JCB

Object Detection Networks and Augmented Reality for Cellular Detection in Fluorescence Microscopy.

Dominic Waithe, Jill Brown, Katharina Reglinski, Isabel Diez-Sevilla, David Roberts, and Christian Eggeling

Corresponding Author(s): Dominic Waithe, Weatherall Institute of Molecular Medicine

Review Timeline:

Submission Date:	2019-03-28
Editorial Decision:	2019-05-14
Revision Received:	2020-05-05
Editorial Decision:	2020-06-09
Revision Received:	2020-06-24
Editorial Decision:	2020-07-07
Revision Received:	2020-07-12

Monitoring Editor: Joerg Bewersdorf

Scientific Editor: Tim Spencer

Transaction Report:

(Note: With the exception of the correction of typographical or spelling errors that could be a source of ambiguity, letters and reports are not edited. The original formatting of letters and referee reports may not be reflected in this compilation.)

DOI: https://doi.org/10.1083/jcb.201903166

May 14, 2019

Re: JCB manuscript #201903166

Dr. Dominic Waithe
University of Oxford
John Radcliffe Hospital
Oxford, Oxfordshire OX3 9DS
United Kingdom

Dear Dr. Waithe,

Thank you for submitting your manuscript entitled "Object Detection Networks and Augmented Reality for Cellular Detection in Fluorescence Microscopy". We apologize for the delay in providing you with a decision. In any case, the manuscript has been evaluated by expert reviewers, whose reports are appended below. Unfortunately, after an assessment of the reviewer feedback, our editorial decision is against publication in JCB.

You will see that, although the two reviewers disagree to some extent on the suitability of this Tools manuscript for the JCB audience, they each raise a number of substantive concerns that preclude publication of the paper, at least in its current form.

Although your manuscript is intriguing, we feel that the points raised by the reviewers are more substantial than can be addressed in a typical revision period. Therefore, if you wish to expedite publication of the current data, it may be best to pursue publication at another journal.

However, we would like to stress that we agree with reviewer #1 (who, as you'll notice, signed his review) and reiterate that JCB is interested in this kind of research and we continue to feel that such studies would be of use to our audience and the cell biology community. Therefore, we would be open to an appeal of our decision and the resubmission to JCB of a significantly revised and extended manuscript that fully addresses the reviewers' concerns and is subject to further peer-review.
Please note, though, that if you choose this course, priority and novelty would be reassessed at resubmission.

Regardless of how you choose to proceed, we hope that the comments below will prove constructive as your work progresses.

You can contact the journal office with any questions, cellbio@rockefeller.edu or call (212) 327-8588.

Thank you for thinking of JCB as an appropriate place to publish your work.

Sincerely,

Joerg Bewersdorf, PhD
Monitoring Editor

JCB

Tim Spencer, PhD
Deputy Editor
Journal of Cell Biology

Reviewer #1 (Comments to the Authors (Required)):

This manuscript describes the characterisation and development of a series of tools for automatically identifying cells in a digital microscope field. The authors methods are focussed on automatic identification of individual cells with a bounding box (a quadrilateral that bounds an object of interest like an individual cell) and using this method to guide automatic data collection of cells in a field. The authors report their development but also include a detailed comparison of recent deep learning-based tools for object identification. Finally the authors also implement a Augmented Reality System for a fluorescence microscope data acquisition system. This apparently is to allow interactive viewing of the analytic results and the data collected in an imaging system.

As the previous paragraph indicates, this manuscript is doing several distinct things that the authors try to sew together. Overall the manuscript should be valuable to the JCB readership. In its current form, there are parts that are excellent, others that are weaker and others that need to be removed entirely. I strongly suggest the authors consider a major rewrite of the paper.

Major points
1. The most valuable part of this paper is a detailed comparison of the performance of Faster-RCNN and YOLO. The authors present a detailed performance characterisation of these tools on several different datasets (however see below). They also compare these to several other existing tools Retina.Net, SSD and others. The first set of figures deliver this characterisation and these data will be of interest to the community. I'll note that this part of the paper is more of a CS paper on cell images-which is OK. JCB should embrace this kind of work!
2. The datasets the authors use are detailed in the Materials and Methods. Some are generated by them, while others are from public resources. The public datasets are not properly cited and don't include unique identifiers or other information that would allow an interested reader to easily source these data. The datasets the authors have generated have been deposited on Zenodo and are accessible through a DOI. These datasets do provide examples of many different types of staining patterns and thus do challenge the object identification tools in different ways. However, the imaging modalities used are quite narrow and limited. Mostly these are thin cells growing in culture on two-dimensional cover slips imaged with a high NA lens in a WF microscope. This is a good start but really focusses on experiments where sparsely plated cells are analysed. This limited set of data characteristics in the test data limit the usefulness of the paper.
3. The authors present a quite detailed description of the implementation of their tools, in particular integration with Micromanager. It's great this is included. The detail is absolutely essential and should be retained in the published manuscript.
4. The major issue with the implementation as presented is that it appears to identify all cells in a field,(except the ones that are cut off on an edge, which is useful). Thus the authors do not show the tool being used to, for example, identify cells with a certain appearance, cell cycle stage, staining density, or other phenotypic characteristics that would be used to target a subselection of cells in a field or on a coverslip. This is in stark contrast to other previous work in this field (e.g.,

