IJHSR
ijhighschoolresearch.org
■ RESEARCH ARTICLE

PuriScan: A Low-Cost Microscope for Pathogenic Detection using Machine Learning and Convolutional Neural Networks

George Cheng

North Carolina School of Science and Mathematics, USA; gcheng713@gmail.com

ABSTRACT: 2.1 billion people worldwide lack access to potable water; roughly 250 million people developed malaria in 2020 alone, and an overwhelming majority of the cases were in underdeveloped African countries. This research aims to remove socioeconomic disparities in drinking water by introducing a novel, cost-effective, and relatively accessible method for pathogen detection, generally focusing on widespread and hazardous bacteria (e.g., *Escherichia coli, Salmonella*) and protozoa (e.g., *Amoeba, Paramecium, Euglena*. By implementing a machine learning (M.L.) algorithm for pathogen identification and consolidating the data into a cloud database to improve the convolutional neural networks (CNN), PuriScan, our app, and product can accurately and rapidly assess the contents of water samples in less than 10 seconds, costing only 20 cents each. Furthermore, with confidence levels of 99% through thousands of selected database images in both pathogenic presence and identification, PuriScan can precisely pinpoint the location of dirty water through a GPS-coordinate system and "map" the feature. Integrating PuriScan's pure engineering with aspects of computerized intelligence can significantly reduce the logistics of traditional water characterization.

KEYWORDS: Systems Software; Machine Learning; Convolutional Neural Network; Detection; Low Cost; Microscope; Water, Applications.

■ Introduction

Literature Overview:

26% of the world's population does not have a safe drinking water supply, and only 1% of the world's water supply is drinkable freshwater.[1,2] Furthermore, more than 3.5 million people die each year due to water-related diseases; water-borne diseases are the leading cause of disease and death around the world.[3] While water potability has been thoroughly examined; many countries simply do not have the resources or funding available to transform non-potable water into potable water. While many solutions are available for water filtration, such solutions are only available locally and do not anticipate the barriers to project implementation. Applied similarly, in many underdeveloped countries in Africa, polluted water simply cannot be fixed through one direct solution; instead, citizens in various municipalities would need to determine whether their water sources are considered safe or not.[4,5]

Recently, the National Academies Press (NAP) released an annual report detailing the current methods of microbial detection in water, stating that most methods are too expensive or unfeasible to be implemented.[6] According to the NAP, filtration methods for water cleanliness are the most congruent and viable; however, other methods from Castillo-Henriquez, who researched possible biosensors for bacterial and viral detection[7] and the World Bank, who suggested commercial chemical methods[8] provided reasoning for the advantages in their respective proposals. Collectively, all parties agree that the prevalence of dangerous pathogens in supposedly "potable" water needs to be urgently addressed.

Current portable microscopy applications fall into three categories: 1) computational sub-micron Raspberry Pi microscopes, 2) low-cost 3D-printed microscopy inverse imaging, and 3) smartphone-compatible UV lenses[9-11]. However, the microscopes listed above have disadvantages that must be addressed for commercial translation. The computational sub-micron microscope is incompatible with smartphones, meaning that many sets of Raspberry Pi technology will have to be purchased and analyzed from the users' devices to test for water purity. Further, with the raw data being Bayer images, it becomes extremely hard to differentiate between bacteria and free-floating particles.[9] For the microscope that provides inverse imaging, optical building for 3D is not necessarily needed for bacterial detection, and the advancement in optical imaging (to train pathogen datasets) may not be holistically accurate.

Further, the construction of the optical microscope will require special parts and modules, further inflating the price needed for a portable microscope.[10] Lastly, POCKET MUSE, a microscope created to fit on a smartphone, could only detect UV frequencies and histology. The objective of this study is to image and predict pathogen levels in drinking water.[11] Other inventions like clip-on microscopes (PNNL) and handheld microscopes (Sigma) do not predict pathogen levels and serve only as a viewing device.

Membrane Filtration:

Membrane Filtration uses layers of semi-permeable material to separate substances when driven across a membrane layer. Membrane Filtration methods can be beneficial for the removal of bacteria, microorganisms, particles, or other organic material.[12] A key advantage of membrane filtration is the compactness of the membrane and the simplicity of separating many contaminants from water. Through multiple construction configurations (plate-frame, tubular, spiral, and hollow), membrane modules allow for the filtration of various materials through their porous composition.[13]

DOI: 10.36838/v5i2.9

Microfiltration, the most prevalent membrane pore arrangement of membrane filtration, has the largest pore size (0.2 µm on 100 kD UF) and typically rejects large particles and various microorganisms.[14] Disadvantages of microfiltration include backwashing (the production of polluted water by intermixing filtered contaminants) and coagulant filters. Composed of ceramics or metals, microfiltration membranes avoid degradation caused by membrane fouling (the process where particles cause severe flux decline).[15] The equation to calculate flux is:

$$1. \quad J = \Delta P / \eta (R_m + R_c)$$

where J is the flux number, ΔP is the pressure difference between both the filtered and unfiltered, η is fluid viscosity, Rm is membrane resistance, and R_C is pushing resistance.[16,17]

$$2. \quad LRV = \log_{10} (S_f / S_p)$$

Where Sf is the bacteria concentration before filtration and S_P is the bacteria concentration after filtration.[17] Comprehensibly, a higher LRV represents higher precision when isolating bacteria. For instance, a filtration of LRV 3 may reject 99.9% of bacteria, but a filtration of LRV 4 may reject 99.99%. The ideal water filtration for bacteria is usually at LRV 6-7, but in many developing countries, the LRV level may only be 3-5; in others, still, the LRV may be nonexistent due to poor infrastructure and lack of LRV-based membrane distribution.

