OpenJuice (OJ): Multispectral sensor spectrophotometer for fruit juice quality inspection

Kenton Kwok

***Abstract*—A multispectral sensor-based transmission spectrophotometer was constructed based on a fruit-juice company's needs to classify and determine the concentration of their beverages. The OpenJuice (OJ) device was designed to be low-cost and made following open-source hardware principles from easily obtainable parts and materials available in typical makerspaces.**

The device was built around the 14-channeled AS7343 visible and near-infrared sensor and a white LED. The device operated in transmission mode, whereby the sensor measures how much light from the LED light is absorbed due to the introduction of the sample into the optical path. Data was collected using an Arduino Nano board and was recorded by a Python program on a personal computer. Testing is quick < 1 second and only < 5 mL of sample is required for measurement.

The performance, including the accuracy, and uncertainties were evaluated based on 4 representative juices (blackcurrant, orange, apple and tomato). Using a nearest-neighbour algorithm on synthetic spectral signals generated from experimentally derived calibration curves, the best juice classification accuracy achieved was 95 % and the mean absolute error on the concentration regression task was 9 % over a test set of $N = 22$ samples.

The bill of materials of OJ costs £54 and all code, design files and construction instructions are provided for replication at https://github.com/kwokkenton/spectrometer and in the Appendix.

I. Introduction

SPECTROMETERS were developed by various scientists such as Newton, Wollaston and Fraunhofer in the 17th century, and has led to tremendous developments in astronomy, atomic physics and chemistry [1]. Modern day spectrophotometers utilise the same principle of splitting light into its constituent wavelength components and are applied in the pharmaceutical, food and beverage industries.

This report details the design, construction, and testing of a low-cost transmission spectroscopy device centred around a filter-based spectral sensor to "test small samples of fruit juices" for a fruit juice manufacturer. As quality control and contamination detection is essential for maintaining product consistency and food safety, instruments like this can help raise standards and potentially reduce product recall rates.

A. Physics of Transmission spectroscopy

Photons passing through a substance undergo absorption, scattering and re-emission probabilistically, depending on the wavelength of light and the chemical or physical composition of the medium. A portion of the light will be transmitted through the sample and is shown in Figure 1. The Beer-Lambert law governs light attenuation through homogeneous and non-turbid substances

$$I_\lambda = I_{0,\lambda} 10^{-\epsilon_\lambda l c} \tag{1}$$

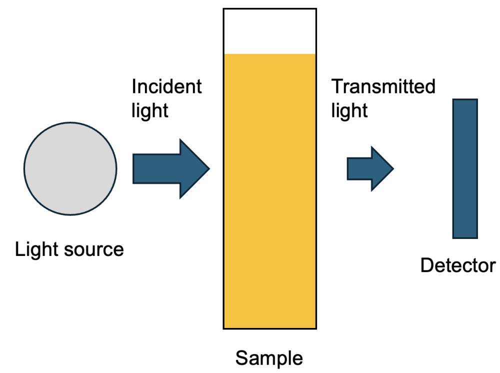

Fig. 1. Schematic of transmission spectroscopy, where the detector is configured to measure the amount of light passing through the sample.

where I is the measured signal and I_0 is the reference signal at sampling wavelengths λ, ϵ_λ is the molar decadic absorption coefficient at λ, l is the optical path length through the sample and c is the molar concentration [2].

B. Detection Mechanism

The multispectral sensor has wavelength-selective optical filters deposited on a silicon chip that function based on thin-film interference as shown in Figure 3. Photons that are captured by the CMOS device get converted into a charge, which is subsequently amplified, digitised and read out as an electric signal. As each sample absorbs light differently across wavelengths, we can use the signal to estimate the identity of the substance.

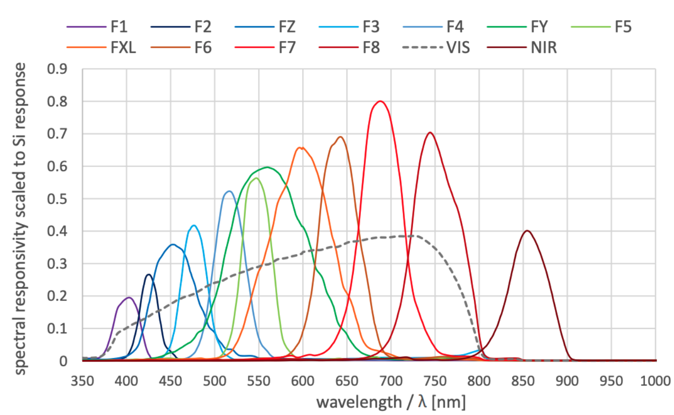

Fig. 2. Transmission efficiencies of the spectral filters on the AS7343. The bandwidths (full-widths at half-maximums) range from 22 - 100 nm. All channels were used except VIS [3].

