


# An Automated Deep Learning Artifact Detection Tool for Quality Control of Whole-Slide Digital Pathology Images

Daniela Rodrigues<sup>1</sup>, Erik Burlingame<sup>1</sup>, Corey Babcock<sup>1</sup>, Vlado Ovtcharov<sup>1</sup>, Stefan Reinhard<sup>2</sup>,  
Therese Waldburger<sup>2</sup>, Daniel Martin<sup>3</sup>, Suzana Couto<sup>3</sup>, Inti Zlobec<sup>2</sup>, Peter Caie<sup>1</sup>

 [drodrigues@indicalab.com](mailto:drodrigues@indicalab.com)  
[info@indicalab.com](mailto:info@indicalab.com)

<sup>1</sup>Indica Labs, Albuquerque, NM, United States  
<sup>2</sup>Institute of Tissue Medicine and Pathology, University of Bern, Switzerland  
<sup>3</sup>Genmab, Plainsboro, NJ, United States

**indica labs**

## INTRODUCTION

Mechanical and digital artifacts have a negative impact on digital pathology workflows. Image focusing issues can be a critical bottleneck during slide digitisation.

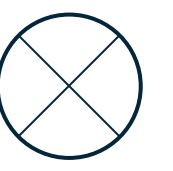

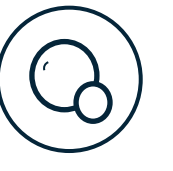
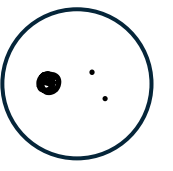

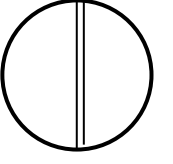
We developed SlideQC to automatically segment tissue artifacts in hematoxylin and eosin (H&E) and immunohistochemistry (IHC) stained whole slide images (WSI).