Besides microfiltration and ultrafiltration, nanofiltration (0.2-2 nm) and reverse-osmosis methods (0.1 nm) can filter chemicals, proteins, particulates, or even smaller microorganisms. Disadvantages such as high cost and manufactured fragility, however, have effectively stopped the development of nanofiltration and reverse-osmosis filtration models.[18,19]

Biosensors:

Biosensors are a relatively new solution for microbial detection in water. These devices involve many transducer circuits and a biological reaction catalyst, normally an enzyme, antibody, or nucleic acid. The biological response attached to the transducer will induce an electrical signal detected by transfer devices such as Bluetooth or near-field communication (NFC) modules. There are six different categories of biosensors. Three of these are currently employed in water detection: electrochemical (signal communication through frequency, current, or voltage),[20] thermometric/fluorescent (visual signal through light waves),[21] and optic biosensors (detection of visual changes and relaying through patterned optic markers) Figure 1.[22] Biosensors, however, present the challenge of energy efficiency when employed and poor sensitivity for the less expensive options available to the public. Biosensor-based diagnostics include relatively poor sensitivity for both qualitative and semi-quantitative results; despite signal amplification or circuit redesign, the flaws in the range and precision of biosensor information, along with the incapacity for long-term use, add up to many disadvantages for biosensor-related applications.[23,24]

Methods

Our Idea - PuriScan:

As of 2018, 83% of adults in both developing and underdeveloped countries own a smartphone.[25] An app would be the most beneficial pathway to make the product accessible and widespread with the modern increase in technology usage and prevalence. We designed and developed PuriScan, a device that enables the user to upload images and take images of water using our 3D printed microscope and upload the image to our custom-designed cloud server where the model is present. After uploading the images, the model can quickly predict and return the results within seconds to the user. The user has the option to add their results to a map that is uploaded to our server. The user can also view the map and see the results of others who have uploaded their results. This design can both act as a lifesaving method for microbial detection and alert regulatory officials to help with their varying water problems.

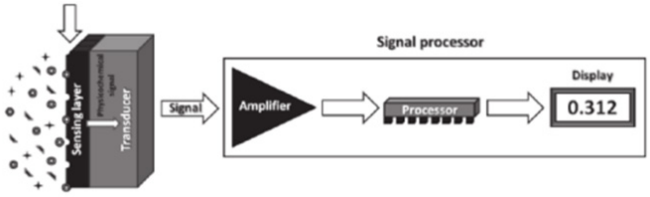

Figure 1: A schematic diagram of a biosensor consisting of the sensing layer, transducer, amplifier, processor, and display (Adapted from Andrade *et al.*).

PuriScan Application Design:

PuriScan was created using the React Native library to design the user interface and functionality of PuriScan. A key advantage of React Native is the universality of the app; React Native allows individuals to code in one programming language, JavaScript, and publish on all devices' download services (as both an APK, APP, or File Formatting). With this universality, PuriScan is accessible to everyone no matter what device they use,[26] which region they are from, or what file they import.

To house the model in the cloud hosting software, Heroku was used to host the model, where the app can request the model by transmitting data through a POST request.[27] In our case, the image submitted will receive results from Heroku after the request has been submitted and the image has been processed. The last feature of the app is the MAP (Machine Assisted Pinning) which is essential for spreading information about the purity of water. Google Maps was used to create an interactable map where the user can zoom, pan, rotate, and tap markers to view information on microbial likelihood.

The data for the location of the markers and the corresponding results were stored in a Firebase database. This Google service allows one to create a cloud database to store information. The app uses Firebase to store the latitude and longitude of the markers along with the results.[28] Information on the longitude and latitude can be downloaded to the user's phone, where the PuriScan can read the data to produce the clickable marker (Figure 2). Figure 3 shows an 8-panel demonstration of the application used and the panels displayed at each interval.

DOI: 10.36838/v5i2.9

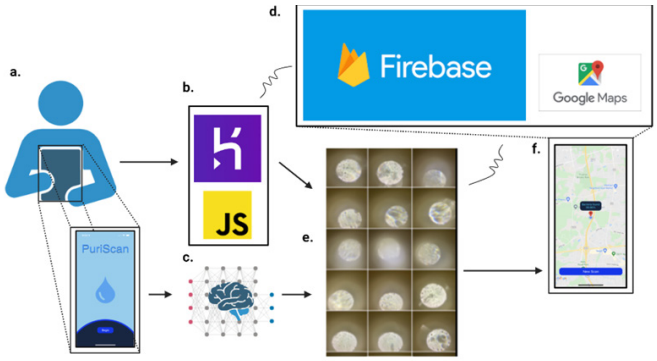

Figure 2: user requesting authorization through Heroku [**b.**] in JavaScript. [**c.**] Convolutional neural networks (CNN) act upon PuriScan authorization and predicts based on thousands of images in the database. [**d.**] Firebase hosting pinpoints user location through GPS coordinates and marks in [**f.**] with [**e.**]'s results. The user will also be able to see the results before submission.

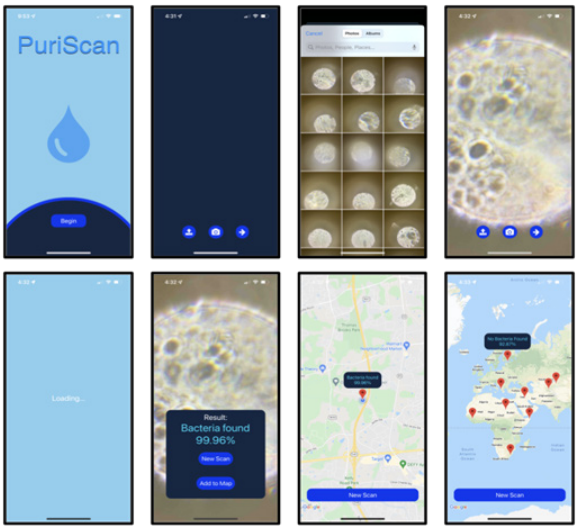

Figure 3: The user interface of the app contains 8 screens where the user can upload an image or take an image, submit the image, and receive the result. They also have the option to add their results to a map that is accessible to all users of the app. The layout is simple and easy to use to ensure accessibility.

