



OPEN ACCESS

EDITED BY

Matthew John Wade,
UK Health Security Agency (UKHSA),
United Kingdom

REVIEWED BY

Jinxin Zhao,
Monash University, China
Manyun Yang,
Cornell University, United States
Yuriy L. Orlov,
I.M. Sechenov First Moscow State
Medical University, Russia

*CORRESPONDENCE

Miguel Fuentes-Cabrera
fuentescabma@ornl.gov

SPECIALTY SECTION

This article was submitted to
Systems Microbiology,
a section of the journal
Frontiers in Microbiology

RECEIVED 01 September 2022

ACCEPTED 21 November 2022

PUBLISHED 05 January 2023

CITATION

Robertson C, Wilmoth JL, Retterer S
and Fuentes-Cabrera M (2023) Video
frame prediction of microbial growth
with a recurrent neural network.
Front. Microbiol. 13:1034586.
doi: 10.3389/fmicb.2022.1034586

COPYRIGHT

© 2023 Robertson, Wilmoth, Retterer
and Fuentes-Cabrera. This is an
open-access article distributed under
the terms of the [Creative Commons
Attribution License \(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use,
distribution or reproduction in other
forums is permitted, provided the
original author(s) and the copyright
owner(s) are credited and that the
original publication in this journal is
cited, in accordance with accepted
academic practice. No use, distribution
or reproduction is permitted which
does not comply with these terms.

Video frame prediction of microbial growth with a recurrent neural network

Connor Robertson¹, Jared L. Wilmoth², Scott Retterer³ and Miguel Fuentes-Cabrera^{3*}

¹Department of Mathematical Sciences, New Jersey Institute of Technology, Newark, NJ, United States, ²Department of Environmental Science and Technology, University of Maryland, College Park, MD, United States, ³Center for Nanophase Materials Sciences, Oak Ridge National Laboratory, Oak Ridge, TN, United States

The recent explosion of interest and advances in machine learning technologies has opened the door to new analytical capabilities in microbiology. Using experimental data such as images or videos, machine learning, in particular deep learning with neural networks, can be harnessed to provide insights and predictions for microbial populations. This paper presents such an application in which a Recurrent Neural Network (RNN) was used to perform prediction of microbial growth for a population of two *Pseudomonas aeruginosa* mutants. The RNN was trained on videos that were acquired previously using fluorescence microscopy and microfluidics. Of the 20 frames that make up each video, 10 were used as inputs to the network which outputs a prediction for the next 10 frames of the video. The accuracy of the network was evaluated by comparing the predicted frames to the original frames, as well as population curves and the number and size of individual colonies extracted from these frames. Overall, the growth predictions are found to be accurate in metrics such as image comparison, colony size, and total population. Yet, limitations exist due to the scarcity of available and comparable data in the literature, indicating a need for more studies. Both the successes and challenges of our approach are discussed.

KEYWORDS

microbial growth, Recurrent Neural Network, deep learning, video frame prediction, machine learning, population growth

1. Introduction

Recurrent Neural Networks, RNNs, are a type of Artificial Neural Network that takes a temporal sequence as inputs, learns the spatiotemporal variations, and predicts future data. RNNs were originally created for tasks related to Natural Language Processing, such as text classification and translation, but were subsequently extended to temporal sequences of images (i.e., videos) by incorporating convolutions. These convolutions combine collections of pixels in previous images to predict each individual pixel in the next image of a sequence, allowing the network to capture spatiotemporal dynamics for prediction. RNNs are now used in a variety of applications such as video captioning and video frame prediction (Shi et al., 2015; Su et al., 2020). In biology and microbiology,

RNNs have recently been used to study proteins (Alley et al., 2019; Le, 2019; Tng et al., 2022), whole bacteria genomes (Alley et al., 2019), and bacterial growth (Cheroutre-Vialette and Lebert, 2002; Baranwal et al., 2022). For example, in the case of proteins, RNNs can be used on large numbers (millions) of amino acid sequences to learn the statistical representations that effectively approximate the encoded protein features, which does not necessarily require structural or evolutionary data (Alley et al., 2019). On the other hand, in Baranwal et al. (2022) and Cheroutre-Vialette and Lebert (2002) an RNN was used to predict growth curves. In a similar manner, RNNs could be trained on multiple images, rather than multiple amino acid sequences, of bacterial cells as they grow under different conditions. To the best of our knowledge, however, RNNs have not been used to perform video frame prediction of bacterial growth. This is the focus of the work presented here: instead of predicting growth curves, we predict fluorescence microscopy images of growth. It's important to mention that growth curves can be obtained directly from such images by simply counting the number of pixels for each bacteria species.

Over the last few years, new advances have made video frame prediction with RNNs more accurate. A comprehensive review on this matter can be found in Oprea et al. (2020). One such advance has been named predRNN (Wang et al., 2021), which has been successfully used for performing video frame prediction of numbers moving across a screen (the so-called “moving MNIST” dataset¹), humans performing different actions (KTH action dataset²), traffic flow, and moisture movement in weather systems (Wang et al., 2021). In each of these scenarios, predRNN outperformed previous RNNs. Here we investigate whether predRNN could be used to perform video frame prediction of microbial growth using videos generated by fluorescence microscopy and microfluidics.

