U-TURNING ON DEEP LEARNING METHODS FOR CELL COUNTING: A BAYESIAN REGRESSION APPROACH

Andrew Pohl[1], Dylan Loader[2], Mingkuan Wu[2], Daniel Yang[2]

[1]Faculty of Kinesiology. [2]Department of Mathematics and Statistics
University of Calgary
andrew.pohl@ucalgary.ca, dylan.loader@ucalgary.ca, mingkuan.wu@ucalgary.ca, daniel.yang@ucalgary.ca

UNIVERSITY OF CALGARY

Introduction

With the invention of high throughput automated image cytometry techniques, the generation of cellular microscopy image data has accelerated drastically and automated methods are required to count cells.

Current methods for cell counting are often adversely affected by deviations in image focus, by the use of differing cell staining techniques and by the fact that cells often overlap each other in images.

Given the recent success of deep learning image processing methods such as the U-Net neural network [1], we hypothesise that prediction models for cell count will improve when images are pre-processed by a U-Net neural network.

Methods

Dataset

The dataset consists of 3600 simulated fluorescence microscope slides obtained using 2 types of stain and 3 levels of imaging blurring, with cell counts ranging from 1 to 100 [2].

The dataset was segmented into three separate sets. A testing set of 1200 images was with-held by the judges. To construct a robust training and validation environment we further segmented the training set into training (1920 images) and validation sets (480 images).

Prediction Models

A Bayesian implementation of polynomial regression was applied to predict the number of cells from either a raw image or U-Net generated image mask.

It was assumed that the cell count followed a rounded, truncated $[0, 1]$ normal distribution such that: $y_{i,j,k} \sim N(\mu_{i,j,k}, \sigma^2)$.

To account for increased cell overlap (and consequently fewer pixels per cell) as the number of cells per image increased, the sum of the normalised pixels x for each image i belonging to stain $j = 1, 2$ and blurring level $k = 1, 2, 3$, along with its square were used as covariates to estimate the corresponding cell count ($y_{i,j,k}$).

Regression coefficients β_1, β_2 and β_3 were allowed to vary on three levels:
1. Single level where regression coefficents were not allowed to vary regardless of stain/blur group.
2. A varying slope model where β_2 and β_3 were allowed to vary based on stain/blur group.
3. A varying slope and intercepts model where all coefficients were allowed to vary given the corresponding stain/blur group.

