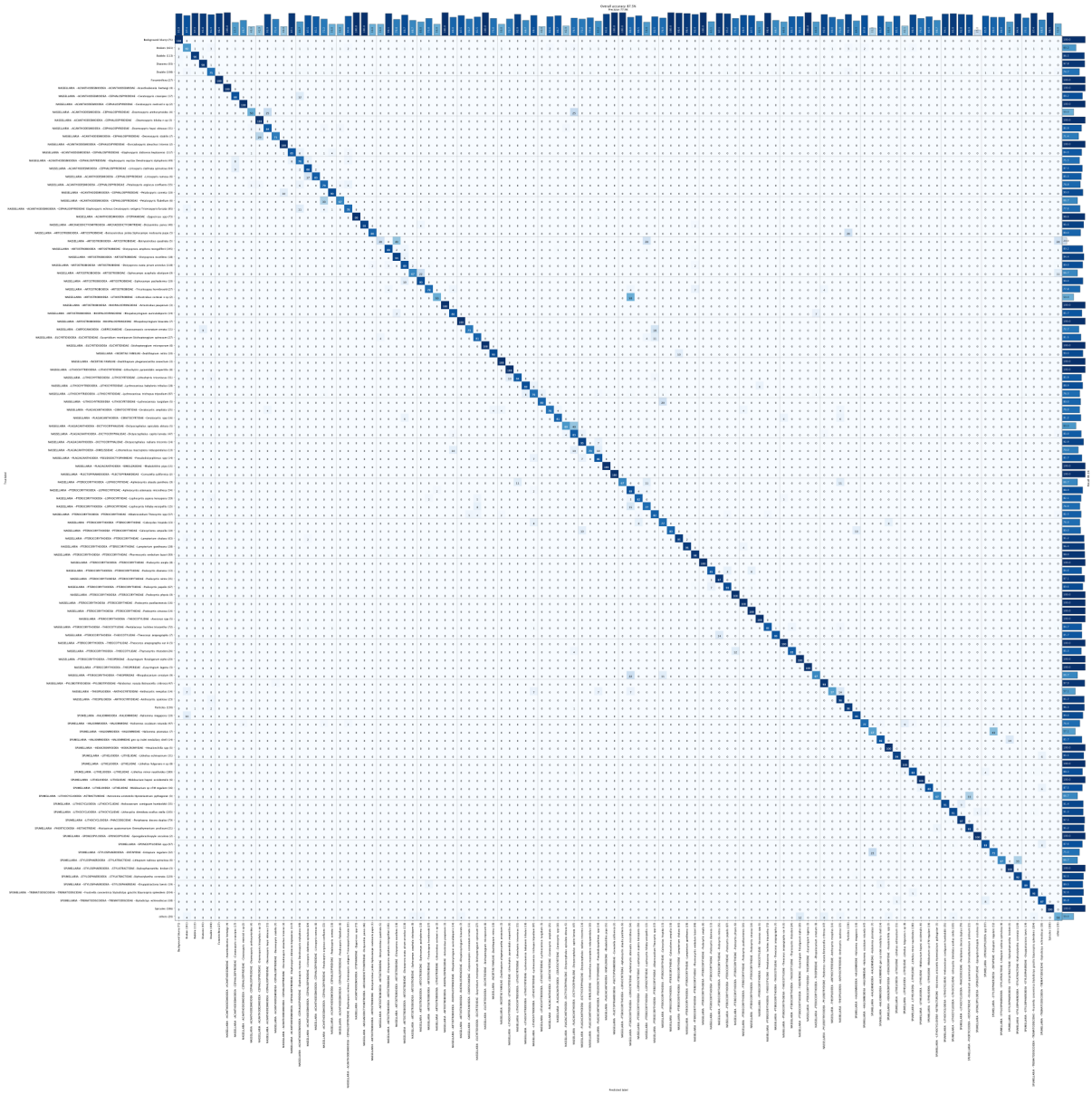
Response to the editor:

Comment: Given the extensive revisions — new object-detection workflow, working test datasets, assembly documentation, consistency fixes, and Referee #2's endorsement — most reproducibility and clarity concerns are now resolved. However, the absence of any quantitative benchmarking or case study evaluating recognition accuracy against existing approaches remains a substantive gap. Minor-to-moderate revisions are recommended: request that the authors add a concise case study or accuracy metrics (e.g., using the newly provided radiolarian/diatom test datasets) to substantiate the tool's practical performance claims, alongside acceptance of the already-resolved Referee #1 and #2 points.