Cost-Effective Microscope:

To complement PuriScan's application, a lightweight, low-cost microscope must be developed and easily accessible for developing countries. Instead of a traditional microscopic lens, PuriScan implements glass beads (1mm, 3mm) for magnification; the concavity of smaller glass beads refracts a smaller portion of the sample, providing a more "zoomed-in" image of pathogens. 3D printing, therefore, provided an avenue to experiment with possible materials or designs regarding microscope infill, structure, translucency, and durability (Table 1).[29]

3D printing also allows us to design on a ubiquitous compatible operating program (STL – Shapr3D) which can be distributed and open-sourced. Further, thermoplastics and thermosetting plastics [made from fused deposition modeling (FDM) printers, stereolithography (SLA) printers, and selective laser sintering (SLS) printers] can be compared with one another based on quality and mass production.[30-32] Moreover, materials like polylactic acid (biodegradable, rigid, brittle), polyethylene terephthalate glycol (PETG) (faster production, high transparency, food-safe), and high impact polystyrene (HIPS) (soluble support, dissolvable) can be compared in durability trait comparisons to determine the best combination of lens and cartridge (Table 2).[33-35]

Table 1: Advantage comparison between 3D printable material and system.

Material	Advantageous Features	Applications
Polylactic Acid (PLA)	- Easily accessible - Biodegradable - Odorless	- Easy for prototyping - Conceptualization
Polyethylene Terephthalate Glycol (PETG)	- Faster production - Chemical resistant - High Transparency	- **Waterproof** - Assimilable components
Nylon	- Durable - Impact-resistant	- Wear-resistant parts
Thermoplastic Polyethylene (TPU)	- Flexible & Stretchable - Vibration Dampening	- Flexible prototypes/parts
High Impact Polystyrene (HIPS)	- Soluble support material - Dissolvable	- Support material

Table 2: Compilation of materials used when manufacturing PuriScan. Polylactic Acid Cost[36]; Polyethylene Terephthalate Glycol (PETG) Cost[37]; Borosilicate Glass Beads Cost[38]; Stickable Mirror Pieces[39].

Materials	Application	Cost (per unit)
Borosilicate Glass Beads	- Magnification (350x)	$0.015
Stickable Mirror Pieces	- Reflection Purposes	$0.004
AND		
Polyethylene Terephthalate Glycol (PETG)	- Waterproof Magnification Support	$0.1
OR		
Polylactic Acid (PLA)	- Magnification Support	$0.07
Total Cost (Per Unit):	→	**$0.0854 – $0.119**

PuriScan Computer Aided-Design:

In the first few months of designing PuriScan, multiple prototypes using 3D-creating software (AutoCAD, Shapr3D, and OpenCAD.io) were created. PuriScan's flexibility was determined through simulations on AutoCAD's personalized simulation software. Various design factors like durability, measurements, size and scaling support (SS) levels were also considered. Finally, PuriScan was printed using different materials (3 of each material) and objectively tested for viability in a water-induced environment (Figure 4).

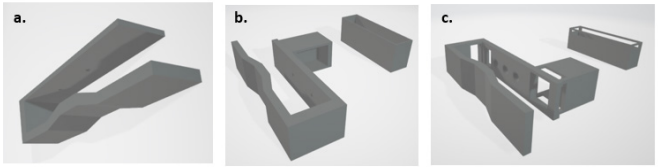

Figure 4: PuriScan [a] was the first design (created in AutoCad & exported to 3D Viewer). This design consisted of 2 holes (1mm x 1mm) for the BioSpec glass bead placement for magnification (350x). PuriScan [b] was the second design (created in Shapr3D and exported to 3D Viewer). This design consisted of 2 holes (measured 1mm x 1mm) for magnification and a cartridge holder for varying water samples. PuriScan [c] was the final design (created in Shapr3D and exported to 3D Viewer). This design consisted of 3 holes (measured 1.5mm x 1.5mm) for magnification and a cartridge holder (with self-filling) for varying water samples. PuriScan [c] also is more lightweight, costs less, and facilitates both natural light and sunlight into the water sample.

3D Printing for PuriScan:

A digital 3D model was converted into an STL file. The model was sliced 2-dimensionally (also known as the G code) to specify infill, durability, orientation, pathway, and predictable support using A.I. known as G-code A.I. The print platform can be started through a nozzle-based system or the melting of plastic or melting of a foundational layer of plastic or model,

essentially carving into the plastic bilayer. Post-processing using varying PSI modulus tested for tensile strength (AccuView^{i} 1000©).

Prototypical Testing:

To measure preliminary data, PuriScan was tested through collected water samples from filtered (F), nonfiltered (NF), and terrestrial (TR) sources to perform an accurate comparison between our machine learning values for PuriScan and laboratory microscope (ECHO) images (n=150, random sampling). To calibrate for microbial size differences, two different datasets for both PuriScan and ECHO were created. Ultimately, the calibrated ECHO datasets were set to match the magnification of PuriScan magnification to form a side-by-side comparison between both datasets and results. The images captured then ran the algorithm coded within the PuriScan App to give the A.I. model prediction on bacterial contamination or lack thereof (Figure 5).

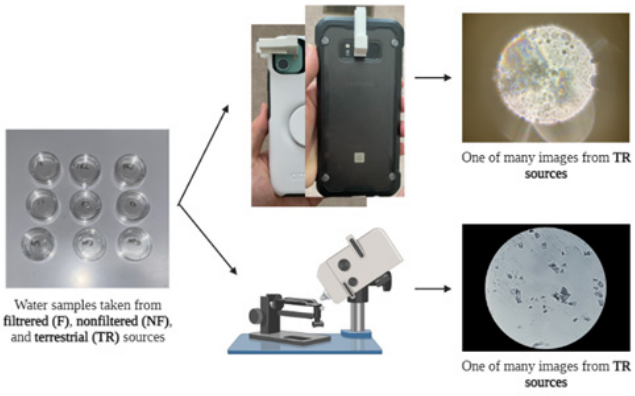

Figure 5: This schematic demonstrates our testing comparisons between PuriScan and conventional microscope methods of **filtered (F)**, **nonfiltered (NF)**, **and terrestrial (TR)** water sources. Since PuriScan scans on a 350x magnification, the bacteria can be more clearly seen in PuriScan compared to the microscope (on 250x magnification).