Training an RNN to perform video frame prediction requires a large number of videos. For example, predicting the random movement of numbers across a screen required 10,000 videos of 20 frames each (Wang et al., 2021). In microbiology, combining microfluidics with fluorescence microscopy has the potential to produce large amounts of videos. Timm et al. (2017) used microfluidics with fluorescence microscopy to investigate the growth of *Pseudomonas aeruginosa*, an antibiotic resistant bacteria that is responsible for many clinical infections (Pang et al., 2019). Using a photolithographic procedure, this group built a chip consisting of an array of microwells with different diameters. The diameters ranged from 5 to 100 μm , and each

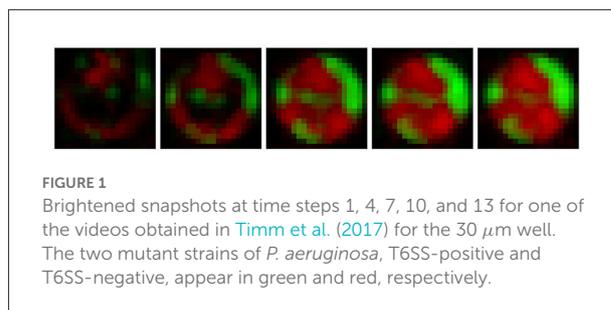


FIGURE 1
Brightened snapshots at time steps 1, 4, 7, 10, and 13 for one of the videos obtained in Timm et al. (2017) for the 30 μm well. The two mutant strains of *P. aeruginosa*, T6SS-positive and T6SS-negative, appear in green and red, respectively.

well had several replicates to ensure the data collected was statistically significant. The chip was seeded with a mixed microbial population that contained two mutant strains of *P. aeruginosa* at different concentrations. One mutant possessed the Type VI Secretion System, called here T6SS-positive, while the other mutant, T6SS-negative, lacked it, which made it more susceptible to attacks by the T6SS-positive mutant. The T6SS involves a needle-like pilus that some bacteria use to attack others (Sana et al., 2017). To monitor the growth of this mixed population, the T6SS-positive and T6SS-negative mutants were tagged with green and red fluorescent protein, respectively, and their growth monitored with fluorescence microscopy. This microwell fabrication and mixed bacterial population seeding on a chip has opened the door to large scale data-driven analysis of bacterial population interactions (Halsted et al., 2016; Hansen et al., 2016). Five snapshots of a microwell video obtained *via* this procedure are shown in Figure 1.

Here we show that predRNN (Wang et al., 2021) can be used to accurately predict future frames for the microwell videos collected in Timm et al. (2017). We show this by directly comparing predicted and original frames as well as by measuring the population curves and individual colony attributes of the predicted and original frames.

2. Methods

Timm et al. (2017) collected videos of microbial growth by generating fluorescence microscopy images of 24 x 24 pixels in RGB (Red, Green, and Blue) format every 30 min for a period of approximately 24 h. In these images, T6SS-positive and T6SS-negative *P. aeruginosa* mutants appeared as green and red, respectively. Of all the well sizes in the micro-fluidic array, we found that the 30 μm well-contained the most useful data. First, this well size allowed for microbial populations to completely fill the well during growth, resulting in a distinct partitioning of the population into colonies of both strains. Second, more videos were collected for this well size than for the other sizes, giving a more complete set for training predRNN. For this reason, we focus the present study on videos collected for the 30 μm wells.

A total of 48 videos of 14 frames were collected for the 30 μm well. In these videos, frames 1–7 showed the most interesting

¹ MNIST is the Modified National Institute of Standards and Technology database of handwritten digits. It is commonly used to test neural network classification performance.

² KTH action is a dataset made up of videos of people moving created by KTH Royal Institute of Technology in Sweden. It is commonly used for testing neural networks in video prediction.

dynamics, which corresponded to the exponential growth of the population from a few initial colonies to a large community that covered the entire well. By contrast, frames 8–14 were practically static, as growth reached a saturation state around frame 8. As a consequence, we discarded frames 8–14 and focused on frames 1 through 7 giving a dataset comprised of 48 videos of seven frames of 24 x 24 pixels in RGB format.

These frames were subjected to a series of four transformations aimed at optimizing the training of the predRNN network. First, we transformed the images from RGB to HSV (Hue, Saturation, Value) format. This transformation aimed to separate the population information contained in each image more clearly than with RGB. Specifically, the presence of either population in a given pixel was recorded by the Value channel (black or not black pixel), the species of the population in the pixel was recorded by the Hue channel (red or green pixel), and the concentration of the species in the pixel was accounted for by Saturation (how bright red or green in the pixel were). Second, each frame was adjusted to match the center of the images with the center of the well. This was done by cropping a square around the brightest pixels of the images averaged over time after triangle thresholding (Zack et al., 1977) which accentuates the brightest pixels in the image. Third, the number of frames per video was increased from 7 to 20 and each frame was expanded from 24 x 24 to 32 x 32 pixels. The number of frames per video was increased using temporal interpolation across frames, and was performed with the scikit-image Python package (van der Walt et al., 2014). The size increase was performed with the bilinear interpolation in the OpenCV library (Bradski, 2000). These three sets of transformations led to a dataset of 48 videos of 20 frames of 32 x 32 pixels recorded in the HSV format. At this point, five of the videos were set aside as a "test" set which would be later used to assess the predictive ability of the neural network on previously unseen data.

