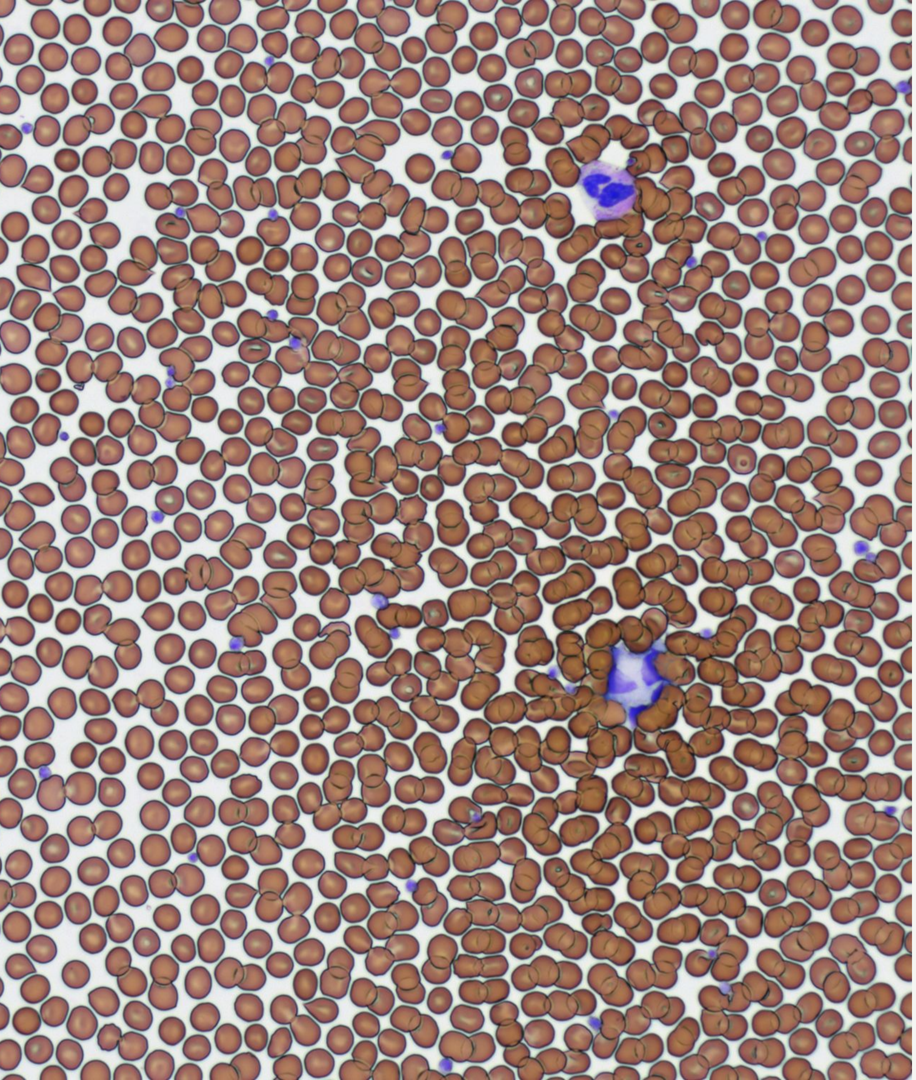


Segmenting Overlapping Red Blood Cells

With Classical Image Processing and Deep Learning

Nils Brünggel, Pascal Vallotton and Patrick Conway

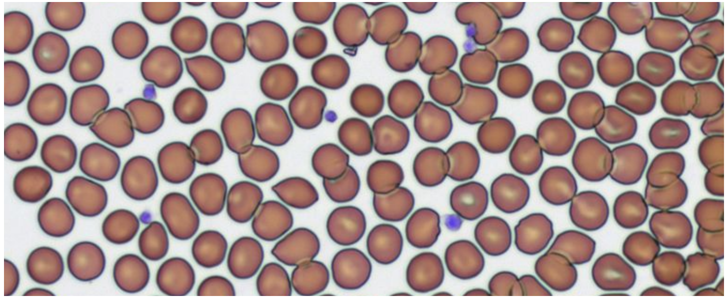
July 23. 2022

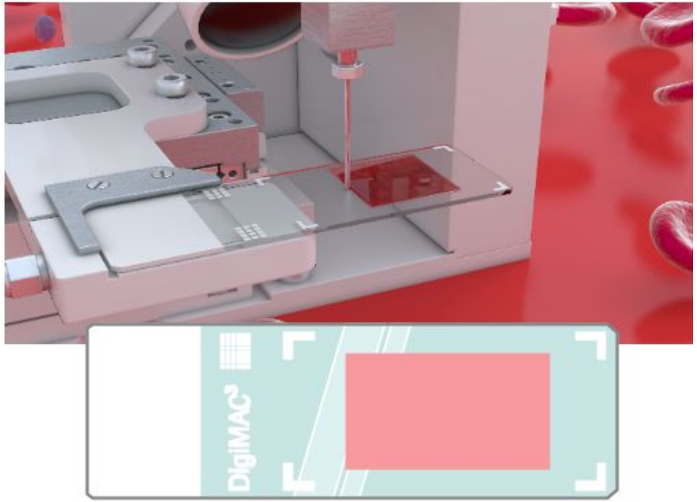
Table of contents

Overlapping Red Blood Cells on Hematology Slides

Overlapping Red Blood Cells on Hematology Slides

- Traditionally blood was dispersed manually on slides in a hematology lab.
- Goal is to create a Monolayer of cells where the cells are mostly freestanding.
- For red blood cells (RBCs) overlaps are still common.
- We used Roche's own Bloodhound®[1] printing and staining technology to create blood slides. But even with this technology RBC overlaps cannot be avoided entirely:





[1] 1 Bruegel, M., George, T. I., Feng, mB., Allen, T. R., Bracco, D., Zahniser, D. J., and Russcher, H. (2018). Multicenter evaluation of the cobas m 511 integrated hematology analyzer. Int J Lab Hem, 40(6):672–682.

Key Idea

Segmenting Red Blood Cells (RBCs)

- Free-standing: Easy because of distinct color and clear edges.
- Overlapping: Difficult to do with image processing. A deep-learning model would work but it requires masks.
- Drawing masks for a good segmentation dataset completely by hand is very tedious.
- Is there a better way?

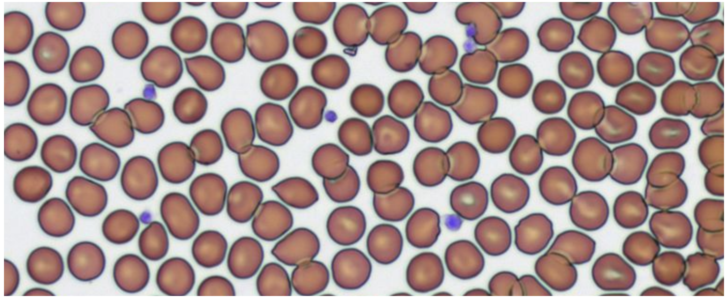
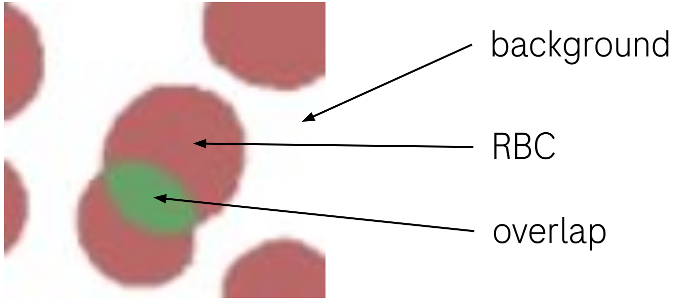
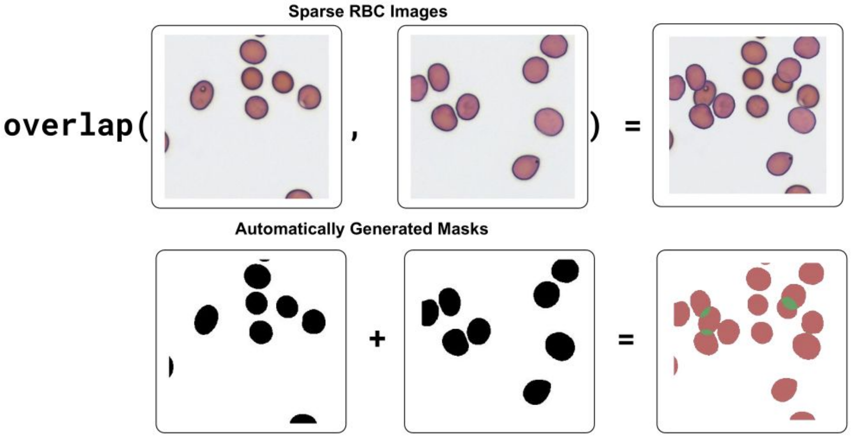


Illustration adapted, original Image by Chris Reed. Taken from [this article].

