Nurdle Count – A machine learning approach to nurdle classification and quantification - Year 1 Quarter 3 Report
(February 3rd, 2025)

PI: Seneca Holland

February 1st, 2025

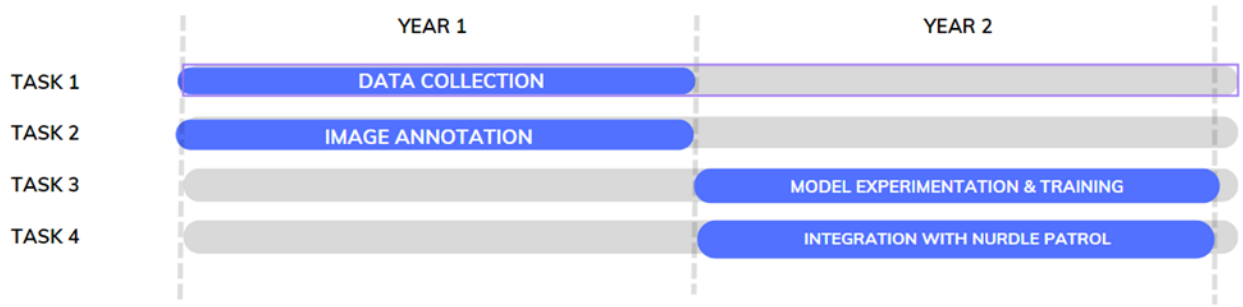
Administration:

The Nurdle Count – A machine learning approach to nurdle classification and quantification was approved for funding on January 8th, 2024, with a requested start date of May 1st, 2024.

Risks and Impacts:

None

Project Tasks:



1) Task 1 - Data collection:

a. **Collect training and test nurdle image data.**
b. **QA/QC collected nurdle image data.**
c. **Research and design AI training methods.**
d. **Develop a standard operating procedure (SOP) for capturing nurdle images.**

Task 1 – Subtasks 1a: Collect training and test nurdle image data

In Year 1, Quarter 1, the research team performed image capturing following the SOP developed for this purpose. Internally, using this SOP, 100 images were captured for Task 2—Image Annotation.

In Year 1, Quarter 2, this process was expanded with the help of middle school citizen scientists who are collecting images of nurdles in their classrooms and submitting them via the Nurdle Patrol Website using the QR code below.

In Year 1, Quarter 3, this process was expanded with the help of undergraduate students who collected images of nurdles in class and submitted them via the Nurdle Patrol Website using the QR code.

Figure 1: Nurdle Count Image Submission QR Code

Students will continue to submit images in Year 1 Quarter 4 to expand the nurdle image library.

Task 1 – Subtasks 1b: QA/QC collected nurdle image data

In Year 1, Quarter 3, the research team continued to execute the image-capturing process as per the SOP specifically developed for this task. Utilizing this SOP, the team successfully captured 500 additional images designated for Task 2—Image Annotation. These images will be considered for initial experimental trials with various AI/ML model methodologies. This approach aims to assess the efficiency of the overall process and identify potential improvements. Through this experimentation, it is expected to uncover insights that will guide enhancements to the methodology and optimize the workflow.

Task 1 – Subtasks 1c: Research and design AI training methods

This task was completed in Year 1, Quarter 1.

Task 1 – Subtask 1d: Develop a standard operating procedure (SOP) for capturing nurdle images.

In Year 1, Quarter 1, and in preparation for collecting training and testing the Nurdle image data, the Nurdle Count team first developed two Standard Operating Procedures (SOPs). After extensive review, two Standard Operating Procedures (SOPs) were created, each tailored to different audiences: internal and external. The internal SOP is designed for use

by the research team, while the external SOP is intended for 8th-grade students. Although both SOPs share similar content and workflow, the external SOP is written in language that is accessible and understandable at an 8th-grade reading level.

In Year 1, Quarter 2, project personnel developed a series of three videos detailing the nurdle capture process and made them available via YouTube for a wider audience. To ensure accessibility to a broader audience, YouTube settings enabled these videos to be viewed by kids and closed captioning was enabled.

