A deep learning workflow for quantification of Micronuclei in DNA damage studies in cultured cancer cell lines: a proof of principle investigation

Anand Panchbhai* [1], Munuse Ceyda Ishanzadeh[2, £], Smarana Pankanti [1], Ahmed Sidali [2], Nadeeen Solaiman [2], Radhakrishnan Kanagaraj[2,3], John J Murphy [2], Kalpana Surendranath [2]*

[1] Logy.AI, Machine Learning Research Division, Indian Institute of Technology Bhilai, Raipur India.
[2] Genome engineering laboratory, University of Westminster, London W1W 6UW, United Kingdom.
[3] School of Life Sciences, University of Bedfordshire, Park Square, Luton LU1 3JU, United Kingdom
[£] Present address: Nuffield Department of Women's & Reproductive Health, Institute of Reproductive Sciences (IRS) Building 8000, Alec Issigonis Way, Oxford Business Park North, Oxford OX4 2HW
* Correspondence: k.surendranath1@westminster.ac.uk; anandp@logy.ai

Simple Summary: This study aims to test a suitable deep learning method for micronucleus detection in images acquired for cytokinesis block micronucleus assay. This study has reached a mean average precision of >90%.

Abstract: The cytokinesis block micronucleus assay is widely used for measuring/scoring/counting micronuclei, a marker of genome instability in cultured and primary cells. Though a gold standard method, this is a laborious and time-consuming process with person-to-person variation observed in quantification of micronuclei. We report in this study the utilisation of a new deep learning workflow for detection of micronuclei in DAPI stained nuclear images. The proposed deep learning framework achieved an average precision of >90% in detection of micronuclei. This proof of principle investigation in a DNA damage studies laboratory supports the idea of deploying AI powered tools in a cost-effective manner for repetitive and laborious tasks with relevant computational expertise. These systems will also help improving the quality of data and wellbeing of researchers.

Keywords: MN, genome instability, DAPI, cancer diagnostics, MN detection, Artificial Intelligence

1. Introduction

DNA damage induced genomic instability is a defining characteristic of human cancers and several other diseases and therefore is at the center of biochemical, molecular and clinical investigations of human diseases [1]. Byproducts of cellular metabolism, replication errors and external sources such as environmental factors and chemotherapeutic drugs promote damage to the double helix. The array of cellular responses activated post DNA damage and outcomes of cell division provide clues about proper cell function and faithful transmission of the genome to daughter cells. Failure to recognise and remove DNA defects is associated with disease pathogenesis particularly in oncogenesis. Micronuclei (MN) are extranuclear bodies originating from chromosome fragments or whole chromosomes that lag in anaphase during cell division. A common effect of mitotic dysfunctions that consequently lead to whole chromosome mis-segregation during cell division is MN [2-5]. MN are not included in the daughter nuclei and serve as valuable biomarkers to evaluate genome integrity. However, experiments require a strenuous process of scoring thousands of binucleated cells and manual counting of MN either directly through the microscope or from the images. This method to study genome instability in a reproducible and reliable manner is time and resource demanding.

Cell detection and tracking has been an age-old problem since the inception of microscopes. Visual inspection by eye has been the common method for detection and counting of cells which is often time consuming and prone to human errors. With the advent of technology and computer vision, automation is being introduced in every aspect of research with applications of Artificial Intelligence (AI) in detection of cells infected in malaria, tuberculosis, and several others [6-14]. However, the adoption of these methodologies in the assessment of DNA damage responses have been comparatively limited.

AI is intelligence demonstrated by a machine unlike the natural intelligence displayed by humans. Deep Learning (DL) is a subset of Machine Learning (ML), which is a subset of AI, DL is inspired by the working of biological neurons. ML has been used in the domain of medical sciences to cater to a number of use cases ranging from detecting cases of COVID-19 from chest X-Rays to diabetic retinopathy [10]. Efficient methods to study DNA damage as well as integration of ML has the potential to yield reliable results and decrease turnaround times in low-resource academic research and hospital settings.