Artificial Overlaps

Naive Approach



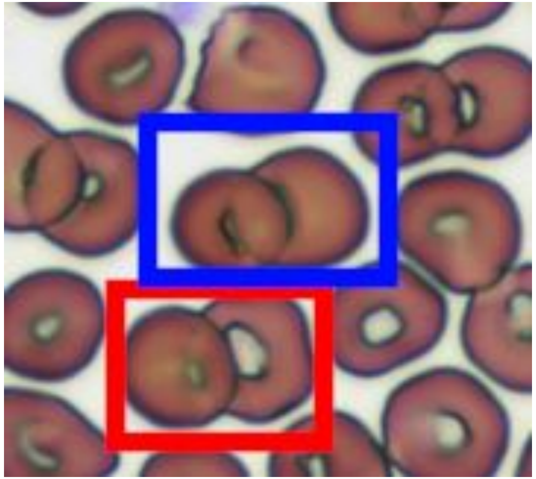


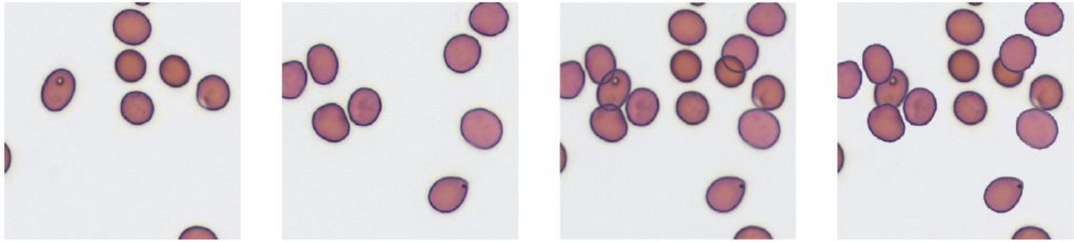
- Key Idea: Use images of free-standing RBCs to generate artificial overlaps and three class masks (BG, RBC and overlap)
- How about overlapping every crop with every other crop and that way creating $\binom{n}{2} \approx \frac{1}{2}n^2$ data points!
- Unfortunately this lead to a quick overfit which did not generalize to real overlaps.

Artificial Overlaps

Different Type of Overlaps

- We found two distinct types of overlaps:
  - Both cells are visible in the overlapping area.
  - Only one cell is visible.
- The generated overlaps still look a bit artificial (see figure 3). It is not binary, there are in between cases, but the extremes seem to dominate.





(a) first image (b) second image (c) min(a, b) (d) b over a

Figure 3: Simulated overlaps: Using two images with non-overlapping RBCs (a and b) we can generate two types of overlaps (image c and d).

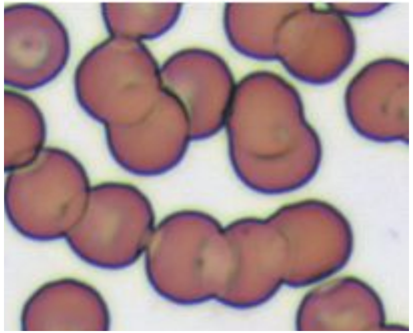
Artificial Overlaps

Algorithm to generate the dataset

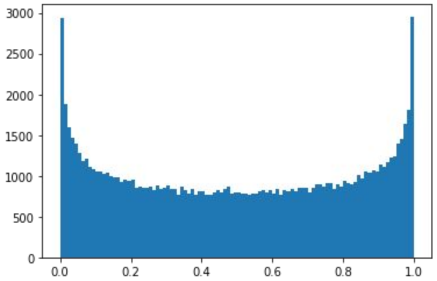
1. Combine each image with another image in the following way (n = 4’600):

$$a \sim Beta(\alpha, \beta)$$
$$i = a \cdot i_{min} + (1-a) \cdot i_{over},$$

2. Simulate Rouleaux by overlapping images with themselves and randomly move the images in a chosen direction (n = 30).
3. Randomly blur half of the generated images.
4. Add manually corrected masks (n = 12).



Rouleaux formation in a printed blood slide.