U-Net Image Preprocessing

A 10 layer encoder-decoder U-Net simplified from the architecture provided by Falk et al. (2019) [1] was trained on a subset of 830 images wtih ground-truth segmentation masks [2].
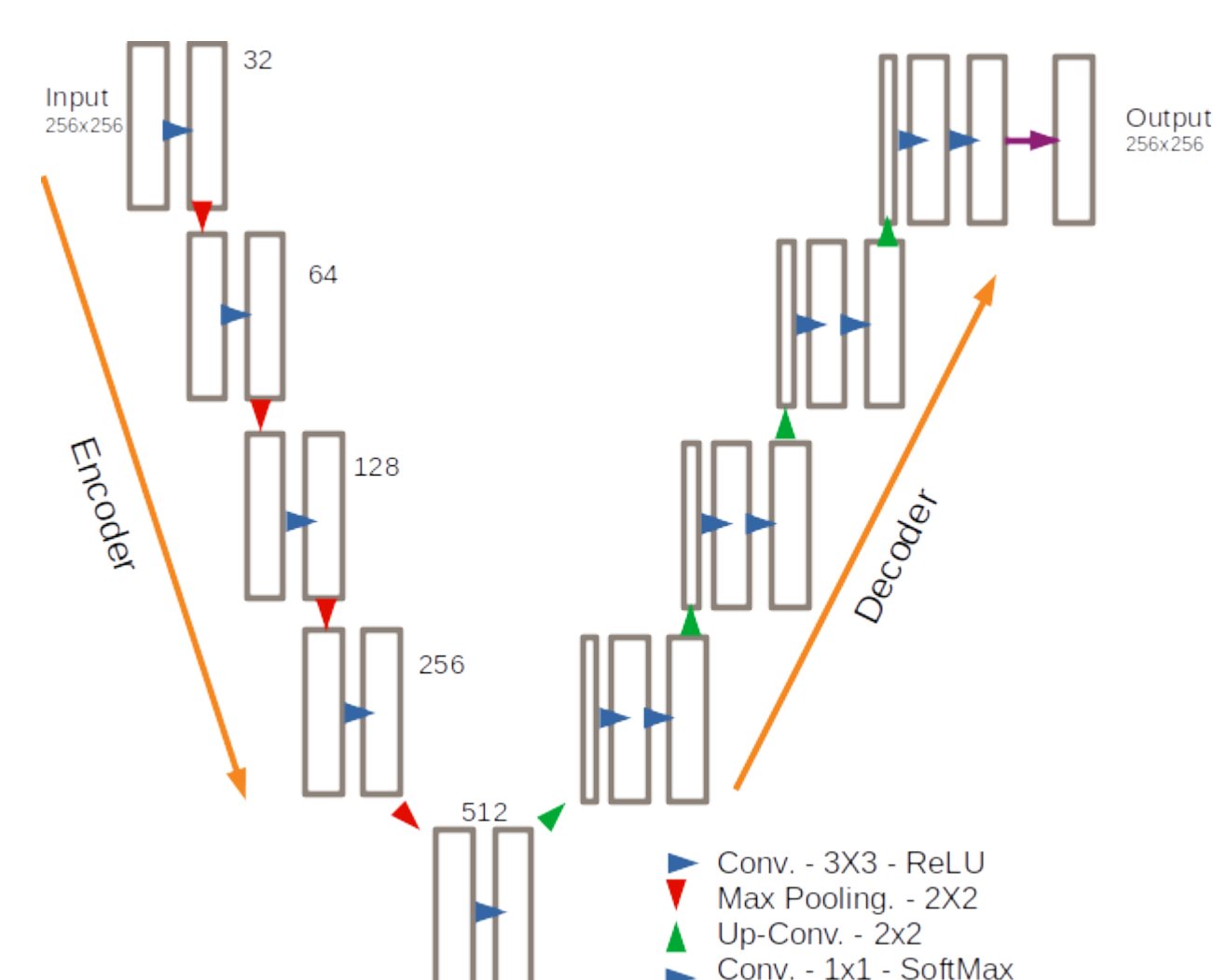

Figure 1: U-Net architecture used for image segmentation.

To increase the training set size (and improve image segmentation performance) a data-generator was constructed which rescaled images to 256x256 pixels and performed random rotations along with horizontal and vertical image shifts.

Results

RMSE on the with-held validation set reduces with each layer of complexity with the most complex varying slope and intercept model producing the most accurate predictions.

The leave one out information criteria [3] does not increase with increasing model complexity suggesting the model is not over-fit to the training data.

Contradictory to our hypothesis RMSE is similar between raw images and when image pre-processing via U-Net was implemented suggesting no improvement in prediction when using RMSE.

Table 1: Performance of regression models on training and validation data set

Model	Number of Parameters	Raw Images RMSE	Raw Images PSIS-LOO	U-Net RMSE	U-Net PSIS-LOO
1: $\mu_{ijk} = \beta_0 + \beta_1 x_{ijk} + \beta_2 x_{ijk}^2$	3	21.77	19265.0 (28.7)	19.80	18244.8 35.9
2: $\mu_{ijk} = \beta_0 + \beta_{1jk} x_{ijk} + \beta_{2jk} x_{ijk}^2$	13	12.45	17568.8 (175.3)	13.39	16551.7 (106.3)
3: $\mu_{ijk} = \beta_{0jk} + \beta_{1jk} x_{ijk} + \beta_{2jk} x_{ijk}^2$	18	1.74	7572.2 (98.2)	1.94	8125.9 (112.2)