In genotoxic studies, there have been proof of concept studies [15-19, 22-24, 25], which try to capitalize on a number of statistical models to detect nuclei, these studies have been of limited use due to lack of differentiation between cell types for detections [19]. The authors in another study have attempted to use off the shelf ML algorithms and tools like CellProfiler to perform segmentation of cell nuclei in microscopic cell images [20]. An observation made pointed out that the algorithms they tested seems to have failed in detecting MN as they tended to be confused with debris and artifacts which comprise a good portion of objects.

A number of research labs around the world depend on manual counting for analysis of genome instability data. This lack of adoption of AI tools may be attributed to lack of awareness, unavailability of necessary resources and a skepticism about the abilities of AI's in general. Answering these factors forms the motivation of this study, where we try to show the efficacy of a deep learning system and elaborate on the cost benefits of adopting such a system in a laboratory setting. Automation and streamlining the studies of genome instability is what the authors believes holds the potential to accelerate the work being done in research labs by many folds. In this study, deep learning-based object

detection models were used to detect MN in microscopic images. Object detection deals with detection of instances of semantic objects in a given digital image. The novel aspects of this paper include the public release of a large, annotated dataset of CBMN assay images (available upon request to corresponding authors), which the authors believe will help in future studies in this domain and a novel AI powered method for the detection and counting of MN from microscopic CBMN assay images.

2. Materials and Methods

2.1 Preparation of slides

2.1. Cell lines and culture conditions

Osteosarcoma cell line U-2 OS (American Type Culture Collection, HTB-96™) and U-2 OS T-REx used in this study were grown as monolayers in T25 and T75 flasks at 37°C with 5% CO2 in a humidified atmosphere. Cells were cultured in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin (P/S) antibiotic (Tetracycline-free serum). Cells were passaged with trypsin and viability was assessed with 0.4% trypan blue solution and the cell density was determined using Denovix Cell DropTM cell counter.

2.2. Reagent for DNA damage assay

Aphidicolin, the DNA polymerase inhibitor was purchased from Sigma and dissolved in DMSO at a concentration of 3 mM.

2.3. Micronucleus assay

U-2 OS wild type and T-REx variants were grown on Poly-D-Lysine coated coverslips (Corning, #354086) in 12-well plates and individually treated with three concentrations [0.1 μM, 0.2 μM, 0.4 μM] of aphidicolin for 24 hrs. Cell culture medium was supplemented with 2 μg/ml of cytochalasin B (Sigma-Aldrich, #C6762) 16 hrs prior to fixation to block cells at cytokinesis stage of cell division. Cells were fixed with PTEMF buffer [20 mM PIPES (pH 6.8), 10 mM EGTA, 0.2% triton X-100, 1mM MgCl2 and 4% formaldehyde] for 10 minutes at RT and mounted with prolong gold antifade mounting medium containing DAPI. Slides were dried, and edges of coverslips sealed and stored at 4°C until further analysis

2.4 Scoring of MN

Slides were analysed with an upright fluorescent microscope (Olympus BX41) equipped with an Elite Micropix camera (Micropix Ltd). Images were acquired at 100x using Cytocam software Version 2.0 (Micropix). Over 1600 DAPI-stained binucleated cells were analysed and distinct MN in the vicinity of these cells were scored manually. Binucleated cells were scored for MN in the presence and absence of aphidicolin, which is extensively used to assess replication stress in cultured cells.

2.2 Deep learning workflow

The overall methodology can be broken down into multiple steps, (step 1) these steps include the preparation of the slides to be observed under a microscope and capturing the images of the slides at a certain magnification, (step 2) annotating the captured images, (step 3) training the AI models based on reference annotation on a part of the annotated data (Train and Validation Set) followed by (step 4) inference results from the AI models on a completely new set of microscopic images (Test Set). Logy.AI Auto ML platform (LAMP v1.0) was used as part of this study. It is a cloud based end-to-end no-code platform for building ML models. LAMP is a scalable web server built on python-3.7 and NGINX which uses TensorFlow v1.15 to train deep learning models based on the input data and corresponding ground-truth. LAMP can be accessed over any standard browser

