Mitotic figure detection in rat liver using supervised deep learning based object detection models

Nijun Wei[1], Jogile Kuklyte[1], Ovidiu Ghita[1], Aleksandra Zuraw[2], Laoise Bissett[1], Lise Bertrand[2], Daniel Rudmann[2], Donal O'Shea[1], Mark Gregson[1]

[1]Deciphex Inc., DCU Alpha Innovation Campus, Dublin, Ireland D11KXN4 [2]Charles River Laboratories Inc., Frederick, MD 21701, US.

Introduction

- Accurate identification of rare events such as mitotic figures, in toxicologic pathology studies provides valuable information about the potential toxic effects of a substance which can help guide decisions about its safety and use. However, pathologists might find it challenging to grade because
  - cellular changes such as mitotic figures are relatively rare in the tissue
  - these cells can be morphologically similar to other types of cells
  - and distributed unevenly within the tissue.
- Artificial intelligence (AI) object detection (OD) models can be trained with peer-reviewed annotations of single cells on whole slide images and can detect these rare events on unseen slides and studies in an efficient and standardised manner.
- In this poster, the experiments and results are shown on the detection of mitotic cells in rat liver tissues. A high level of generalisation capability is shown with F1 scores average of 0.94 achieved on the five blinded studies.

Methods

⇒ Dataset

- 99 hematoxylin and eosin (H&E) slides of rat livers including 567 annotations across
  - 8 toxicological studies
  - 2 scanners (Fig 2)
  - 20x and 40x magnifications (Fig 3)
- All annotated tiles have been fully checked for mitotic figures in order to identify false positives during evaluation.

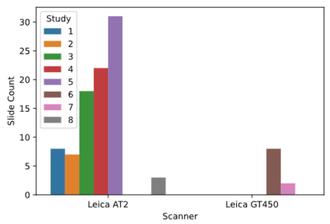

Fig.1 Distribution of image slides from 8 annotated studies across 2 scanners.

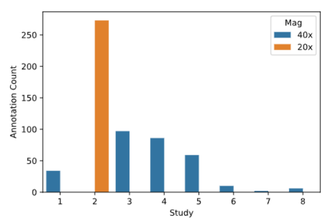

Fig.2 Distribution of annotations in 8 annotated studies including both 20x and 40x slide images.

⇒ Method

- Three classic OD algorithms were investigated here including FrCNN [1], RetinaNet [2] and EfficientDet [3].
- Experiments on training with 20x image tiles or 40x image tiles were explored.

⇒Evaluation

- Metrics used for evaluation were specificity, sensitivity and F1-score which are calculated with a default probability threshold of 0.4 and a Intersection over Union (IoU) threshold of 0.7.
- 15% of data was set aside for validation at the tile level.
- Generalisation capabilities were quantified by study-preserved validation approach, where full studies were set aside for evaluation.
- Qualitative evaluation was performed by the pathologists at the slide level.

Model	Specificity	Sensitivity	F1-Score	AP
FrCNN	0.949	0.949	0.949	0.972
EfficientDet	0.619	0.987	0.761	0.664
RetinaNet	0.92	0.988	0.952	0.936

Tab.1 Tile level evaluation results on mitosis detection of three OD models.

Results

⇒ Faster R-CNN (FrCNN) model gave the best detection accuracy out of the three algorithms explored based on an average precision (AP) of 0.972 from the 15% validation result (Tab 1).

⇒ Validation results on 15% unseen tiles for models trained at 20x and 40x magnification respectively showed the model trained at 40x magnification achieved a better performance with a F1-score of 0.957 (Tab 2). One study scanned at 20x magnification had to be excluded from the 40x experiments.

⇒ FrCNN model trained at 40x magnification with the non-max suppression (NMS) threshold of 0.1 and IoU threshold of 0.9 performed the best in the experiments.

Slide Magnification	Specificity	Sensitivity	F1-Score
20x	0.905	0.76	0.826
40x	0.971	0.943	0.957

Tab.2 Tile level evaluation results on mitosis detection of three OD models.

⇒ The results showed high generalisation capabilities of the current approach - where the results were evaluated on unseen studies. Five experiments with different studies excluded showed that the average F1 score is 0.94 (Fig 3-4).

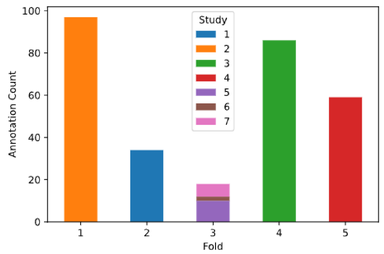

Fig.3 Study distribution for 5 validation experiments.

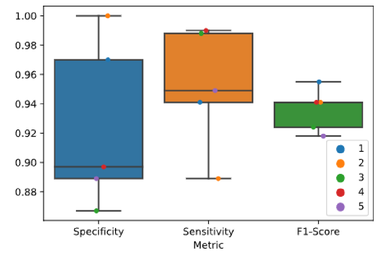

Fig.4 Metric values for the evaluation across 5 experiments

⇒Mitotic figures detected represented different phases of mitotic figures, e.g. prophase, metaphase, anaphase and telophase (Fig 5). False positive and false negative examples are shown in Fig 6 and Fig 7.

