Cell RAFTs: Applying RAFT to Cell Microscopy

Ziyad Rahman, Srijana Niraula, Liam Jones, Jennifer Mitchel

Wesleyan University

QAC Apprenticeship, Department of Molecular Biology and Biochemistry, Wesleyan University, Middletown, CT 06459

Introduction

The wound healing assay experiment is used to study cell motility. Practically , it involves mimicking a wound by separating cells and taking high resolution videos in the hours or days after separation. To study cell movement, biologists dye the cells and use computer algorithms to track their movement. However, certain dyes prove difficult for traditional algorithms to track.

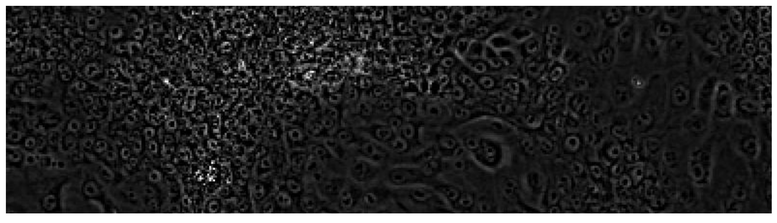

Fig. 1 Phase Contrast Channel of Nuclei-Dyed Cells

This study investigates an alternate method, Recurrent All-Pairs Field Transforms (RAFT), which uses deep learning to extract trajectories for the cells.[1] We ask if we can use nuclei-dyed cells and their trajectories obtained through traditional methods (the Farneback algorithm[2]) as the training data. Given the large sizes of these videos (~8GB-32GB), computers may not be able to load the entire video into memory at once. As such, we compare different preprocessing approaches against each other as well.

Data and Preprocessing

Liam Jones provides the nuclei-dyed training data and Srijana Niraula provides the cytoplasm-dyed data that we will use to check our model. Both videos are split into three channels: one phase contrast and two fluorescent. Using the fact that the phase contrast channels "look" the same between videos, we can train our model to extract optical flow information directly from it.

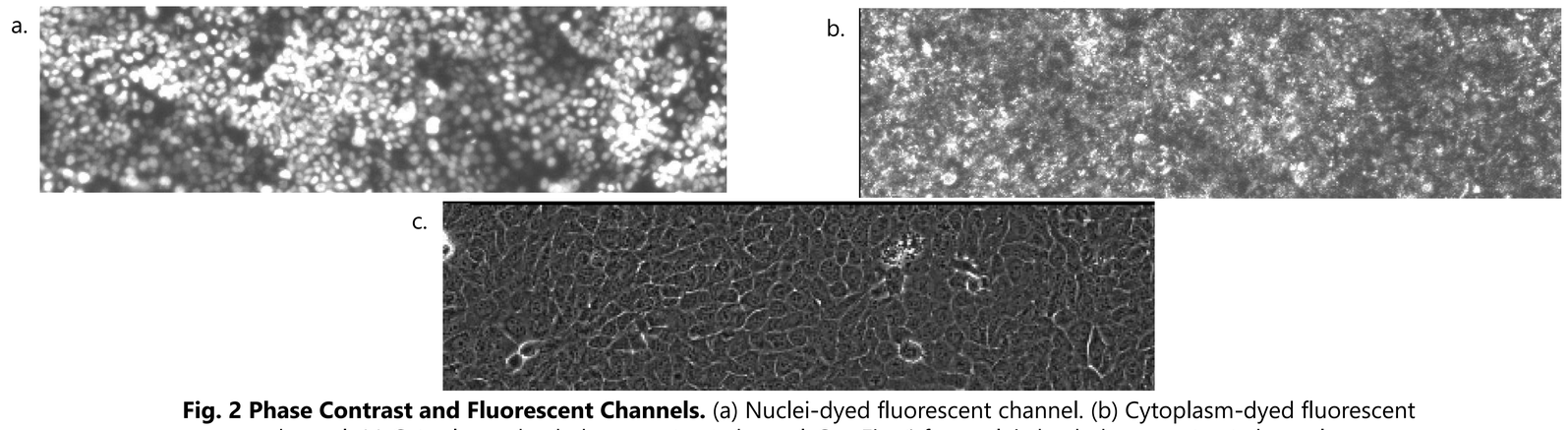

Fig. 2 Phase Contrast and Fluorescent Channels. (a) Nuclei-dyed fluorescent channel. (b) Cytoplasm-dyed fluorescent channel. (c) Cytoplasm-dyed phase contrast channel. See Fig. 1 for nuclei-dyed phase contrast channel.

