

Turbidivision: a machine vision application for estimating turbidity from underwater images (#101992)

1

First submission

Guidance from your Editor

Please submit by **25 Jul 2024** for the benefit of the authors (and your token reward) .



Structure and Criteria

Please read the 'Structure and Criteria' page for guidance.



Author notes

Have you read the author notes on the [guidance page](#)?



Raw data check

Review the raw data.



Image check

Check that figures and images have not been inappropriately manipulated.

If this article is published your review will be made public. You can choose whether to sign your review. If uploading a PDF please remove any identifiable information (if you want to remain anonymous).

Files

Download and review all files from the [materials page](#).

3 Figure file(s)

2 Table file(s)



Structure and Criteria

Structure your review

The review form is divided into 5 sections. Please consider these when composing your review:

1. **BASIC REPORTING**
2. **EXPERIMENTAL DESIGN**
3. **VALIDITY OF THE FINDINGS**
4. General comments
5. Confidential notes to the editor

 You can also annotate this PDF and upload it as part of your review

When ready [submit online](#).

Editorial Criteria

Use these criteria points to structure your review. The full detailed editorial criteria is on your [guidance page](#).

BASIC REPORTING

-  Clear, unambiguous, professional English language used throughout.
-  Intro & background to show context. Literature well referenced & relevant.
-  Structure conforms to [Peerj standards](#), discipline norm, or improved for clarity.
-  Figures are relevant, high quality, well labelled & described.
-  Raw data supplied (see [Peerj policy](#)).

EXPERIMENTAL DESIGN

-  Original primary research within [Scope of the journal](#).
-  Research question well defined, relevant & meaningful. It is stated how the research fills an identified knowledge gap.
-  Rigorous investigation performed to a high technical & ethical standard.
-  Methods described with sufficient detail & information to replicate.

VALIDITY OF THE FINDINGS

-  **Impact and novelty is not assessed.** Meaningful replication encouraged where rationale & benefit to literature is clearly stated.
-  All underlying data have been provided; they are robust, statistically sound, & controlled.
-  Conclusions are well stated, linked to original research question & limited to supporting results.



The best reviewers use these techniques

Tip

Example

Support criticisms with evidence from the text or from other sources

Smith et al (J of Methodology, 2005, V3, pp 123) have shown that the analysis you use in Lines 241-250 is not the most appropriate for this situation. Please explain why you used this method.

Give specific suggestions on how to improve the manuscript

Your introduction needs more detail. I suggest that you improve the description at lines 57- 86 to provide more justification for your study (specifically, you should expand upon the knowledge gap being filled).

Comment on language and grammar issues

The English language should be improved to ensure that an international audience can clearly understand your text. Some examples where the language could be improved include lines 23, 77, 121, 128 – the current phrasing makes comprehension difficult. I suggest you have a colleague who is proficient in English and familiar with the subject matter review your manuscript, or contact a professional editing service.

Organize by importance of the issues, and number your points

1. Your most important issue
2. The next most important item
3. ...
4. The least important points

Please provide constructive criticism, and avoid personal opinions

I thank you for providing the raw data, however your supplemental files need more descriptive metadata identifiers to be useful to future readers. Although your results are compelling, the data analysis should be improved in the following ways: AA, BB, CC

Comment on strengths (as well as weaknesses) of the manuscript

I commend the authors for their extensive data set, compiled over many years of detailed fieldwork. In addition, the manuscript is clearly written in professional, unambiguous language. If there is a weakness, it is in the statistical analysis (as I have noted above) which should be improved upon before Acceptance.

Turbidivision: a machine vision application for estimating turbidity from underwater images

Ian M Rudy¹, Matthew J Wilson^{Corresp. 2}

¹ Department of Math and Computer Science, Susquehanna University, Sellinsgrove, Pennsylvania, United States

² Freshwater Research Institute, Susquehanna University, Sellinsgrove, Pennsylvania, United States

Corresponding Author: Matthew J Wilson
Email address: wilsonmatt@susqu.edu

The measurement of turbidity serves as a key indicator of water quality and purity, crucial for informing decisions related to industrial, ecological, and public health applications. As existing processes require both additional expenses and additional steps to be taken during data collection, we seek to generate accurate estimations of turbidity from underwater images. Such a process could give new insight to historical image datasets and provide an alternative to measuring turbidity when lower accuracy is acceptable, such as in citizen science and education applications. We used machine vision to create an image classification and regression model trained on image data and their corresponding turbidity values recorded from a turbidimeter. To create a robust model, we collected data for model training from a combination of in situ field sites and lab mesocosms across suspended sediment and colorimetric profiles, with and without a Secchi disk for visual standard. Our resulting classification model is highly accurate with 100% of predictions within one class of the expected class, and 84% of predictions matching the expected class. Regression results provide a continuous value that is accurate to ± 0.7 Formazin Nephelometric Units (FNU) of true values below 2.5 FNU and $\pm 33\%$ between 2.5 and 55 FNU; values that are less accurate than conventional turbidimeters but comparable to field-based test kits frequently used in classroom and citizen science applications. To make the model widely accessible, we have implemented it as a free and open-source user-friendly web, computer, and Google Play application that enables anyone with a modern device to make use of the tool, the model, or our repository of training images for data collection or future model development.