As the resulting 43 videos still made a relatively small dataset for training predRNN, we performed a fourth set of transformations that consisted of flipping, rotating, blurring, and adding Gaussian noise to all the videos. These transformations were performed with the AtomAI Python package (Ziatdinov et al., 2021) developed for material science applications of deep learning. These final transformations gave a dataset of 392 videos with 20 frames of 32 x 32 pixels in the HSV format. Figure 2 shows schematically the transformations and the workflow for preparing the input dataset.

In order to train the network, the 392 video dataset was randomly split into training and validation sets in a 80/20 proportion which are used in the training process. This is a customary practice in machine learning meant to verify that the model can perform well on new data (i.e., data not used to tune the parameters of the model). Specifically, the training set is used to update the "weights" or parameters of the neural network. The model is then tested on the validation set to get a measure of how well it is performing on new data. The process is then repeated.

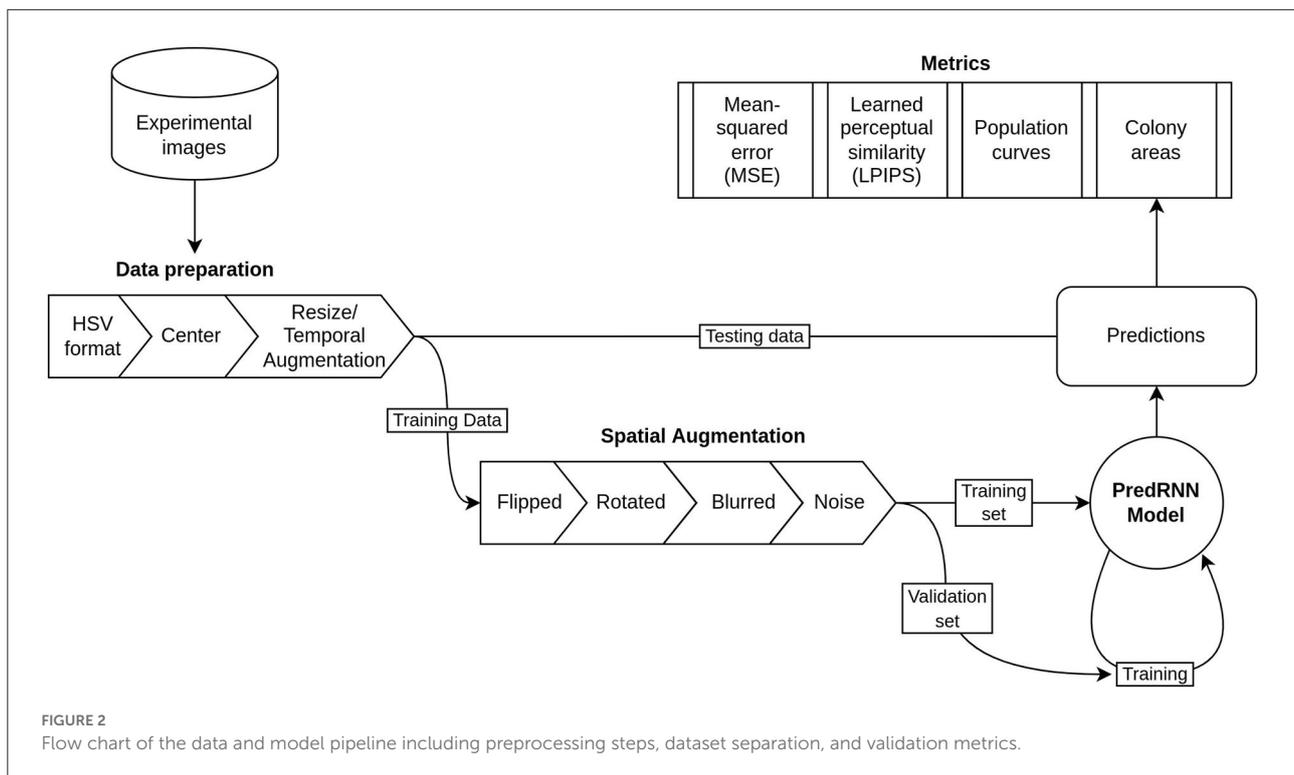
To assess the final quality of the trained network after training, a test set, which is withheld earlier and never seen by the network during training, is used. This procedure is shown schematically in Figure 2.

During the training procedure, the first 10 frames of a video sequence were used as inputs and the remaining 10 were used for predictions. In what follows, the original 11–20 frames will be referred to as groundtruth (the images we want the network to output), whereas those predicted by the network will be referred to as predicted frames.

There are a variety of network parameters, called hyperparameters, that can be adjusted for improving the training process including size and step of convolutions, parameter normalization, patch sizes of the input images, random sampling of the training data, number of total parameters, learning rate, size of training data batches, and total number of training rounds or "epochs." Together, these represent over 100 million possible parameter permutations, each of which would require retraining the model (a process that could take hours using a GPU accelerated computer). Due to this prohibitive complexity, we used the standard parameters from the original PredRNN model and varied only the layer size and learning rate. We tested the models performance with each parameter set by iterating across a range of layer sizes and learning weights. To evaluate the success of each set of hyperparameters, we evaluated the mean-squared error (MSE) of the predictions with the groundtruth, which is a pixel-wise comparison of two images and the learned perceptual image patch similarity (LPIPS). The LPIPS metric was recently developed as an image similarity metric to mimic the perception of the human eye (Zhang et al., 2018). It should be noted that the smaller the MSE and LPIPS values are, the more similar the predicted frame is to the corresponding groundtruth. We ultimately determined to use layers of size 32 and a learning rate of 0.0003. A link to *GitHub* repository that contains the trained network, as well as the list of optimal hyperparameter values, is included in the Supporting Information.