U-Net segmentation had the effect of making differences between the 3 levels of blurring and two different stains become more apparent.
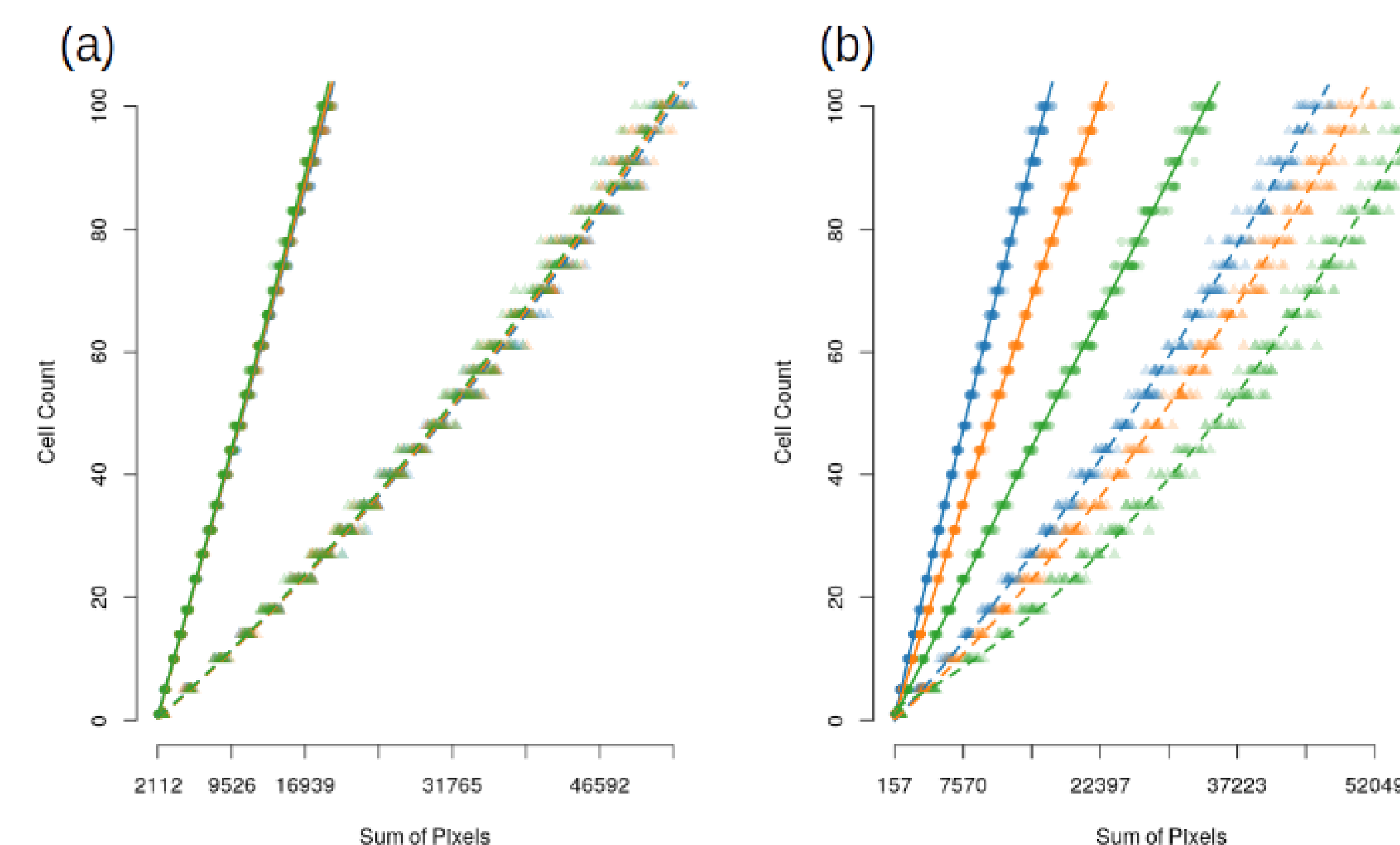

Figure 2: Relationship between sum of pixels and cell count for (a) raw images and (b) images after U-Net segmentation processing. Low blur is indicated in blue, medium in orange and high in green with nuclei stained images represented by circles and cell body stain represented by triangles. Fitted regression lines for the varying slopes and intercept model are included.