Response: We do understand the editor concerns, especially regarding the benchmarking part, however, we want to argue that the technological advancement and interest of our manuscript does not lie in the fact that we use neural network for recognition of microfossils, but in the fact that we have adapted and developed new software to easily use these artificial intelligence tools for automated microscopy and recognition of microfossils (the RaCAM_IS and RaCAM_OD workflow) and on top of that, that they both work on a tiny device that can be mounted on any microscope and used on field. The absence of a concise case study regarding accuracy metrics of the neural network is due to the fact that those are not related to the camera but inherent to the CNN models themselves. The radiolarian case study that we use here was developed in Tetard et al, 2024, and will work exactly the same on the camera, not better nor worst, than on any computer, as the camera here just act as a micro-computer to use already existing and trained CNN. We made the choice to use common CNN models (developed in Marchant et al., 2020) for the Image segmentation workflow, and very common YOLO models (that are now commonly used in micropaleontology) precisely so that people can re-use their existing databases and trained models directly on the camera, without the need to re-train new models for this purposes. As a case study, we presented the use of the camera on field, in Antarctica, but comparing the accuracy metrics of image recognition models would give the same as on any computer using the “saved model files” as they are independent from the camera.

To still address this comment, we run a test using the provided Eocene Radiolarian CNN model. We did the acquisition of FOVs images using our Zeiss microscopy platform (Zeiss Axio Imager M2 microscope and Zeiss 712 Mono camera). We then acquired the same FOVs images using the RaCAM Camera. Both batches of images were processed and recognised (imagej segmentation and recognition using the same trained CNN model) on the Dell computer associated with the Zeiss microscope, and directly on the RaCAM, respectively. The results described here are now added to the manuscript:



The Eocene Radiolarian CNN model has an overall accuracy of 87.5 % (see confusion matrix attached and published in Tetard et al., 2023 (**Tetard, M., Carlsson, V., Meunier, M., and Danelian, T.: Merging databases for CNN image recognition, increasing bias or improving results? Mar.**

Micropaleontol. 185, 102296, https://doi.org/10.1016/j.marmicro.2023.102296, 2023). Eleven fields of view were acquired (the same FOVs were imaged for both systems, but due to very small differences in resolution and mount between the cameras, the magnification was very slightly different, and the acquisition on the zeiss system resulted in a few more individual specimen images after segmentation than with the RaCAM (273 vs 226 individual specimen images total)). Of the 273 individual specimen images acquired with the Zeiss microscope and camera, 234 were correct (**86% overall accuracy**). Of the 226 individual specimen images acquired with the RaCAM, 192 were correct (**85% overall accuracy**). Both systems thus perform very similarly, as expected by the overall accuracy predicted by the trained CNN itself (87.5%). This also shows that again, the system that the CNN is working on has no effect on the accuracy, but that the RaCAM sensor and specification can cope with hi-end system regarding this very specific task at least.

Comment: The following editor's concerns should be weighed against the new material:

Technology vs. outcome framing: The manuscript still centers on describing a hardware/software platform rather than delivering a benchmarked scientific outcome. The addition of a second workflow (object detection) strengthens versatility but does not by itself provide comparative performance data.

Response: We tried to address this comment and the benchmarking issue by taking new FOV images of Eocene radiolarian slide both using the RaCAM camera and a high-end zeiss camera, and compare the identification accuracy of our Eocene Radiolarian CNN (The one given as an example in our study) for both systems as described in the comment above. The CNN is the same, but this would compare the sensor / resolution / whole system accuracy as a proof-test of the camera itself, not the independent CNN models. This is now included in the comparison section between both system (linked with existing Fig.6). The FOVs used in Figure 6 has now been replaced by an example of FOV used in this test comparison, to show what the FOV used look like, and avoid using a new set of images from another age (previously Paleocene radiolaria which can become confusing for the reader).

Comment: Resolution/downsampling mismatch: The high native 12.3 MP sensor is still ultimately reduced to a much smaller processing resolution — the rationale for foregrounding native resolution as a key advantage remains unaddressed by the revisions.

