Guiding the Management of Cervical Cancer with Convolutional Neural Networks

Stephen Pfohl (spfohl@stanford.edu), Oskar Triebe (oskar2@stanford.edu), Ben Marafino (marafino@stanford.edu)

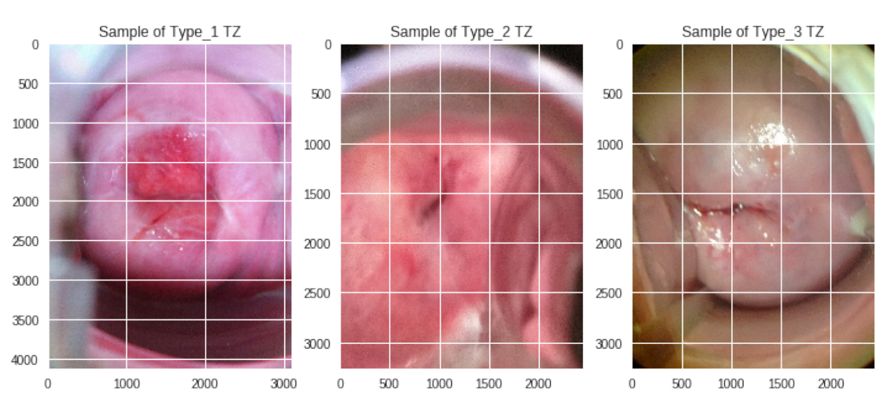
Abstract

Cervical cancer is the fourth most common cancer worldwide, with roughly 600,000 new cases and 300,000 deaths annually [1]. While cervical cancer may be prevented with timely screening and treatment, the selection of the most effective management protocol is dependent on type of the transformation zone (TZ) of the cervix, which may be determined through colopscopy. However, determining the type of a patient's TZ can be difficult, even for trained providers using colposcopy, and no computer vision-based algorithms currently exist to address this problem.
To confront this issue, we propose a convolutional network pipeline capable of determining the TZ type using the Intel/MobileODT database of images recently released on Kaggle.

With a simple transfer learning approach that builds off the Resnet architecture, we were able to achieve **position 51 out of 753 teams** on the Kaggle leaderboard with a **loss of 0.64971** on the leaderboard test set.