https:doi.org/10.1038/nbt.3146; https://doi.org/10.1038/nmeth.1558 and many others). The authors should openly clarify the limitations of their method and whether it could be used for automatically identifying cells with a specific set of characteristics. If they do this, the authors will substantially increase the value of their tool for the JCB readership.
5. The authors could consider the value of testing the utility of models they calculate on one imaging system for data on another. For example, does the model calculated on the Olympus microscope work on the Zeiss microscope. Information about this performance would be good for the JBC readership.
6. All of the implementations the authors describe depend on JPG images. The exact level of compression used and its effect on the performance of the algorithms should be stated. The authors might consider an implementation that directly reads image data from the original data-files as they come off the microscopes. Again this is an implementation issue that would be very useful for the JCB readership.

Jason Swedlow, Univ of Dundee.

Reviewer #2 (Comments to the Authors (Required)):

This paper introduces a pipeline to autmatically detect and display cells in a fluorescence microscopy setup. For that, the authors compare two different state-of-the-art 2D object detection networks on six different cell datasets in 3D, for which the authors propose to track 2D detections using the SORT algorithm. The authors further introduce AMCA ("Automated Microscopy Control Algorithm"), a method for joint imaging and cell detection.

Unfortunately, this paper has several shortcomings due to which acceptance in the Journal of Cell Biology can not be recommended:

1. Limited technical novelty

The presented work uses off-the-shelf algorithm for object detection and tracking. The technical contribution is very small, and indeed limited to the writing of glue-code that brings existing components together.

2. Insufficient experimental evaluation

For a paper lacking a technical innovation, a thorough experimental evaluation is imperative. However, the authors only compare the two investigated object detection networks (Faster-RCNN and YOLO) under questionable conditions (e.g., with and without horizontal flip augmentations). No comparison is performed to competing methods for 3D cell detection methods. It remains unclear how the findings reported in this paper help fellow scientists to decide which method to use for 3D cell detection in fluorescence microscopy.

3. Unsupported claims

The authors contrast their work to [3] and point out that they "lack flexibility and a skilled analyst is often required to tweak the parameters", suggesting that the proposed method mitigates this problem. However, this is not shown. In particular, it is unclear how to find hyperparamters of the proposed method. Furthermore, the proposed method requires a substantial amount of annotated

training data. It is not shown whether and how well the proposed method generalizes between different cell types and imaging methodologies.

The "Automated Microscopy Control Algorithm" does not seem to be primarily a "control algorithm". In fact, the "passive" mode explicitly images the whole volume first, and then performs an offline analysis.

MRC WIMM Centre for Computational Biology
University of Oxford
John Radcliffe Hospital
Headington Oxford OX3 9DS

Dear Dr Tim Spencer and Prof. Joerg Bewersdorf,

I request that you reconsider for publication in JCB, our paper:

Object Detection Networks and Augmented Reality for Cellular Detection in Fluorescence Microscopy.

Previously we submitted the paper to you and, despite rejecting the paper at the time, you encouraged us to appeal the decision and resubmit a much revised version of the paper. We have taken a year to expand and rework the article and believe we have made a much improved article and have in the process fully addressed the original points made by the reviewers. In the meantime, we have not submitted the article elsewhere. We believe that the optimum readership for the paper is that of microscopy users whom perform research using cellular imaging and we believe JCB to be a perfect journal for this readership.

The proposed technique has the potential to empower users to perform automated imaging with conventional microscopy equipment and without needing high-level technical knowledge due to the use of machine learning algorithms that learn from example. This is novel work as it makes this type of approach much more readily available and accessible. This system allows readers the ability to increase the number of samples they perform during an experiment, increasing the statistical power and reproducibility of their research. Furthermore we combine our tool with augmented reality for microscopy (a first for fluorescence microscopy), creating a powerful and interactive tool using affordable and readily available computational and optical elements.

A pre-print of the paper has proven popular, with an altmetric score of 46 (https://www.biorxiv.org/content/10.1101/544833v2) and 1243 downloads of the paper.

Please find below this text our specific rebuttal to the original comments. We hope the reviewers and editors will enjoy this much improved work and we look forward to hearing from you soon.

Yours Sincerely,

Dr. Dominic Waithe

MRC WIMM Centre for Computational Biology. Weatherall Institute of Molecular Medicine. University of Oxford, Oxford, OX3 9DS

phone: +44 (0) 1865 222517 mobile: +44 (0) 7450 216944 email: dominic.waithe@imm.ox.ac.uk

---

The MRC Weatherall Institute of Molecular Medicine is a strategic alliance between the Medical Research Council and the University of Oxford

Original Review (black), rebuttal (blue)

Reviewer #1 (Comments to the Authors (Required)):

This manuscript describes the characterisation and development of a series of tools for automatically identifying cells in a digital microscope field. The authors methods are focussed on automatic identification of individual cells with a bounding box (a quadrilateral that bounds an object of interest like an individual cell) and using this method to guide automatic data collection of cells in a field. The authors report their development but also include a detailed comparison of recent deep learning-based tools for object identification. Finally the authors also implement a Augmented Reality System for a fluorescence microscope data acquisition system. This apparently is to allow interactive viewing of the analytic results and the data collected in an imaging system.