Bathtub Distribution, 100k samples

Model and Data Augmentation

Model and Data Augmentation

- U-Net from fastai[1]
    - ResNet50[2] as encoder.
    - No dropout layer, relies on PixelShuffle ICNR[3] upsampling.
- Data Augmentation
    - Default data augmentation, with the exception of setting the maximum rotation to 5 degrees and enable fips (horizontal and vertical).
- Many thanks to the fastai team! The fastai library saved me a lot of time.

[1] Howard, J. and Gugger, S. (2020). fastai: A Layered API for Deep Learning. arXiv: [2002.04688](https://arxiv.org/abs/2002.04688).

[2] Kaiming, H., Zhang, X., Ren S., and Sun J. (2015) Deep Residual Learning for Image Recognition [arXiv: 1512.03385](https://arxiv.org/abs/1512.03385)

[3] Aitken, A., Ledig, C., Theis, L., Caballero, J., Wang, Z., and Shi, W. (2017). Checkerboard artifact free sub-pixel convolution. page 16.

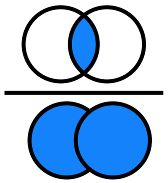
Results

Segmentation Performance on Test Set

Intersection Over Union (IOU)

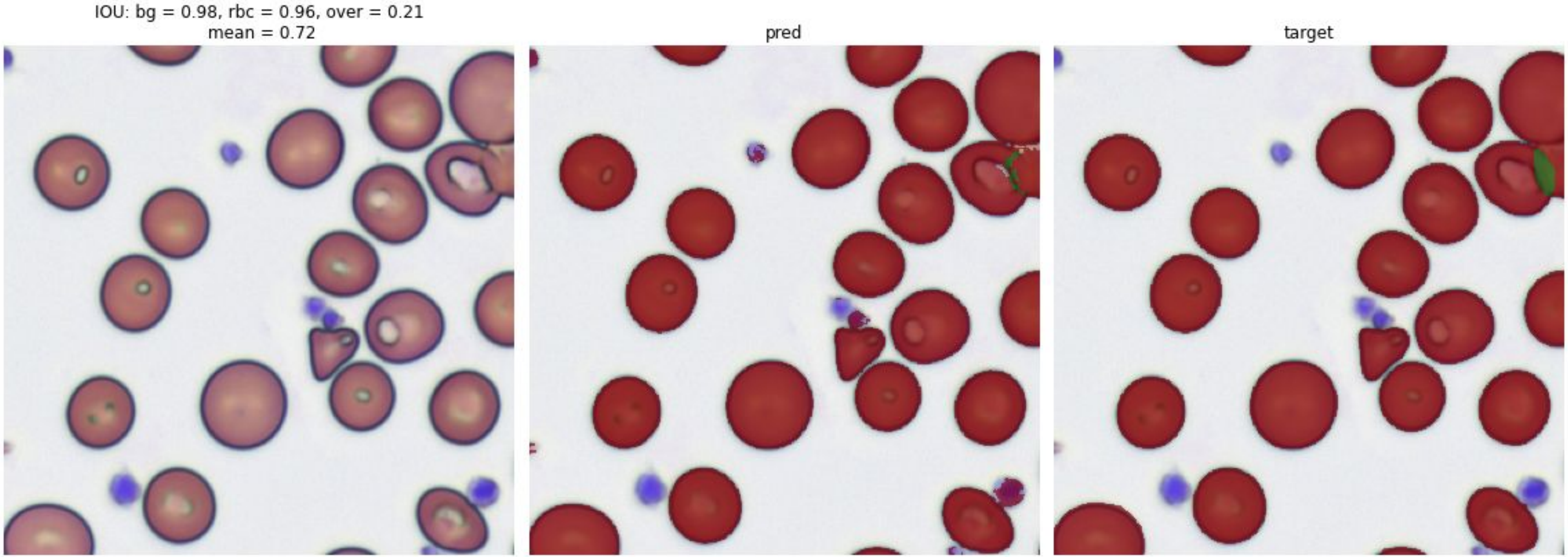
	background	rbc	overlap	mean
mean	0.96	0.96	0.65	0.86
std dev	0.02	0.02	0.11	0.04
min	0.9	0.89	0.21	0.72
max	0.99	0.98	0.79	0.91

IOU =



- The test set consists of 36 RBCs images from normal slides. The masks were drawn by hand.
- The IOU does not tell the full story. Let's look at predicted segmentation masks!