These videos are:

Part 1 – Setting up Nurdles in Nurdle Count: https://youtu.be/99pSZEfB37g
Part 2 – Capturing Pictures for Nurdle Count: https://youtu.be/rLRbYLwNVVg
Part 3 - Nurdle Count Image Submission: https://youtu.be/TyTd6OBw9HA

In Year 1, Quarter 3, these videos and materials were leveraged to collect nurdle image data and collect feedback and improve the Nurdle Count application.

Task 2 – Image Annotation

In Year 1, Quarter 3, to enhance the quality assurance (QA) and quality control (QC) of images collected for training images for Nurdle Count, the Nurdle Swipe feature was developed and successfully integrated into Nurdle Patrol. This interactive tool allows subject matter experts to validate Nurdle images with a simple yes/no swiping action, providing an intuitive and efficient way to review and confirm image accuracy.

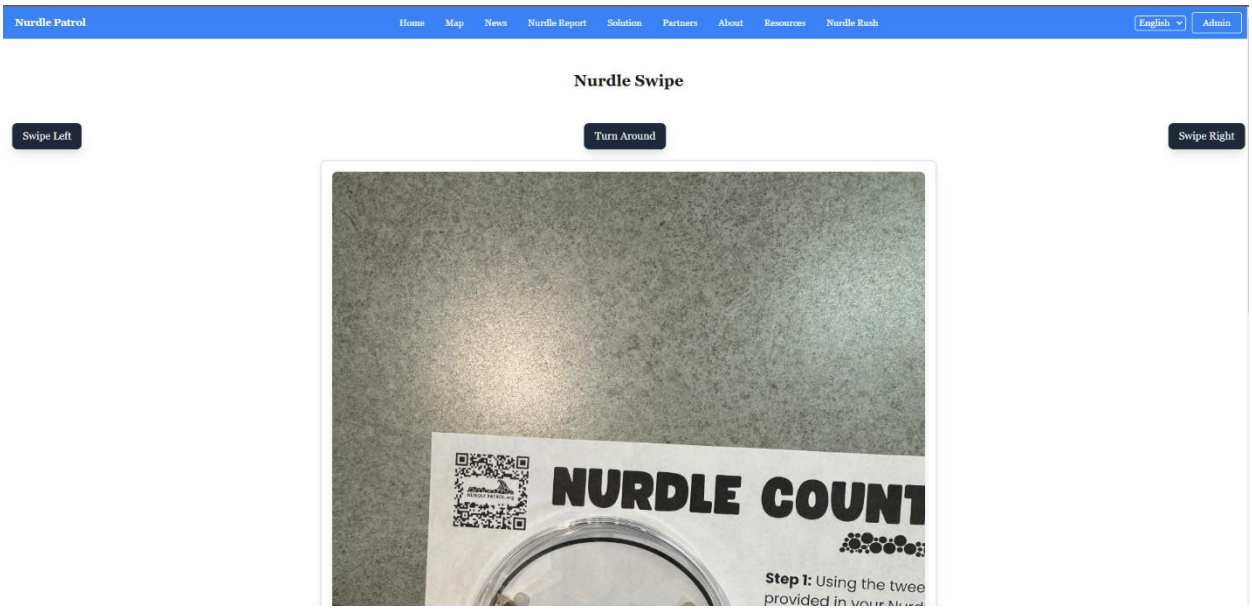


Figure 2: Nurdle Swipe - Swiping Right for a Qualified Image

Nurdle Swipe works in the following way:

For a Qualified Image:

Action:

- Swipe right on the image.

Purpose:

- This action designates the image as qualified, indicating it meets the criteria set forth for acceptance.

Outcome:

- The image is saved as 'Qualified' and moved to the corresponding folder or category within the system.