like Chrome, Safari or Edge. The intuitive interface of LAMP is specifically designed to eliminate the need of any technical expertise. Post the preparation of data (step 1), it is uploaded to the LAMP where it can then be staged for annotations (step 2). Annotation on images is a way to provide ground-truth to the LAMP system on which the AI models are trained. LAMP provides a simple drag and drop interface to create boxes on the regions of interest, here MN as shown in figure 1c. LAMP supports online learning which can help the annotator by automatically suggesting possible annotations. Once the data is annotated it is split into training, validation and testing sets as per the requirements. LAMP then initiates the ML training process (step 3) and continuously provides learning statistics and on-the-go annotations on the test set at a regular number of epochs. Once the training process is over it produces a readily deployable ML model as a TensorFlow-Serving-Container which is a snapshot of the learned parameters that can then be immediately used for generating (step 4) inferences. These ML models can be easily integrated into the workflow of a data heavy laboratory processing 1000s of images per second.

2.3 Training the AI model

A dataset containing 1642 manually annotated images was used for this study (CBMNA2021, https://bbbc.broadinstitute.org/BBBC039). The images were captured at a magnification of 100 x with an overall magnification of 1000x. The training was carried out by LAMP. The model used for this study is a modified version of Faster-RCNN coupled with coco-resnet-101 [21] for feature extraction. The images annotated contained only one label namely micronucleus.

2.3.1 Method of evaluating the proposed flow:

The authors aim to showcase the inherent ability of the proposed model to learn and identify the morphological features of MN. Instead of just mentioning the results for one such instance with a predefined dataset distribution into train, validate and test sets, a multifold approach was chosen. A n-fold validation study helps us gauge the performance of the framework instead of showing the results for one instance. Thus, generated results should thus be considered to hold a more generic value showcasing the learning ability of this framework in the task of MN detection. For a 3-fold validation (n=3), the annotated dataset is divided into 3 parts randomly of equal size. Two parts are used for training purposes and the results are quoted on tests on the third part. Each part is tested once with the other two being used for training. The results shown here are an average of the 3 cross-validation results.

2.3.2 Evaluation metrics

To evaluate the performance of our model on the data and how well the predictions are made we use a set of standard evaluation metrics. IoU is used to measure the overlap between the predicted object boundary and the ground truth object boundary to decide if the predicted object boundary is true positive or false positive or false negative by setting a threshold. In this experiment the IoU threshold was set to 0.5 empirically.

Precision, recall and mean average precision (mAP) were used as metrics to gauge the results of this study. mAP is a popular metric to gauge the performance of an object detection model. It is calculated by finding the area under the precision-recall curve.

3. Results

Micronuclei, the well-studied biomarker of genome instability can be induced due to dysfunctional genes and genotoxic agents. For imaging of MN, U-2OS cells were cultured with low concentrations of aphidicolin known to induce replication stress, the key driving factor of chromosome instability in dividing cells. The presence of binucleated cells represent cytokinesis block and aphidicolin treated cells displayed increased number of MN as compared to untreated cells (Fig 1). Little to no cellular debris were identified in the areas analysed under microscope. Three or more MN were excluded from scoring to avoid artefacts due to catastrophic cellular events such as necrosis. To further ensure the specificity of our results, only DAPI-stained binucleated with distinct micronuclei in the vicinity were scored (Fig 2)