To begin, we scale each image's x and y direction by 1/4, 1/6 and 1/8 (saving them as separate videos). Then, we preprocess the nuclei-dyed and cytoplasm-dyed videos by increasing contrast and applying blurs. We can use the Farneback method on the nuclei-dyed videos, producing optical flow data for the fluorescent channels. Now, we add the channels together, mimicking the phase contrast channel. Finally, we split the videos into frame pairs associated with their flow. This provides the training and validation data. To accomplish this, we made use of Python and the Open-CV and Numpy libraries.

The RAFT Model

Traditional models like Farneback essentially take look at every pixel's intensity to determine movement.[2] In contrast, RAFT generates a feature map (which highlights edges and shapes) and draws correlations between shapes to determine movement.[1]

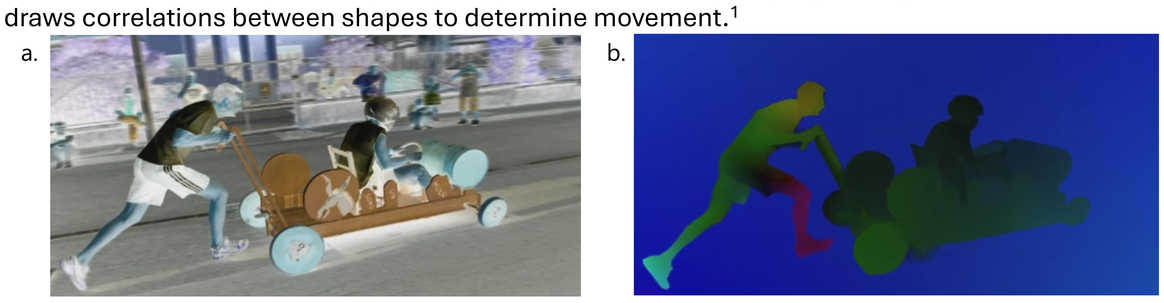

Fig. 3 RAFT Output.[1] (a) Processed image provided to RAFT for prediction. (b) RAFT's optical flow output. Visible in this is the feature map it creates, removing noise and only keeping core shapes.

Machines have difficulty constructing shapes out of cells since they are constantly changing shape. This is likely why RAFT has not been widely used for cell microscopy. However, the Python library PyTorch provides a RAFT model trained on videos of hot-air balloons that performed decently on our data, prompting us to investigate this model further.

Training Data Groups

The cell microscopy videos used for training are extremely large (~8GB-32GB). Our lab's standard is to scale them down by a factor of 1/4 (by the x and y directions), but this means we can only load a single frame into RAM at once (since forwards and backwards propagation also use RAM). As such, we investigate the impact of further scaling on the images – examining the tradeoff between larger batch sizes and resolution. Thus, we train six models (in addition to using the pre-trained model).

Scale Factor	Batch Sizes
1/4	1 frame
1/6	1 frame, 4 frames
1/8	1 frame, 4, frames, 6 frames

Fig. 4 Table of Training Data Scale Factors and Batch Sizes

We chose not to go further than the 1/8 scale factor since that would lose too much information. Further, six frames appeared to be the upper limit for how many frames we could load into RAM even at the smallest resolution.

Model Comparison and Conclusion

We'll label each model by their scale factor and their batch size. For example, 1/4 = 0.25, so the 1/4 scale factor with a batch size of one is model 25b1. We find that model 167b1 (1/8 scale factor) has the lowest loss. Via the Kruskal-Wallis test, we find that the difference in loss is statistically significant. We also found that while batch size and scale-factor were significant towards lower loss values, their interaction was not.