As the previous paragraph indicates, this manuscript is doing several distinct things that the authors try to sew together. Overall the manuscript should be valuable to the JCB readership. In its current form, there are parts that are excellent, others that are weaker and others that need to be removed entirely. I strongly suggest the authors consider a major rewrite of the paper.

Major points
1. The most valuable part of this paper is a detailed comparison of the performance of Faster-RCNN and YOLO. The authors present a detailed performance characterisation of these tools on several different datasets (however see below). They also compare these to several other existing tools Retina.Net, SSD and others. The first set of figures deliver this characterisation and these data will be of interest to the community. I'll note that this part of the paper is more of a CS paper on cell images-which is OK. JCB should embrace this kind of work!
DW: Thank you for the kind comments. We have now included RetinaNet in this detailed comparison in addition to YOLOv3 which was released around the time of the submission of this paper. My comparing these different algorithms on the data we get a clear idea of the strengths and weaknesses of each one.

2. The datasets the authors use are detailed in the Materials and Methods. Some are generated by them, while others are from public resources. The public datasets are not properly cited and don't include unique identifiers or other information that would allow an interested reader to easily source these data. The datasets the authors have generated have been deposited on Zenodo and are accessible through a DOI. These datasets do provide examples of many different types of staining patterns and thus do challenge the object identification tools in different ways. However, the imaging modalities used are quite narrow and limited. Mostly these are thin cells growing in culture on two-dimensional cover slips imaged with a high NA lens in a WF microscope. This is a good start but really focusses on experiments where sparsely plated cells are analysed. This limited set of data characteristics in the test data limit the usefulness of the paper.

The public dataset in Zenodo includes all of the data used in the study that was generated by ourselves and the images we used from the external data. We had to adapt and crop the annotation of the external data for our study and so we include the data together. The Zenodo DOI does allow the dataset and any revisions to be directly linked. Our approach does mainly focus on cells on coverslips as the reviewer points out, which have limited depth, but this is an extremely common imaging modality and we believe we will have good uptake from the technique. We believe the neuroblastoma and peroxisome datasets are especially challenging and close to confluence but agree z-depth is limited had we have not shown any tissue imaging. We have also now included a total of 10 datasets (four more than before) which represent different stainings, cell types, microscopes and objectives. We see the technique in the future working in close coordination with super-resolution microscopy and dynamic cell imaging techniques. For these and many other techniques Widefield microscopy is often used as a preliminary imaging step before activation of these advanced modalities and is used to find and localise cells. These techniques tend to focus on monolayers because of the difficulties involved in imaging super-resolution at depth with these techniques. We believe therefore that our technique is of value in its own right as well as supporting automated acquisition for these more advanced methods. We have made this justification clearer in the article text.

3. The authors present a quite detailed description of the implementation of their tools, in particular integration with Micromanager. It's great this is included. The detail is absolutely essential and should be retained in the published manuscript.
Thank you for the kind comments. We have and will continue to work on the documentation for our technique on the associated github resource (https://github.com/dwaithe/amca)

4. The major issue with the implementation as presented is that it appears to identify all cells in a field,(except the ones that are cut off on an edge, which is useful). Thus the authors do not show the tool being used to, for example, identify cells with a certain appearance, cell cycle stage, staining density, or other phenotypic characteristics that would be used to target a subselection of cells in a field or on a coverslip. This is in stark contrast to other previous work in this field (e.g., https:doi.org/10.1038/nbt.3146; https://doi.org/10.1038/nmeth.1558 and many others). The authors should openly clarify the limitations of their method and whether it could be used for automatically identifying cells with a specific set of characteristics. If they do this, the authors will substantially increase the value of their tool for the JCB readership.
With this paper we wanted to focus on sampling all the cells, despite appearance on the coverslip. This allows us to generate a representative distribution of the cells present. From this distribution it is then possible to make decisions about which cells to image and the coordinates of the cells can be easily established. This is a good approach as it means that the algorithm will not be potentially used to cherry pick cells from the outset. We believe generating a full distribution of cells is a priority over identifying cells with a particular appearance. However, to reflect the interest of the reviewer and subsequent readers we have in Figure 4K-M now applied the YOLOv2 algorithm to discern multiple phenotypes in two independent datasets. This will exactly allow for the exact kind of phenotypic localisation that the reviewer one would like to see.

5. The authors could consider the value of testing the utility of models they calculate on one imaging system for data on another. For example, does the model calculated on the Olympus microscope work on the Zeiss microscope. Information about this performance would be good for the JBC readership.
We have now compared the performance of images taken on a different microscope and compared the performance at different scales, including this data in Figure 4A-G. The technique works well as long as the cells have a similar visual appearance, which in micsroscopy is very much dependent on the NA of the objective used.