Data

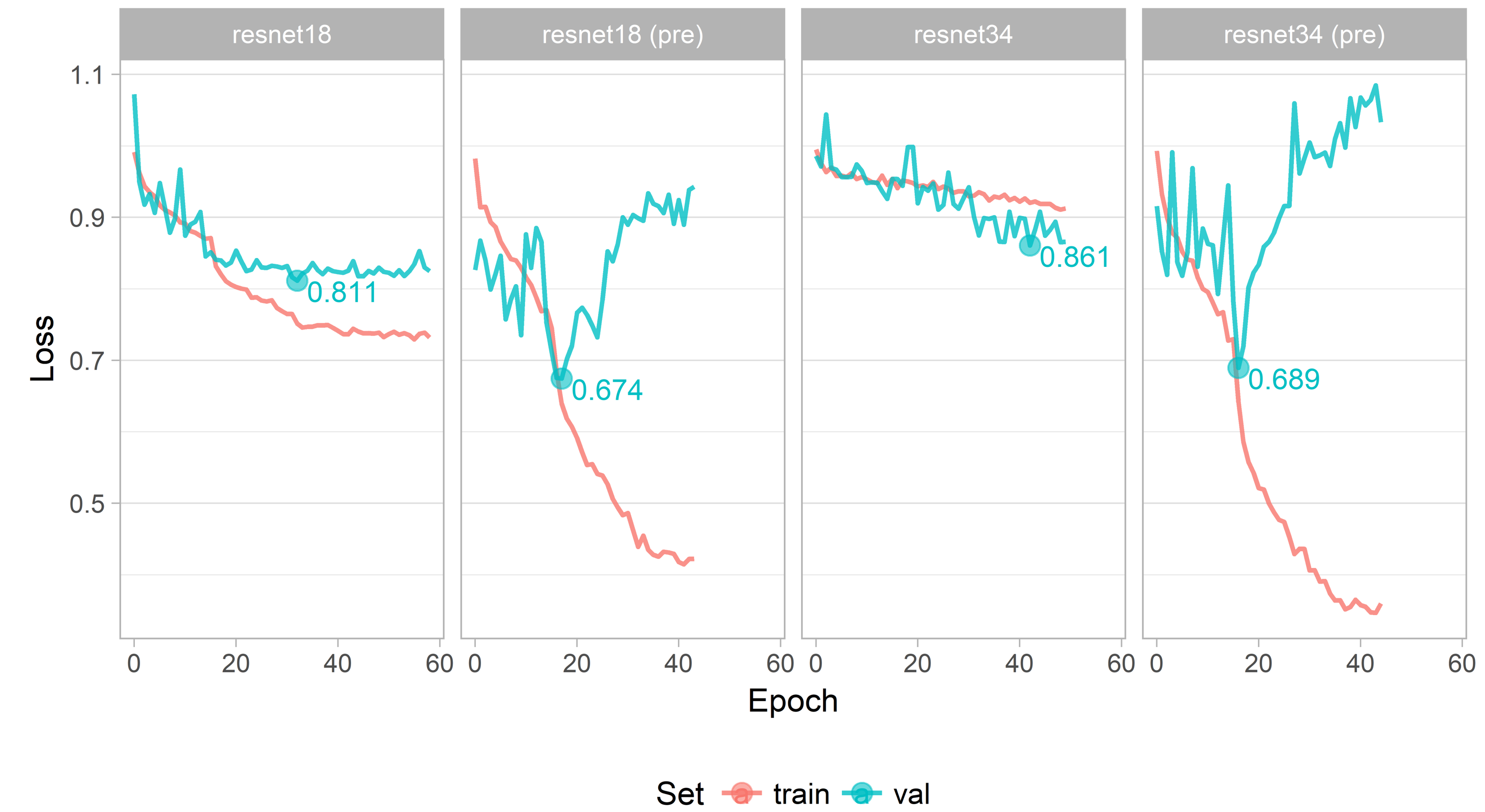
	Type 1	Type 2	Type 3
Train data (30/50/20)	250	781	450
Additional train data	1,191	3,567	1,976
Total	8,215 images		
Test set	512 images (unlabeled)		