3. Results and discussion

Figure 3 shows the changes in the training loss (the MSE of "predictions" using the training data) and the MSE and LPIPS of the validation set, with the number of epochs (the number of iterations used to train the network). It is seen that after 50,000 epochs the network had converged to a somewhat steady state in the training loss as well as the MSE and LPIPS validation metrics. This means that when given an input of 10 frames from one of the training videos, the network will accurately output the next 10 frames (good training loss) and when given 10 frames from one of the validation videos the network will output a prediction of 10 frames that are similar pixel-wise (good MSE)

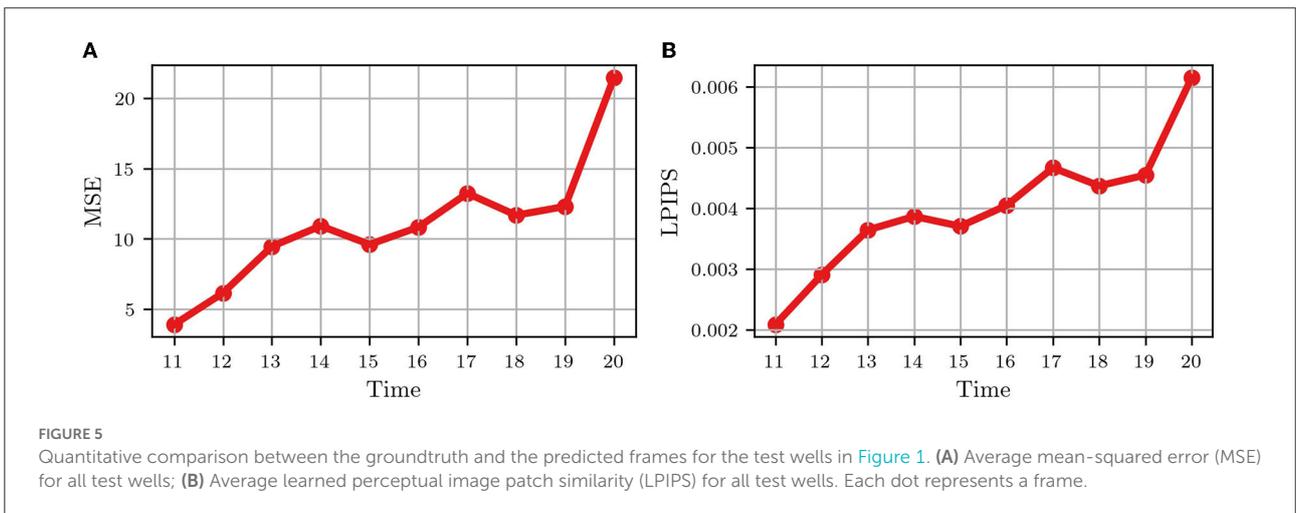
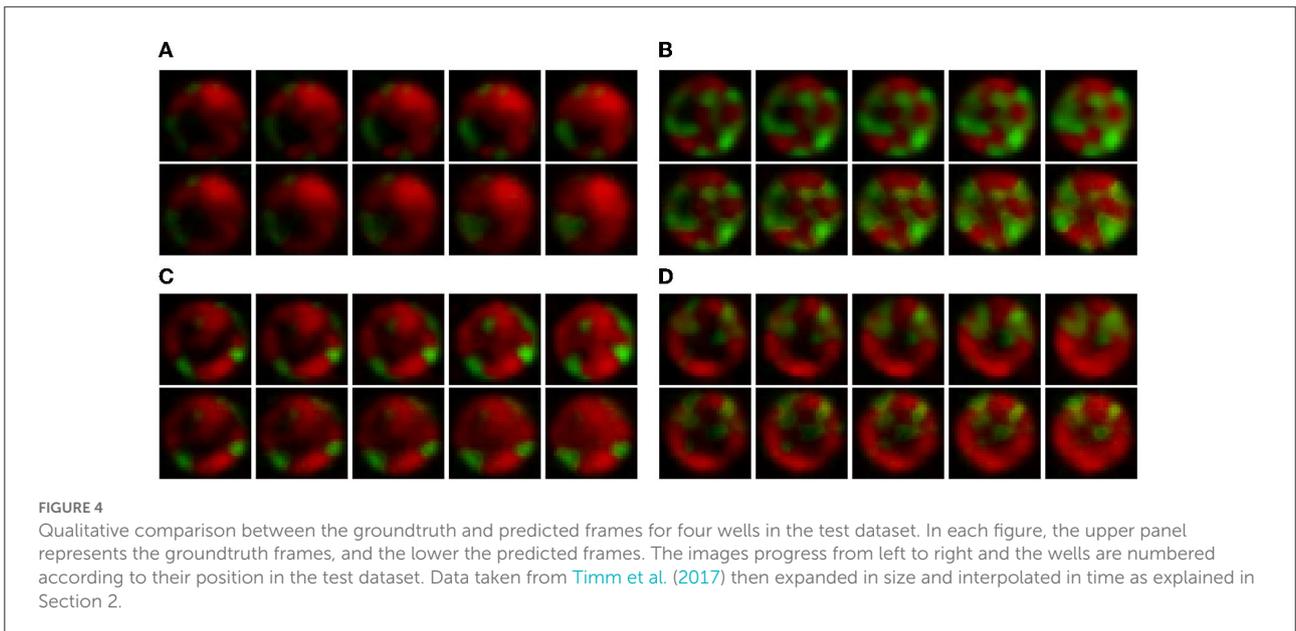
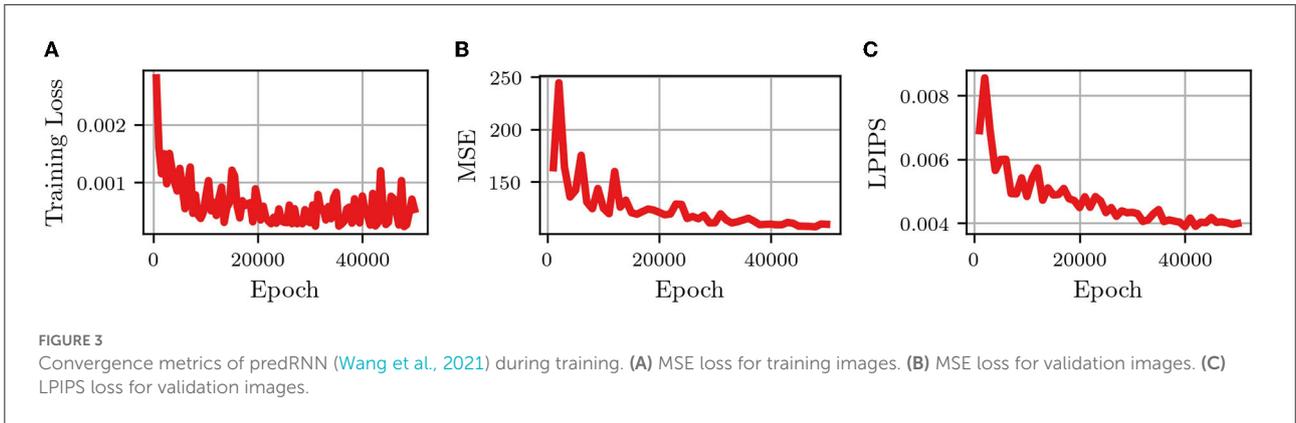