Data Augmentation:

The dataset used for training and validating the model was from the Woods Hole Open Access Server (WHOAS), consisting of microscopic images of microbes in water. For reference to "clean water," we collected purified water and imaged it according to the same background color as WHOAS to simulate the same atmosphere for that group.[40]

Data augmentation is useful in creating new modified images from the original dataset; the process of data augmentation takes advantage of existing pictures by rotating, shifting, skewing, and enlarging images. Creating new images using data augmentation allows PuriScan to learn how to classify similar pictures. Data augmentation, seen in Figure 6, accounts for variations in images such as rotations which expands the dataset and benefits the accuracy of the model.[41]

Convolutional Neural Networks:

Convolutional Neural Networks (CNNs) is a branch of deep learning for classifying images; CNNs take in an image as an input where they can analyze the image and assign weight that correlates to the "importance" (based upon logistical regression values) of features. The building blocks of CNN consist of nodes, essentially acting as the decision processes of human neurons. The central concept of a CNN is to eliminate all features that are not beneficial for classification, resulting in the most essential features through a combination of convolutional and pooling layers. These features consist of edges, color, and orientation; the last layer is a dense layer with n nodes, where n represents the number of labels---for our CCN model, a binary classification model, n would be 2 since there are 2 labels.[42] TensorFlow was used for CNN implementation in PuriScan.

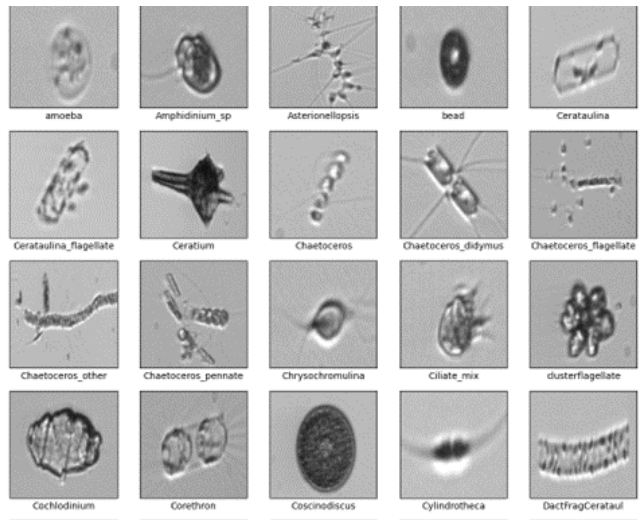

Figure 6: The data augmentation from the most prevalent (and database-plentiful) classes of bacteria, protozoa, and other microorganisms found in water by using Woods Hole Open Access Server (WHOAS).

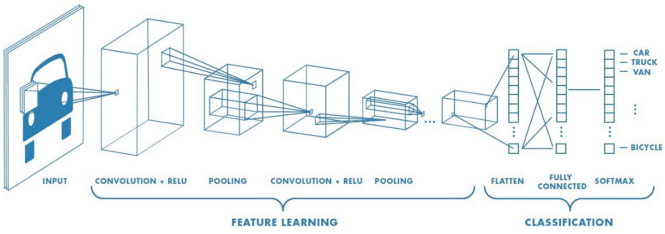

Figure 7: A general representation of the layers within a convolution neural network. The feature learning layers consist of CNN layers and pooling layers (adapted from Saha *et al.*).

Logistical Regression:

Logistical regression in machine learning is an algorithm for binary classification where the predicted values are re-arranged to fit between 0 and 1 (Figures 7, 8). As a result, the predicted values are transformed from an integer indicating a label to a decimal value representing the probability of the correct prediction.[43]

```
model = tf.keras.models.Sequential()
model.add(Conv2D(32, (2, 2), activation='relu', input_shape=(100,100,3)))
model.add(MaxPool2D((2, 2)))
model.add(Conv2D(16, (2, 2), activation='relu'))
model.add(MaxPool2D((2, 2)))
model.add(Conv2D(8, (2, 2), activation='relu'))
model.add(MaxPool2D((2, 2)))
model.add(Conv2D(4, (2, 2), activation='relu'))
model.add(MaxPool2D((2, 2)))
```

Figure 8: Representation of Logistical Regression in PuriScan

■ Results and Discussion

PuriScan's Results and Discussion section can be separated into 3 distinct categories:

1. Characterization of PETG, PLA, HIPS, and Nylon morphology and durability

DOI: 10.36838/v5i2.9

2. Characterization of PuriScan's A.I. confusion, convolution, and confidence (3Cs)

3. Statistical significance based on PuriScan Data for major Bacteria groups

Testing for UTC, Yield Strength, Young's Modulus, and Elongation % is important because longevity in pathogen-containing slides can help reduce inaccurate readings, the need for replacement, or slide-host contamination.

UTC, Yield Strength, Young's Modulus, and Elongation % Analyses:

Durability, or Ultimate Tensile Strength (UTS), can be determined through PSI modules (AccuView^i 1000© tensile machine). UTS is the maximum stress a material can withstand in tension before breaking entirely. HIPS had the largest UTS, followed by Nylon, PLA, and PETG. All the prototypes were laid horizontally to reduce area exposure and weakness points (Figure 9).

Yield Strength can be determined through PSI modules (AccuViewi 1000© tensile machine). Yield strength is the maximum stress a material can withstand tension before any denaturant. PLA had the most considerable Yield Strength, followed by Nylon, PETG, and HIPS. Again, all prototypes are laid horizontally as a constant factor to reduce exposure to weak points and to keep record data without biases (Figure 10).