Response: We are unsure about this comment. The high native 12.3 MP sensor is actually used to its full capability and generates images of about 4000x3000 px. During image processing we routinely decrease the jpeg quality to 90%, to save space and processing time as it has been proven to have no effect on the automated recognition accuracy (In Tetard et al., 2023). However, this is a choice that can be made by the operator and is not mandatory, in the software, this can be adjusted and the maximum quality can be chosen, as well as saving the images in color too. The format in which the images can be saved can be changed in the code too, to tiff, of png, but jpeg is commonly used in automated microscopy study, due to the amount of generated data.

Comment: Tool fragmentation (ImageJ + Python): This remains true for the RaCAM_IS workflow, though the new RaCAM_OD workflow uses YOLO models entirely within Python, which partially addresses this critique for one of the two now-available pipelines.

Response: The rationale for using a combination of ImageJ + Python for the RaCAM_IS workflow is again to enable compatibility with existing system and workflow, instead of having to rebuild new databases and new models for the same purposes of image segmentation. As we think than object detection will be more and more common in the near future and outgrow image segmentation, we decided to develop a new software for this purpose, but couldn’t unfortunately invest more time in the image segmentation workflow as our solution is working already.

Comment: Case study / benchmarking gap: This is the most significant remaining gap. Despite the new object-detection workflow, live-detection demo, and field deployment on SWAIS2C, the manuscript still lacks a quantitative case study comparing RaCAM's performance (accuracy, throughput, misclassification rate) against an existing method or ground-truth dataset. The video demonstrations and expanded GitHub materials substantially improve

reproducibility and usability, but do not substitute for a benchmarked performance evaluation.

Response: Following this suggestion and the ones above, we decided to run a test for comparing this system against our existing automated platform at ESNZ, as described above. Again, the technological advancement does not lie in the fact that we use neural network for recognition of microfossils, but in the fact that we have adapted and developed new work to easily use these artificial intelligence tools for automated microscopy and recognition of microfossils (the RaCAM_IS and RaCAM_OD workflow) and on top of that, that they both work on a tiny device that can be mounted on any microscope.

Reviewer 1

Comment: The manuscript introduces a camera designed for automated microscopy, utilising Raspberry Pi hardware, a 3D-printed case, and free software. This system acquires images, segments the object of interest within the field of view, and identifies it using a convolutional neural network (CNN). The innovative feature of this solution is that all processing occurs on the Raspberry Pi, requiring only a monitor or screen for operation.

Response: Thank you very much to the reviewer for this review. We are very happy that the reviewer took time to follow all instructions to install and use a RaCAM camera. This feedback is extremely valuable to us. Our first comment is that the workflow that the reviewer tested is the “image segmentation workflow” that uses a python interface for image acquisition, then every particle is segmented using an imageJ script, and then each image is recognised using a trained CNN.

Since publication, we also developed an object detection workflow, more user friendly as it only required a onnx model (yolo11n for example) to run inference on whole FOV directly acquired by the camera, in less than a second, and also inference directly on the live preview. This workflow will be refered to as Object detection workflow. We would be very happy if the reviewer was also able to test this second approach, which is much easier to install, faster to run, and work on for people not used to such automated approaches. We should be able to upload the new workflow and instruction soon on the github repository. A video of it can be seen at: https://www.youtube.com/watch?v=xw3MP37GsEo

Comment: This approach offers several advantages, including a low-cost solution accessible to individuals worldwide who need to count and identify biological objects. The low cost stems from the 3D-printable case and the Raspberry Pi itself. Notably, the image resolution from the high-quality camera sensor (4056x3040, 12 MP) exceeds that of many commercial cameras, such as the Leica ICC50 W/E (2592x1944, 5 MP). Additionally, the segmentation and artificial intelligence software tools provide a significant advantage for anyone looking to take their first steps toward automation.

Response: Thank you to the reviewer for this comment. Of course, we camera being low cost, as we said, it cannot perform as well as hi-end camera regarding video acquisition of fast moving object for example, but for regular microscopy, focusing on high resolution capability is usually better than video quality (again, depending on the application).