Test Set Worst



Test Set Best

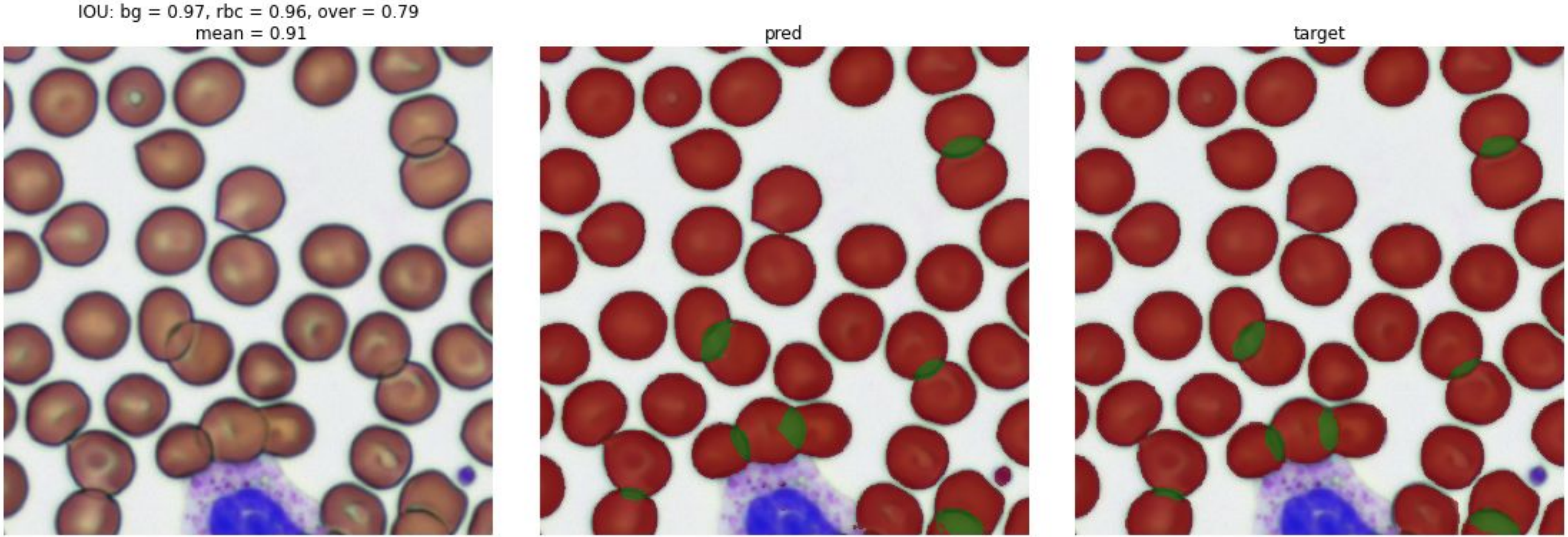