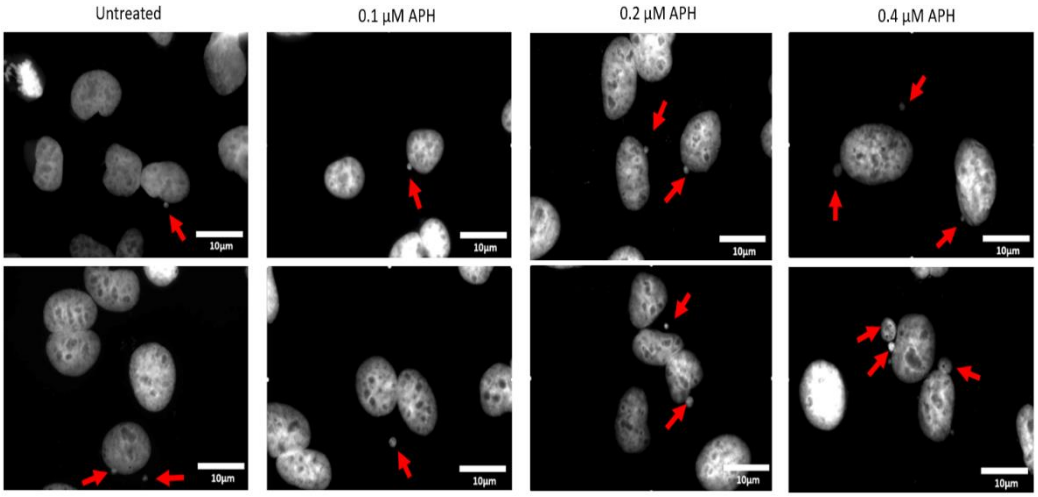

Figure 1. Formation of micronuclei after aphidicolin-induced replication stress. Representative images in grey scale indicate nuclei of human osteosarcoma cells cultured in the presence or absence of 0.1, 0.2 and 0.4 µM DNA polymerase inhibitor aphidicolin. Binucleated cells indicate cytokinesis block by the addition of cytochalasin B, red arrows indicate micronuclei.

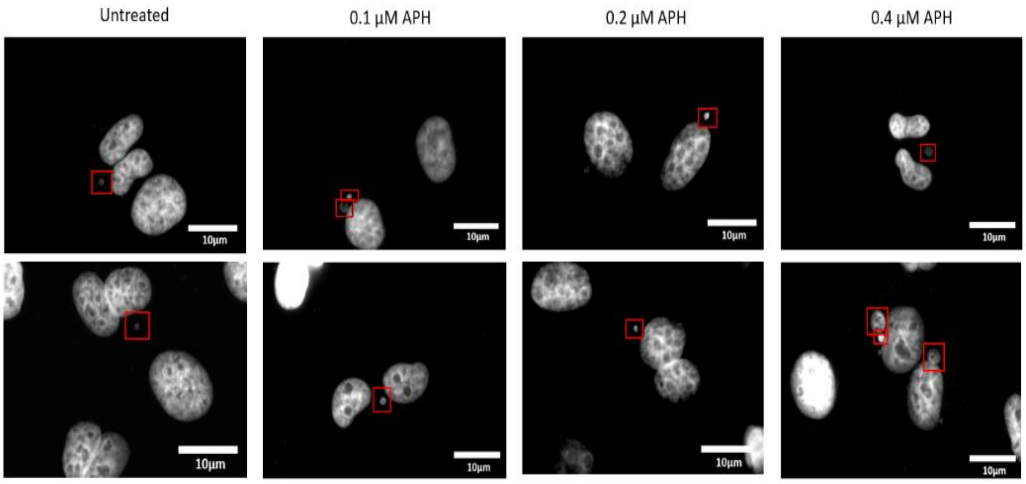

Figure 2. Scoring of micronuclei in a cytokinesis-block micronucleus assay. Cells were cultured on coverslips and treated with 0.1- 0.4 μM aphidicolin for a period of 48 hrs before staining with DNA binding blue, fluorescent dye DAPI. Representative grey scale images indicate cells blocked at anaphase stage of cell cycle with 2 daughter nuclei of similar size in proximity to each other. Red boxes indicate, micronuclei identified by manual/ AI scoring.

Overall, the proposed workflow for deep learning framework proposed here exhibited acceptable performance in the detection of MN. This being an object detection task, mAP was considered as the primary metric. The AI model was able to reach a mAP of 0.91 at a confidence threshold of 0.9, a higher threshold can be correlated with a higher confidence. The results are as follows: sensitivity: 86%, precision: 78%; mAP: 0.93(93%) . From figure 3, it becomes clear that the mAP, sensitivity and specificity remained almost the same or increased as the threshold was increased, which signifies the reliability of the AI decisions.