Turbidivision: a machine vision application for estimating turbidity from underwater images

Ian M. Rudy¹, Matthew J. Wilson²

¹ Math and Computer Sciences, Susquehanna University, Selinsgrove, PA, United States

² Freshwater Research Institute, Susquehanna University, Selinsgrove, PA, United States

Corresponding Author:

Matthew J. Wilson

514 University Ave, Selinsgrove, PA, 17870, United States

Email address: wilsonmatt@susqu.edu

Abstract

The measurement of turbidity serves as a key indicator of water quality and purity, crucial for informing decisions related to industrial, ecological, and public health applications. As existing processes require both additional expenses and additional steps to be taken during data collection, we seek to generate accurate estimations of turbidity from underwater images. Such a process could give new insight to historical image datasets and provide an alternative to measuring turbidity when lower accuracy is acceptable, such as in citizen science and education applications. We used machine vision to create an image classification and regression model trained on image data and their corresponding turbidity values recorded from a turbidimeter. To create a robust model, we collected data for model training from a combination of in situ field sites and lab mesocosms across suspended sediment and colorimetric profiles, with and without a Secchi disk for visual standard. Our resulting classification model is highly accurate with 100% of predictions within one class of the expected class, and 84% of predictions matching the expected class. Regression results provide a continuous value that is accurate to ± 0.7 Formazin Nephelometric Units (FNU) of true values below 2.5 FNU and $\pm 33\%$ between 2.5 and 55 FNU; values that are less accurate than conventional turbidimeters but comparable to field-based test kits frequently used in classroom and citizen science applications. To make the model widely accessible, we have implemented it as a free and open-source user-friendly web, computer, and Google Play application that enables anyone with a modern device to make use of the tool, the model, or our repository of training images for data collection or future model development.

Introduction

In fields ranging from environmental science to public health, assessing water quality is vital. These assessments of water quality often begin with measuring turbidity, which can impact water clarity (Davies-Colley and Smith, 2001) and potability (LeChevallier et al., 1981). Turbidity is defined as a measure of particles suspended in water, contributing to a murky or cloudy appearance caused by the scattering of light. In nature, many of these particles are agitated sediment, such as clays and soils or suspended organic matter, such as plant debris or microscopic organisms and largely depend on surrounding land use (Moreno Madriñán et al., 2012). Pollutants such as industrial and agricultural runoff can also be linked to turbidity (Rügner et al., 2013; WHO, 2017). High turbidity can have negative impacts on aquatic life, since it often shares an inverse correlation with dissolved oxygen levels, creating an inhospitable environment for many taxa (Talke et al., 2009). In addition, increasing turbidity can cause nonlethal effects on aquatic taxa through altered predator-prey interactions (Abrahams and Kattenfeld, 1997; Ferrari et al. 2010), decreasing light penetration and reduced photosynthesis (Moore et al., 1997), and physical stress such as gill deformities (Lowe et al. 2015). Beyond affecting aquatic ecosystems and taxa, overly turbid water is unsuitable for consumption by humans (Muioio et al., 2020) and livestock (Umar et al., 2014), as well as acting as an indicator of bacterial contamination (Gharibi et al., 2012). For water treatment plants, this necessitates the use of filtration processes that can remove suspended particles and bring turbidity within

acceptable levels (WHO, 2017). Furthermore, monitoring turbidity can be helpful in tracking sediment runoff and pollution in bodies of water, providing valuable insights for environmental management and conservation efforts (Owens et al., 2005).

Typically, turbidity is measured with a turbidimeter which shines a source of light – either white or infrared – into the fluid and a probe which measures the resulting scatter of light. Meters that measure in Formazin Nephelometric Units (FNU) shine an infrared light into the solution and measure the scatter at a 90-degree angle of incidence. Other common units for turbidity include Nephelometric Turbidity Units (NTU), which are measured with white light at a 90-degree angle of incidence, and Formazin Attenuation Units (FAU), which are measured with infrared light at a 180-degree angle of incidence (Anderson, 2005). Turbidimeters are generally quite expensive, costing hundreds or even thousands of dollars. For example, the LaMotte 2020i turbidimeter used during this project has an MSRP of \$1,449 USD (LaMotte, 2024). This expense can be a significant barrier to citizen science initiatives, or labs with a limited budget, decreasing the ability of laypeople to contribute to water quality monitoring efforts. The converse can also be true, that lowering costs and creating easy and consistent methods of data collection increase the quality of data and participation in data collection (Zheng et al., 2018; Lee et al., 2020). Measuring turbidity with a turbidimeter can also be time-consuming, as they necessitate the transport of the meter itself, or the collection of individual water samples for later measurement. Furthermore, historical data from underwater image sets often lack turbidity measurements, making it impossible to retroactively assess water quality for historical datasets.

With the expansion and increasing accessibility of machine vision models, image analysis for water quality has become an emerging field of research. Current approaches include machine vision in conjunction with existing analytical tools (Yan et al., 2024), model development and image analysis from controlled environments (Nazemi Ashani et al., 2024), and remote sensing (Leeuw et al., 2018). The focus of this work has largely been in economically important fields such as aquaculture (Li and Du, 2022) and wastewater treatment (Mullins et al., 2018). However, there has been little work with model development from in situ images for specific water quality parameters, such as turbidity.