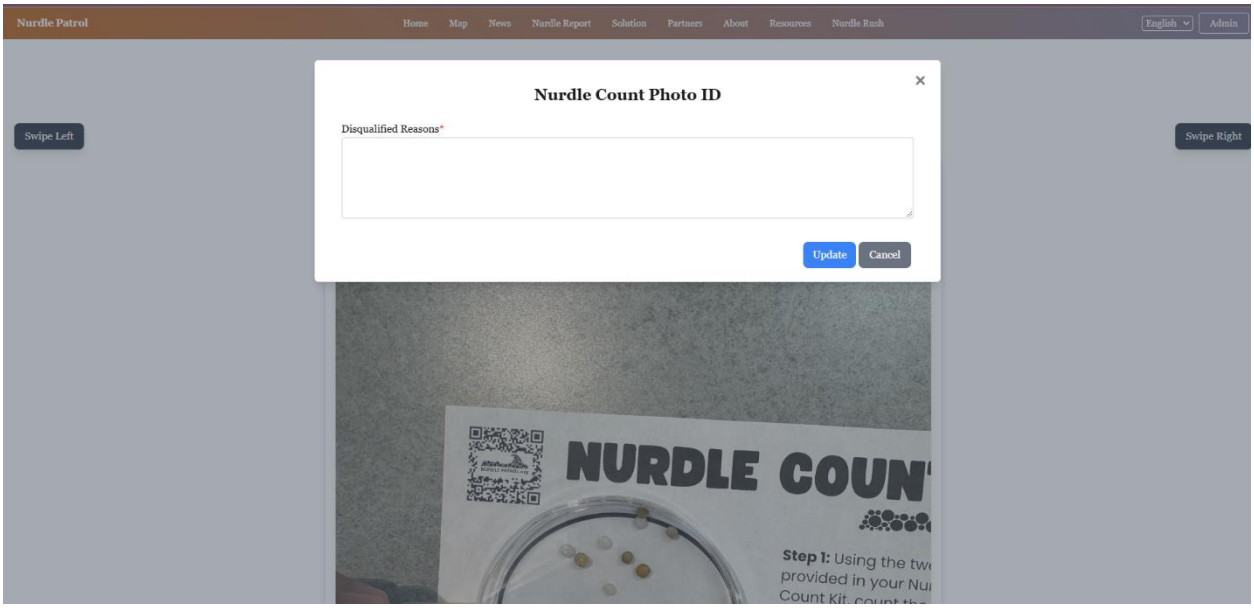


Figure 3: Nurdle Swipe - Swiping Left for an Unqualified Image

For an Unqualified Image:

Action:

- Swipe left on the image.

Requirement:

- Upon swiping left, you must provide a specific reason why the image is considered unqualified. This will typically involve selecting a reason from a predefined list or entering a custom reason in a text field that appears.

Purpose:

- This action ensures that each unqualified image is documented with a rationale, which aids in future review processes and helps in improving the criteria or the image capturing process.

Outcome:

- The image is saved as 'Unqualified' along with the reason provided and marked to a separate category for unqualified images.

For Reviewing Previous Images:

Action:

- Use the 'Turn Around' option to navigate back to previous images.

Purpose:

- This feature allows you to reconsider your earlier decisions. It is useful if it is believed that an error was made during the initial review of images.

Outcome:

- Re-evaluate and possibly change the qualification status of previously reviewed images.

Instructions for Use:

1. Navigate to the image you wish to reconsider by using the provided controls to scroll through the image history.
2. Perform a new swipe action (right or left) based on your re-evaluation and update any reasons if necessary.

Work on this task for the next quarter will be to build a workflow to transfer qualified images to CVAT for annotation and to annotate new images coming from Nurdle Patrol.

Task 3 - Model experimentation and training: ***to be completed in year 2***

a. Train Nurdle Count AI.
b. Collect feedback and improve the AI model.

Task 4 - Integration with Nurdle Patrol: ***to be completed in year 2***

a. Implement the Nurdle Count feature on NurdlePatrol.org.
b. Implement the Nurdle Count feature in the Nurdle Patrol Apple iOS mobile application.
c. Implement the Nurdle Count feature in the Nurdle Patrol Android mobile application.
d. Publish AI model to the public.

Summary:

We have made substantial progress in Year 1, Quarter 3 with Tasks 1 and 2, with a marked increase in nurdle image data collection and foundational steps in image annotation, including the successful deployment of a self-hosted CVAT platform for streamlined annotation and the development of Nurdle Swipe for the classification of images for the image library. In the next quarter, the focus will intensify on expanding the nurdle image dataset through additional contributions from middle school citizen scientists, which will enhance the diversity and environmental representation of the dataset. Concurrently, the project team will prioritize the development of a more rigorous QA/QC SOP for annotation, implementing advanced filtering and thresholding techniques to improve data reliability and optimize the model's accuracy in subsequent training phases. This approach aims to ensure the model is trained on high-quality, consistently annotated images, laying a stronger foundation for robust model development.

Obstacles: None