## SlideQC DEVELOPMENT

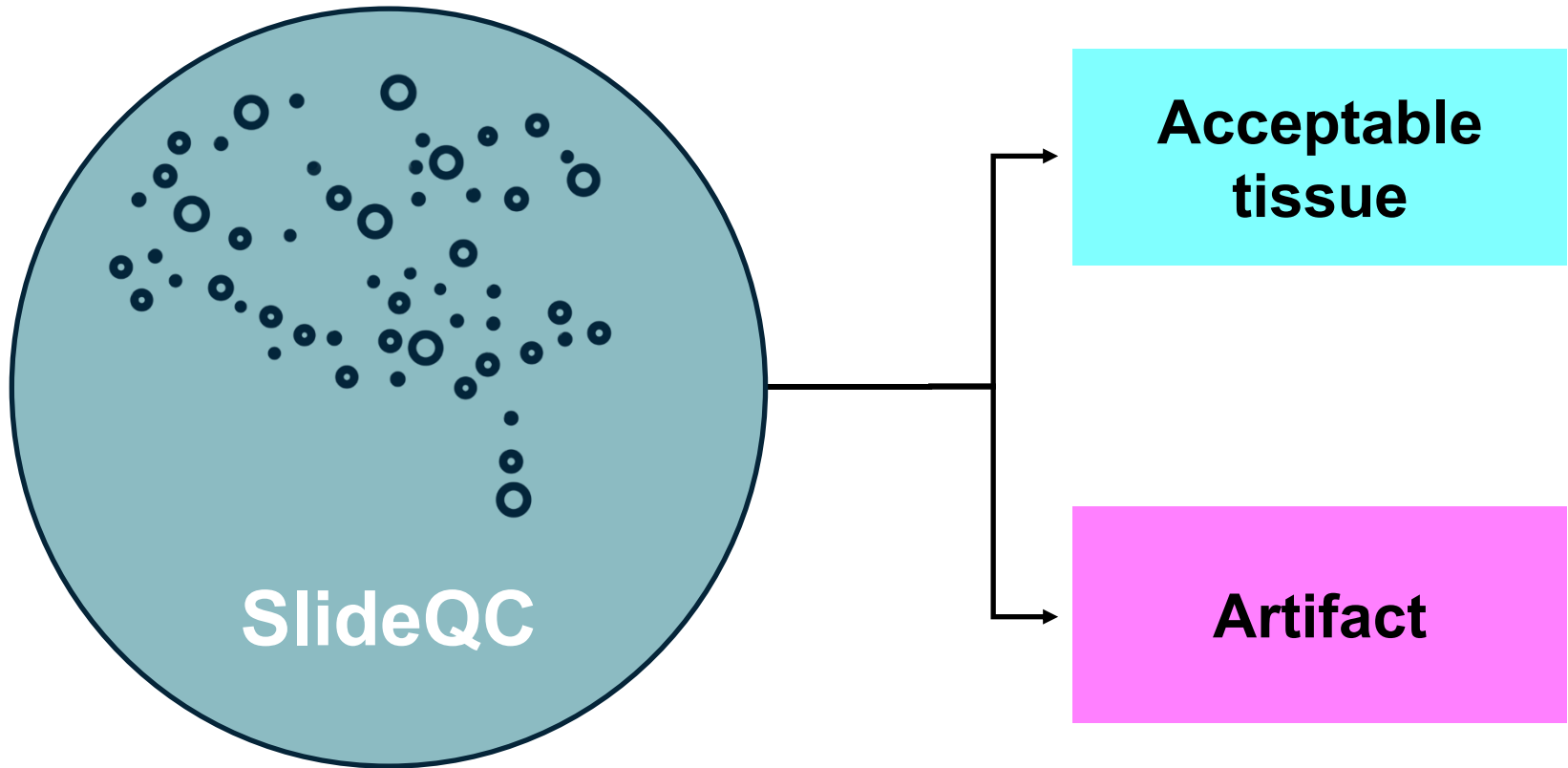
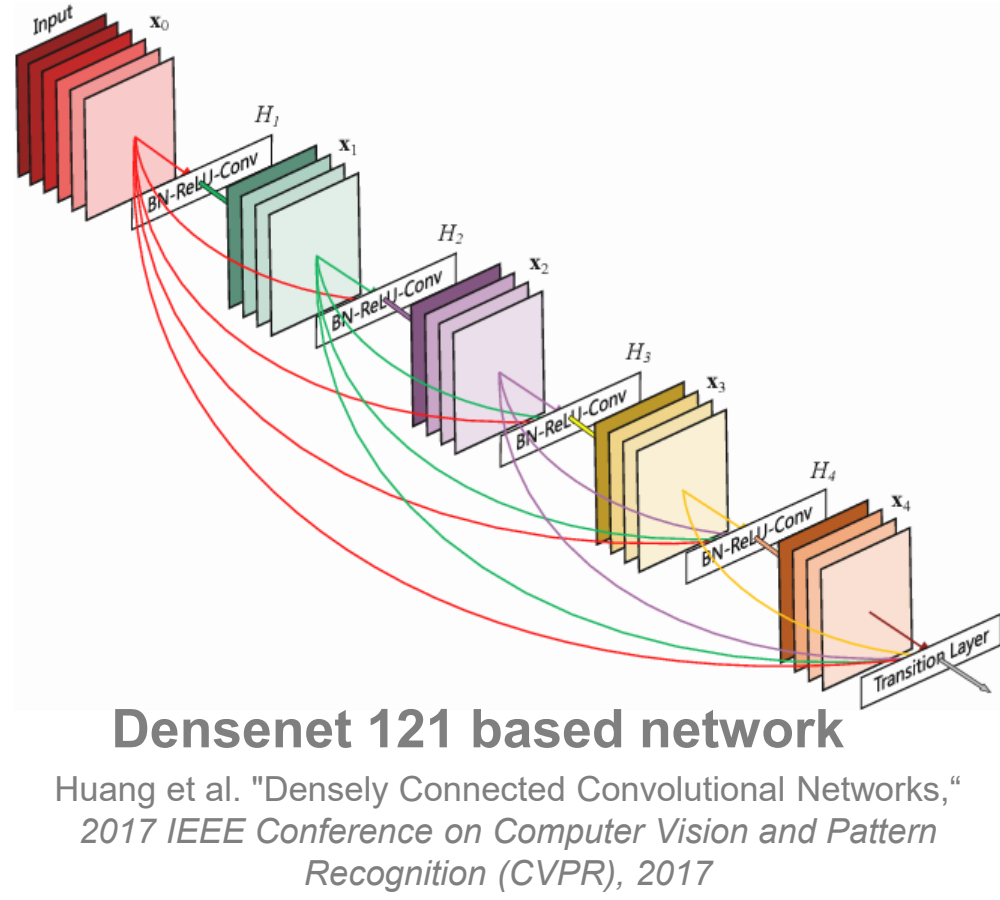
2499 Annotations comprising a total artifact area of 52.15 mm<sup>2</sup>