Our primary goal was to develop a machine vision model capable of estimating turbidity from underwater images. Such a model would offer a cost-effective alternative to traditional turbidimeters, making water quality monitoring more accessible for those who may not have the necessary funds to purchase expensive equipment. Given the affordability and widespread availability of waterproof digital cameras, including smartphones, this approach has the potential to democratize non-critical water quality assessment (e.g., Zheng et al., 2022). Using images can simplify the process for field sampling by requiring less equipment and eliminate sample processing. It also allows existing underwater image data without turbidity readings to be retroactively analyzed. We also evaluate the feasibility and accuracy of machine vision for predicting turbidity levels in water bodies. By assessing how well the model performs, this initiative could pave the way for innovative applications in research and citizen science, as well

as gaining insights into historical data, fostering greater engagement and participation in environmental monitoring and conservation efforts.

Materials & Methods

Photo/Data Collection

To develop this model, we paired two components in data collection: an underwater photograph, and a measurement of the turbidity of the water source in the image. The turbidity was measured in FNU, and the readings were taken with a LaMotte 2020i Turbidity Meter. Two cameras were used for photography, an Olympus TG-6, and a Sony HDR-AS30V, with photos taken in the “Auto” photo mode. For each photo we recorded associated data for location, whether the water was flowing or not (y/n), whether a Secchi disk was present in the photo (y/n), camera information such as ISO, shutter speed, focal length, F-stop, and white balance (when available), and the substrate present in the image (when possible).

We collected field photos from water sources including rivers, lakes, ponds, and the ocean at turbidity levels between 0 and 55 FNU and calibrated the turbidity meter at each sampling location. The images were collected from bodies of water in Pennsylvania and Maryland, USA. Of the field photos taken, 298 photos were from lotic, 30 from lentic, 87 from brackish, and 25 from marine sources. We fixed the camera to the end of a metal rod and submerged it underwater either with or without a 4.2 cm Secchi disk in frame at either 12 or 23 cm from the camera. Images were collected by taking individual photographs while the camera was submerged and by saving a selection of frames from a video taken while the camera was submerged. We took care to avoid disturbing sediment to ensure the FNU reading that was recorded matched the FNU of the water in the image. We took controlled experimental photos in two different systems following the same procedure as field photos, an opaque 20 L bucket to compare with lentic field photos and a 100 L acrylic fish tank with an additional sheet of acrylic placed vertically in the tank center and pumps to create a recirculating system to compare with lotic photos from the field. For 38% of all photos we included the 4.2 cm diameter Secchi.

The standard procedure for data collected from the bucket followed a graduated increase in turbidity with photos taken after each sediment addition. We added sediment in 4-gram increments and stirred to homogenize. When the water settled, we would take another turbidity reading, and additional photos. We repeated this process multiple times until the turbidity reached the upper limit of our defined range (55 FNU). To increase the robustness of the model against colorimetric changes, we also collected photos with colored ink (J. Herbin: Perle Noire, Rouge Hematit, Gris Nuage, and Vert Empire) added to the water in increments of 0.01mL/L, both with and without sediment. We repeated the same process in the recirculating aquarium with gravel added to the bottom of the tank for background texture. In total, we used 675 images in model development: 440 from the field, and 235 from the lab (114 from the bucket and 121 from the aquarium).

Data Preparation

Using the master image dataset, the images were broken into 11 groups (classes) based on the corresponding FNU values: 0-0.49, 0.5-0.99, 1-2.49, 2.5-4.99, 5-9.99, 10-14.99, 15-20.99, 21-28.99, 29-36.99, 37-44.99, 45-55. We split data into training, testing, and validation groups without duplicate images within any two or all three of them. Approximately 75% of the images were used in training, 15% for testing, and 10% for validation. This training/testing/validation split follows the standard folder structure for classification datasets required for model training. All python scripts made for the project were run using Jupyter Lab, and the notebook files that contain them are included with the source code. The master dataset, including all images and associated data can be found in the GitHub repository or dataset archive (Rudy & Wilson, 2024a).

Model Training

We used Ultralytics YOLOv8 classification model for training (Jocher et al., 2023). YOLOv8 provides five different size models to use as the basis for transfer learning when training a model. From smallest to largest they include, Nano, Small, Medium, Large, and Extra Large. There are different sets of these models for object detection, instance segmentation, pose estimation, object detection with oriented bounding boxes, and classification tasks, all pretrained on the ImageNet dataset. We used the Large classification model as the starting point for transfer learning. Using transfer learning and starting with a pretrained model allowed us to increase the speed of the training process by eliminating the need to build up weights suited for classification, opting instead to modify a set of pre-existing weights to better suit the new task of estimating turbidity. The final model was trained to use 320x320 pixel images, resized from the dataset, and was trained for 15 epochs, with a batch size of 16. The final model was chosen as the model with the highest fitness score from a group of 50 models trained with the specified parameters, each with different hyperparameters mutated using the AdamW optimizer and called using the tuning function provided by Ultralytics (Jocher et al., 2023). The hyperparameters of the final model are provided alongside the archived source code (Rudy & Wilson, 2024b). Training all 50 models during the final tuning process took approximately 2.2 hours using an Nvidia RTX 4080 model GPU.