and similar by human examination (good LPIPS) to the actual next 10 frames in the video.

Figure 4 shows the video frame predictions for four different wells from the test dataset. It can be observed that the groundtruth frames (upper panel) and the predicted frames (lower panel) are visually very similar. To quantify this similarity, we computed the per-frame and average MSE and LPIPS. Figure 5 shows the per-frame MSE and LPIPS, respectively, and both increase with time. In other words, the predicted and groundtruth frames differ more for predictions performed later in time. This is an expected result because errors accumulate as more frames are predicted. For example, frame 12 is predicted from frame 11, and any error in predicting frame 11 carries on to frame 12. Despite this, the accuracy of the predictions is satisfactory and comparable to the results obtained in Wang et al. (2021) for the moving MNIST database. Further, the average MSE and LPIPS values obtained for the test set are 13.7 and 0.002, which are smaller than the values 48.4 and 0.071 obtained in Wang et al. (2021).

The results above are encouraging as they suggest that video frame prediction of microbial growth can indeed be performed with predRNN. However, aside from directly comparing images, neither MSE nor LPIPS are particularly insightful in the context of microbial experiments. It is instead preferable to compare images with metrics that provide quantitative information on microbial growth and which are commonly used by microbiologists.

Many previous studies for microbial growth use population growth curves for the respective species as quantitative metrics (Timm et al., 2017). As discussed previously, Timm et al. used green and red fluorescent proteins to tag the T6SS-positive and T6SS-negative mutants of *P. aeruginosa*, respectively. The growth of each mutant, and of the whole community, was then monitored by plotting the intensity of the fluorescent signal vs. time. Here we do not measure the intensity of the fluorescent signal, but instead extract population curves directly from the groundtruth and predicted frames by summing the pixels for green and red. Figure 6 shows the population curves calculated for the four test wells. In general, the agreement between prediction and groundtruth is better for the T6SS-positive (green) than for the T6SS-negative (red) mutant. The reason for this is not entirely clear to us and the scarcity of data has posed a challenge for additional analysis. However, our hypothesis is that as T6SS-positive attacks, the growth of T6SS-negative is more susceptible to fluctuations, which causes predRNN to have more difficulties to learn the spatiotemporal evolution from frame to frame. This can be observed most acutely in the population curves of test well 2 which show an almost equal number of the two strains, causing the most aggressive of fluctuations in population. This particular case is investigated further below. Finally, it is also seen that the T6SS-positive mutant grows less than the T6SS-negative. Understanding the reasons behind this observation is beyond the scope of this work, but some studies have suggested that possessing the T6SS pilus puts a significant