Young's Modulus of Elasticity can be determined through PSI modules (AccuViewi 1000© tensile machine). Young's Modulus is a measure of elasticity compared to one another; it is the ratio of stress to which the object is exposed to the strain produced. Overall, PLA had the highest Modulus of Elasticity, followed by Nylon, PETG, and HIPS. All prototypes were laid horizontally as a constant factor to reduce exposure to weak points (Figure 11).

Elongation % demonstrates the percentage of elongation based on millimeter measures before and after the material breaks. Elongation % is a post-measurement stress factor. PETG had the most extended elongation difference, followed by PLA, Nylon, and HIPS. All prototypes were laid horizontally as a constant factor to reduce exposure to weak points (Figure 12).

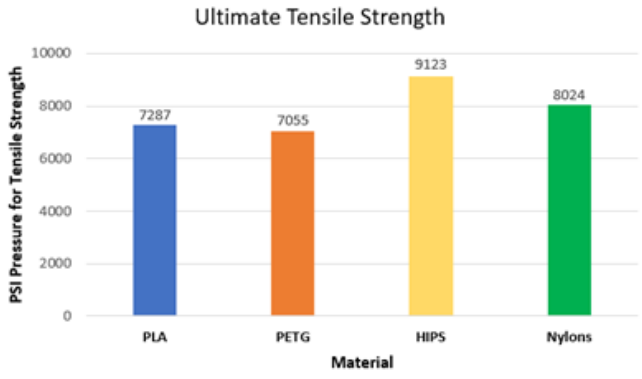

Figure 9: UTS test comparing PLA, PETG, HIPS, and Nylon materials

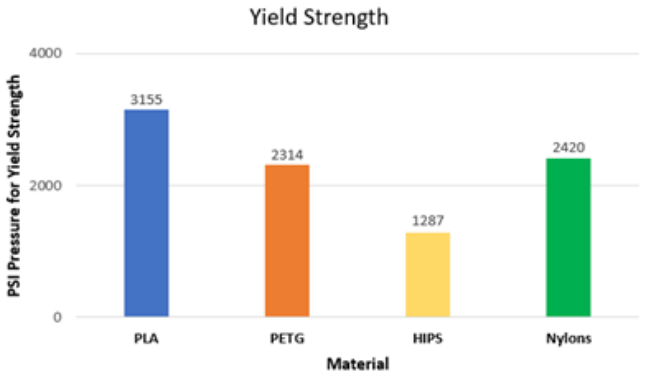

Figure 10: UTS test comparing PLA, PETG, HIPS, and Nylon materials.

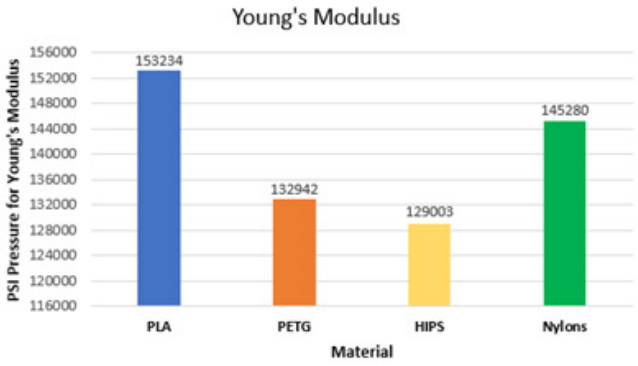

Figure 11: Young's Modulus test between PLA, PETG, HIPS, and Nylon materials.

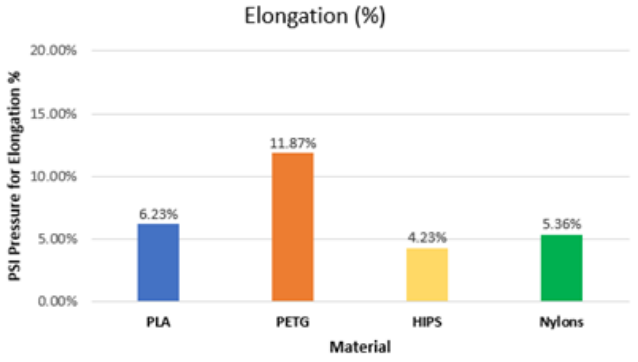

Figure 12: Elongation % measurement between PLA, PETG, HIPS, and Nylon materials.

Discussion of UTS, Yield Strength, Young's Modulus, and Elongation:

AccuView^i 1000 PSI Tensile Tester© provides a venue to examine the different materials through an objective lens. Based on the data for Nylon (8024, 2420, 145280, and 5.36% for UTS, Yield Strength, Young's Modulus, and Elongation %, respectively), Nylon seems the most plausible material to use because of its optimal ranking in all categories. Meanwhile, PETG (7055, 2314, 132942, and 11.87% for UTS, Yield Strength, Young's Modulus, and Elongation %, respectively) seems the least likely material to employ because of its extreme tendency for deformity and elasticity. Both PLA (7287, 3155,

153234, and 6.23%) and HIPS (9123, 1287, 129003, and 4.23%) seem like moderate options.

However, when considering the applications of viability, PETG and PLA seem the most plausible because of their cost and features. For instance, PETG is the only waterproof plastic, proven non-toxic through many literature studies detailing its usage, and mass manufacturable. PLA, meanwhile, has widespread use in prototyping and is extremely popular because of its strength and simplicity---yielding a simple design. In conclusion, PETG and PLA were chosen for the future manufacture of PuriScan because of their availability and real-world implications.

Confusion of PuriScan Predication Labels:

A confusion matrix of the machine learning model displaying true positives, true negatives, false positives, and false negatives is shown (Figure 13). True positives and negatives represent the model correctly predicting whether bacteria are present (positive) and not are present (negative). False-positive and -negatives represent the model incorrectly being present (positive) and not being present (negative).