Regression Prediction

Once the final classification model was trained, we performed inference on each image in the master dataset, which saves the confidence values between 0 and 1 of the top five most confident classes for each image in individual text files. From this we applied the confidence scores for each image and set all blank entries to a value of 0. We then compared these values to measured FNU by fitting a multiple linear regression model using the Weighted Least Squares method, with the predicted classes as the independent variables, and the measured value as the target variable to create a continuous estimate of turbidity in addition to the defined bins.

Results

The YOLOv8 training process generated a confusion matrix to visualize the performance of the model when inference was run on the validation set. An ideal model would have a perfect correlation, where each predicted class matches the actual class. For our model, predictions from the validation set were tightly clustered around the true classes, with 100% of predictions within one class of the expected class, and 84% of predictions matching the expected class (Figure 1).

By comparison, the regression had an R^2 of 0.975, with the parameters (x) and coefficients (β) of the regression equation of form $y = \beta_0 + \beta_1x_1 + \dots + \beta_{11}x_{11}$ shown in Table 1, with additional regression metrics in Table 2. While the accuracy of the regression model is not significantly greater than the classification model, the numerical, rather than categorical, estimates provide a more broadly useful result for further analysis (Figure 2). Given that our model predicts turbidity values in a range from 0 FNU to 55 FNU, assessing its accuracy based on the root mean square error (RMSE) does not make sense, as an error of approximately 2 FNU at the high end of the range is considerably less impactful than an error of 2 FNU at the low end. This is why relative root mean square error (RRMSE) and relative standard deviation (RSD) were calculated. While there is no accepted rule for RRMSE or RSD, the lower it is the more accurate the model is, and our 18.46% could generally be considered good. We found that from 0 to 2.5 FNU, 95% of our model's predictions fell within ± 0.7 FNU of the true values, and that from 2.5 to 55 FNU, 95% of our model's predictions fell within $\pm 33\%$ of the true values. In addition, there was no change in model accuracy when trained based on metadata, such as water body type or Secchi disk presence. We also visually compared subsets of the predicted and actual values against image appearance to confirm there were no patterns between image background and accuracy (Figure 3).

Discussion

In this study, we evaluated the ability of a machine vision model to estimate turbidity and the accuracy of that model. While the accuracy of the model is lower than a physical meter, it is comparable to field test kits typically used in citizen science programs, free, and a more accessible way to measure turbidity for those willing to accept the reduced accuracy compared to a traditional turbidimeter. In addition, this model is accurate enough to provide insights on turbidity conditions from historical data - underwater image datasets that never had turbidity recorded alongside the images - which would otherwise be impossible to go back and obtain. As an open-source program, the model can also be integrated into other systems which could be used for any variety of other potential data processing tasks related to underwater imagery. Even though the linear regression model is no more accurate than the classification, a single numerical value with a known confidence level is often more useful for analysis.

Web App

The Turbidivision web app combines the classification and regression models into one web-based program. It can run on any modern web browser, including those on desktop/laptop computers, as well as on phones and other mobile devices. Once the classification model is

downloaded, the code runs fully on the client machine; the server is only used to provide the initial download of webpage code and assets, which should remain cached in the browser for future use. After processing each image, the model outputs confidence values for each class, and the linear regression converts that to a discrete numerical estimate. For each image processed, the name of the image, confidence values for each class, and output of the linear regression are added to a CSV file. The web app has a skeuomorphic GUI designed to look like a turbidimeter, but a version using basic unadorned html elements is available for increased compatibility. The user simply needs to click on the file input button, which opens the file browser, and select the image(s) they want to process. The images will then be processed on the user's computer, and a button to save the output CSV file will appear once the processing is complete. Binary distributions of the web app for Windows (.exe), Linux (AppImage), and Android (.apk), as well as the files for the web app for serving locally, are available (Rudy & Wilson, 2024c). In addition, the Android application has been released on the Google Play store as "Turbidivision" and the web app is available for use (Rudy & Wilson, 2024d).

Future Directions

One potential improvement for future research could be changing the underlying model. We are predicting a continuous variable (turbidity), but instead of directly predicting a continuous variable, we first predicted a discrete variable, a class, and then used that to run a secondary regression model. If we were instead to create a neural network that ends in a fully connected layer which outputs a single continuous variable, from 0 to 55, we could effectively run regression directly on the images. However, such a process would require much more data and the creation of a custom convolutional neural network. The exact amount of data needed to train such a model would depend upon the architecture of the model itself, along with the desired accuracy. In addition, since regression is continuous, and not based on classification, direct model comparison is not possible. However, if we assume that quantizing turbidimeter measurements based on the reported accuracy of the meter provides an adequate analogue to number of classes, we can then calculate a minimum recommended amount of images needed for training such a regression model.