Across 302 H&E and IHC stained WSI from more than 9 tissue types

A Densenet based network was trained to classify slides as acceptable tissue or artifact.

-  Out of focus
-  Pen marks
-  Air bubbles
-  Dust
-  Tissue folds
-  Coverslip Issues

Thyroid  
Stomach  
Uterus  
Skin  
Ovary  
Colon  
Breast  
Prostate  
Lung

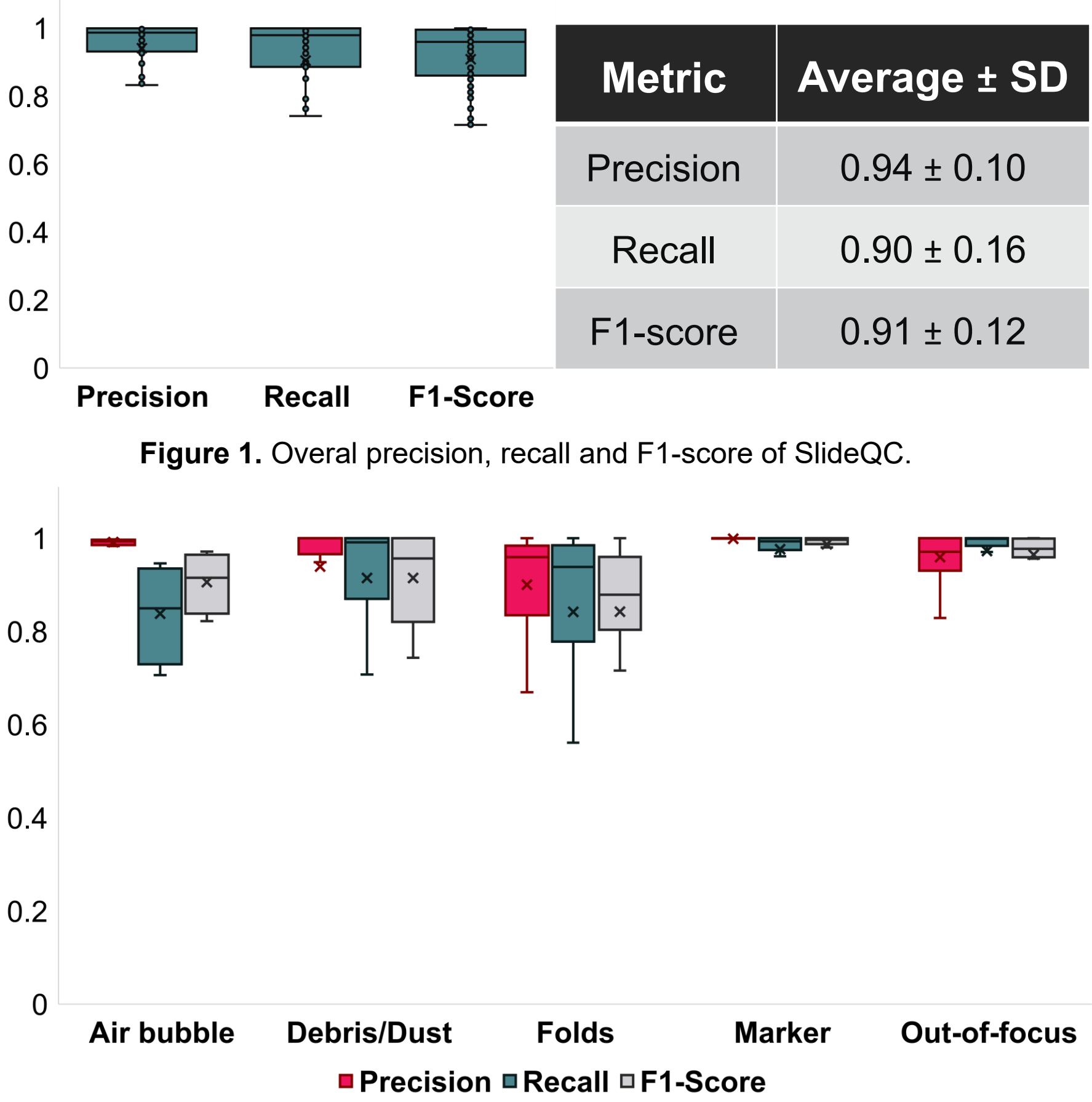


The training set was supplemented with a set of 2048 synthetically generated out-of-focus images.