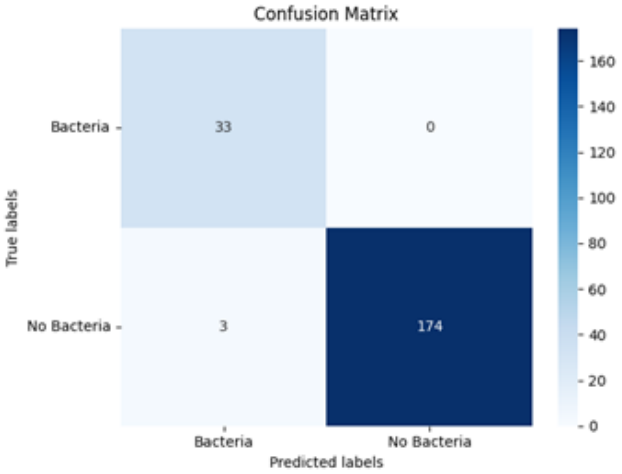

Figure 13: Demonstration of the confusion matrix for PuriScan's code, containing true positives, true negatives, false positives, and false negatives.

Convolution of PuriScan - Layering and Parameters:

Layering convolutions are especially important in determining the parameters the machine learning algorithm may have, and the effects convolution may have on the accuracy and confidence of the model. In total, we had three iterations of Conv2D pooling (resulting in different transformations and scaling), three iterations of MaxPooling (resulting in various transformations and scaling), and one interaction of Flatten, Batch Normalization, and Density consideration. In total, we augmented 3734 new images (Table 3).

Table 3: Demonstration of the confusion matrix for PuriScan's code, containing true positives, true negatives, false positives, and false negatives.

Layer (type)	Output Shape	Parameter #
Conv2D	(None, 99, 99, 32)	416
MaxPooling2D	(None, 49, 49, 32)	0
Conv2D	(None, 48, 48, 16)	2064
MaxPooling2D	(None, 24, 24, 16)	0
Conv2D	(None, 23, 23, 8)	520
MaxPooling2D	(None, 11, 11, 8)	0
Conv2D	(None, 10, 10, 4)	132
MaxPooling2D	(None, 5, 5, 4)	0
Flatten	(None, 100)	0
Batch Normalization	(None, 100)	400
Dense	(None, 2)	202

Confidence of PuriScan:

Confidence in PuriScan is characterized by a logistical regression format in which the machine learning algorithms will calculate the "probability" of the likelihood of a specific categorization of the mentioned image. One hundred eighty-one images had confidence levels between 0.91-1.0, 9 images had confidence levels between 0.82-0.91, 7 images had confidence levels between 0.73-0.82, 0 images had confidence levels between 0.64-0.73, and 3 images had confidence levels between 0.55-0.64 (Figure 14).

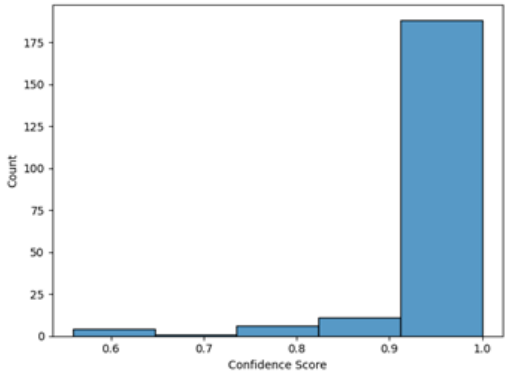

Figure 14: A confidence chart demonstrating a compilation of scores from 200 randomly chosen images from all testing groups (TR, NF, and F). These images were selected from convolutional layers marked as "important," seen in previous parametric equation labeling.

Discussion of PuriScan's Confusion, Convolution, and Confidence (3Cs):

PuriScan's confusion had only three false negatives, garnering a relatively accurate prediction based on confusion data. Meanwhile, 174 were true negatives (which correlated), and 33 were true positives (which correlated as well). Overall, the pathogen detection in our 200-sample test was reliable, our ± 1.5% margin of error demonstrates the reliability of using PuriScan, and the accessibility of PuriScan (returning results in less than 5 seconds) demonstrates the implications of PuriScan for real-world application. PuriScan's confidence (181 characterized in 0.91-1.0, 9 characterized in 0.82-0.91, 7 characterized in 0.73-0.82, 0 characterized in 0.64-0.73, and 3 characterized in 0.55-0.64) demonstrates confidence values in PuriScan's machine learning decision tree. The higher the confidence values, the more likely PuriScan will select these values. The confidence values for PuriScan were likely extremely accurate because of the large number of datasets available and the intricacy of our CNN algorithm.

Compared to existing products like CleanWater A.I. and Intel A.I., PuriScan is roughly 5,200 times cheaper,[44] just as accurate (98-99%)45, and possesses a cloud storage service to track locations of water contamination. Furthermore, PuriScan is roughly 24 times faster in assembly, 113 times faster in manufacturing, and provides real-time analysis, just like CleanWater A.I..[46] Using glass beads rather than convex lenses conceptualized the first microscopes, however, expanding beyond the

DOI: 10.36838/v5i2.9

"archaic" theory has been only accomplished by a few. Their companies did not receive any funding to continue their project.

In conclusion, PuriScan performed well when compared to itself and others. Through our machine learning testing (confusion, convolutional, and confidence), we validated the integrity of PuriScan and the algorithm on which PuriScan runs.

Statistical Analysis:

PuriScan's statistical analysis groups can be separated into major groups of bacteria (for which we have over 500 samples of "true" samples). These samples can be plotted against microscopic detection and PuriScan detection to generate an ANOVA-2-way testing platform to generate an r^2 value (Figure 15).

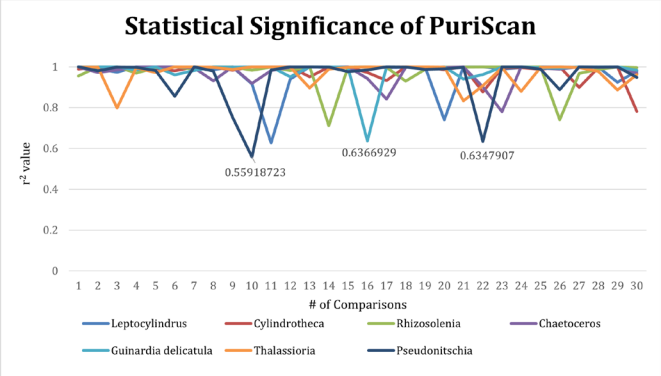

Figure 15: A statistical analysis based on our 7 selected bacterial factors.