Image 1 with ~ mean IOU

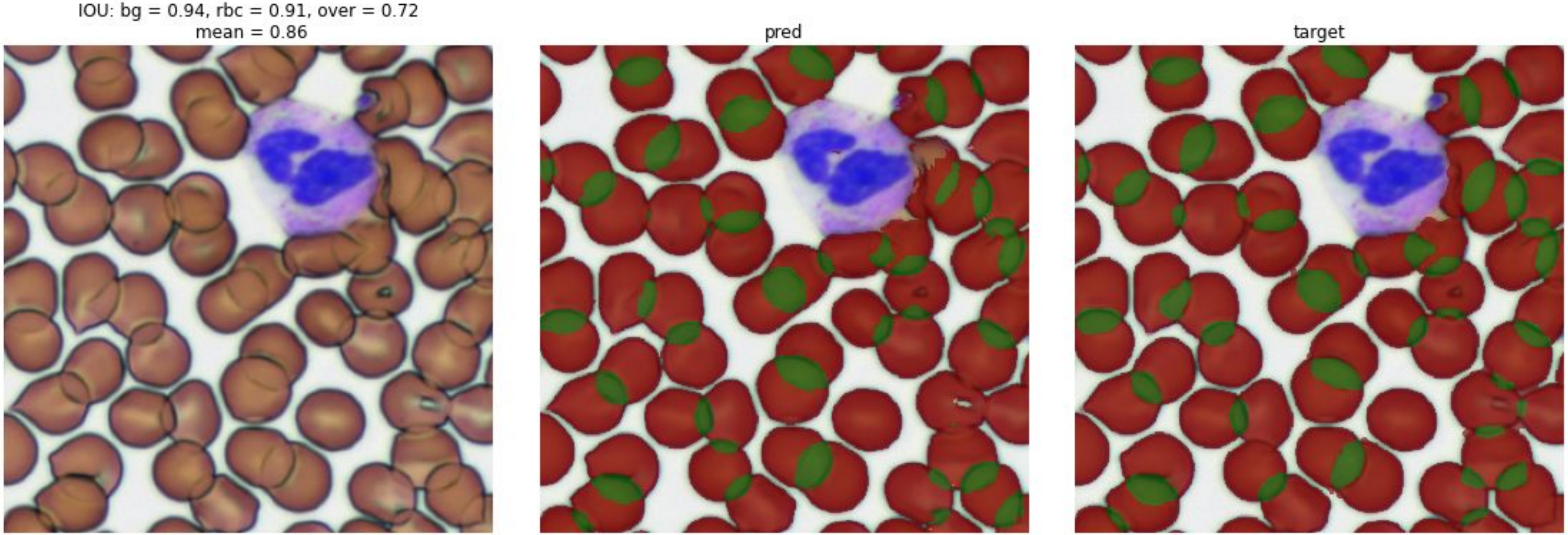