Given that the LaMotte 2020i turbidimeter is accurate to within ± 0.05 FNU between 0-2.5 FNU (LaMotte, 2024), we can say that there are approximately 25 unique ranges of distinct accuracy within this range ($2.5/0.1$); and since it has a reported accuracy of $\pm 2\%$ from 2.5-100 FNU, we can split it into 79 unique ranges of distinct accuracy between 2.5 and 55 FNU ($2.5 \times 1.04^{count} > 55$, *solved for count*). The sum of these two values, 104, gives us a rudimentary analogue to the number of classes needed. If we assume such a regression model would require a similar amount of data as a classification model and a recommended minimum of 150 images per class (Shahinfar et al., 2020), we can use our analogue class count multiplied by 150 pictures per class to estimate an absolute minimum of 15600 images, ideally with the turbidity values of these images evenly distributed between 0 and 55 FNU. Such a model would not be able to make use of the development speed increases of transfer learning, as all weights would need to be trained

from scratch. This training would require a larger dataset, potentially larger than mentioned above, as well as more time required for model training. If such a model were developed with a similar architecture to YOLOv8, processing times for the end user should be comparable to the model created in this project.

Within the existing modelling framework, more images, but fewer than those needed for a custom convolutional neural network, would improve the training process and should be able to achieve greater accuracy. More images would allow the model to learn the features of each class better, and a greater variety of images, such as in location, water quality, and lighting conditions, should help the model gain more resilience against features not related to turbidity. A larger dataset could also allow for the creation of more classes (smaller bins), which would allow the model to make more granular predictions.

Conclusions

Measuring turbidity can be critically important for contexts ranging from human health to food webs. Our goal was to determine efficacy, accuracy, and precision of a machine vision model for measuring turbidity from underwater images. We successfully demonstrated the potential of modern machine vision techniques as viable for estimating the turbidity of natural bodies of water, with an accuracy comparable to commercial test kits up to 55 FNU and down to near 0 FNU (e.g. LaMotte Turbidity Test Kit #7519-01). The model we created offers an accessible alternative to traditional turbidimeters and can provide turbidity measurements within an acceptable margin of error in many applications.

The application developed as part of this research project is also the first photo-based turbidity measuring tool accessible to the public. To make the model widely accessible, we implemented it as a free, user-friendly web application that would enable anyone with a modern web-enabled device to make use of the tool. Additionally, the web application is designed to allow the data to be processed on the user's device, keeping their data secure and avoiding unnecessary use of internet bandwidth by preventing the need to upload their files to a server. The app is compatible with a wide range of devices and has a simple user-interface, allowing anyone to easily benefit from the results of this research. In addition, its use of images as an input allows users to retroactively gain insights on turbidity from historical underwater image datasets and understand past trends.

Acknowledgements

We thank Nabeel Siddiqui for his early feedback on model approach and for connecting computer science to freshwater research at Susquehanna University.

References

- Abrahams, M.V. and M.G. Kattenfeld. 1997. The role of turbidity as a constraint on predator-prey interactions in aquatic environments. *Behavioral Ecology and Sociobiology*, 40:169-174. <https://doi.org/10.1007/s002650050330>
- Anderson, C.W. 2005. Turbidity: U.S. Geological Survey Techniques of Water-Resources Investigations. Book 9, Chapter A6.7. <https://doi.org/10.3133/twri09A6.7>
- Davies-Colley, R.J. and D.G. Smith. 2001. Turbidity, suspended sediment, and water clarity: a review. *Journal of the American Water Resources Association*, 37:1085-1101. <https://doi.org/10.1111/j.1752-1688.2001.tb03624.x>
- Ferrari, M.C., K.R. Lysak, and D.P. Chivers. 2010. Turbidity as an ecological constraint on learned predator recognition and generalization in a prey fish. *Animal Behaviour*, 79:515-519. <https://doi.org/10.1016/j.anbehav.2009.12.006>
- Gharibi, H., M.H. Sowlat, A.H. Mahvi, H. Mahmoudzadeh, H. Arabalibeik, M. Keshavarz, N. Karimzadeh, and G. Hassani. 2012. Development of a dairy cattle drinking water quality index (DCWQI) based on fuzzy inference systems. *Ecological Indicators*, 20:228-237. <https://doi.org/10.1016/j.ecolind.2012.02.015>
- Jocher, G., A. Chaurasia, and J. Qiu. 2023. Ultralytics YOLO. Version 8.1.8. [Computer software]. <https://github.com/ultralytics/ultralytics>.
- LaMotte. 2024. 2020i Portable Turbidity Meter. Retrieved 24 April 2024: <https://lamotte.com/2020i-portable-turbidity-meter#specifications>.
- LeChevallier, M.W., T.M. Evans, and R.J. Seidler. 1981. Effect of turbidity on chlorination efficiency and bacterial persistence in drinking water. *Applied and Environmental Microbiology*, 42:159-167. <https://doi.org/10.1128/aem.42.1.159-167.1981>
- Lee, K.A., J.R. Lee, and P. Bell. 2020. A review of citizen science within the earth sciences: potential benefits and obstacles. *Proceedings of the Geologists' Association*, 131:605-617. <https://doi.org/10.1016/j.pgeola.2020.07.010>
- Leeuw, T. and E. Boss. 2018. The HydroColor app: Above water measurements of remote sensing reflectance and turbidity using a smartphone camera. *Sensors*, 18:256-271. <https://doi.org/10.3390/s18010256>