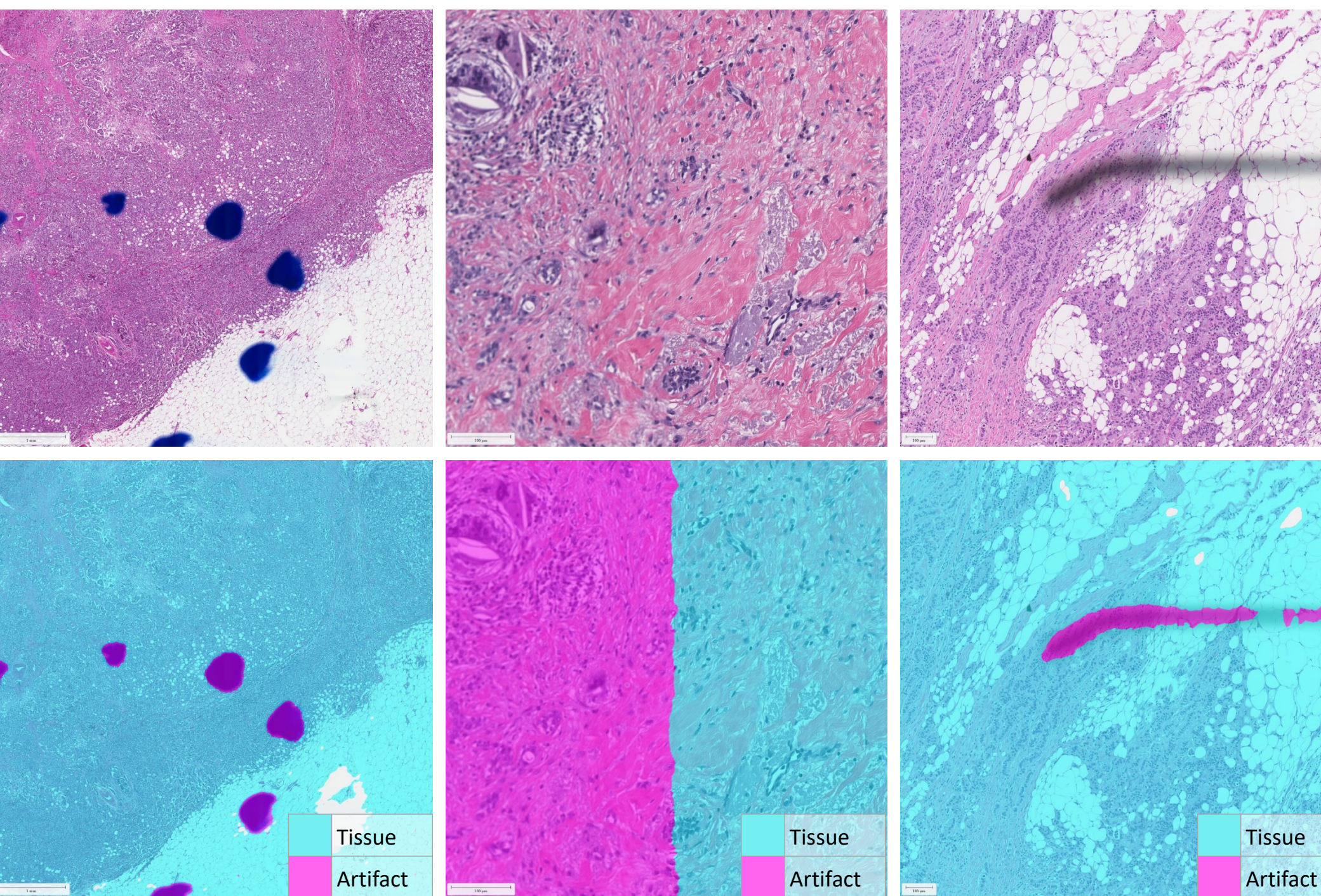
## RESULTS

SlideQC showed high precision, recall and F1-score over pixel-level annotations on external test sets