II. Device Design

The brief required the creation of a spectrophotometer that was low cost, easy to construct, easy to use and fit for purpose. The following section attempts to justify the key choices for making the device.

A. Component choice

With the myriad ways of building a spectrometers using dispersive elements such as prisms, diffractive elements (gratings), a multispectral sensor module was chosen as a good balance between cost, complexity and ease of construction.

The sensor chosen was the commercially available AS7343, which detects 14 distinct spectral channels from 380 -900 nm with 11 in the visible range, 1 in the infrared range and 1 clear channel. We used a sensor responsive in the visible range due to its low cost and suitable wavelength range, as fruit juices differ in absorbance in the visible range (as evidenced by the human eye) [3].

The Arduino Nano 33 BLE was chosen as the MCU board due to its wide availability, good developer ecosystem and support, as well as low cost.

A standard white light emitting diode (LED) was chosen to match the wavelength range of the sensor. As white LEDs have two distinct peaks (one from the blue LED emission and also another from the phosphor emission), a drawback here would be the lack of smoothness compared to Xenon or tungsten lamps in high-end devices, which reduces sensitivity in particular spectral regions.

The body was chosen to be 3D printed using Fused Deposition Modeling (FDM), as these types of printers are sufficiently accurate for our parts and the fact that they are widely available in most makerspaces [4]. The material chosen was polylactic acid (PLA) for its strength, low cost and good accuracy. Although PLA begins softening at 50 °C, this was not a concern for us as we expected to operate at room temperature. Black PLA was chosen to reduce the amount of light reflected by the surfaces, which will interfere with the signals detected.

A 1 cm path length plastic cuvette (instead of quartz) was chosen for cost and availability and also as our measurements did not require UV measurements [5].

B. Mechanical design

The device was chosen to have a basic functional geometry, with the illumination, sample and detector collinear on the principal axis, similar to Figure 1. The LED was designed to be at the same height as the detector, centred at the cuvette to ensure that the optical distance travelled was the shortest path. For protection, the device was housed in a case that also contained the Arduino board and the connection wires. The custom-made LED breadboard was fastened to the device to ensure mechanical robustness. The size of the entire device was 5 cm x 13 cm x 5.5 cm, which is small, portable and can be placed on any desk.

C. User interface design

For usability, a lid was was included at the top so that the cabling is hidden from the user during routine use. The window reveals the compartment that holds the cuvette, which makes it simple and intuitive to use.

The device was intended to be interfaced with a personal computer via a micro-USB cable. As personal computers are ubiquitous, this minimises the cost of the system itself and does not require additional peripherals or displays.

The user records data through the software via a graphical user interface (built with PyQt) for maximal usability. The interface exports a .csv file for compatibility, which can be analysed subsequently in a program of their choice. Calibration datafiles are preloaded for automated classification and regression of juices and their concentrations.

Item	Quantity	Supplier	Cost (£)
Arduino Nano 33 BLE	1	Arduino	24.90
AS7343 Multi-Spectral Sensor	1	Pimoroni	18.90
M/F Dupont Jumper wires	4	SBComponents	< 1
Electric jumper wires	2	SBComponents	< 1
M2 self tapping screws	4	Screwfix	< 1
170 point solderless mini breadboard	1	4tronix	1.50
72 x 47mm Copper Prototyping PCB	1	Switch Electronics	0.91
Resistor 150 Ω	1	RS Electronics	< 1
3mm White LED	1	Pimoroni	0.13
1 cm pathlength macro Cuvette	1	ThermoFisher	< 1
3D printed frame, black PLA	80.34 g	BambuLabs	< 1
MicroUSB cable	1	Amazon	< 1
			53.54

TABLE I
BILL OF MATERIALS OF THE SPECTROMETER (PARTS)

Bambu Labs FDM Printer e.g. X1 Carbon, A1
Phillips screwdriver
Electric Drill with 2.3 mm drill bit
Soldering Iron
Leadless solder
Pliers

TABLE II
ADDITIONAL TOOLS AND MATERIALS REQUIRED

III. Evaluation

The device's ability to measure spectral signals was tested using a set of 4 juices detailed in Table III.