- Li, D. and L. Du. 2022. Recent advances of deep learning algorithms for aquacultural machine vision systems with emphasis on fish. *Artificial Intelligence Review*, 55:4077-4116. <https://doi.org/10.1007/s10462-021-10102-3>
- Lowe, M.L., M.A. Morrison, and R.B. Taylor. 2015. Harmful effects of sediment-induced turbidity on juvenile fish in estuaries. *Marine Ecology Progress Series*, 539:241-254. <https://doi.org/10.3354/meps11496>
- Moore, K.A., R.L. Wetzel, and R.J. Orth. 1997. Seasonal pulses of turbidity and their relations to eelgrass (*Zostera marina* L.) survival in an estuary. *Journal of Experimental Marine Biology and Ecology*, 215:115-134. [https://doi.org/10.1016/S0022-0981\(96\)02774-8](https://doi.org/10.1016/S0022-0981(96)02774-8)
- Moreno Madriñán, M.J., M.Z. Al-Hamdan, D.L. Rickman, and J. Ye. 2012. Relationship between watershed land-cover/land-use change and water turbidity status of Tampa Bay major tributaries, Florida, USA. *Water, Air, and Soil Pollution*, 223:2093-2109. <https://doi.org/10.1007/s11270-011-1007-2>
- Mullins, D., D. Coburn, L. Hannon, E. Jones, E. Clifford, and M. Glavin. 2018. A novel image processing-based system for turbidity measurement in domestic and industrial wastewater. *Water Science and Technology*, 77:1469-1482. <https://doi.org/10.2166/wst.2018.030>
- Muoio, R., C. Caretti, L. Rossi, D. Santianni, and C. Lubello. 2020. Water safety plans and risk assessment: A novel procedure applied to treated water turbidity and gastrointestinal diseases. *International Journal of Hygiene and Environmental Health*, 223:281-288. <https://doi.org/10.1016/j.ijheh.2019.07.008>
- Nazemi Ashani, Z., M.F. Zainuddin, I.S. Che Ilias, and K.Y. Ng. 2024. A combined computer vision and convolution neural network approach to classify turbid water samples in accordance with national water quality standards. *Arabian Journal for Science and Engineering*, 49:3503-3516. <https://doi.org/10.1007/s13369-023-08064-5>
- Owens, P.N., R.J. Batalla, A.J. Collins, B. Gomez, D.M. Hicks, A.J. Horowitz, G.M. Kondolf, M. Marden, M.J. Page, D.H. Peacock, and E.L. Petticrew. 2005. Fine-grained sediment in river systems: environmental significance and management issues. *River Research and Applications*, 21:693-717. <https://doi.org/10.1002/rra.878>
- Rudy, I. and M. Wilson. 2024a. Underwater Image/Turbidity Dataset (1.0.0) [Data set]. <https://doi.org/10.5281/zenodo.10951021>

- Rudy, I. and M. Wilson. (2024b). Turbidivision: Estimating Turbidity with Underwater Photos - GitHub Repository (1.0.2). <https://doi.org/10.5281/zenodo.10982478>
- Rudy, I. and M. Wilson. (2024c). Turbidivision Web App and Binaries (1.0.2). <https://doi.org/10.5281/zenodo.11175905>
- Rudy, I. and M. Wilson. (2024d). Turbidivision live web app. <https://turbidi.vision/>.
- Rügner, H., M. Schwientek, B. Beckingham, B. Kuch, and P. Grathwohl. 2013. Turbidity as a proxy for total suspended solids (TSS) and particle facilitated pollutant transport in catchments. *Environmental Earth Sciences*, 69:373-380. <https://doi.org/10.1007/s12665-013-2307-1>
- Shahinfar, S., P. Meek, and G. Falzon. 2020. “How many images do I need?” Understanding how sample size per class affects deep learning model performance metrics for balanced designs in autonomous wildlife monitoring. *Ecological Informatics*, 57:101085. <https://doi.org/10.1016/j.ecoinf.2020.101085>
- Stormsson, C., G. Jocher, and W. Xin. 2024. Image Classification Datasets Overview. Ultralytics. Retrieved 10 April 2024: <https://docs.ultralytics.com/datasets/classify/>.
- Talke, S.A., H.E. de Swart, and V.N. De Jonge. 2009. An idealized model and systematic process study of oxygen depletion in highly turbid estuaries. *Estuaries and Coasts*, 32:602-620. <https://doi.org/10.1007/s12237-009-9171-y>
- Umar, S., M.T. Munir, T. Azeem, S. Ali, W. Umar, A. Rehman, and M.A. Shah. 2014. Effects of water quality on productivity and performance of livestock: A mini review. *Veterinaria*, 2:11-15.
- World Health Organization. 2017. Water quality and health - review of turbidity: Information for regulators and water suppliers. Retrieved 10 April 2024: <https://iris.who.int/bitstream/handle/10665/254631/WHO-FWC-WSH-17.01-eng.pdf>.
- Yan, J., J. Lee, L. Liu, Q. Duan, J. Lei, Z. Fu, C. Zhou, W. Wu, and F. Wang. 2024. A novel method for multi-pollutant monitoring in water supply systems using chemical machine vision. *Environmental Science and Pollution Research*, 31:26555–26566. <https://doi.org/10.1007/s11356-024-32791-3>
- Zheng, F., R. Tao, H.R. Maier, L. See, D. Savic, T. Zhang, Q. Chen, T.H. Assumpção, P. Yang, B. Heidari, and J. Rieckermann. 2018. Crowdsourcing methods for data collection in geophysics: State of the art, issues, and future directions. *Reviews of Geophysics*, 56:698-740. <https://doi.org/10.1029/2018RG000616>