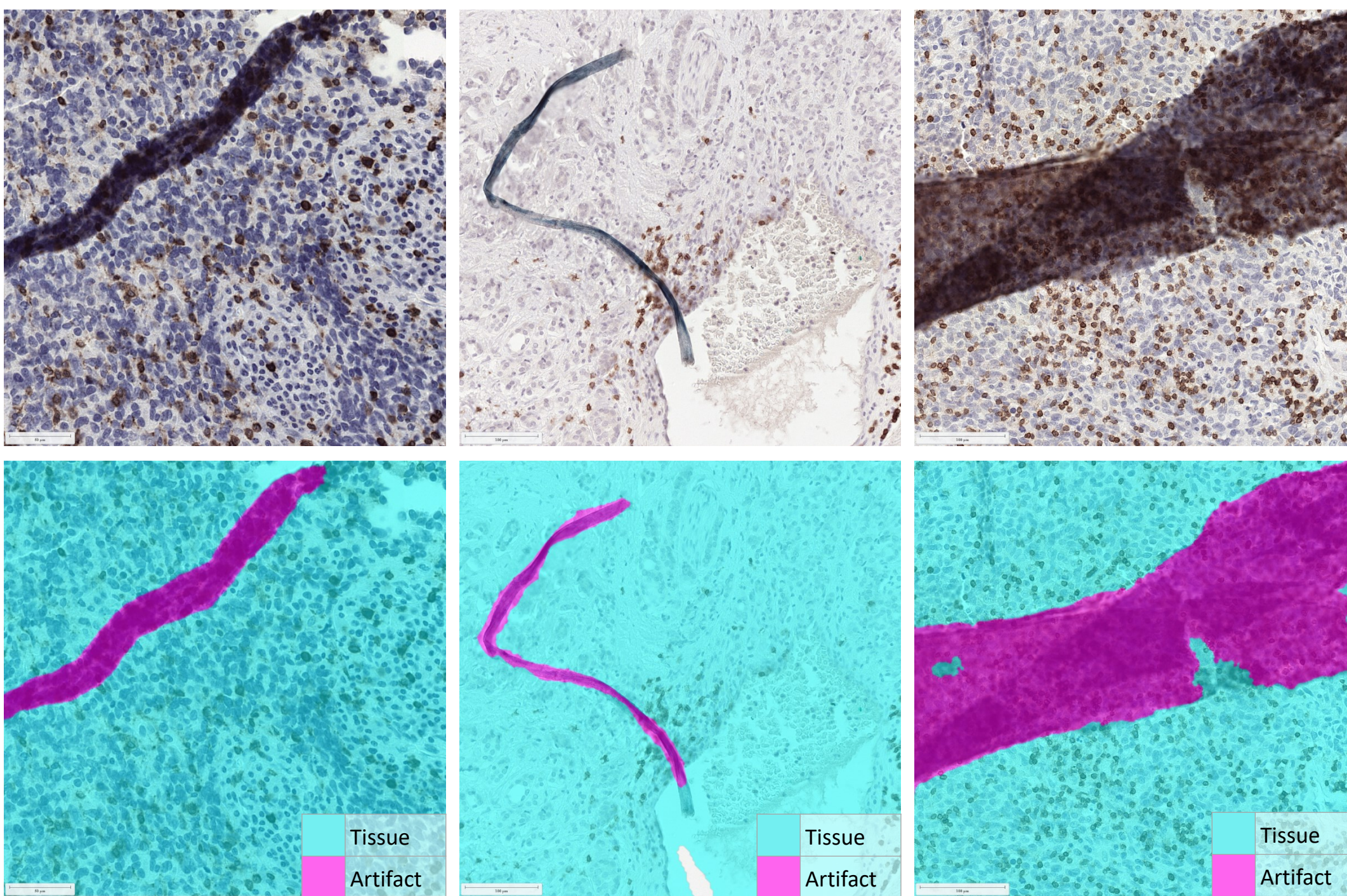
SlideQC performance was evaluated on 432 annotations across 79 external H&E (HistoQCRepo) and IHC (LYON19) test images.