Comment: I believe this manuscript has the potential to be relevant for scientific investigations within the journal's scope. As I followed the individual steps, I encountered some technical issues. Specifically, I got stuck during the segmentation process: the program displays "processing," but nothing happens because AutoDiato_RaCAMx40.ijm is designed for radiolaria, not for pollen. I also attempted to download the radiolarian dataset from Carlsson et al. (2023), but the images were already segmented, preventing me from testing the segmentation and identification.

Reponse: Thanks you very much for showing us this mistake. Actually, the file that you are supposed to copy is not AutoDiato_RaCAMx40.ijm but AutoRadio_RaCAMx10.ijm (L172), we made an error here. The fact that nothing was happening in ImageJ is probably because the script path was thus wrong. In order to help reviewers and scientists building their own camera, we will integrate a complete example with already acquired FOV of radiolarian images, so people can train using the workflow with the provided CNN and provided ImageJ script for segmentation. We will also now provide this CNN separately as it weights 97 out of the 98 mo of the RaCAM_software.zip, to prevent having to download it again everytime we update the software.

Again, the provided images, imageJ script for segmentation, and trained CNN (for Eocene radiolarian) should be regarding as an example of capability for the camera, to show how the workflow works, and also for people to use these as a template / base to adapt it for their group of interest. As our team has been working on images of Radiolaria, Pollen, Foraminifera and diatoms, we could also add these imagej segmentation script to the “RaCAM_software.zip” to help people getting started.

If this issue if still encountered, we propose the referee to contact us directly so we can more easily figure out the issue.

Comment: This issue could arise for anyone, so I would like to request some sample images of Radiolaria that cover the entire field of view and work with the provided codes. Additionally, it would be helpful to include examples of other proxies to support the strong statements in the last sentence of the abstract. This would demonstrate that your tool can handle "any micropaleontological image," not just radiolaria.

Reponse: Indeed, this is exactly why you trying to use our tutorial and finding error is a very precious feedback for us as, we did it from scratch several times, but still missed such error in the tutorial. For the segmentation workflow, models are heavy and cannot be easily stored in our github (about 100mo each). However, we remind here that models are “fixed” and won’t perform better or worst on the RaCAM than on any computer running ParticleTrieur, which is used to train them. Our research group has trained models for various groups over the past year (Pollen, Diatoms, Radiolaria, Chironomids) and we provide one that was already published (Eocene Radiolaria) as an example here. For other groups models, we refer the reader to original publication (Such as Tetard et al., 2020 for Modern Radiolaria; Bourel et al., 2020 for Pollens; Another pollen CNN will be published soon for another study, another foraminifera CNN will be available soon (publication submitted), and a chironomid model will be submitted soon associated with another publication). Again, our present manuscript attempt to demonstrate the use of a new tools, what are the possibilities and how to use it, but we cannot provide numerous models for each microfossil groups.

Comment:

- The sentence at line 295 is correct, but it comes across as somewhat strong. The development of the segmentation macro and the ONNX model must take place before fully utilising RACAM. Additionally, the phrase “user-friendly interface” at line 295 seems a bit misleading, as the requirement to close the preview window in order to capture an image, as mentioned at line 235, is somewhat impractical.

Reponse: Thank you, we have edited this sentence. Indeed, closing the preview window is mandatory and operate as a separate process that we could not overcome in our development as attempting to capture an image while the preview is operating would return a “busy” argument.

Comment:

- Based on the RaCAM output shown in Figure 5, you could create a RaCAM input schema that would be applicable for any proxy. This schema can include the following components:
    - Hardware: List the items you need to purchase or print (excluding the monitor and keyboard). Please note that you should specify HDMI micro cables, not mini, as mentioned on line 110. Additionally, include M2 screws and their respective lengths.
    - Software: This consists of three parts:

- Your RaCAM_software.zip file

- A segmentation macro from ImageJ

- A CNN ONNX model, which must be trained beforehand using the ParticleTrieur program. Alternatively, this model can be created following the last part of your GitHub instructions, which appears below the program screenshot. Please note that this part is not clearly described in the manuscript.