6. All of the implementations the authors describe depend on JPG images. The exact level of compression used and its effect on the performance of the algorithms should be stated. The authors might consider an implementation that directly reads image data from the original data-files as they come off the microscopes. Again this is an implementation issue that would be very useful for the JCB readership.
Although JPG files are used as an intermediate file format for the training (with quality at max), the real-time classification is performed on raw image-data from the microscope and then images are saved as TIFF files (either ImageJ or OME-TIFF). A work-flow has now been produced which involves using OMERO to manage and store the data and annotations and OMERO or Fiji/ImageJ to annotate the data. This means that original image meta-data is preserved and regions of interest are embedded within the subsequent TIFF files. Details of this have been added to the manuscript methods and full details are available through the github resource (https://github.com/dwaithe/amca).

Jason Swedlow, Univ of Dundee.

Reviewer #2 (Comments to the Authors (Required)):

This paper introduces a pipeline to autmatically detect and display cells in a fluorescence microscopy setup. For that, the authors compare two different state-of-the-art 2D object detection networks on six different cell datasets in 3D, for which the authors propose to track 2D detections using the SORT algorithm. The authors further introduce AMCA ("Automated Microscopy Control Algorithm"), a method for joint imaging and cell detection.

Unfortunately, this paper has several shortcomings due to which acceptance in the Journal of Cell Biology can not be recommended:

1. Limited technical novelty

The presented work uses off-the-shelf algorithm for object detection and tracking. The technical contribution is very small, and indeed limited to the writing of glue-code that brings existing components together.

Thank you for your comments. In the past decade significant advances have occurred in the field of machine learning and computer vision, largely due to advances in deep learning. These advances now make it possible to solve certain problems in microscopy and bioimaging, which were not possible before. Rather than reinventing the wheel, this paper showcases and compares some of the best examples of object detection algorithm and discusses their application in microscopy acquisition. This work is very valuable and is very important given the influx of new techniques. We have now expanded our comparison to four of the leading algorithms. These algorithms function in real-time, in concert with the microscope, and is one of the first examples of this work. We believe this article has real appeal to the audience of JCB as we compare and make available these high-level techniques into a usable and practical form. We have now optimised the whole system to work on inexpensive and readily available computer development boards, which dramatically reduce the cost of the system. Furthermore, we have developed and simplified the system (working with the vendors) so that it will work in the Python coding language and so it no longer requires expensive Labview licenses to run. The paper also asks important questions about the optimisation and training of the algorithms and solves several important challenges involved in this process. Some of these achievements include optimising the tracking of objects in 3-D, optimising the training regime for each algorithm, developing an augmented reality system for the microscope and also an end-to-end guide for how to repeat it. We have reworked the text to better highlight many of the achievements.

2. Insufficient experimental evaluation

For a paper lacking a technical innovation, a thorough experimental evaluation is imperative. However, the authors only compare the two investigated object detection networks (Faster-RCNN and YOLO) under questionable conditions (e.g., with and without horizontal flip augmentations). No comparison is performed to competing methods for 3D cell detection methods. It remains unclear how the findings reported in this paper help fellow scientists to decide which method to use for 3D cell detection in fluorescence microscopy.

In this paper we are showcasing a new way of isolating and identifying cells using the object detection functionality. The algorithms extract volumes rather than pixel level masks. This makes the training very simple (just a bounding box in 2-D) and makes the system very flexible for adjustment to new data. Conventional signal processing algorithms do not predict bounding boxes and are not specifically designed to generalise well for different dataset. Our method is perfect for acquisition of image volumes, which is the focus of this study, for segmentation and subsequent detailed analysis there are many solutions in the signal-processing and machine learning domains, however these are slower and also require often much more data to train. Once appropriate volumes have been acquired on the microscope and the underlying distribution visualised then fine-grain pixel-level segmentation could be applied and or high-resolution imaging subsequently. I have revised the text to make this point more clear. In addition rather than two algorithms we now compare four leading and competing object detection algorithms.

3. Unsupported claims

The authors contrast their work to [3] and point out that they "lack flexibility and a skilled analyst is often required to tweak the parameters", suggesting that the proposed method mitigates this problem. However, this is not shown. In particular, it is unclear how to find hyperparamters of the proposed method. Furthermore, the proposed method requires a substantial amount of annotated training data. It is not shown whether and how well the

proposed method generalizes between different cell types and imaging methodologies.

The "Automated Microscopy Control Algorithm" does not seem to be primarily a "control algorithm". In fact, the "passive" mode explicitly images the whole volume first, and then performs an offline analysis.
The performance of these algorithms has been tuned in terms of how much training is required and also how many images is likely to be required based on the image quality. Although object detection networks have hyper parameters, these do not need to be changed by subsequent users. It is possible that through following a simple protocol that a user could retrain the network for their own use without knowledge of the hyper parameters. In the revised manuscript I have shown more evidence for how well the algorithm generalises for the different datasets. The ‘passive’ mode is for, as the reviewer suggests, just acquiring volumes, but the ‘active’ does the analysis and acquisition in real-time and in concert. We have now removed the ‘passive’ form from the manuscript text to make it clearer to the reader the benefits of the fully-automated technique. In addition we have improved the detailed description of the annotation and training procedure which can be found on the github repository, where it can be constantly improved and refined over-time.