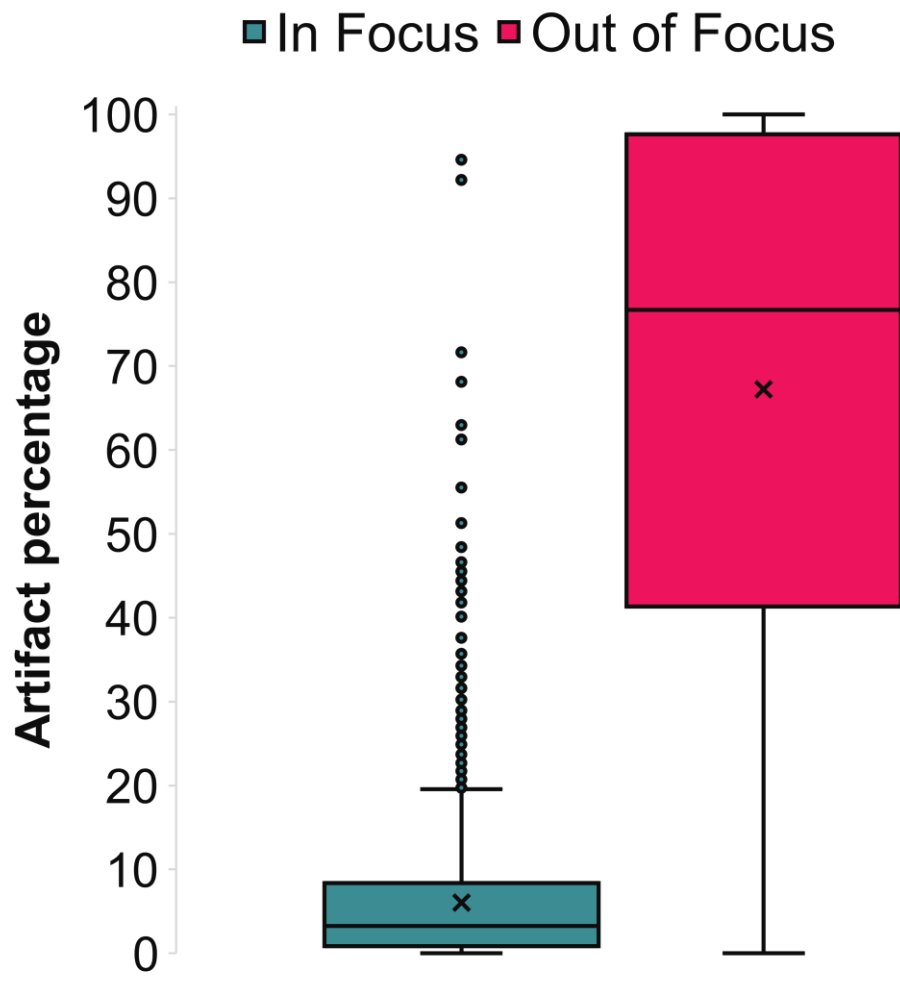
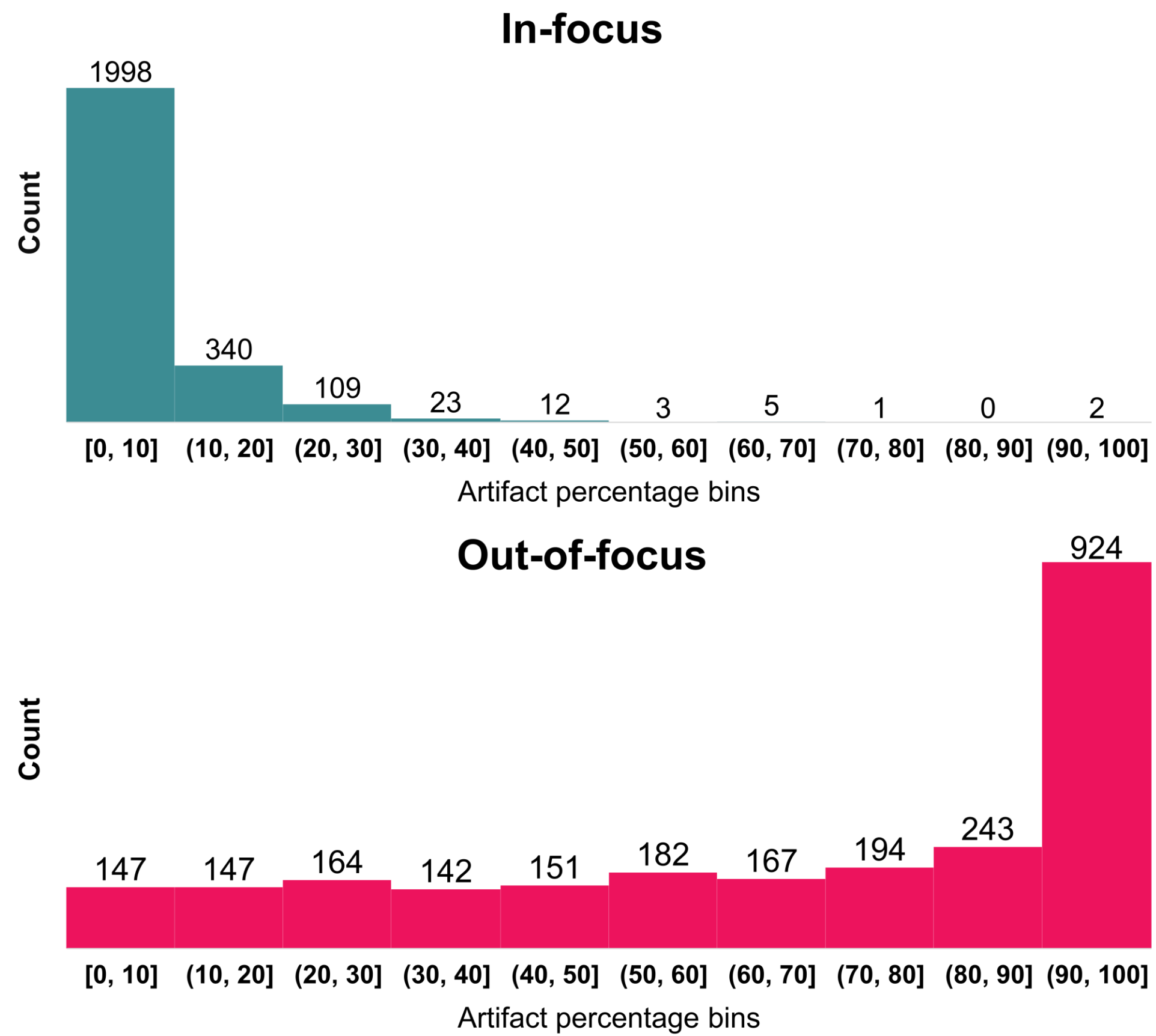
Markup Results on H&E images (HistoQC Repo)