We offer three analyses that pertain to three different use-cases of the device. Firstly we describe the measurement of a single spectral signal, then the construction of a standard calibration curve, then finally the performance when using the device directly as an automated juice classifier and concentration regressor.

Juice	Characteristics
Tomato	Bright red, turbid
Blackcurrant	Purple, translucent
Orange	Orange, turbid
Apple	Brown, clear

TABLE III
JUICES AND THE DEVICE'S MEASUREMENT RANGE AND UNCERTAINTY

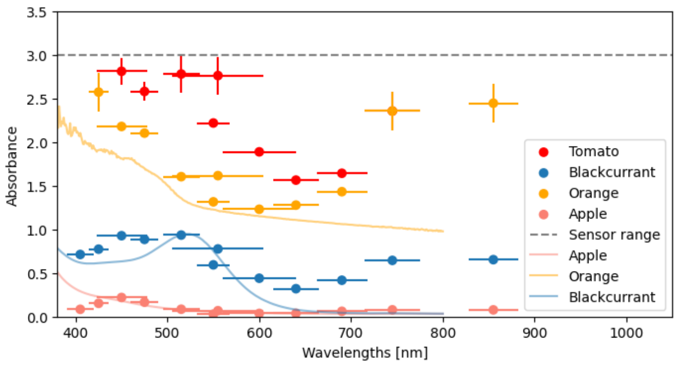

Fig. 3. Plots of absorbance of the 4 undiluted fruit juices measured with the OJ multispectral device (scatterpoints), with points centred at the central wavelength of the channels. The solid lines refer to the continuous spectra collected by the commercial UV-vis spectrophotometer. Due to the dynamic range of the sensor, the maximum absorbance is 3.0 (1000 counts max), which is shown by the grey dotted line.

A. Measuring spectral absorbance

It is customary to convert raw measurements I_λ (in counts) from the spectral sensor into absorbance to make signals comparable across devices. The spectral absorbance A_λ of a sample is given by the formula

$$A_\lambda = -\log_{10}(I_\lambda / I_{0,\lambda})$$

It can be shown that the uncertainty in the absorbance is the sum of the fractional uncertainties of the reference and measured signals

$$\Delta A = \frac{1}{\ln 10}\sqrt{(\Delta I/I)^2 + (\Delta I_0/I_0)^2}$$

We choose ΔI to be the standard deviation of the counts, which is an appropriate frequentist estimate of uncertainty of the signal value.

Figure 3 shows the plotted values and the associated uncertainties. As expected, the more turbid the juice, the higher the absorbance. As the absorbance increases, the value's uncertainty also increases as the counts are low. The uncertainty in the wavelength dimension is plotted to illustrate the representative bandwidth of the.

To qualitatively investigate the accuracy of our instrument, the spectral absorbances of several measurements using a commercial UV-VIS spectrometer (Agilent Cary 3500 multicell, bandwidth 2 nm) are superimposed on Figure 3. It can be observed that the trends and magnitude of the absorbances are similar, though the peaks (see blackcurrant) are much less pronounced. This agrees with the fact that the spectral filters are broadband of the spectral filters.

B. Estimating sample concentration with calibration curves

A calibration curve is a plot of the substance's absorbance when its concentration is varied. Its inverse can effectively be used as a lookup table to determine the concentration, provided the absorbance. The objective of this section is to determine a suitable calibration curve and its uncertainty.