June 9, 2020

Re: JCB manuscript #201903166R-A

Dr. Dominic Waithe
University of Oxford
John Radcliffe Hospital
Oxford, Oxfordshire OX3 9DS
United Kingdom

Dear Dr. Waithe,

Thank you for submitting your revised manuscript entitled "Object Detection Networks and Augmented Reality for Cellular Detection in Fluorescence Microscopy". The manuscript has been seen by the original reviewer #1 whose full comments are appended below. While the reviewer continues to be overall positive about the work in terms of its suitability for JCB, some important issues remain.

As you will see, the reviewer (who has signed his report) feels that the main message of the paper is not sufficiently clear, in part due to the very dense nature of the manuscript and I'm afraid that we agree. The reviewer has provided some suggestions indicating steps that can be taken to clarify the message and conclusions of the paper while also streamlining the presentation. We hope that you will be able to revise the manuscript according to the reviewer's suggestions.

When revising, please note that JCB does not allow any text in the supplementary materials (beyond the supplementary figure/table legends) but you may place detailed descriptions of methodology and/or modeling in the materials and methods section (there is no limit to the word/character count of the materials and methods but try to be as concise as you can without omitting essential information).

Our general policy is that papers are considered through only one revision cycle; however, given that the suggested changes are relatively minor we are open to one additional short round of revision.

Please submit the final revision as quickly as time allows (within one month, preferably, but if lab closures due to COVID-19 prevent you from completing the revisions in this time frame just let us know and we can work out a suitable revision schedule) along with a cover letter that includes a point by point response to the remaining reviewer comments.

Thank you for this interesting contribution to Journal of Cell Biology. You can the journal office with any questions at cellbio@rockefeller.edu.

Sincerely,

Joerg Bewersdorf, PhD
Monitoring Editor
Journal of Cell Biology

Tim Spencer, PhD
Executive Editor
Journal of Cell Biology

Reviewer #1 (Comments to the Authors (Required)):

This revised paper from Dominic Waithe and colleagues explores the application of model-based object identification in digital imaging of cells and potentially of tissues. The author has substantially revised and updated the manuscript and at a high level the additional data and work is quite useful and well received. In its current form the manuscript is extremely dense and forces the reader to work too hard to extract the essential messages. While I am quite positive about the findings reported in this paper, I suggest a full-scale rewrite and a substantial shortening of the Intro and Results sections to focus on the major points of significance in the paper.

There are three major points in this manuscript:

1. A comparison of the performance of various deep learning networks in object identification in microscopy images of single cells. One point that is not emphasised nearly enough is that this tool is running in near real-time, or nearly enough so that it can be integrated into a data acquisition workflow. This is stated but not emphatically enough.
2. The implementation of an augmented reality system for visual feedback of identified cells through a microscope binocular.
3. The use of low-cost computer and GPU for applying these models in near real time.

These are three related topics which are currently embedded in an enormous amount of extraneous information. While I think the tools reported here are quite valuable, the paper needs to be substantially rewritten, and in particular shortened and focussed in order to make these essential points clear to the interested reader. In the following I have highlighted several sections of text that can either be moved to supplementary methods, converted to supplementary tables, stated with reference to previous work or even in some cases simply to be removed altogether. The goal should be to make the most important concepts in the paper much clearer in the Abstract, Intro, Results and in the first paragraph or two of the Discussion.

Major points:

1. The first section of the results highlights the performance characteristics of the different networks compared in this paper. The text on pp. 2 - 3 is way too wordy and uses a substantial amount of undefined jargon in its description. Examples are ("one-shot detector" and several others). This could be condensed to 3 paragraphs max that highlight the aim and results in summary form. Page 13 is especially dense and the characterisation of network performance should be described in summary form. Forcing the reader to negotiate the names of all the different datasets and the performance characteristics is just too hard. Moreover, the section on Data Augmentation compares a technique that is extremely well established in machine learning in the cellular imaging community. Anne Carpenter's lab has published extensively on this and there seems no reason to include this level of detail. Similarly, sections on "training across multiple classes" and performance of "YOLOv2..." recite techniques that have been looked at extensively in

previous papers (w/r/t microscopy, see https://www.biorxiv.org/content/10.1101/085118v1; as this on biorxiv, maybe a short statement is worthy, but not the long description. See also the recent paper from Horvath and Carpenter (https://www.cell.com/cell-systems/pdfExtended/S2405-4712(20)30117-4) and the also CellPose from Stringer et al (https://www.biorxiv.org/content/10.1101/2020.02.02.931238v2)). In short, the use of models pre-trained with ImageNet or the use of combinations of training sets is well-established in microscopy and referencing these papers and saying the work here confirms these (as of yet to my knowledge not peer-reviewed preprints) would be sufficient. The goal should be to get the first section of the results down to 2 paragraphs max that highlight the choice of network for object identification. Nothing more is needed.
2. This revised manuscript now has quite good examples of the application of the methods in Fig.4 and the author is to be commended on this additional information. Nonetheless the section on the "automated screening using autodetection" is again very dense and could be shortened to focus on the essential points. The development of the AMCA method seems extremely powerful and to my knowledge is a completely novel implementation. The author should highlight this advance much more clearly by emphasising the implementation in a small form factor low-cost platform. I imagine this advance will be of extreme interest to many members of the development and commercial imaging communities.
3. The author has added substantial characterisation of DAPI stain cells and reports times for acquiring data that appear somewhat slow. While the author correctly states that his methods produce significant acceleration because of the ability to focus on exact cell positions it would be useful to know whether there are potential accelerations possible in future developments.