Discussion of Statistical Analysis:

PuriScan's statistical analysis showed an extreme correlation between PuriScan's confidence levels and ECHO's analysis. None of the r^2 correlation values was below 0.5, and over 94% of r^2 correlation values were within the range of 0.99-0.999.

From these seven groups of major threatening bacteria, *Leptocylindrus, Cylindrotheca, Rhizosolenia, Chaetoceros, Guinardia delicatula, Thalassioria*, and *Pseudonitschia*, PuriScan reveals with correct certainty that the percentage of confidence and r2 correlation.

Conclusion

Summary:

PuriScan has successfully incorporated a Machine Learning algorithm (with CNN and Deep Learning algorithm), functional magnifyable prototype/cartridge, and MAP feature. PuriScan revitalized a pre-existing method of magnification (with our novel design for the borosilicate glass beads container and lens) for microbe magnification. PuriScan's filaments (PLA and PETG) demonstrated high UTC, Yield, Young's Modulus, and Elongation % Values: for PETG, 7055, 2314, 132942, and 11.87%; while for PLA, 7287, 3155, 153234, and 6.23% for all above factors, respectively. PuriScan's pure confusion and confidence values, based on logistical analysis, were 99.2% & 99%, respectively, and PuriScan's margin of error was roughly ±1.5% deviation. PuriScan's general trend had a 94% correlation rate between PuriScan's microscope prototype and the laboratory ECHO microscope. PuriScan is expected to cost roughly $0.10 for the cost of filament (PLA or PETG), borosilicate glass beads (1mm x 1mm), and stickable mirror pieces (5 per prototype).

Future Directions:

PuriScan can be expanded to encompass viruses or other pathogens soon. While new beads (of smaller sizes) will have to be made, the existing backbone machine-learning algorithm for bacteria and protozoa is already present; thus, changing the parameters to fit viral detection will not be as complicated.

PuriScan can also be patented and branched, as it is a novel design for the iPhone clip-on. Additionally,

PuriScan can be distributed in developing countries that do not yet have existing clean water infrastructure. Lastly, PuriScan can be programmed into its own water-based A.I. detection program, which will be exclusively for detecting pathogens in water.

Acknowledgments

The author thanks the Biomedical Engineering Department (BME) of North Carolina State University for allowing us to collect & image water samples and translate the data into our machine learning algorithm. We thank Dr. Suleman Asif for his help in 3D printing. The author also thanks Dr. Rohan Shirwalker from North Carolina State University for helping us understand convolutional neural networks and Dr. Wubin Bai from the University of North Carolina Chapel Hill for helping us with computer-aided design (CAD).

References

i. Twin-column materials tensile tester: Mecmesin: Tensile tester. Twin-column materials tensile tester | Mecmesin | Tensile tester. (n.d.). https://www.tensiletester.com/products/omnitest-twin- column-materials-tensile-tester

1. "Drinking water." 2022. WHO | World Health Organization.
2. "Earth's Freshwater." 2012. National Geographic Society. https://www.nationalgeographic.org/media/earths-fresh-water/.
3. "World Water Day Reminds Us of the Value of a Precious Resource | United Nations." n.d. the United Nations
4. "Water in developing countries." 2017. Global Affairs Canada.
5. "WHO: Waterborne Disease is World's Leading Killer." 2009. Voice of America.
6. "Reading: Indicators for Waterborne Pathogens." n.d. The National Academies Press. 2022. https://nap.nationalacademies.org/read/11010/chapter/6#167
7. Castillo-Henriquez L., Brenes-Acuna M., Castro-Rojas A., Cordero-Salmeron R., Lopretti-Correa M., and Vega-Baudrit J. R. "Biosensors for the Detection of Bacterial and Viral Clinical Pathogens." n.d. MDPI. 2020. https://www.mdpi.com/1424-8220/20/23/6926/htm.
8. "Quality Unknown: The Invisible Water Crisis." 2019. World Bank Group. https://www.worldbank.org/en/news/feature/2019/08/20/quality-unknown.
9. Aidukas, T., Eckert, R., Harvey, A.R., Waller L., Konda P. C. "Low-cost, sub-micron resolution, wide-field computational microscopy using opensource hardware." Sci Rep 9, 7457. 2019. https://doi.org/10.1038/s41598-019-43845-9.
10. Diederich, B., Lachmann, R., Carlstedt, S., Marsikova B., Wang H., Uwurukundo X., Mosig A. S., & Heintzmann R. "A versatile and customizable low-cost 3D-printed open standard for microscopic imaging." Nat Commun 11, 5979. 2020. https://doi. org/10.1038/s41467-020-19447-9.
11. Liu, Y., Rollins, A.M., Levenson, R.M., Fereidouni F., Jenkins M. W. "Pocket MUSE: an affordable, versatile and high-performance