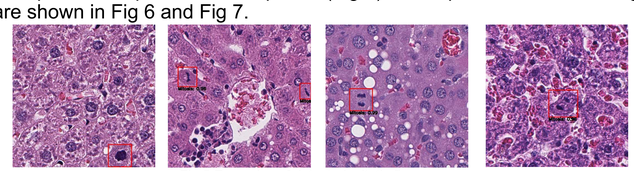

Fig.5 Examples of detected mitotic figures.

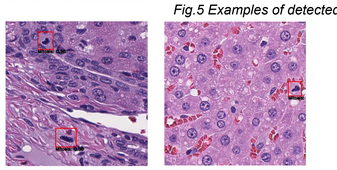

Fig.6 Examples of false positive detections.

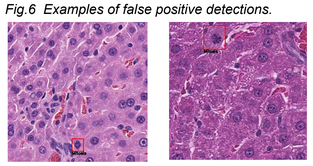

Fig.7 Examples of false negative detections.

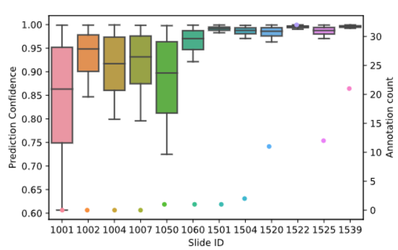

Fig.8 Distribution of confidence levels of predicted mitotic figures versus number of mitotic figures used in training accumulated per slide along x-axis.

⇒Figure 8 shows the distribution of the confidence levels of the mitotic figure detections. The slides that were not used for training (1001-1007) have higher percentages of lower confidence than the one that have at least one mitotic figure annotated and included in the training data (1050-1539). Lower probabilities were mostly false positives, that can be further filtered out by increasing the confidence threshold.

⇒Figure 9 shows results on two groups of slides - the ones that do not have reported findings, and the ones that have "Increased Mitosis" reported by pathologists (primary and peer-reviewed). Number of mitotic figures detected by the algorithm correlated with pathologists reported findings.

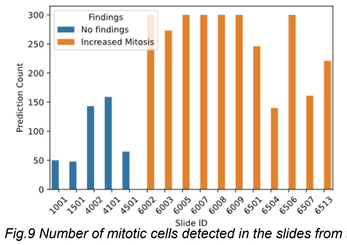

Fig.9 Number of mitotic cells detected in the slides from two groups: without findings reported, and with "Increased Mitosis" finding reported. The study was not included in the dataset used for training.

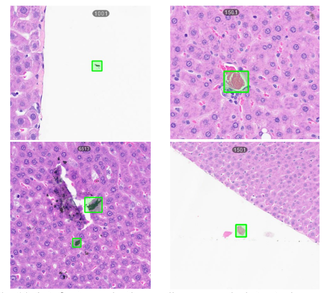

Fig.10 Artefacts and other cells were misdetected as mitosis.

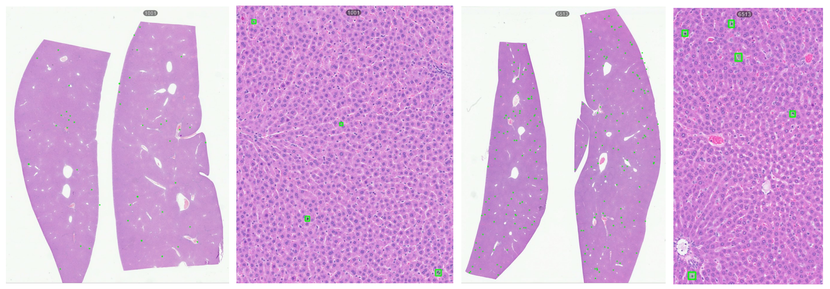

Fig.11 Prediction masks on slide 1001 (two images on left) and 6513 (two images on right) in this study.

Conclusions

⇒ The tile-level study preserved validation results show a good average F1-score of mitotic figure detection can be achieved while some over-detections are observed at the whole slide level. Experiments to reduce the false detections using thresholding of the prediction probability are planned in the future work.

⇒ Further investigations are required, including future model iterations, based on slide and study level evaluation, in order to get a more generalisable view on model performance.

⇒ Accurate detection of mitotic figures is the first step to get the computer-assisted mitotic count and other quantitative measures in order to help toxicologic pathologists rapidly focus on areas of interest within tissues.

Future Directions

⇒ Perform evaluation on data from different organisations.

⇒ Mitotic figure detection in other organs and other species.

⇒ Applying the same approach to difference cellular lesions, e.g. single cell necrosis.

References

[1] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," Advances in Neural Information Processing Systems, 2015, Accessed: Aug. 16, 2022. [Online]. Available: http://image-net.org/challenges/LSVRC/2015/results.

[2] T.-Y. Lin, P. Goyal, R. Girshick, K. He, and P. Dollár, "Focal Loss for Dense Object Detection," 2017, doi: 10.1109/ICCV.2017.324.

[3] M. Tan, R. Pang, and Q. v. Le, "EfficientDet: Scalable and efficient object detection," Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 10778–10787, 2020, doi: 10.1109/CVPR42600.2020.01079.

Contact Information

Dr. Nijun Wei ✉ nijun.wei@deciphex.com

https://deciphex.com

https://twitter.com/deciphex

Acknowledgements

We thank Dr. Maurice Cary for pathology insights.