An automatic out of focus detection model for enabling the assessment of liver biopsies.

Authors : Kezia Hobson[1], Dylan Windell[1], Paul Aljabar[1], George Ralli[1], Robert Goldin[1,2], Kenneth Fleming[1,3] and Caitlin Langford[1].

Affiliations:
1. Perspectum Ltd, Oxford, UK
2. Section for Pathology, Imperial College, London, UK
3. Emeritus Fellow, Green Templeton College, University of Oxford, UK

Background: As digitised workflows become predominant in clinical trials, new artefacts are introduced not observed in traditional glass workflows. Focus artefacts occur when glass slides are scanned to produce Whole Slide Images (WSIs). Quality Control (QC) of WSIs is required to enable robust reporting by pathologist. Manually checking WSIs for artefacts is time consuming and subject to error, thus automated QC is preferable. To address this, we developed an automated out of focus (OoF) detection algorithm for liver biopsies.

Methods: A dataset of 100 in-focus tiles from 51 liver core WSIs were extracted randomly. Each tile was subsequently converted to greyscale and 3 levels of blurring were simulated using a gaussian filter, resulting in a final training (70%) and testing (30%) set of 400 tiles (288 Masson's Trichrome, 112 H&E). The entropy, skew and kurtosis of the tile was calculated before and after applying a further gaussian blur, as well as the change in entropy. These metrics were used to train a binary random forest classifier, which assigns each tile as OoF or not.

A validation study was run on 50 WSIs from 3 liver clinical trials (2 x MASLD/MASH and 1 mixed liver injury) and an additional 4 wedge biopsies. Ground truth labels were assigned to WSIs manually. Tiles including foreground tissue were classified as in focus or OoF and the resulting percentage was assigned a WSI level label. WSIs with >50% OoF tiles were flagged as inadequate quality for reporting.

Results: The algorithm had a high prediction accuracy (98%) on the test set and validation study (98.1% with 0% false negative rate). Two incorrectly predicted tiles in the test data were highly fibrotic (liver capsule) where low levels of gaussian blur was applied and incorrectly classified as in focus. Five WSIs in the validation study were >50% OoF with 49 WSIs containing varying amounts of OoF but all below the threshold.

Conclusion: Our Machine Learning (ML) model can correctly detect tiles that contain OoF regions on liver biopsies with high accuracy across differently stained WSIs from various laboratories. Due to the accuracy, this can be utilised as an automated QC step on WSIs for the diagnosis and monitoring of MASLD and MASH in clinical trials. While we trained on MA**S**LD/MASH data, this ML algorithm could be applied to a wide range of clinical trials.