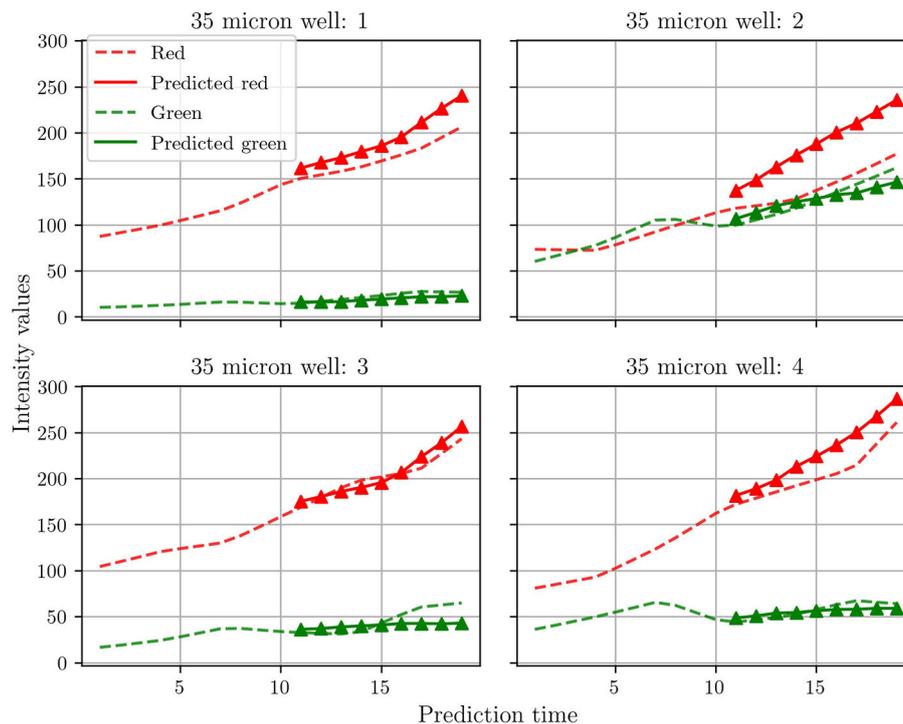


FIGURE 6

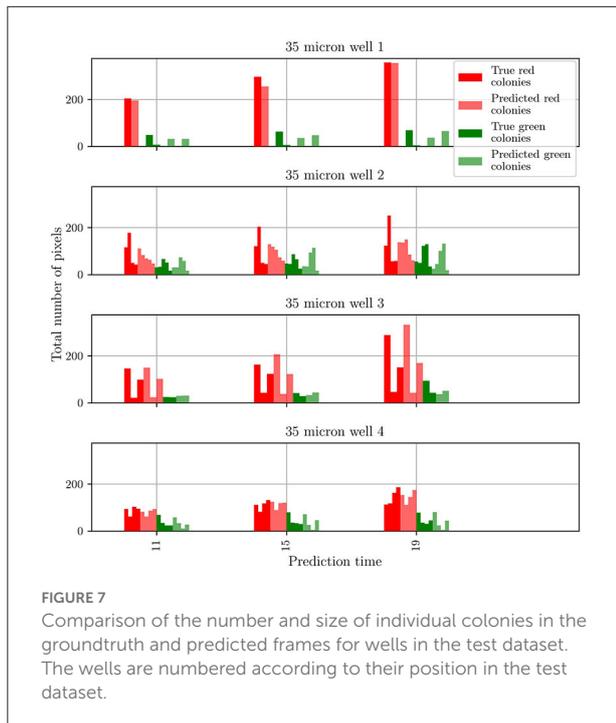
Comparisons between the population curve extracted from the groundtruth frames and the corresponding predicted population curves. See text for an explanation of how the population curves were extracted.

energy burden on the T6SS-positive strain, which ultimately limits its growth (Wilmoth et al., 2018).

To gain insights into the discrepancy between predictions and observations for test well 2, we extracted two more metrics from the predicted and groundtruth frames. These metrics are the number of colonies for each mutant and the number of pixels comprising each colony. The results obtained for all four test wells are shown in Figure 7. For clarity, the data is shown only for three different frames taken from the whole sequence of groundtruth and corresponding predicted frames. For both mutants, and for wells 1, 3, and 4 the agreement between prediction and groundtruth is in general very good. Minor differences are seen for the T6SS-positive mutant: for well 1, the number of predicted colonies are overestimated, and for well 4, frames 15 and 19 show one colony less for the predicted frame than for the groundtruth one. However, the differences are more significant for test well 2, especially for the T6SS-negative mutant (red). In this case the number of T6SS-negative colonies in the groundtruth and predicted frames, and the corresponding number of pixels in the colonies, are different. On the other hand, for the T6SS-positive mutant, the predictions are very good.

From these results, it is clear that the origin of the differences in the population curves for test well 2 are rooted on differences

at the level of individual colonies. To further understand this, we refer to Figure 4: It is seen that, in general, the growth in test wells 1, 3, and 4 is characterized by the partitioning of the population into a large T6SS-negative colony and a few smaller T6SS-positive ones. By contrast, the growth in test well 2 is characterized by the appearance of small colonies for both mutants. It is unknown what might have caused this distinct growth pattern, but visual inspection of the original dataset (48 videos of 14 frames) reveals that this pattern is uncommon. The dataset is thus imbalanced, which leads to a bias in the training of predRNN, hampering its ability to accurately predict the growth for test well 2, specifically for the T6SS-negative mutant. Indeed, for this well the T6SS-positive mutant predictions are accurate because the growth pattern is similar to the patterns observed in other wells. However, for the T6SS-negative mutant, the growth pattern in test well 2 is unusual and the network is not capable of predicting it since it was not trained with many similar cases. Imbalanced datasets are a major yet subtle issue in machine learning, as accurate predictions depend on the specifics of each dataset. In our case, the difficulty is compounded by the scarcity of video data. For example, in this case any effort to balance the dataset by removing videos with uncommon growth patterns would reduce the size of the dataset even further. Finally, our trained



model is not expected to accurately predict microbial growth for other well sizes than the 30 μm one, because the training dataset did not include them. To enable that, one would have to retrain the model by including those wells, but unfortunately, as already mentioned, this is not possible due to the scarcity of data.