Response: The micro vs mini hdmi mistake has been corrected. We thank the referee for this very good suggestion, we have prepared a figure summarising all hardware and software component, hopefully this will make the whole procedure easier to understand. Regarding the comment about training a model using the last part of our GitHub instructions: Actually, this last part gives instructions on how to use an existing trained model. We did not detail the model training is this manuscript, as it aims to present the Camera and how we put the hardware and software piece together, and as the model training part was already detailed in Marchant et al., 2020, and at https://particle-classification.readthedocs.io/en/latest/ We have edited our instruction to explain this.

Comment:

- Add to your GitHub instructions that, after installing Miniconda 3, it is necessary to close the window to initialise conda

Reponse: We added it, thanks.

Comment:

- Program Behaviour:
    - The requirement to close the preview window in order to capture an image is a significant issue, especially since there is no "X" button in the corner to close it.

Reponse: The preview window should have a x button to close it. As visible in fig.5. We recently updated the camera and found that the preview window would open in the top left corner with no visible button bar. An error likely due to a recent system update. To sort it out, you can click on Alt or the Windows/Raspberry key on your keyboard, and you should be able to drag the preview window out of the corner by click and drag it with the left mouse button (or click on alt+f4 to close it). There might be some compability issue with recent updates. To prevent this we have updated the RaCAM software to force the preview in a window that can be moved around and will appear in a specific position. We will upload it shortly.

Comment:

- The JPEG output format contains compression artefacts. It would be beneficial to add another option for saving data, ideally using an image format like OME-TIFF that includes standardised metadata for microscopy.

Reponse: Tests have been conducted in “Tetard, M., Carlsson, V., Meunier, M., and Danelian, T.: Merging databases for CNN image recognition, increasing bias or improving results?, Mar. Micropaleontol., 185, 102296, https://doi.org/10.1016/j.marmicro.2023.102296, 2023”. Showing no difference in CNN accuracy between using the same FOV image saved as TIFF, JPEG, and 90% quality JPEG, so we really encouraged using JPEG with the camera to use files as light as possible for the camera. The possibility to save in tiff format in the RaCAM software can however be changed easily in the code, and we will try

to add this possibility in the near future, together with the PNG format. Please let us know if you feel like other option are missing from the interface.

Comment:

- The magnification drop-down menu should include both objective and projective information; my microscope has an objective of 50x and a projective of 0.5.

Response: The magnification dropdown menu is just used for naming convention of the acquired image, to include it in the metadata of the image. Even if your magnification is not in the dropdown menu, you should be able to manually write it in the field. We will add some common magnification in the dropdown menu.

Comment:

- Reopening the program resets the settings from the last session – again, not very user-friendly.

Reponse: We thank the referee for this excellent suggestion, I was actually also having a hard time re-entering all the parameters everytime I used the camera on field. We had a button to save most of the parameters (the one that are numerical, as well as the core and sample names); and a button to load it instantly. Instructions were added to detail this.

Comment:

- I understand that the program has five functions: Live View, Snapshot (Image Acquisition), Image Processing, Image Recognition, and Census Data. However, only four red buttons are available. What happens if all functions are set to "yes"? I would expect that if one function is set to "yes," the others would automatically be set to "no."

Response: The workflow actually work differently, and we thank the reviewer for this, we will edit the text and software instructions to make it clearer. The live view only show a live preview, no matter if Image Acquisition, Image Processing, and Image Recognition are set to “yes”, or “no”. These parameters only affect what the software will do regarding snapshot and batch processing. If only image acquisition is set to “yes”, it will just take a picture. If image recognition is also set to yes, “snapshot” will take a picture and process it. If image recognition is also set to yes, “snapshot” will take an image, segment it (processing) and identify every particle segmented. If image acquisition is set to “no” you can’t do a “snapshot”, but you can “batch process” a core folder, containing sample subfolders, containing FOV images, either by only doing image processing, only image recognition, or both image processing and recognition. We will update the instructions to make it clearer.

Thanks again for all the constructive comments.

Reviewer 2

Comment: The manuscript presents RaCAM, a low-cost and accessible Raspberry Pi-powered camera system that integrates automated image acquisition, segmentation, and CNN-based recognition into a single portable microscopy workflow. The authors provide open-access hardware designs and software that allow the system to operate independently of a computer, including in field conditions, as demonstrated during the SWAIS2C voyage. To my knowledge, this is the first fully self-contained microscopy platform that combines image capture, processing, and automated identification on a micro-computer. This technology has the potential to make automated microscopy and micropalaeontology more accessible both in field settings and in laboratories with limited resources.