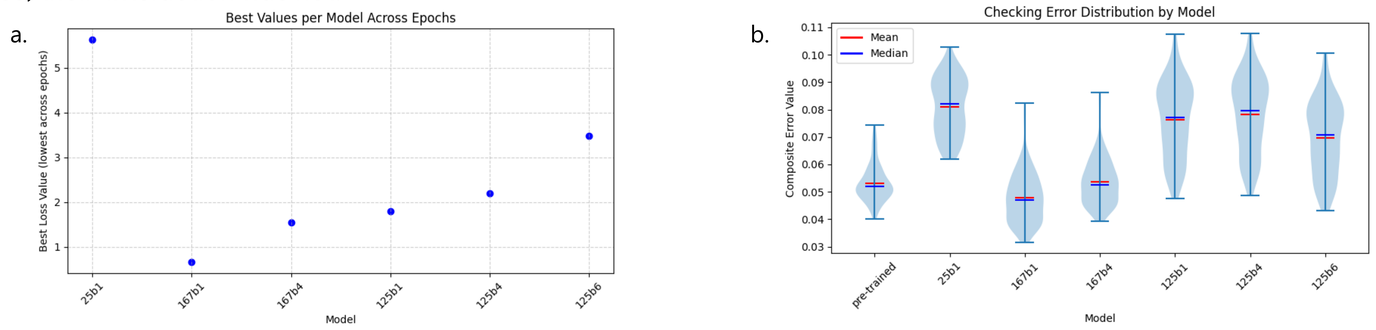

Fig. 5 Best Training and Checking Loss. (a) Scatterplot of best training loss per model. (b) Boxplot of checking loss distributions by model.

We also found that lower resolutions performed better during the checks (using the cytoplasm dyed-data). Due to the lack of ground truth, we used photometric warp and backwards flow to determine accuracy. We suspect that the lower resolutions obfuscate the cell edges, allowing the model to create a feature map of sliding groups of cells, which functionally create edges.

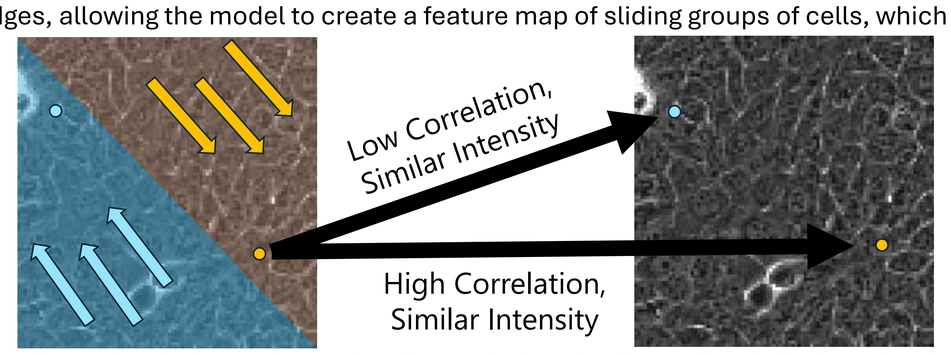

Fig. 6 Suspected Reasoning for Lower Resolution Models Performing Better. The similar intensity of both points would lead the Farneback algorithm to deduce these points are moving in the same direction, but RAFT's correlation approach differentiates them.

Limitations and Future Work

First, the Mitchel Lab computer does not have a GPU, so code was optimized for CPU in case future training was required. This restricted the number of epochs and samples we could use. However, GPUs have less VRAM, which can further restrict the size of the training data, so work needs to be done in creating smaller training data if the lab begins using a GPU. Second, the available training data largely used the same types of cells (MCF10A vs MCF10A-Rab5a) which may have biased the model towards their specific movements. Future work should use more varied training data. Third, there was no ground truth for the cytoplasm-dyed videos, which could have increased the accuracy of our checks when comparing the models. Hand-tagging flow can provide accurate ground truth.

Acknowledgments

This work was funded by the Wesleyan Hazel Quantitative Analysis Center. Thank you to those in the Mitchel Lab for their help on this project, especially Liam Jones and Srijana Niraula for providing their wound assay videos. A special thanks to Jennifer Mitchel for her continued guidance and support.

References

1. Zachary Teed, , and Jia Deng. "RAFT: Recurrent All-Pairs Field Transforms for Optical Flow".*CoRR* abs/2003.12039 (2020).
2. Farnebäck, Gunnar. 'Two-Frame Motion Estimation Based on Polynomial Expansion'. In *Image Analysis*, edited by Josef Bigun and Tomas Gustavsson, 363–70. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003.

Code and Data

The code used for this poster is available at *https://github.com/ZSR3004/cellraft*.