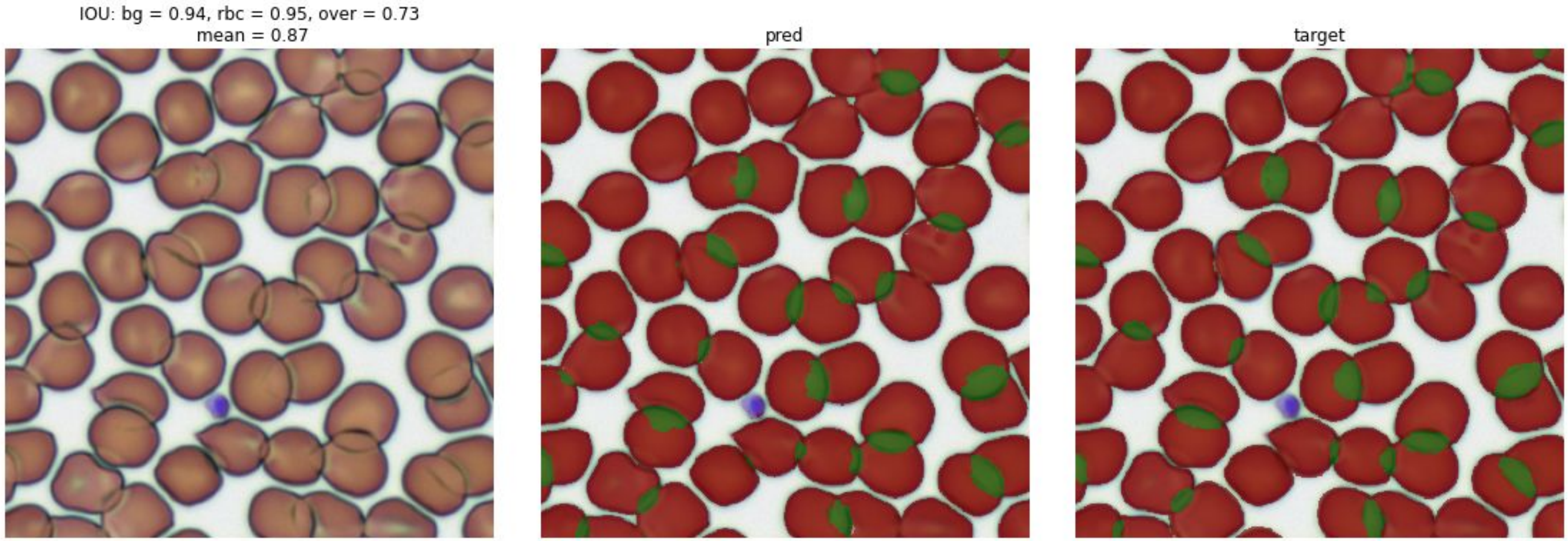
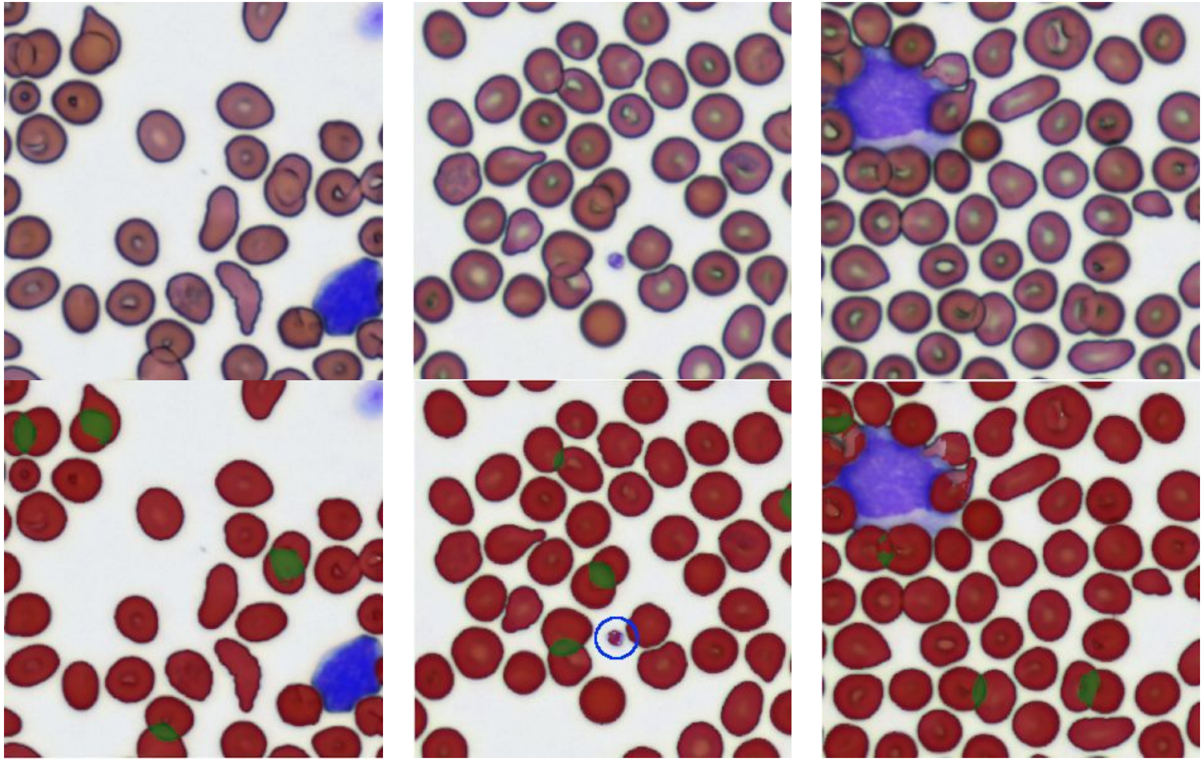