4. Conclusion

We trained an RNN, named predRNN (Wang et al., 2021), to perform video frame prediction of microbial growth for a population containing two mutants of *P. aeruginosa*. For training, we used 48 videos that were previously collected by Timm et al. (2017) with microfluidics and fluorescence microscopy.

To assess the quality of the predictions, we used image-to-image metrics, in particular MSE and LPIPS, as well as microbially significant metrics such as population curves and the characteristics of individual colonies. It was found that predRNN (Wang et al., 2021) can, in general, predict correctly the growth of this population. However, we also noticed that in some cases the predicted growth could be inaccurate. In this particular setting, these inaccuracies were found to be caused by an imbalanced training dataset, which contained more wells with faster T6SS-negative mutant growth than the T6SS-positive mutant. As a consequence, the network tended to favor the growth of the T6SS-negative mutant

over the other mutant, which led to large errors in frame prediction for wells where both mutants happened to grow similarly. This emphasizes the need to create large databases of microbial growth, which could be accomplished by combining microfluidics and fluorescence microscopy. A recent study, however, offers a new possibility for increasing the dataset: In Pawlowski et al. (2022) used deep learning style transfer to create a dataset of synthetic images of realistic microbial growth. This technique could be useful for augmenting the dataset for video frame prediction. Nonetheless, despite the challenges associated with the size of the dataset, our results provide promising steps toward the possibility of performing autonomous experiments in microbiology, where time spent in image acquisition could be saved by predicting microbial growth with RNNs.

The successful application of this recurrent model architecture for the included fluorescence images also suggests future applications in predictions of cell morphology such as shape and size (Li et al., 2021; Way et al., 2021), in different spatiotemporal resolutions, or in predicting compositions of more heterogeneous mixtures of diverse bacterial strains (e.g., each strain could be labeled with a unique fluorescent protein; Larsen et al., 2015). Although these applications lie within the scope of the predRNN model architecture, successful predictions would require training using the a large dataset of high resolution images with consistent time and space scales (see Supplementary material for more details on this limitation). For example, to predict cell morphology changes, we could follow the same procedure as documented in this manuscript with high resolution images of cell populations rather than fluorescence microscopy images of populations. The metrics of success would also need to be adjusted to measure characteristics of cell shape or size rather than population or colony level measurements. Each of these applications promise a new range of machine learning based predictive analytics beyond the scope of current quantitative image analysis methods.

Licences and permissions

This article has been authored by UT-Battelle, LLC under Contract No. DE-AC05-00OR22725 with the U.S. Department of Energy. The United States Government retains and the publisher, by accepting the article for publication, acknowledges that the United States Government retains a non-exclusive, paid-up, irrevocable, world-wide license to publish or reproduce the published form of this article, or allow others to do so, for United States Government purposes. The Department of Energy will provide public access to these results of federally sponsored research in accordance with the DOE Public Access Plan (<http://energy.gov/downloads/doe-public-access-plan>).

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

Author contributions

MF-C designed the research. CR trained the recurrent neural network and analyzed the results. JW and SR provided and explained the video data. All authors contributed to writing and editing the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

The machine learning modeling research was conducted at the Center for Nanophase Materials Sciences, which is a US Department of Energy Office of Science User Facility at Oak Ridge National Laboratory. CR was funded by U.S. Department of Energy, Office of Science, Office of Workforce Development for Teachers and Scientists, Office of Science Graduate Student Research (SCGSR) program award. The SCGSR program is administered by the Oak Ridge Institute for Science and Education (ORISE) for the DOE. ORISE is managed by ORAU under contract number DESC0014664. The authors would like