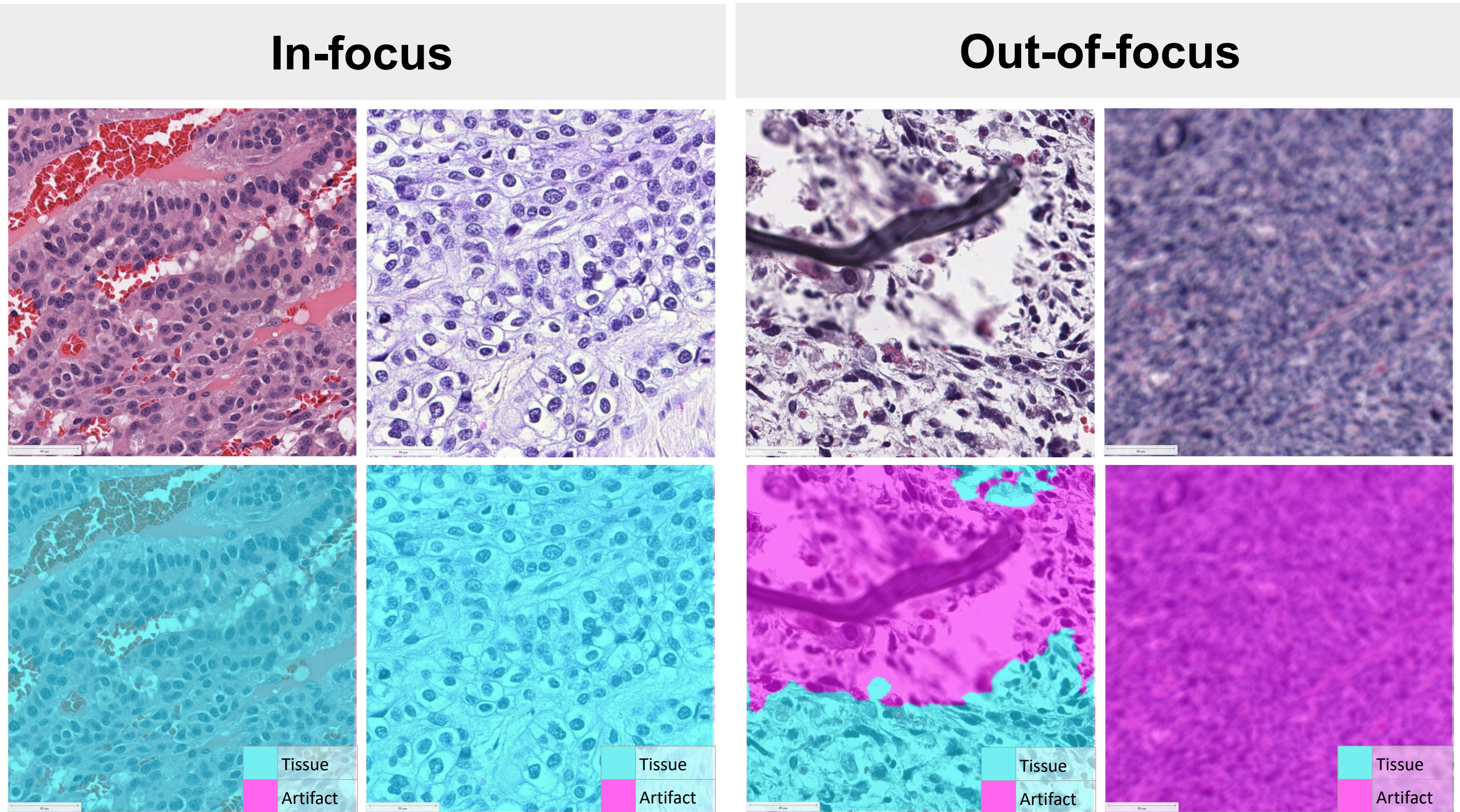
Markup Results on IHC images (LYON19)



SlideQC showed a good ability to distinguish out-of-focus from in-focus labelled patches from the TCGA@Focus dataset



The median percent of tissue area classified as artifact within the 2461 patches labelled as out-of-focus was 76.7 [IQR 41.3 – 97.6] and for the 2493 in-focus labelled patches was 3.3 [IQR 0.8 – 8.4].



## CONCLUSIONS

SlideQC can alleviate the bottleneck of manual quality control in both clinical and research based digital pathology workflows, thereby bringing efficiency gains to both fields. SlideQC achieved high precision, recall, and F1-score in H&E and IHC external test cohorts. Furthermore, SlideQC showed the ability to distinguish out-of-focus from in-focus patches in the TCGA@Focus dataset. By identifying and reporting the percentage of artifacts on each slide, SlideQC can provide an automated, quantifiable quality control procedure.

SlideQC is commercially available in HALO AI and in the HALO AP® platform. SlideQC is for Research Use Only, not intended for clinical diagnostic use.