410

411 Zheng, S., H. Li, T. Fang, G. Bo, D. Yuan, and J. Ma. 2022. Towards citizen science: On-site
 412 detection of nitrite and ammonium using a smartphone and social media software. Science of the
 413 Total Environment, 815:152613. <https://doi.org/10.1016/j.scitotenv.2021.152613>

Table 1(on next page)

Coefficients, 95% confidence interval, and significance of multiple linear regression by FNU class.

The lower and upper intervals (0.025 and 0.975, respectively) represent the upper and lower bounds of the 95% confidence interval for each FNU class. All p values were below the measurable limit and represented with 0 by the regression output.

FNU Class	Coefficient	Standard Error	Lower interval (0.025)	Upper interval (0.975)	t	P> t
Intercept	-60.90	11.61	-83.69	-38.11	-5.25	0
00-0.49	61.34	11.66	38.45	84.24	5.26	0
00.5-0.99	61.57	11.61	38.77	84.37	5.30	0
01-2.49	62.48	11.61	39.69	85.28	5.38	0
02.5-4.99	65.53	11.62	42.72	88.35	5.64	0
05-9.99	67.76	11.62	44.95	90.58	5.83	0
10-14.99	73.91	11.63	51.08	96.74	6.36	0
15-20.99	77.63	11.63	54.78	100.48	6.67	0
21-28.99	85.64	11.66	62.75	108.53	7.35	0
29-36.99	93.29	11.66	70.40	116.17	8.00	0
37-44.99	99.40	11.60	76.61	122.19	8.57	0
45-55	114.32	11.76	91.24	137.41	9.72	0

1

Table 2 (on next page)

Statistical summary of the multiple linear regression model for predicted versus actual turbidity.

Summary abbreviations include root mean square error (RMSE), relative root mean square error (RRMSE), and degrees of freedom (DF).

Statistic	Value
Adj. R-squared	0.975
RMSE (standard deviation)	2.055
RRMSE (relative standard deviation)	18.46%
F-statistic	2349
Log-Likelihood	-1416.2
AIC	2856
Skew	0.432
Kurtosis	11.997
Number Observations	662
DF Residuals	650
DF Model	11

Figure 1

Confusion matrices of model inference compared against validation set.

Panels represent predicted turbidity by (A) normalized values and (B) prediction counts against actual turbidity readings from in situ measurements. Darker blues represent closer match with proportional match or counts given within the heatmap.