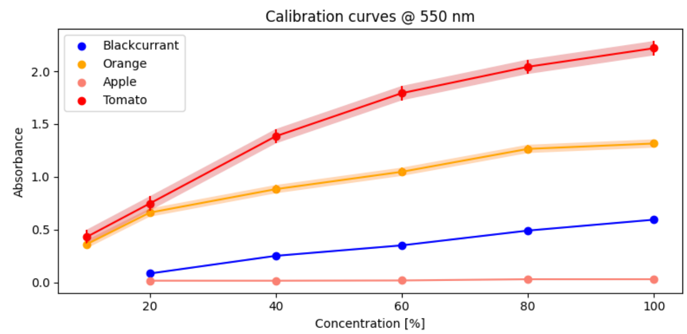

Fig. 4. Example calibration curves based on a single channel at 550 nm for the 4 juices. The uncertainties plotted are based on the difference between a linear model

Bambu Labs FDM Printer e.g. X1 Carbon, A1
Phillips screwdriver
Electric Drill with 2.3 mm drill bit
Soldering Iron
Leadless solder
Pliers

TABLE IV
ADDITIONAL TOOLS REQUIRED

The standard way of obtaining a calibration curve is to determine the magnitude of suitable feature, such as by fitting a curve to a prominent absorption peak [6]. As no obvious peaks are present in our coarse spectral measurements, we used each channel of the spectral sensor obtain separate calibration curves. Figure 4 shows this for the 4 juices at a single wavelength.

The uncertainty of the liquid handling process is much smaller than the variation in the absorbance measurement. As an example, using a pipette with error of ± 0.2 mL and creating 40 mL total volume of juice at 10% concentration gives a final measurement of (10.0 ± 0.6) %.

So far, we only have dealt with **aleatoric** uncertainty, which is the intrinsic randomness in measurements. Determining the uncertainties on the curves represent **epistemic** uncertainty [7], which is dependent on the model we assume between the datapoints. The Beer-Lambert law predicts that absorbance should vary linearly with concentration. This, however was not observed in the 2/4 juices that were turbid, and so prohibited a linear fit.

The following section presents methods of using these calibration curves without the Beer-Lambert assumption of linearity for juice classification and concentration regression.

C. Simultaneously inferring sample identity and concentration

Two algorithms are proposed to determine concentration, depending on whether the juice's identity has been provided or not in advance.

1) Using calibration curves as lookup tables (LUTs): Assuming we have been provided the juice's identity, we can use each channel's calibration curve as separate LUTs (after interpolation) to estimate the concentration c_i, which can then be averaged together.

Algorithm	Classification accuracy	Regression MAE %
Calibration curve LUT	/	11 ± 12
NN: Spline	0.86	10 ± 10
NN: Linear + Exclusion	0.95	10 ± 10
NN: Spline + Exclusion	**0.95**	$\mathbf{9 \pm 9}$

TABLE V

RESULTS FROM AUTOMATED JUICE CLASSIFICATION AND CONCENTRATION REGRESSION FROM 2 ALGORITHMS (LUT, NN) AND THEIR VARIATIONS

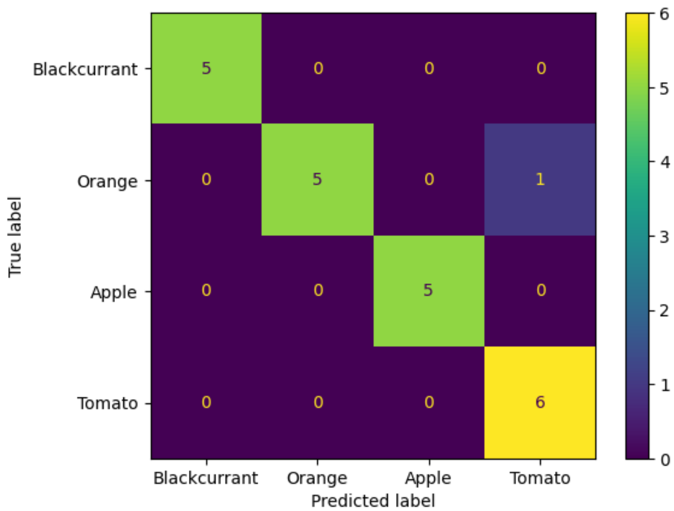

Fig. 5. Classification results for NN: Spline + Exclusion. Most juices were classified correctly, except one orange sample that was confused as Tomato.

2) 1-nearest neighbour (NN): If we haven't been given the juice's identity, it is not possible to know which calibration curve to use. Instead, we use a 1-nearest neighbour algorithm [8] using synthetic data as follows

1) Data preparation: Using the calibration curves, generate synthetic absorbance spectral measurements A_λ interpolating between the minimum and maximum concentration value every 2%. Tabulate these results and associated labels of juice identity j and concentration c.
2) Classification: Use the 1-nearest-neighbours algorithm to find the spectral curves with the closest Euclidean distance to the newly provided absorbance spectra amongst the tabulated results to infer the label (j, c).