Results

Train and Validation Loss

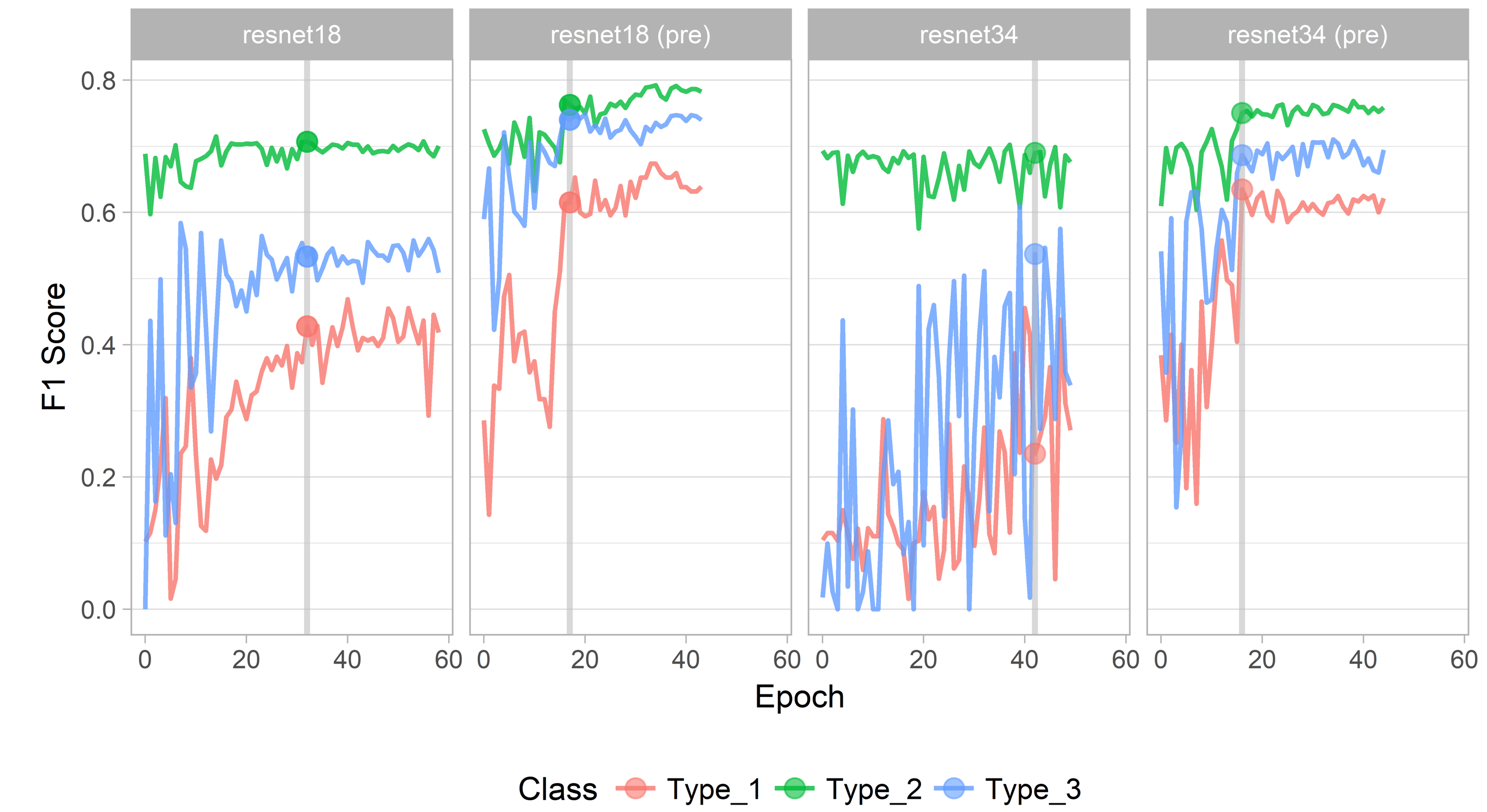
Models initialized with pretrained weights perform best, but start to strongly overfit after epoch 20.
The randomly initialized ResNet34 might profit from further training.



Fitted on train set (30% + additional) and evaluated on validation set (50%)

Validation F1 Scores

The models converge to a new optimum of class level F1 scores at optimal validation loss (grey line), particularly pretrained models. Further training does not improve the scores significantly.

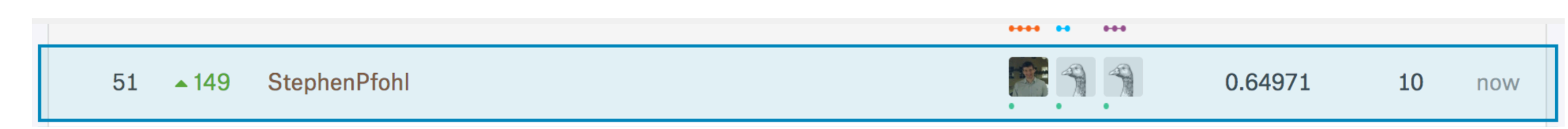


Fitted on train set (30% + additional) and evaluated on validation set (50%)