fluorescence microscope using a smartphone." Commun Biol 4, 334. 2021. https://doi.org/10.1038/s42003-021-01860-5.
12. "Membrane Filtration." n.d. National Drinking Water Clearing House. 1999. https://www.mrwa.com/WaterWorksMnl/Chapter%2019%20Membrane%20Filtration.pdf.
13. Berk, Zeki. "Membrane Configuration - an overview | Science Direct Topics." Food Process Engineering and Technology. 2009. https://ww.sciencedirectcom/topics/engineering/membrane-configuration
14. Charcosset, Charlie. "Microfiltration - an overview | ScienceDirect Topics." Membrane Processes in Biotechnology and Pharmaceutics. 2020. https://www.sciencedirect.com/topics/chemical-engineering/microfiltration
15. Ariono D., Aryanti P. T. P., Subagjo S., and Wenten I. G. "The effect of polymer concentration on flux stability of polysulfone membrane." AIP Conference Proceedings, vol. 1788, no. 1, 2017, pp. 1-10. DOI: https://doi.org/10.1063/1.4968301
16. "Membrane Resistance - an Overview | ScienceDirect Topics." n.d. 2020. https://www.sciencedirect.com/topics/chemical-engineering/microfiltration
17. Sun, Dongyang. "Pull in and Push Out: Mechanisms of Horizontal Gene Transfer in Bacteria." Frontiers. 2018. https://www.frontiersin.org/articles/10.3389/fmicb.2018.02154/full
18. Batra, Ekta. 2021. "Microfiltration in Food Industry." PMG Engineering.
19. Yang Z., Zhou Y., Feng Z. Y., Rui X. B., Zhang T., and Zhang Z. "A Review on Reverse Osmosis and Nanofiltration Membranes for Water Purification." NCBI. 2019. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6723865/.
20. Bhalla N., Jolly P., Formisano N., and Estrela P. "Introduction to biosensors - PMC." NCBI. 2016. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4986445/.
21. MacDonald, Anna. "Types of Biosensors and Evolving Applications: Glucose Monitoring to Personalized Health." Technology Networks. 2017.
22. "Biosensor - Principle, Components, Types & Their Applications." n.d. ElProCus. https://www.elprocus.com/what-is-a-biosensor-types-of-biosensors-and-applications/.
23. Fletcher M., Camara M., Barett D., and Williams, P. "Biosensors for qualitative and semiquantitative analysis of quorum sensing signal molecules." PubMed. 2014. https://pubmed.ncbi.nlm.nih.gov/24818910/.
24. Andrade C., Oliveira M., Faulin T., and Hering V. "Biosensors for detection of Low-Density Lipoprotein and its modified forms." 2011. DOI: 10.1007/978-1-4939-0473-0_20. PMID: 24818910.
25. Klapper, Leora. "Mobile phones are key to economic development. Are women missing out?" Brookings. 2019.
26. Svensson, Oskar, and Marcus Presa. n.d. "React Native and native application development." DiVA-Portal. 2022.
27. Ibanez, Lazaro. "Development: Difference between the HTTP requests POST and GET." Lazaro Ibanez. 2022.
28. Vrijraj, Singh. "Introduction to Firebase." Medium. 2018.
29. "How Much Does A 3D Printing Filament Cost? 3D Printer Filament" 2022. https://www.fusion3design.com/how-much-does-3d-printing-filament-cost/.
30. "Basics of Fused Deposition Modelling (FDM) - PMC." NCBI. 2020.
31. Xu X. Y., Robles-Martinez P., Christine M. M., Joubert F., Goyanes A., Basit A. B., and Gaisford S. "Stereolithography (SLA) 3D Printing of an Antihypertensive Polyprintlet: Case Study of an Unexpected Photopolymer-Drug Reaction." Additive Manufacturing. 2020. https://doi.org/10.1016/j.addma.2020.101071.
32. Gueche Y. A., Sanchez-Ballester N. M., Cailleaux S., Bataille B., and Soulairol I. "Selective Laser Sintering (SLS), a New Chapter in the Production of Solid Oral Forms (SOFs) by 3D Printing." NCBI. 2021. https://www.ncbi.nlm.nih.gov/pmc/articles/PMC8399326/.
33. Boschmans B., Royen P. V., and Vaeck L. V. "Use of Monoatomic and Polyatomic Projectiles for the Characterization of Polylactic Acid by Static Secondary Ion Mass Spectrometry." Rapid Communications in Mass Spectrometry. 2005.
34. Ksawery S., Credo W., and Osiński D. "Selected Mechanical Properties of PETG 3-D Prints." Procedia Engineering, vol. 177, pp. 455–61. 2017. https://doi.org/10.1016/j.proeng.2017.02.245.
35. Vinita V. S., Shelke S. D., Celestine A. N., and Beckingham B. S. 2021. "Self-Healing in High Impact Polystyrene (HIPS) Composites via Embedded Non-Toxic Solvent-Filled Microcapsules." Journal of Applied Polymer Science, vol. 139, no. 2, pp. 51463. 2017. https://doi.org/10.1002/app.51463.
36. Roberson, David. "How much does 3D printing cost?" Ultimaker. 2021.
37. "How Much Does a 3D Printing Filament Cost?" 3D Printer Filament. 2017.
38. "Amazon.com: Adamas-Beta Solid Round Clear Glass Boiling Stones Beads, 2mm Diameter, Approx 1000 Beads: Industrial & Scientific." 2022.
39. "Amazon.com: 2400pcs Self-Adhesive Mini Silver Square Glass Decorative Craft DIY Accessory Mirrors Mosaic Tiles 5mm by 5 Mm: Arts, Crafts & Sewing." 2022.
40. Sosik H. M., Peacock E. E., Brownlee E. F. "2014 labeled IFCB images." 2014. https://hdl.handle.net/1912/7350
41. Gandhi, Arun. "Data Augmentation | How to use Deep Learning when you have Limited Data." Nanonets. 2021.
42. Saha, Sumit. "A Comprehensive Guide to Convolutional Neural Networks — the ELI5 way." Data Science. 2021.
43. "Unsupervised Feature Learning and Deep Learning Tutorial." Unsupervised Feature Learning and Deep Learning Tutorial. 2022.
44. "AI-Driven Test System Detects Bacteria in Water." Intel. 2018.
45. Ma, Peter. "Clean Water AI." Devpost. 2018. https://devpost.com/software/clean-water-ai.
46. "Clean Water AI." Hackster.io. 2018.

Author

George Cheng is a junior at North Carolina School of Science and Mathematics in Durham, North Carolina. He has worked on other sensor technologies and wearable devices before and started PuriScan in 2021. He hopes PuriScan can provide equitable access to water-based detection methods, especially those living in water-limited areas in the world.

DOI: 10.36838/v5i2.9