Image 2 with ~ mean IOU



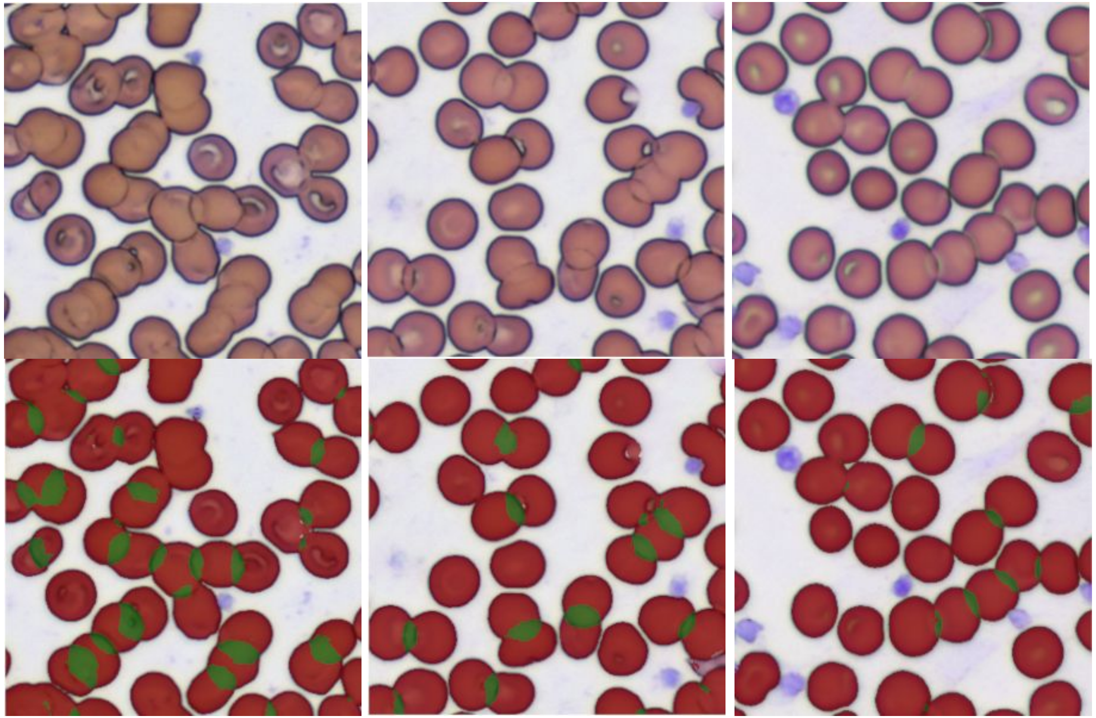
Qualitative Results: Normal Slide



a b c

	rbc	bg	over	mean
a	0.98	0.95	0.82	0.92
b	0.97	0.97	0.78	0.91
c	0.92	0.93	0.53	0.8

Qualitative Results: Rouleaux



a b c

"Rouleaux are stacks or aggregations of RBCs that form because of the unique discoid shape of the cells in vertebrates. … This is a nonspecific indicator of the presence of disease.

Conversely, the presence of rouleaux is a cause of disease because it will restrict the flow of blood throughout the body"

Shortened definition from Wikipedia.

	rbc	bg	over	mean
a	0.96	0.88	0.56	0.8
b	0.95	0.91	0.59	0.82
c	0.98	0.95	0.5	0.81

Conclusion

Conclusion

- State of the art U-Net architectures work well, if you have the right masks.
- Solving easy cases with simple image processing can bootstrap the learning process.
  - Experiment! The first attempt might not generalize.
  - An effective tool for mask review / editing is essential: We found that a simple web app in combination with GIMP works well.
- A single metric such as IOU is not enough, instead one need to systematically review predicted masks: For example it is desirable to separate barely touching RBCs by just one pixel to simplify subsequent instance segmentation.
- Further research: Use a deformable CNN[1] that can learn to adapt its receptive field. Zhang et al[2] could show that it reduces typical errors made by a CNN such as the one mentioned above.

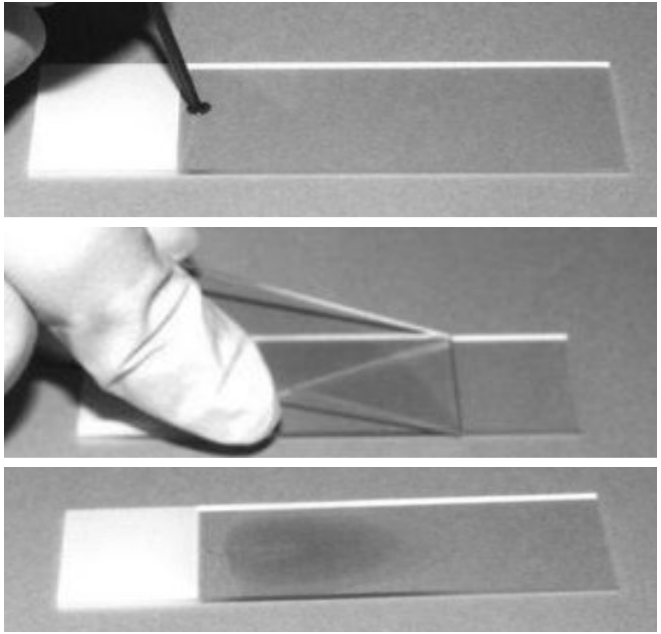
[1] Dai, J., Qi, H., Xiong, Y., Li, Y., Zhang, G., Hu, H., and Wei, Y. (2017). Deformable Convolutional Networks. arXiv:1703.06211

[2] Zhang, M., Li, X., Xu, M., and Li, Q. (2020). Automated Semantic Segmentation of Red Blood Cells for Sickle Cell Disease. IEEE J. Biomed. Health Inform., 24(11):3095–3102

Doing now what patients need next

Backup Slides

Manual Wedge Smear



Manual wedge smear[2]

[2] Images taken from a presentation by Hadeel Al Sadoun