial agents such as essential oils (López et al. 2005; White et al. 2001). Fungal or microbial area is usually estimated by measuring the colony radii or area at specific time intervals: this method assumes structural uniformity on the colony, but significant deviations from the circular shape lead to biased measurements. This can be mitigated by averaging measurements made in more than one direction over the colony at the cost of increasing the time allocation for analyzing each specimen, which leads to the need of more time efficient measurement methods (Alvarez 2004). An alternative, more precise approach to measure growth areas is to process images with image processing software, such as ImageJ (Abrámoff et al. 2004). Several computer-based image analysis methods have been successfully implemented for this purpose in the last years (Duan, Wang, and Chen 2012; Medeiros et al. 2016): these methods remain time-consuming and require skilled personnel to analyze the individual images, warranting the development of more efficient methods with minimal human supervision. High accuracy and operator-independence for detecting the biological growth from its background (growth media) are significant criteria to be met as previous work has shown that automatic image analysis can outperform visual examination and provide higher sensitivity to subtle differences (Murphy, Velliste, and Porreca 2003).

Machine learning based supervised image classification is a routine task in several disciplines, from remote sensing (Foody and Mathur 2004) to tumor detection in ultrasound images (Nascimento et al. 2016). Support vector machines (SVM) are a powerful tool often used in image processing techniques (Melgani and Bruzzone 2004), with easy implementation and fast computational times, generally with high accuracy. SVMs are commonly chosen due to their flexibility in hyperplane establishment to split between categories through kernelization and because they give consistent results once the hyperplane boundaries have been set (Shawe-Taylor and Cristianini 2004).

Our study on the suitability of SVM on growth and kinetics in agar cultures sets a novel approach that will reduce the needed time allocation to supervised measurement processes, while also providing the consistent quality and capability of more efficient image processing when dealing with substantial image archives. The
case study presented here deals with assessment of fungal growth area on a Petri dish, as compared between manual identification and SVM supervised classification.

Material and methods

Study subjects

Three fungal species with distinct growth patterns were used in this study: *Coniophora puteana* (BAM 112, Germany) with clear hyphae interface, *Gloeophyllum trabeum* (BAM 109, Germany) that usually shows diffuse boundary with growth media and *Rhodonia placenta* (BAM 113, Germany), which is easily confounded with the agar, since the hyphae are rather narrow and show diffuse features.

Growth media containing 5% malt and 2% agar were prepared, and 20 ml poured into 90mm Petri dishes. Petri dishes, two for each species, were inoculated with fungi by setting a spherical 0.275 cm$^2$ cutout in the center of the dish, and grown at 20°C and 70% humidity until each colony had reached the maximum areal cover in the Petri dish. The fungal growth was followed by taking a time series of digital images every second day after inoculation.

Images

21 pictures were taken as JPEG files, seven for each species, at different growth stages with a Canon EOS 700D camera in manual mode. Shutter speed, aperture, ISO and color temperature were fixed for every image (1/20", f/5, ISO 1600 and 3400K respectively). The camera was set in a custom made soft-box to prevent stray light from disturbing the image, with white walls on the sides and a black surface both on the top and bottom sides. The light source was a LED lamp (660 lumens, 3000K temperature) pointed towards a white wall at a 60° angle and a distance of 21 cm from the box (Figure 1). The Petri dish was put upside down inside the soft-box to avoid the images being affected by presence of vapor condensation on the lids.
A randomly chosen image for each fungal species was processed to collect training samples of parts of the image where the edge of the Petri dish, the background, the agar and the fungi were represented, each one at random: these image samples only contained pixels of their pertaining category.

Red (R), green (G) and blue (B) channels were extracted from reference images (with values from 0 to 255 for each channel) to create a table where each unique R, G and B value combination was assigned to its corresponding category (i.e. agar medium, edge of the Petri dish, fungi and background).

**Statistical analyses**

All statistical analyses were performed in R (R Core Team 2016), using the packages “imager” (Barthelme 2016) and “e1071” (Meyer et al. 2015). These points were used to train a SVM model, where the class (agar, Petri dish, fungus or background) was set as a categorical response variable and the three color channels (R, G and B) were set as numeric predictor variables.

The model robustness was assessed in two different ways: first, we estimated the misclassification rates of the training points to assess the model’s internal validity. Second, we applied the SVM model to the test images, and compared these measurements to a manual estimation performed by 4 experienced technicians using a combination of Adobe Photoshop CS5 and ImageJ.