There is a lot to like in this manuscript and the author should be commended for all the hard work. In its current incarnation the manuscript is too dense and needs a significant revision to focus on the extremely important points that represent the novel impactful components of the author's work.

Some comments on the rebuttal:

1. "My [sic] comparing these different algorithms on the data we get a clear idea of the strengths and weaknesses of each one."
Reply:
a. Proofread your text and do not allow typos in your responses.
b. Yes, but the text is too dense to access the conclusion of these comparisons!

2. "We have now compared the performance of images taken on a different microscope and compared the performance at different scales, including this data in Figure 4A-G. The technique works well as long as the cells have a similar visual appearance, which in micsroscopy [sic] is very much dependent on the NA of the objective used.
a. Proofread your text and do not allow typos in your responses.
b. This comparison is lost in the way too dense paragraph on pp 15-16. It is critical this is brought out in more detail.

Jason Swedlow
Univ of Dundee

MRC WIMM Centre for Computational Biology
University of Oxford
John Radcliffe Hospital
Headington Oxford OX3 9DS

Dear Dr Tim Spencer and Prof. Joerg Bewersdorf,

Thank you for considering our paper entitled:

Object Detection Networks and Augmented Reality for Cellular Detection in Fluorescence Microscopy.

Thank you for the thorough review we received on 12th June. We have taken the points onboard and have revised the manuscript accordingly.

Please find below this letter our specific rebuttal to the comments. We hope the reviewers and editors will enjoy this revised work and that we have fulfilled the necessary changes required to publish in JCB.

We look forward to hearing from you soon.

Yours Sincerely,

Dr. Dominic Waithe

MRC WIMM Centre for Computational Biology. Weatherall Institute of Molecular Medicine. University of Oxford, Oxford, OX3 9DS

phone: +44 (0) 1865 222517 mobile: +44 (0) 7450 216944 email: dominic.waithe@imm.ox.ac.uk

Original Review (black), rebuttal (blue)

Reviewer #1 (Comments to the Authors (Required)):

This revised paper from Dominic Waithe and colleagues explores the application of model-based object identification in digital imaging of cells and potentially of tissues. The author has substantially revised and updated the manuscript and at a high level the additional data and work is quite useful and well received. In its current form the manuscript is extremely dense and forces the reader to work too hard to extract the essential messages. While I am quite positive about the findings reported in this paper, I suggest a full-scale rewrite and a substantial shortening of the Intro and Results sections to focus on the major points of significance in the paper.

There are three major points in this manuscript:

1. A comparison of the performance of various deep learning networks in object identification in microscopy images of single cells. One point that is not emphasised nearly enough is that this tool is running in near real-time, or nearly enough so that it can be integrated into a data acquisition workflow. This is stated but not emphatically enough.
2. The implementation of an augmented reality system for visual feedback of identified cells through a microscope binocular.
3. The use of low-cost computer and GPU for applying these models in near real time.

These are three related topics which are currently embedded in an enormous amount of extraneous information. While I think the tools reported here are quite valuable, the paper needs to be substantially rewritten, and in particular shortened and focussed in order to make these essential points clear to the interested reader. In the following I have highlighted several sections of text that can either be moved to supplementary methods, converted to supplementary tables, stated with reference to previous work or even in some cases simply to be removed altogether. The goal should be to make the most important concepts in the paper much clearer in the Abstract, Intro, Results and in the first paragraph or two of the Discussion.

Thank you for your comments and your praise of my work. I have taken each point onboard and have addressed each of them. Overall I have reduced the number of words from >14000 to around 10000 and have refocused the text considerably. I apologise for the elaborate detail in certain places, I agree that the level of detail had reduced the fluency, making it hard to read. Please see below, alongside your specific points, the areas I have addressed.

Major points:

1. The first section of the results highlights the performance characteristics of the different networks compared in this paper. The text on pp. 2 - 3 is way too wordy and uses a substantial amount of undefined jargon in its description. Examples are ("one-shot detector" and several others). This could be condensed to 3 paragraphs max that highlight the aim and results in summary form.

I have reduced the introduction length substantially and refocused it onto the main outcomes of the paper reducing the jargon heavy aspects.

Page 13 is especially dense and the characterisation of network performance should be described in summary form. Forcing the reader to negotiate the names of all the different datasets and the performance characteristics is just too hard. Moreover, the section on Data Augmentation compares a technique that is extremely well established in machine learning in the cellular imaging community.