Using a separate test dataset collected on a different day, we determined the **generalisation error** of the device and associated computational algorithms.

We experimented with linear interpolation (Linear) and cubic splines (Spline) for both methods and documented the results in Table V. We also found that excluding the channels centred at 405, 425 and 855 nm improved performance (Exclusion) as for some juices light was completely attenuated by the sample and so the device's raw counts was 0.

The best classification results were high at 0.95, while the mean absolute error (MAE) was 9 %. The error in the MAE was skewed by an incorrect prediction of the sample class (Orange, Tomato), which resulted in a large difference in the estimation of the concentration. This was expected as both are turbid liquids with an orange/ red colour when diluted.

D. Constraints and Improvements

The current device exhibits high performance to the design specification. Several mechanical improvements that would improve usability include making snap-close mechanisms on the lid and the base plate and to add further compartments to hide the electrical components from view. In addition, the LED breadboard was tricky to construct and a next iteration could use an LED module instead for quicker construction.

A constraint of the system is that if juices display **metamerism** in the multispectral channels, it could be hard, even impossible to distinguish between their identities. This was seen when the device misclassified orange and tomato juice. One improvement could be adding reflectance signal measurement by measuring backscattered light from front-illumination of the sample. This could improve classification accuracy for opaque juices. A big upgrade would be to opt to make a grating spectrometer, which is cheaper than a prism-based spectromer, but would increase the complexity of the system. This however would increase the spectral resolution and can potentially lead to higher performance.

The next constraint would be the low signal around short wavelengths, partially due to weak illumination and also that it is highly scattered by the samples tested. An improvement would be to incorporate a high-dynamic range acquisition algorithm (by either varying integration time, or using a variable resistor with the LED) to improve the signal quality.

IV. CONCLUSION

The OpenJuice spectrophotometer device was constructed based on a visible/ near-infrared multispectral sensor and white LED to determine the identity and concentration of fruit juices. All code, data and documentation was made open-source to facilitate easy reproduction.

APPENDIX A
DOCUMENTATION

The following documentation details the construction and use of the OpenJuice device. It has been written with replicability in mind. **Software installation instructions are located in the readme in the GitHub repository**, whilst **electronic and mechanical construction, and practical usage instructions are located here**.

A. Assembly instructions

1) Make the mechanical enclosure: Using the 3D printer of choice (suggested Bambu Labs X1/ A1), print the components using the included .stl files in black PLA. Use a 0.4 mm nozzle, set to 15% sparse infill using a gyroid pattern and print with black PLA. No brims were required. There are four parts, the main body, the lid, the window cover, and the back plate for the sensor, which is shown in Figure 6.

2) Make LED board using a the stripboard: Cut down the piece of copper prototyping PCB to 2.5 x 3.5 cm using a saw. Place the LED diode in the middle and connect the resistor and wires appropriately according to the circuit diagram. Solder the components in place to the prototyping stripboard. Drill two holes (size 2.3 mm) according to Figure 9. This is shown in 8 (a,b).

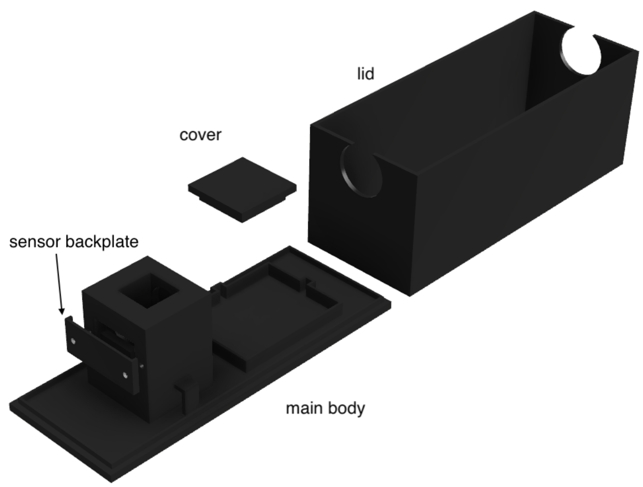

Fig. 6. 3D printed Mechanical enclosure for the device, showing 4 separate components.