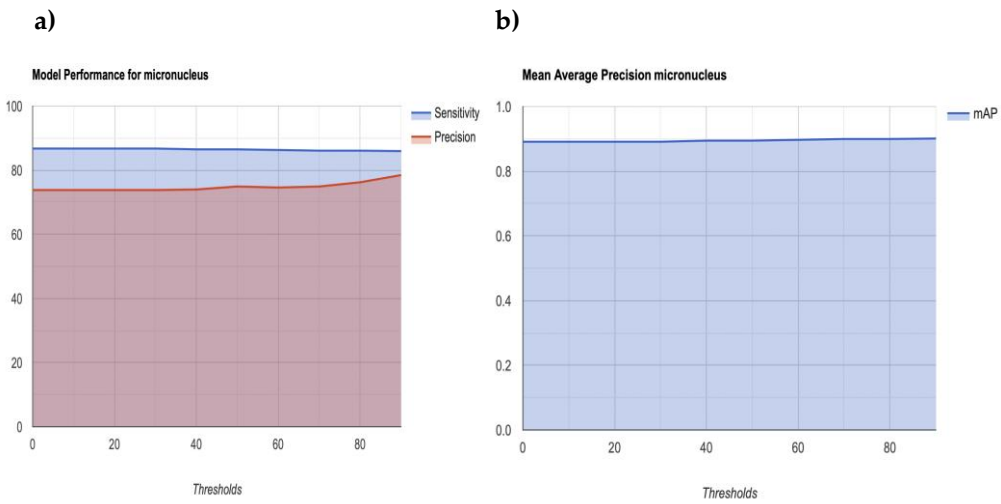

Figure 3. Depiction of 3a) Sensitivity/precision vs confidence threshold of the Logy AI tool. 3b) mAP vs confidence threshold

.

In an independent experiment, annotators were requested to annotate sets of images with the help of the trained AI model as a suggestion tool. Overall, the AI-human duo annotated more MN than their manual counterpart. On a set of 200 images with manual annotation, the human annotator detected 171 MN whereas the AI-human duo was able to detect more MN. On changing the quality of the images in terms of clarity and resolution, no significant drop in performance was observed. It is tempting to suggest that the AI tool may be more sensitive and shows potential to complement manual counting with respect to speed and accuracy. However, further studies are required to establish the efficacy of such AI tools in counting MN, and to ensure that the AI tool is performing effectively and scoring MN accurately.

4. Discussion

In vitro and *in vivo* cytogenetic tests are used in epidemiological studies and in the investigation of biomarkers of human diseases. The widely used micronucleus test method requires manual scoring of MN. Previous studies [19] have mentioned the difficulty faced by researchers in the detection of MN in fluorescent CBMN assay images. These results were broadly driven by the difficulty of differentiating MN from debris, other artifacts, and human error. The proposed method in this study was able to detect

MN with a high efficacy. The results presented here show significant potential to decrease the amount of manual work currently undertaken by researchers working in laboratories involved in delineating the link between genomic instability and molecular basis of diseases. Such methodologies when integrated will play an essential role in standardizing and speeding up the process of detection and counting of MN.

The AI model was trained with relatively small data. Improvements have been observed in previous AI studies with enhanced data quality and size. The proposed AI model was also tested on a similar public dataset BBBC039[1], it was found that the model was able to detect MN in those images too. As the dataset was not originally annotated for MN the authors could not quote a value on the model's accuracy.

As the pace of sequencing genomes increases, the quantities of raw data produced by next generation sequencing methods has increased massively [18]. Over the past 4 decades many software tools - bioinformatics tools - have been developed that enable the analysis and investigation of sequenced data. ML techniques could be utilised to further enhance existing bioinformatics tools and to develop entirely new classes of bioinformatics tools that leverage auto-ml platforms such as Logy.AI. In theory, any repetitive activity which involves a human having to analyse images and count specific shapes - or more generally *patterns* (which could include textual genetic sequences) - could benefit from being automated using an ML platform like Logy.AI. In fact, this may indeed become a necessity as data volumes continue to increase and human resources remain in short supply. By using no code ML training, the writing of custom application specific software is eliminated. The advantages of this include reduced costs, reduced IT infrastructure, and reduced IT training costs as non-programmer bio-scientists are able to train AI's without learning to code.