Data augmentation is normally done in a certain way, but we wanted to really highlight that additional vertical flipping, which is not performed normally in photography, is very pertinent and effective within microscopy. I had over-compensated with detail, but have now revised and shortened the text so that the salient points are clearer.

Anne Carpenter's lab has published extensively on this and there seems no reason to include this level of detail. Similarly, sections on "training across multiple classes" and performance of "YOLOv2..." recite techniques that have been looked at extensively in previous papers (w/r/t microscopy,
see https://www.biorxiv.org/content/10.1101/085118v1; as this on biorxiv, maybe a short statement is worthy, but not the long description. See also the recent paper from Horvath and Carpenter (https://www.cell.com/cell-systems/pdfExtended/S2405-4712(20)30117-4) and the also CellPose from Stringer et al (https://www.biorxiv.org/content/10.1101/2020.02.02.931238v2)). In short, the use of models pre-trained with ImageNet or the use of combinations of training sets is well-established in microscopy and referencing these

papers and saying the work here confirms these (as of yet to my knowledge not peer-reviewed preprints) would be sufficient. The goal should be to get the first section of the results down to 2 paragraphs max that highlight the choice of network for object identification. Nothing more is needed.

I have reduced the length of the text dramatically and have included the references you mention. In addition, I reduced the number of graphs in Figure 2 and placed the removed graphs into the supplementary materials. This represents a separation of the summary data from the fine-grain analysis and so works well.

2. This revised manuscript now has quite good examples of the application of the methods in Fig.4 and the author is to be commended on this additional information. Nonetheless the section on the "automated screening using autodetection" is again very dense and could be shortened to focus on the essential points. The development of the AMCA method seems extremely powerful and to my knowledge is a completely novel implementation.

I have reduced the level of detail in Figure 4 removing much of the biological background for that experiment. This level of detail was too much for this section and distracted the reader from the overall message of the article.

The author should highlight this advance much more clearly by emphasising the implementation in a small form factor low-cost platform. I imagine this advance will be of extreme interest to many members of the development and commercial imaging communities.

I have made additional references to this development in the introduction/results and conclusion. Thank you for recognising it, we believe that it is an important advance also.

3. The author has added substantial characterisation of DAPI stain cells and reports times for acquiring data that appear somewhat slow. While the author correctly states that his methods produce significant acceleration because of the ability to focus on exact cell positions it would be useful to know whether there are potential accelerations possible in future developments.

Thank you for bringing this to my attention. I have clearly specified now the expected frame-rate when using the detection network with the Jetson computer. I have also discussed the future of this technique in the discussion section and where I expect future developments to be made.

There is a lot to like in this manuscript and the author should be commended for all the hard work. In its current incarnation the manuscript is too dense and needs a significant revision to focus on the extremely important points that represent the novel impactful components of the author's work.

Thank you for the comments and thorough breakdown.

Some comments on the rebuttal:

1. "My [sic] comparing these different algorithms on the data we get a clear idea of the strengths and weaknesses of each one."
Reply:
a. Proofread your text and do not allow typos in your responses.
b. Yes, but the text is too dense to access the conclusion of these comparisons!

Thank you for the comment. Hopefully I have now sufficiently addressed the density of the text, making the conclusions clearer. Apologies for the typographical error.

2. "We have now compared the performance of images taken on a different microscope and compared the performance at different scales, including this data in Figure 4A-G. The technique works well as long as the cells have a similar visual appearance, which in micsroscopy [sic] is very much dependent on the NA of the objective used.
a. Proofread your text and do not allow typos in your responses.
b. This comparison is lost in the way too dense paragraph on pp 15-16. It is critical this is brought out in more detail.

Thank you for the comment. I have reduced the discussion on the multi-channel analysis as this was largely uneventful and have refocused on the comparison of objectives. Apologies once more for the typographical error, I will be more careful in future.

Jason Swedlow
Univ of Dundee

July 7, 2020

RE: JCB Manuscript #201903166RR

Dr. Dominic Waithe
Weatherall Institute of Molecular Medicine
John Radcliffe Hospital
Oxford, Oxfordshire OX3 9DS
United Kingdom

Dear Dr. Waithe:

Thank you for submitting your revised manuscript entitled "Object Detection Networks and Augmented Reality for Cellular Detection in Fluorescence Microscopy Acquisition and Analysis". We would be happy to publish your paper in JCB pending final revisions necessary to meet our formatting guidelines (see details below).

To avoid unnecessary delays in the acceptance and publication of your paper, please read the following information carefully.

A. MANUSCRIPT ORGANIZATION AND FORMATTING:

Full guidelines are available on our Instructions for Authors page, http://jcb.rupress.org/submission-guidelines#revised. **Submission of a paper that does not conform to JCB guidelines will delay the acceptance of your manuscript.**

1) Text limits: Character count for Tools is < 40,000, not including spaces. Count includes title page, abstract, introduction, results, discussion, and acknowledgments. Count does not include materials and methods, figure legends, references, tables, or supplemental legends. You are well under this limit but please bear it in mind when revising.