3) Fasten sensor and illumination: Using 2 self tapping screws, secure the LED board to the standoffs on the side of the cuvette holder on the main body closer to the centre via the two drilled holes. The LED should be inserted into a hole in the cuvette holder. This is shown in Figure 8 (b,c).

Place the multispectral sensor into the indent in the opposite side of the cuvette holder and secure to the main body using the back plate and 2 self tapping screws. This is shown in Figure 8 (d,e,f).

4) Electrical Connections: Ensure that the Arduino has associated headers attached and plug it into the solderless breadboard. Orient the Arduino and breadboard such that the microUSB port faces away from the cuvette. Follow Figure 7 to connect the pins of the spectral sensor to the Arduino, tucking the wires into the wire holder at the side of the cuvette holder. Solder the wires on the LED board to the appropriate pins on the Arduino shown in Figure 7. Plug the micro USB cable into the Arduino and cover the device with the lid and window cover. The fully constructed assembly is shown in Figure 8 (g, h).

5) Software installation: Clone the GitHub repository.

Install the Arduino IDE via the internet on a personal computer. Connect the Arduino Flash the file ‘spectrometer.ino‘ onto the Arduino board and run the file. Open the serial monitor and then you should see the white LED turn on and sensor values being printed. If not, check all the connections are made correctly.

Ensure that Python is installed on a personal computer and then install the relevant requirements. Modify the script such that the Nano 33 BLE is addressed by the app via the a suitable COM port. Run *gui.py* for the applet. This has been tested on MacOS Sequoia 15.1.1 but is expected to work on other Operating Systems.

B. Regular Usage instructions

The following instructions are for regular usage of the device. A screenshot of the GUI is provided in Figure 10.

1) Run the graphical user interface *gui.py*
2) Pipette 4.5 mL of the water in the cuvette, put cuvette into the device as shown in 8 (i).
3) Cover instrument up so no ambient illumination is captured by the device
4) Set as a ‘blank’ measurement on the GUI
5) Remove cuvette via the window, rinse and dry if necessary
6) Repeat steps 2) and 3) for your sample of choice
7) Change display to toggle the display of raw counts versus the signal normalised by the blank measurement
8) The system will automatically classify the identity and find its concentration of the juice based on its absorbance using the **NN: spline + exclusion** algorithm
9) To save data, press Save data to CSV.
10) Repeat steps 5) and 6) for new measurements

C. Calibration

A method to generate the calibrations for the app has been provided in the GitHub repository. These are the instructions followed so that the user can replicate the study.

Select the juices relevant to the user, here blackcurrant, apple, orange juices were chosen (supplier: Sainsbury’s, UK).

1) Perform serial dilutions of the juices in question at concentrations 10, 20, 40, 60, 80% and keep the undiluted juice at 100%. We chose to make up 40 mL of the liquid in a falcon tube and varied the relative volumes of water and juice to perform dilutions.
2) Save each sample to a csv following instructions in ’Regular usage instructions’
3) Using the data in the csv, generate a database of synthetic spectral signals following the code in *model.py*
4) Use new database for juice classification and regression

D. Troubleshooting

1) If the LED doesn’t turn on: Check the LED’s poles are connected the right way round, the positive pole should be connected to the digital pin via a wire and the negative pole should be connected to ground on the Arduino via a jumpter wire.
2) If device doesn’t turn on: Flash the Arduino script *spectrometer.ino* again onto the Arduino board and open the serial monitor to see whether there is a signal. Check that the Arduino’s indicator lights are on.
3) If the python GUI does not display anything, check the COM port is correctly configured, which is shown in the Arduino IDE.