Performance on Test Set

Model	F_1 - Type 1	F_1 - Type 2	F_1 - Type 3	Loss	Accuracy
ResNet-18	0.40	0.67	0.55	0.87	0.60
ResNet-18 (pre)	**0.62**	**0.76**	0.66	**0.72**	**0.71**
ResNet-34	0.22	0.66	0.42	0.93	0.55
ResNet-34 (pre)	0.58	0.73	**0.67**	**0.72**	0.69

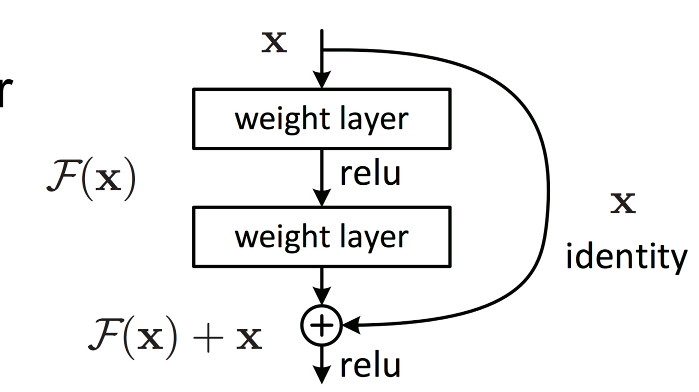
Kaggle Leaderboard



Artifacts and Heterogeneity

Models and Experimental Setup

- ResNet with 18 or 34 layers
- Initialize weights from a model pre-trained on Imagenet or train network from scratch
- Train all layers of the initialized network
- Perform hyper-parameter selection over validation loss
  - Learning rate
  - L_2 regularization
  - Learning rate decay
  - Augmentations
- Compute test performance for model with lowest validation loss



Augmentations

TRAIN

Pad

Flip

Rotate

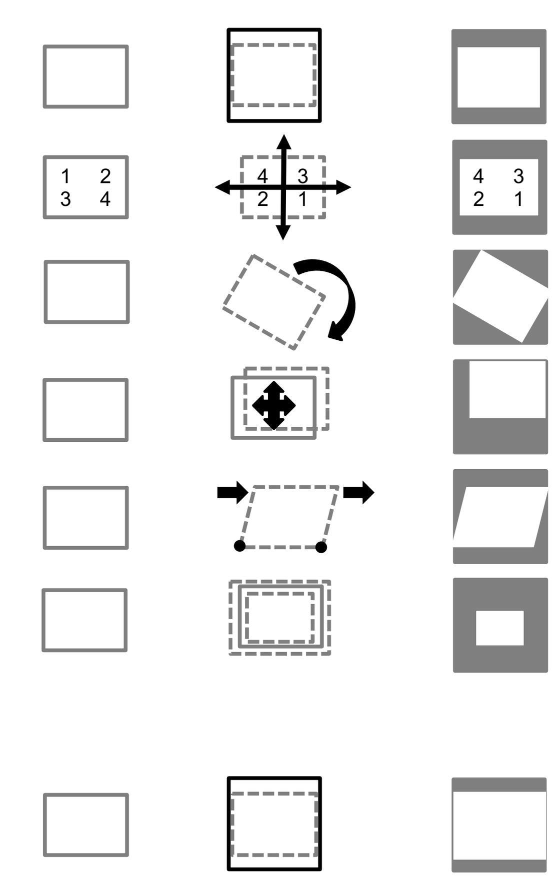
Translate

Shear

Zoom

VAL/TEST

Pad



Conclusion

- Standard transfer learning pipelines are sufficient for developing a high performing cervical classification algorithm
- Class-level F_1 is correlated with the prevalence of the class in the data and thus imbalanced learning approaches may be appropriate
- Generalization performance would likely improve upon addressing quality issues in the data

References

•B. Stewart and C. Wild, editors. World Cancer Report 2014. World Health Organization.
•Kaggle. Intel & MobileODT Cervical Cancer Screening:Which cancer treatment will be most effective?
•K. He, X. Zhang, S. Ren, and J. Sun. Deep Residual Learning for Image Recognition. arXiv: 1512.03385,2015