References

- Alley, E. C., Khimulya, G., Biswas, S., Al Quraishi, M., and Church, G. M. (2019). Unified rational protein engineering with sequence-based deep representation learning. *Nat. Methods* 16, 1315–1322. doi: 10.1038/s41592-019-0598-1
- Baranwal, M., Clark, R. L., Thompson, J., Sun, Z., Hero, A. O., and Venturelli, O. S. (2022). Recurrent neural networks enable design of multifunctional synthetic human gut microbiome dynamics. *eLife* 11, e73870. doi: 10.7554/eLife.73870
- Bradski, G. (2000). The opencv library. *Dr. Dobbs J.* 25, 120–123.
- Cheroutre-Vialette, M., and Lebert, A. (2002). Application of recurrent neural network to predict bacterial growth in dynamic conditions. *Int. J. Food Microbiol.* 73, 107–118. doi: 10.1016/s0168-1605(01)00642-0
- Halsted, M., Wilmoth, J. L., Briggs, P. A., Hansen, R. R., Briggs, D. P., Timm, A. C., et al. (2016). Development of transparent microwell arrays for optical monitoring and dissection of microbial communities. *J. Vacuum Sci. Technol. B* 34, e06K103. doi: 10.1116/1.4962739
- Hansen, R. H., Timm, A. C., Timm, C. M., Bible, A. N., Morrell-Falvey, J. L., Pelletier, D. A., et al. (2016). Stochastic assembly of bacteria in microwell arrays reveals the importance of confinement in community development. *PLoS ONE* 11, 1–18. doi: 10.1371/journal.pone.0155080
- Larsen, P., Dai, Y., and Collart, F. R. (2015). Predicting bacterial community assemblages using an artificial neural network approach. *Methods Mol. Biol.* 1260, 33–43. doi: 10.1007/978-1-4939-2239-0_3
- Le, N. Q. K. (2019). Fertility-gru: identifying fertility-related proteins by incorporating deep-gated recurrent units and original position-specific scoring matrix profiles. *J. Proteome Res.* 18, 3503–3511. doi: 10.1021/acs.jproteome.9b00411
- Li, Y., Nowak, C. M., Pham, U., Nguyen, K., and Bleris, L. (2021). Cell morphology-based machine learning models for human cell state classification. *NPJ Syst. Biol. Appl.* 7, 1–9. doi: 10.1038/s41540-021-00180-y
- Oprea, S., Martinez-Gonzalez, P., Garcia-Garcia, A., Castro-Vargas, J. A., Orts-Escolano, S., Garcia-Rodriguez, J., et al. (2020). A review on deep learning techniques for video prediction. *IEEE Trans. Pattern Anal. Machine Intell.* 2004, 5214. doi: 10.48550/arXiv.2004.05214
- Pang, Z., Raudonis, R., Glick, B. R., Lin, T. J., and Cheng, Z. (2019). Antibiotic resistance in *Pseudomonas aeruginosa*: mechanisms and alternative therapeutic strategies. *Biotechnol. Adv.* 37, 177–192. doi: 10.1016/j.biotechadv.2018.11.013
- Pawlowski, J., Majchrowska, S., and Golan, T. (2022). Generation of microbial colonies dataset with deep learning style transfer. *Sci. Rep.* 12, 1–12.
- Sana, T. G., Lugo, K. A., and Monack, D. M. (2017). T6ss: the bacterial "fight club" in the host gut. *PLoS Pathogens* 13, e1006325. doi: 10.1371/journal.ppat.1006325
- Shi, X., Chen, Z., Wang, H., Yeung, D. Y., Wong, W. K., and Woo, W. C. (2015). Convolutional LSTM network: a machine learning approach for precipitation nowcasting. *Advances in Neural Information Processing Systems* 18.
- Su, J., Byeon, W., Kossai, J., Huang, F., Kautz, J., and Anandkumar, A. (2020). Convolutional tensor-train LSTM for spatio-temporal learning. *Advances in Neural Information Processing Systems* 33, 13714–13726.
- Timm, A. C., Halsted, M. C., Wilmoth, J. L., and Retterer, S. T. (2017). Assembly and tracking of microbial community development within a microwell array platform. *J. Visualized Exp.* 124, 55701. doi: 10.3791/55701

to thank Andrea Timm for providing the experimental videos used for training predRNN. All the simulations were run in CADES, <https://cades.ornl.gov/>.

Conflict of interest

MF-C and SR are employees of Oak Ridge National Lab which is managed by UT-Batelle.

The remaining 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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fmicb.2022.1034586/full#supplementary-material>

- Tng, S. S., Le, N. Q. K., Yeh, H. Y., and Chua, M. C. H. (2022). Improved prediction model of protein lysine crotonylation sites using bidirectional recurrent neural networks. *J. Proteome Res.* 21, 265–273. doi: 10.1021/acs.jproteome.1c00848
- van der Walt, S., Schonberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., et al. (2014). Scikit-image: image processing in python. *PeerJ* 2, e453. doi: 10.7717/peerj.453
- Wang, Y., Wu, H., Zhang, J., Gao, Z., Wang, J., Yu, P. S., et al. (2021). PredRNN: a recurrent neural network for spatiotemporal predictive learning. *CoRR* 2021, e09504. doi: 10.48550/arXiv.2103.09504
- Way, G. P., Kost-Alimova, M., Shibue, T., Harrington, W. F., Gill, S., Piccioni, F., et al. (2021). Predicting cell health phenotypes using image-based morphology profiling. *Mol. Biol. Cell* 32, 995–1005. doi: 10.1091/mbc.E20-12-0784
- Wilmoth, J. L., Doak, P. W., Timm, A., Halsted, M., Anderson, J. D., Ginovart, M., et al. (2018). A microfluidics and agent -based modeling framework for investigating spatial organization in bacterial colonies: the case of *Pseudomonas aeruginosa* and h1-type vi secretion interactions. *Front. Microbiol.* 9, 33. doi: 10.3389/fmicb.2018.00033
- Zack, G. W., Rogers, W. E., and Latt, S. A. (1977). Automatic measurement of sister chromatid exchange frequency. *J. Histochem. Cytochem.* 25, 741–753. doi: 10.1177/25.7.70454
- Zhang, R., Isola, P., Efros, A. A., Shechtman, E., and Wang, O. (2018). The unreasonable effectiveness of deep features as a perceptual metric. *CVPR* 2018, 3924. doi: 10.48550/arXiv.1801.03924
- Ziatdinov, M., Ghosh, A., Wong, T., and Kalinin, S. V. (2021). Atomai: a deep learning framework for analysis of image and spectroscopy data in (scanning) transmission electron microscopy and beyond. *arXiv* 2021, 7485. doi: 10.48550/arXiv.2105.07485