References

[1] N. C. Thomas, “The early history of spectroscopy,” *Journal of Chemical Education*, vol. 68, no. 8, p. 631, Aug. 1991, publisher: American Chemical Society. [Online]. Available: https://doi.org/10.1021/ed068p631

[2] T. G. Mayerhöfer, S. Pahlow, and J. Popp, “The Bouguer-Beer-Lambert Law: Shining Light on the Obscure,” *Chemphyschem*, vol. 21, no. 18, pp. 2029–2046, Sep. 2020. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC7540309/

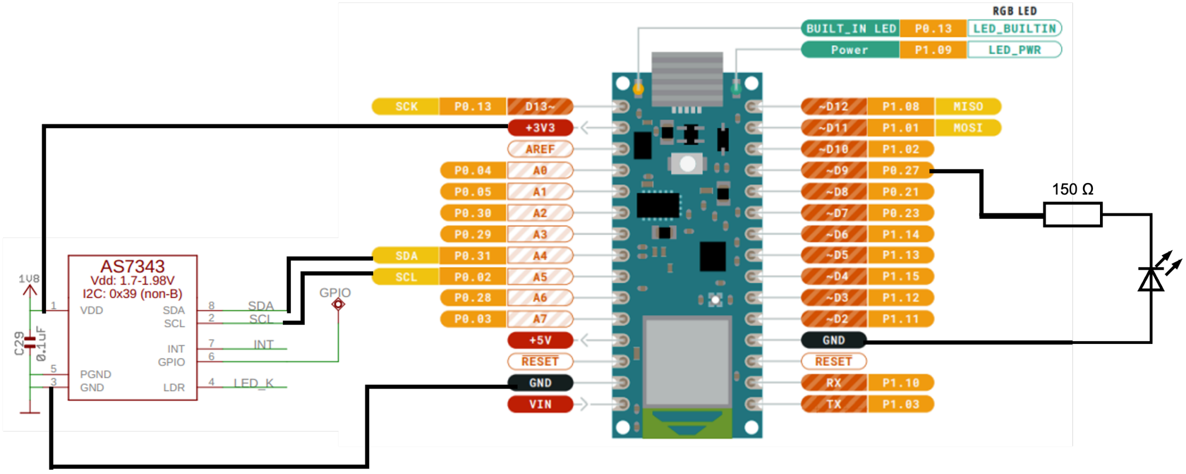

Fig. 7. Circuit schematic showing the connections of the AS7343 board and the custom LED with the Arduino. The Arduino is then connected via a microUSB cable to a personal computer for power and data transfer.

[3] "ams as7343 spectral sensor ambient light, color, spectral proximity sensors - ams-osram - ams." [Online]. Available: https://ams-osram.com/products/sensor-solutions/ambient-light-color-spectral-proximity-sensors/ams-as7343-spectral-sensor

[4] "Choosing a 3d printer (including fdm vs sla) — makerspace." [Online]. Available: https://www.ucreatestudio.is.ed.ac.uk/support_docs/choosing_3d_printer_including_fdm_vs_sla

[5] "Which cuvette is the right one? glass vs. plastic, vis vs. uv, micro-volume vs. macro-volume - eppendorf united kingdom." [Online]. Available: https://www.eppendorf.com/gb-en/lab-academy/lab-solutions/other/which-cuvette-is-the-right-one-glass-vs-plastic-vis-vs-uv-micro-volume-vs-macro-volume/

[6] V. Barwick, *Preparation of Calibration Curves*, ser. Valid Analytical Measurement, 2003, no. LGC/VAM/2003/032.

[7] A. D. Kiureghian and O. Ditlevsen, "Aleatory or epistemic? Does it matter?" *Structural Safety*, vol. 31, no. 2, pp. 105–112, Mar. 2009. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167473008000556

[8] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, vol. 13, no. 1, pp. 21–27, 1967.