Fungal area cover estimates based on the SVM model were developed after classifying all pixels in the image, and then using the next formula:

\[ A = \frac{F}{F+P} \times \left( \pi \times r^2 \right) \]

Where A= Fungal area cover (cm\(^2\)), F= pixels identified as fungi, P= pixels identified as growth medium and r= radius of the Petri dish (cm).
Results

The reference table containing R, G and B channels of the image resulted in 9836 unique RGB combinations, 1251 for the agar, 1670 for the background, 5758 for the three fungus species and 1157 for the Petri dish. The resulting SVM model shows high prediction accuracy, with a per-pixel misclassification rate of 3.5% in the calibration dataset (Table 1). When compared to independent technicians, SVM showed an average coefficient of determination ($R^2$) of 0.97 and a mean error of 0.72 cm$^2$, while the average $R^2$ between researchers was 0.99 (Figure 2) and the mean error ranged from 0.45 cm$^2$ to 1.65 cm$^2$.

Table 1. Per-pixel misclassification rates of the Support Vector Machine based supervised classification model

<table>
<thead>
<tr>
<th></th>
<th>Agar</th>
<th>Background</th>
<th>Fungus</th>
<th>Petri</th>
</tr>
</thead>
<tbody>
<tr>
<td>Agar</td>
<td>1062</td>
<td>134</td>
<td>4</td>
<td>0</td>
</tr>
<tr>
<td>Background</td>
<td>173</td>
<td>1536</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Fungus</td>
<td>16</td>
<td>0</td>
<td>5748</td>
<td>8</td>
</tr>
<tr>
<td>Petri</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>1149</td>
</tr>
</tbody>
</table>

Figure 2 Scatterplot showing concordance between SVM measured area and the four researchers measured areas (in cm$^2$), represented as different shaped points. Straight line represents a perfect fit.

The main advantage of using SVM models to measure microbial or fungal growth is that processing images with the SVM model requires an average of 32 seconds per image, compared to approximately 5 minutes per
picture needed for manual measurements: furthermore, the only process where the researcher has to take part on are the first few minutes of the process, when the script is set to run. This means substantial time savings in personnel hours at a low cost on accuracy. In addition, the human performance over a set of 21 images is an overestimate of the accuracy, given factors such as stress or fatigue that appear when larger sets of images are to be processed (Gonzalez et al. 2011; Healy et al. 2004).

Attempts to use algorithms in order to increase precision on microbial growth measurement have been tried in the recent years: Medeiros et al. (2016) applied a per-image thresholding method that allowed precise measurements of fungal area cover based on macro programming. This method, while being highly accurate, requires manually processing every image individually, thus requiring substantial amounts of time to process large datasets. On the other hand, Heaton et al (2010) and Vidal-Diez de Ulzurrun (2015) have successfully applied increasingly automated computer methods to measure hyphal growth properties. To our knowledge, however, there are no available procedures to measure microbial area cover through automatized image analyses: our study opens up new avenues where microbial growth kinetics can be measured on large datasets with minimal use of time resources (Figure 3).

However, this approach has its limitations: given that the model is based in color intensity (as given by the red, green and blue combinations), changes in color temperature (both in the light source and the in-camera settings) can affect the model’s predictive ability, even rendering it unusable. It is therefore of paramount importance that this method is applied with caution, keeping all conditions equal to avoid random variability that will affect image parameters, such as a different light source or camera setting to avoid spurious results.

In case the SVM model made inaccurate measurements of fungal coverage areas in particular images, it is a simple task to detect outliers in a growth curve (i.e. by unexpected growth patterns such as an unexpected

Figure 3 Example of fungal growth curves measured by the SVM model over 21 days. The dotted line represents full growth as estimated visually. Producing this plot required approximately 5 minutes of processing, with 30 seconds of human supervision.
decrease in areal coverage) or by visually assessing the classification on a classified reconstruction of the image (Figure 4). This can be done either by manually measuring the correct area, or by removing that data point, provided that the remaining measurements have been taken at frequent enough time intervals to reveal the growth phases.

![Figure 4 Example of a classified image produced by the SVM model (left) and the original image (right)](image)

In addition, the model is highly sensitive to differences in culture media: new media (e.g. blood agar) should be included into the model before performing any prediction. Given that a new SVM model development is a quick process once the model development script is prepared, it may be best to develop independent models for highly differing agar types. Likewise, the three fungal species demonstrated in this study are characterized by a white/cream color: fungi with different colors require development of new predictive models. We strongly recommend testing the suitability of the model for each new set of samples by manually measuring a subset of images and comparing the measured fungal area cover to the area estimated by the model.

**Conclusion**

Our study shows that automated estimation of fungal or microbial area cover in Petri dishes can be performed, as long as conditions are kept constant during the experiment. To ensure data quality, a subset of images needs to be manually measured.

The model to estimate fungal area in Petri dishes is available in GitHub: there is a file containing the model, a script file with an explanation of the predicting procedure and a README file. Basic R language skills are needed to perform the prediction.

**Acknowledgments**
We would like to thank Gregory Taff for fruitful discussions and proofreading.

This work was supported by University of Eastern Finland project 931673/2016 and the Tekes funded research project “SafeWood” 2723/31/2015. ABL received economic support from Teollisuusneuvos Heikki Väänänen's fund and the Kaute foundation.
References