I have only minor comments for this manuscript, and believe that once these are addressed, this manuscript will be a publishable piece. Some comments are suggestions only, but may simply assist with making the manuscript a little more reader friendly (especially for those who are not as computer literate as the authors).

Overall Comments:

This is a solid technical note, it is clear, well-structured and provides good detail for reproducibility. My main comment would be to ensure consistency between the github files and the technical note instructions, although I do understand that the instructions on github would be the working file. My comments are very minor, and should not prevent this technical note being published.

Response: We thank the reviewer for his review of our manuscript and all his comments. Regarding the first comment about consistency between the github files and technical instructions from the manuscript, we apologise for the confusion. As we continue to work and improve the camera and associated software (now providing a second completely new and working workflow), some files and naming are now different than the original submission. All files, names, and sections are now consistent between the revised manuscript and github folder.

Comment: I have 2 main comments:

1. Please ensure consistency between the file names and folder names referenced in the manuscript and those available through the GitHub repository. Readers are likely to use both resources simultaneously during installation, and discrepancies may cause unnecessary confusion.
2. A simple setup guide that lists all hardware and software, and a very simple instruction guide would be useful. A diagram of setup would be great to go along with it. I felt it was simple enough, but too much reading to get to the actual putting everything together.

Response: Consistency is now ensure between the manuscript and github folders. A BOM list for all necessary components is now included in the github repository, as all as a schematic in supplementary figure as suggested by reviewer 1. As suggested here, we will also provide a simple instruction guide on how to use the RaCAM software, both in the manuscript and in the github repository. The manuscript was also emended to make it more user-friendly to read.

Comment: Minor Comments:

Lines 34-35 – Not really sure what you mean by this sentence, mainly the “from existing” part.

Response: This was corrected by “Some initiatives focused on building a whole Raspberry Pi-assisted microscope from mass-produced parts that can be easily bought or 3D-printed.

Comment: Line 69 – straight-out – remove hyphen.

Response: This is now corrected

Comment: Figures 1& 2 Captions: could you please add “using the Fusion360 3D-design software” to make this standalone.

Response: This suggestion was added to the manuscript.

Comment: Line 85 – Could you please clarify if Computer Module 5 is referring to the Raspberry Pi 5 board?

Response: It actually was, but was corrected as the CM5 refers to another product.

Comment: Line 114 – are the USB-A ports on the Raspberry Pi Screen, or whatever screen you are using? It makes sense when reading the whole paper, but perhaps it could be made really clear here, just for the readers who are using this as a set-up manual.

Response: We agree with the referee, this refers to USB-A ports on the raspberry pi board. It is now corrected in the text.

Comment: Line 117 – is this the correct filename? I can see RaCAM_IS_software.zip and RaCAM_OD_software.zip.

Response: We apologise for this confusion, this is an example of the files that were corrected and added on the github repository and in the updated version of the manuscript, but differs from the original submission. The original “RaCAM_software.zip” now refers to “RaCAM_IS_software.zip”. The “RaCAM_OD_software.zip” refers to the second workflow developed after the original submission. We updated the name of the software to avoid confusion when talking about the camera itself (the RaCAM), and the software used to

perform automated microscopy (previously the RaCAM software; now RaCAM_IS and RaCAM_OD). This is now corrected everywhere in the manuscript.

Comment: Line 131 – it might be worth saying that it doesn't work on a windows or mac platform at the moment, so people are clear. As I said before, people will likely read this as a set-up manual, and will probably need everything spelled out.

Response: We completely agree with the reviewer on this point and have added this in the text. We are currently working on a version of the RaCAM_IS and RaCAM_OD software that will be able to perform the batch processing of images on a Mac and Windows platform but that are ultimately not related to the present manuscript.

Comment: Line 290 – "operated on field" should read "operated in the field"

Response: We thank the reviewer for this mistake, it is now corrected.

Comment: Usability does not appear to be an issue, but I am sure there will be suggestions for improvements as people begin using the system. After a couple of minor edits, this is a publishable piece, and I look forward to seeing the system used
in labs and in the field.

Response: Again we thank the reviewer for his review of the manuscript.