There are several issues when humans manually count and annotate micronuclei in fluorescent CBMN assay images. There is also the additional problem of researcher *bias* - where one researcher would count a specific artifact as a positive and another would count it as a negative. This would lead to inconsistent data, as different results would be recorded depending on who was doing the counting during a particular time period. The bias present in a researcher can also be passed on to the ML (known as *machine bias*) [19]. Thus, one aspect to consider is that during training, more than one human could assist in training the AI, providing a 'consensus' of what is positive and negative. It is important to prevent human biases to enter the ML process as that would then distort results. In addition, after the AI has been trained, it is still important to randomly sample 'counted' images to determine if they have been correctly counted, and then to 'tweak' the AI if required - essentially providing it with more learning and thus more future accuracy. Regular testing of the workflow is essential, this ensures that the system is performing to acceptable standards and thresholds. For example, if a different resolution of image is presented, or different colour tones are used, this could cause the AI to misinterpret the image - and thus miscount. Hence retesting and calibration is important - assumptions should be rechecked at regular intervals.

AI model training by LAMP was performed on an Intel®Xeon @ 2.70GHz machine (aws p2.xlarge [26]) with an NVIDIA®Tesla K80 [27] with 12 GB of memory running Ubuntu 18.04 [28] optimized for deep-learning. It took 4 hours to train each model. The total computational cost of training the model was ~10USD*[2]. The resultant model can be run on a personal computer and the running cost post the creation of the AI model is thus very minimal. This demonstrates that platforms like LAMP are very cost effective in building

[1] https://bbbc.broadinstitute.org/BBBC039

[2] The price mentioned here only comprises the computational cost of AWS as on 11th July 2021.

such AI models on demand for labs of all sizes and domains. Tools like LAMP make this process much more accessible and affordable than before.

5. Conclusions

In the current work, we demonstrate how advancements in the field of computer science, especially in AI [29-31], can be leveraged in genome instability studies to build viable tools. The work is presented as fully working in terms of functional proof of concept and is intended to provide motivation to encourage further work in building such AI powered tools. The AI models proposed here were able to achieve convincing results for the detection of MN in CBMN assay images with a mean average precision of 90%. There is a potential use of this as a suggestive tool to speed up the counting of MN which is predominantly carried out manually today.

ML techniques could be utilised to further enhance existing bioinformatics tools and to develop entirely new classes of bioinformatics tools that leverage AI as well as research in biological sciences. In theory, any repetitive activity which involves a human having to analyse images and count specific shapes- or more generally *patterns* not limited to images could benefit from being automated using AI. The sheer speed of building and deploying such cost-effective systems will play a pivotal role in revolutionizing modern research. This will aid the professionals to work on more pressing matters and let the machine do the mundane tasks both more accurately and reliably.

Author Contributions: Conceptualization, Kalpana Surendranath (K.S) and Anand Panchbhai (A.P); writing-original draft preparation, K.S., A.P., Munuse Ceyda Ishanzadeh (M.C.I);, DNA damage studies, A.S.,N.S.S.,R.K.; Testing the model and visualization M.C.I and A.P. supervision, J.J.M. and K.S. All authors have read and agreed to the published version of the manuscript.

Funding: Work of the Genome Engineering laboratory is supported by the Quintin Hogg Trust and the external donor fund of Raj Sitlani acquired through the development team (Simay Sali Sevik and Jordan Scamell) of the University of Westminster.

Data Availability Statement: Two datasets were analysed as part of this study CBMNA2021 was specifically curated for the purpose of this study and is available from the corresponding authors on request. The second dataset used in this study was released as part of another study and can be found at https://bbbc.broadinstitute.org/BBBC039.

Acknowledgments: We gratefully acknowledge the collaborative efforts of members of Genome Engineering laboratory in the establishment of CRISPR knock out models. We thank Khalid Akram for extending his support in critical reading of the manuscript.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, or in the decision to publish the results.

References

1. Hoeijmakers JH. DNA damage, aging, and cancer. N Engl J Med. 2009 Oct 8;361(15):1475-85. doi: 10.1056/NEJMra0804615. Erratum in: N Engl J Med. 2009 Nov 5;361(19):1914. PMID: 19812404
2. Bonassi S et al. The HUman MicroNucleus project on eXfoLiated buccal cells (HUMN(XL)): the role of life-style, host factors, occupational exposures, health status, and assay protocol. Mutat Res. 2011 Nov-Dec;728(3):88-97. doi: 10.1016/j.mrrev.2011.06.005. Epub 2011 Jul 8. PMID: 21763453.