2) Figure formatting: Scale bars must be present on all microscopy images, including inset magnifications. Molecular weight or nucleic acid size markers must be included on all gel electrophoresis. Please double-check and make sure that all images have scale bars and/or size markers.

3) Statistical analysis: Error bars on graphic representations of numerical data must be clearly described in the figure legend. The number of independent data points (n) represented in a graph must be indicated in the legend. Statistical methods should be explained in full in the materials and methods. For figures presenting pooled data the statistical measure should be defined in the figure legends. Please also be sure to indicate the statistical tests used in each of your experiments (both in the figure legend itself and in a separate methods section) as well as the parameters of the test (for example, if you ran a t-test, please indicate if it was one- or two-sided, etc.). Also, if you used parametric tests, please indicate if the data distribution was tested for normality (and if so, how). If not, you must state something to the effect that "Data distribution was assumed to be normal but this was not formally tested."

4) Materials and methods: Should be comprehensive and not simply reference a previous publication for details on how an experiment was performed. Please provide full descriptions (at least in brief) in the text for readers who may not have access to referenced manuscripts. The text should not refer to methods "...as previously described."

5) Please be sure to provide the sequences for all of your primers/oligos and RNAi constructs in the materials and methods. You must also indicate in the methods the source, species, and catalog numbers (where appropriate) for all of your antibodies.

6) Microscope image acquisition: The following information must be provided about the acquisition and processing of images:
a. Make and model of microscope
b. Type, magnification, and numerical aperture of the objective lenses
c. Temperature
d. imaging medium
e. Fluorochromes
f. Camera make and model
g. Acquisition software
h. Any software used for image processing subsequent to data acquisition. Please include details and types of operations involved (e.g., type of deconvolution, 3D reconstitutions, surface or volume rendering, gamma adjustments, etc.).

7) References: There is no limit to the number of references cited in a manuscript. References should be cited parenthetically in the text by author and year of publication. Abbreviate the names of journals according to PubMed.

8) Supplemental materials: There are strict limits on the allowable amount of supplemental data. Articles/Tools may have up to 5 supplemental figures. At the moment, you are below this limit but please bear it in mind when revising.
Please also note that tables, like figures, should be provided as individual, editable files. A summary of all supplemental material should appear at the end of the Materials and methods section.
Note: Although we plan to include your adaptor/plate schematics as a separate supplementary file (i.e. not supplementary figures), you should still include a brief summary of the contents of the file just as you must do for the other supp figures and table.

9) eTOC summary: A ~40-50 word summary that describes the context and significance of the findings for a general readership should be included on the title page. The statement should be written in the present tense and refer to the work in the third person. It should begin with "First author name(s) et al..." to match our preferred style.

10) Conflict of interest statement: JCB requires inclusion of a statement in the acknowledgements regarding competing financial interests. If no competing financial interests exist, please include the following statement: "The authors declare no competing financial interests." If competing interests are declared, please follow your statement of these competing interests with the following statement: "The authors declare no further competing financial interests."

11) A separate author contribution section is required following the Acknowledgments in all research manuscripts. All authors should be mentioned and designated by their first and middle initials and full surnames. We encourage use of the CRediT nomenclature (https://casrai.org/credit/).

12) ORCID IDs: ORCID IDs are unique identifiers allowing researchers to create a record of their various scholarly contributions in a single place. At resubmission of your final files, please consider providing an ORCID ID for as many contributing authors as possible.

B. FINAL FILES:

Please upload the following materials to our online submission system. These items are required prior to acceptance. If you have any questions, contact JCB's Managing Editor, Lindsey Hollander (lhollander@rockefeller.edu).

-- An editable version of the final text (.DOC or .DOCX) is needed for copyediting (no PDFs).

-- High-resolution figure and video files: See our detailed guidelines for preparing your production-ready images, http://jcb.rupress.org/fig-vid-guidelines.

-- Cover images: If you have any striking images related to this story, we would be happy to consider them for inclusion on the journal cover. Submitted images may also be chosen for highlighting on the journal table of contents or JCB homepage carousel. Images should be uploaded as TIFF or EPS files and must be at least 300 dpi resolution.

It is JCB policy that if requested, original data images must be made available to the editors. Failure to provide original images upon request will result in unavoidable delays in publication. Please ensure that you have access to all original data images prior to final submission.

The license to publish form must be signed before your manuscript can be sent to production. A link to the electronic license to publish form will be sent to the corresponding author only. Please take a moment to check your funder requirements before choosing the appropriate license.

Thank you for your attention to these final processing requirements. Please revise and format the manuscript and upload materials within 7 days. If complications arising from measures taken to prevent the spread of COVID-19 will prevent you from meeting this deadline (e.g. if you cannot retrieve necessary files from your laboratory, etc.), please let us know and we can work with you to determine a suitable revision period.

Please contact the journal office with any questions, cellbio@rockefeller.edu or call (212) 327-8588.

Thank you for this interesting contribution, we look forward to publishing your paper in Journal of Cell Biology.

Sincerely,

Joerg Bewersdorf, PhD
Monitoring Editor
Journal of Cell Biology

Tim Spencer, PhD
Executive Editor
Journal of Cell Biology