With Bayesian mixed level regression models we are able to sample from the posterior predictive distribution for the number of cells given each image in the test set and generate prediction intervals. The prediction intervals for images within the validation set were well calibrated with 95.8% of 95% prediction intervals containing the true cell count and intervals having an average width of 6 cells.
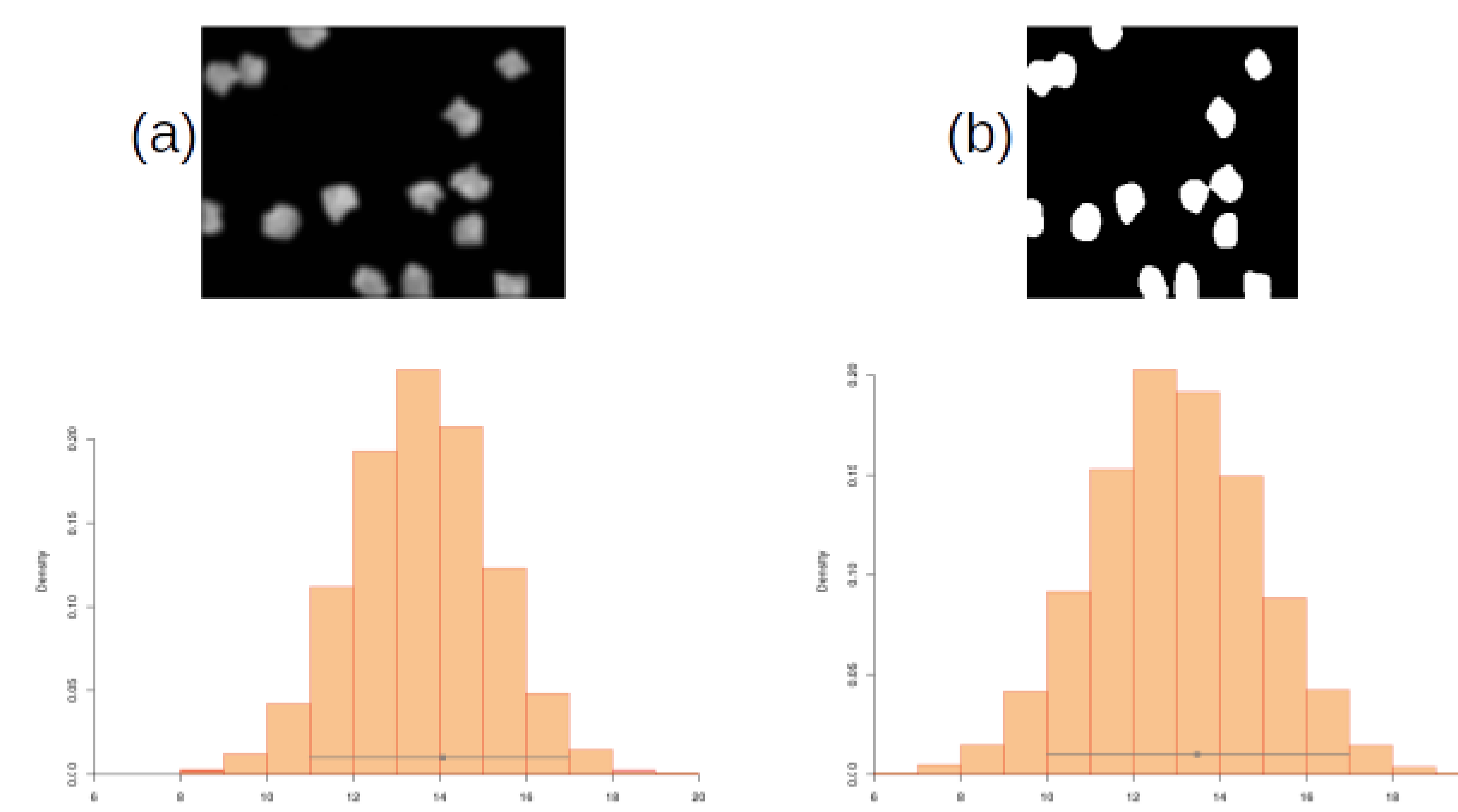

Figure 3: Posterior predictive distribution for an image containing 14 cells when generated via the raw image (a) or U-Net segmentation mask (b).

Conclusions

A multilevel polynomial regression model was effective in providing fairly accurate estimates of cell count and well calibrated prediction images with low computational cost while maintaining parameter imperturbability.

There was no apparent benefit to applying a segmentation mask generated by a U-Net Neural Network with regards to improving prediction performance.

Acknowledgements

References

[1] T. Falk, D. Mai, R. Bensch, O. Cicek, A. Abdulkadir, Y. Marrakchi, and A. Bohm. U-net: deep learning for cell counting, detection, and morphometry. *Nature Methods*, 16(1), 2019.

[2] V. Ljosa, K. L. Sokolnicki, and A. E. Carpenter. Annotated high-throughput microscopy image sets for validation. *Nature Methods*, 9(7), 2012.

[3] A. Vehtari, A. Gelman, and J. Gabry. Practical bayesian model evaluation using leave-one-out cross-validation and waic. *Statistics and Computing*, 27(5):1413–1432, 2017.