3. Capper D et al. DNA methylation-based classification of central nervous system tumours. Nature. 2018 Mar 22;555(7697):469-474. doi: 10.1038/nature26000. Epub 2018 Mar 14. PMID: 29539639; PMCID: PMC6093218.
4. Kwon, M., Leibowitz, M.L. & Lee, JH. Small but mighty: the causes and consequences of micronucleus rupture. Exp Mol Med 52, 1777–1786 (2020). https://doi.org/10.1038/s12276-020-00529-z
5. Samwer M, Schneider MWG, Hoefler R, Schmalhorst PS, Jude JG, Zuber J, Gerlich DW. DNA Cross-Bridging Shapes a Single Nucleus from a Set of Mitotic Chromosomes. Cell. 2017 Aug 24;170(5):956-972.e23. doi: 10.1016/j.cell.2017.07.038. PMID: 28841419; PMCID: PMC5638020.
6. Delahunt, C. B.; Jaiswal, M. S.; Horning, M. P.; Janko, S.; Thompson, C. M.; Kulhare, S.; Hu, L.; Ostbye, T.; Yun, G.; Gebrehiwot, R.; Wilson, B. K.; Long, E.; Proux, S.; Gamboa, D.; Chiodini, P.; Carter, J.; Dhorda, M.; Isaboke, D.; Ogutu, B.; Oyibo, W.; Villasis, E.; Tun, K. M.; Bachman, C.; Bell, D.; Mehanian, C. In *Fully-automated patient-level malaria assessment on field-prepared thin blood film microscopy images*, 2019 IEEE Global Humanitarian Technology Conference (GHTC), 17-20 Oct. **2019**; pp 1-8.
7. Manku, R.; Sharma, A.; Panchbhai, A., *Malaria Detection and Classification*. **2020.**
8. Lakhani, P.; Sundaram, B., Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. *Radiology* **2017,** *284* (2), 574-582.
9. Alzubaidi, L.; Fadhel, M. A.; Al-Shamma, O.; Zhang, J.; Duan, Y., Deep Learning Models for Classification of Red Blood Cells in Microscopy Images to Aid in Sickle Cell Anemia Diagnosis. *Electronics* **2020,** *9* (3), 427.
10. Karbhari, Y.; Basu, A.; Geem, Z. W.; Han, G.-T.; Sarkar, R., Generation of Synthetic Chest X-ray Images and Detection of COVID-19: A Deep Learning Based Approach. *Diagnostics* **2021,** *11* (5), 895.
11. Krause, J.; Gulshan, V.; Rahimy, E.; Karth, P.; Widner, K.; Corrado, G. S.; Peng, L.; Webster, D. R., Grader Variability and the Importance of Reference Standards for Evaluating Machine Learning Models for Diabetic Retinopathy. *Ophthalmology* **2018,** *125* (8), 1264-1272.
12. Lucius, M.; De All, J.; De All, J. A.; Belvisi, M.; Radizza, L.; Lanfranconi, M.; Lorenzatti, V.; Galmarini, C. M., Deep Neural Frameworks Improve the Accuracy of General Practitioners in the Classification of Pigmented Skin Lesions. *Diagnostics* **2020,** *10* (11), 969.
13. Jacobs, C.; van Ginneken, B., Google's lung cancer AI: a promising tool that needs further validation. *Nature Reviews Clinical Oncology* **2019,** *16* (9), 532-533.
14. Tschandl, P.; Rinner, C.; Apalla, Z.; Argenziano, G.; Codella, N.; Halpern, A.; Janda, M.; Lallas, A.; Longo, C.; Malvehy, J.; Paoli, J.; Puig, S.; Rosendahl, C.; Soyer, H. P.; Zalaudek, I.; Kittler, H., Human–computer collaboration for skin cancer recognition. *Nature Medicine* **2020,** *26* (8), 1229-1234.
15. Dias, R.; Torkamani, A., Artificial intelligence in clinical and genomic diagnostics. *Genome Medicine* **2019,** *11* (1), 70.
16. Rodrigues, M. A.; Probst, C. E.; Zayats, A.; Davidson, B.; Riedel, M.; Li, Y.; Venkatachalam, V., The in vitro micronucleus assay using imaging flow cytometry and deep learning. *npj Systems Biology and Applications* **2021,** *7* (1), 20.
17. Wills, J. W.; Verma, J. R.; Rees, B. J.; Harte, D. S. G.; Haxhiraj, Q.; Barnes, C. M.; Barnes, R.; Rodrigues, M. A.; Doan, M.; Filby, A.; Hewitt, R. E.; Thornton, C. A.; Cronin, J. G.; Kenny, J. D.; Buckley, R.; Lynch, A. M.; Carpenter, A. E.; Summers, H. D.; Johnson, G.; Rees, P., Inter-laboratory automation of the in vitro micronucleus assay using imaging flow cytometry and deep learning. *bioRxiv* **2021**, 2021.05.05.442619.
18. Alafif, T.; Qari, S.; Albassam, A.; Al Refaei, A., *Deep Transfer Learning for Nucleus and Micronucleus Recognition*. **2020**.
19. Caicedo, J. C.; Roth, J.; Goodman, A.; Becker, T.; Karhohs, K. W.; Broisin, M.; Molnar, C.; McQuin, C.; Singh, S.; Theis, F. J.; Carpenter, A. E., Evaluation of Deep Learning Strategies for Nucleus Segmentation in Fluorescence Images. *Cytometry Part A* **2019,** *95* (9), 952-965.