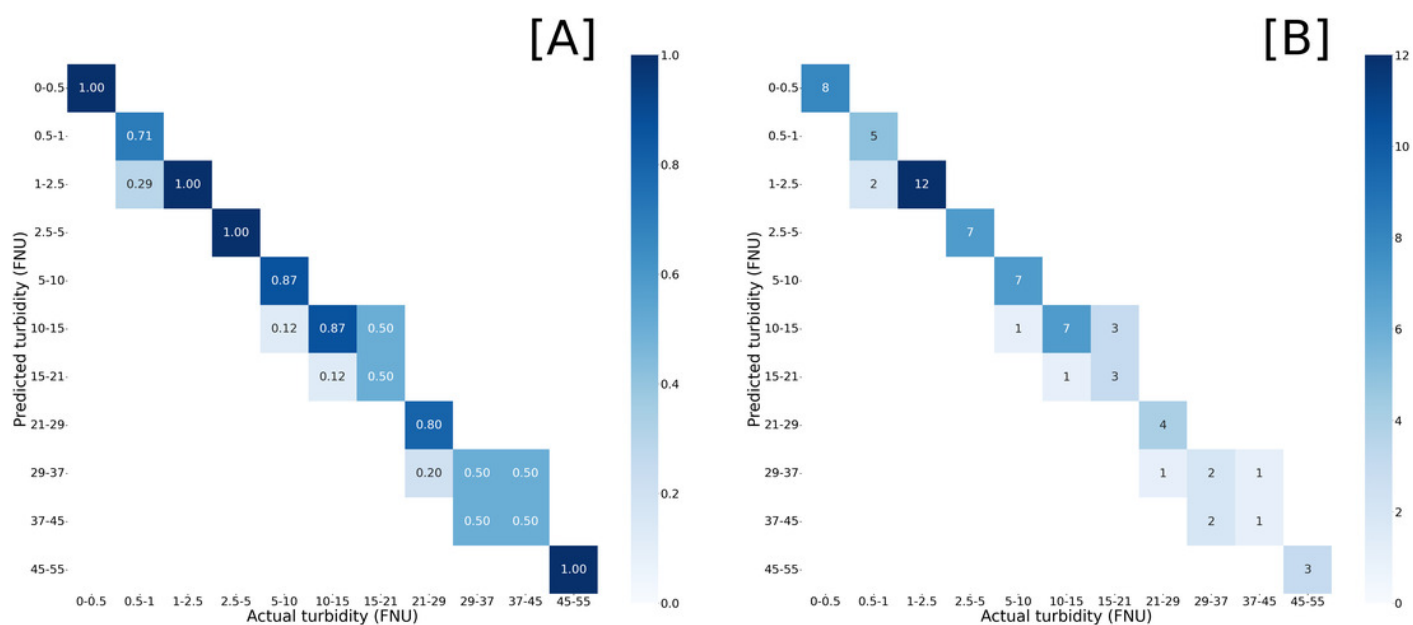


Figure 2

Regression model of predicted turbidity against in situ turbidity.

The dashed red line represents $x = y$ where predicted values would be a perfect match for in situ measurements. Each blue dot is a prediction of the linear regression based on the classification model confidence values of each image from the dataset.

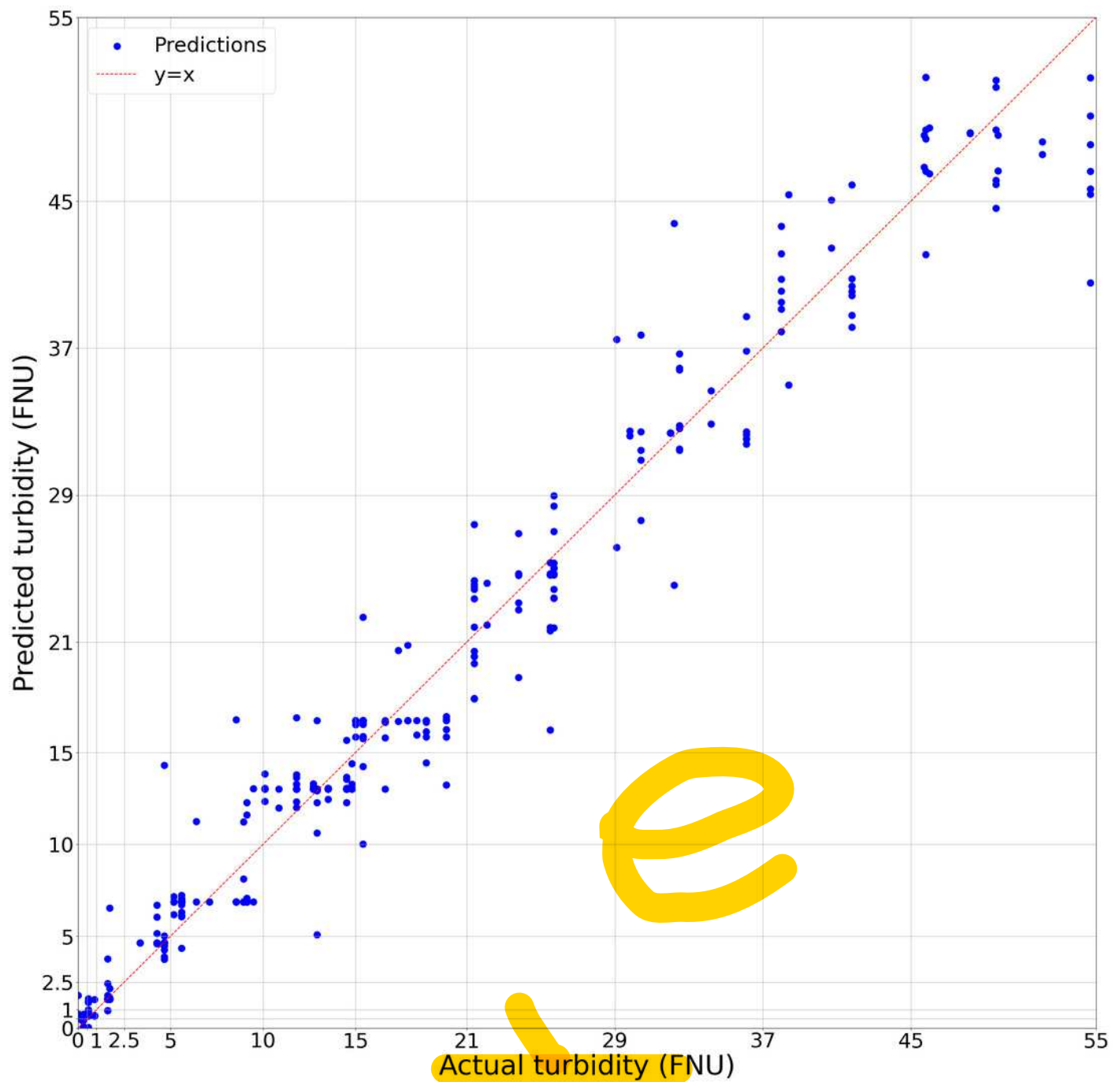


Figure 3

Sample of predicted and measured turbidity values from field and lab images.

Images selected represent the diversity of locations, Secchi presence, substrate, angle, and dyes included in image collection and model training.