20. Carpenter, A. E.; Jones, T. R.; Lamprecht, M. R.; Clarke, C.; Kang, I. H.; Friman, O.; Guertin, D. A.; Chang, J. H.; Lindquist, R. A.; Moffat, J.; Golland, P.; Sabatini, D. M., CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biology* **2006,** *7* (10), R100.
21. Ren, S.; He, K.; Girshick, R.; Sun, J., Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **2017,** *39* (6), 1137-1149.
22. Reuter, J. A.; Damek V. S.; Michael P. S., "High-throughput sequencing technologies." *Molecular cell* 58, no. 4 **2015**, 586-597.
23. Tarca, A. L.; Vincent J. C; Xue-wen, C; Roberto, R; Sorin, D., "Machine learning and its applications to biology." *PLoS computational biology* 3, no. 6 **2007,** e116.
24. Su HH, Pan HW, Lu CP, Chuang JJ, Yang T. Automatic Detection Method for Cancer Cell Nucleus Image Based on Deep-Learning Analysis and Color Layer Signature Analysis Algorithm. Sensors (Basel). 2020 Aug 7;20(16):4409. doi: 10.3390/s20164409. PMID: 32784663; PMCID: PMC7472205.
25. Bahreyni Toossi MT, Azimian H, Sarrafzadeh O, Mohebbi S, Soleymanifard S. Automatic detection of MN by cell microscopic image processing. Mutat Res. 2017 Dec;806:9-18. doi: 10.1016/j.mrfmmm.2017.07.012. Epub 2017 Aug 12. PMID: 28934716.
26. https://docs.aws.amazon.com/AWSEC2/latest/APIReference/Welcome.html (accessed Jul 26, 2021).
27. Nickolls, J. Scalable Parallel Programming with CUDA Introduction. In 2008 IEEE Hot Chips 20 Symposium (HCS); IEEE, 2008; pp 1–9.
28. Sobell, M. G. A practical guide to linux: Commands, editors, and shell programming https://www.usenix.org/legacy/publications/login/2010-04/openpdfs/bookreviews1004.pdf (accessed Jul 26, 2021).
29. Ting, D. S. W.; Liu, Y.; Burlina, P.; Xu, X.; Bressler, N. M.; Wong, T. Y. AI for Medical Imaging Goes Deep. Nat. Med. 2018, 24 (5), 539–540.
30. Fazal, M. I.; Patel, M. E.; Tye, J.; Gupta, Y. The Past, Present and Future Role of Artificial Intelligence in Imaging. Eur. J. Radiol. 2018, 105, 246–250.
31. Kundu, S.; Elhalawani, H.; Gichoya, J. W.; Kahn, C. E., Jr. How Might AI and Chest Imaging Help Unravel COVID-19's Mysteries? Radiol Artif Intell 2020, 2 (3), e200053.