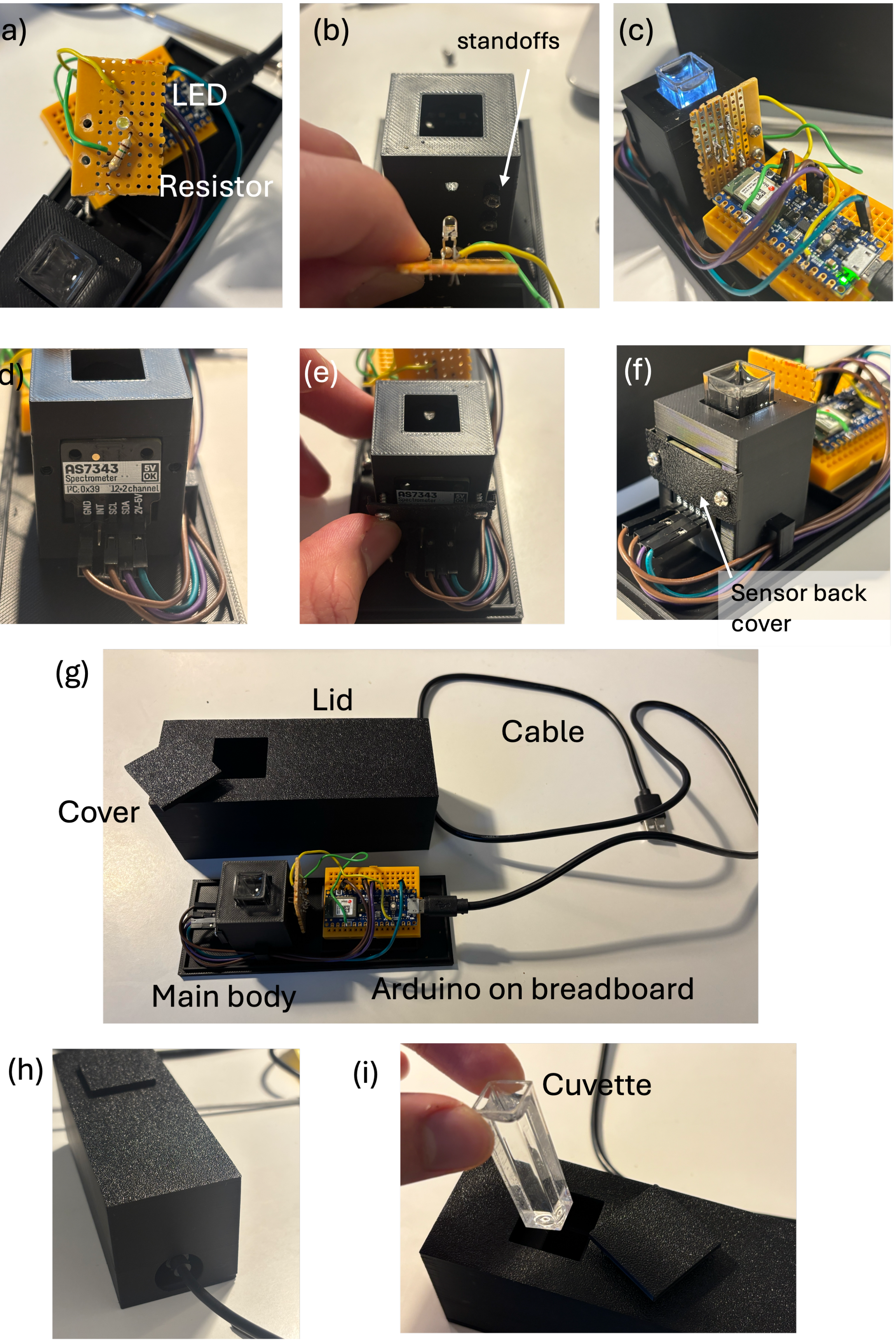

Fig. 8. (a,b,c) Shows the connection of the custom LED board to the main body. (d,e,f) shows the fastening of the multispectral sensor to the main body using with the sensor back cover. (g) shows the fininished assembly. (h) shows the outlet of the microUSB cord from the device. (i) shows example usage of the device via inserting a cuvette via the window on the lid and subsequently covering it wih the square cover.

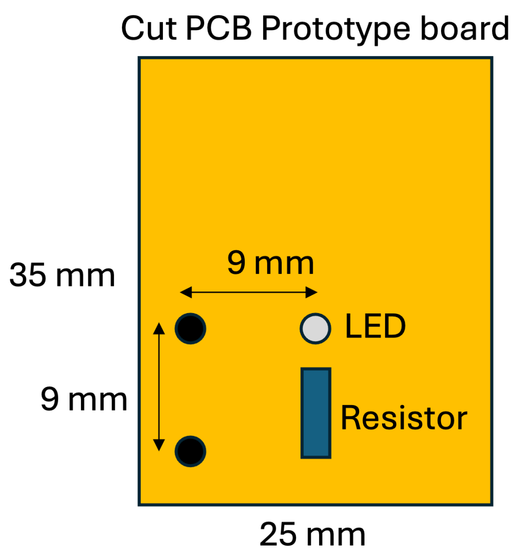

Fig. 9. LED breadboard construction (viewed from front)

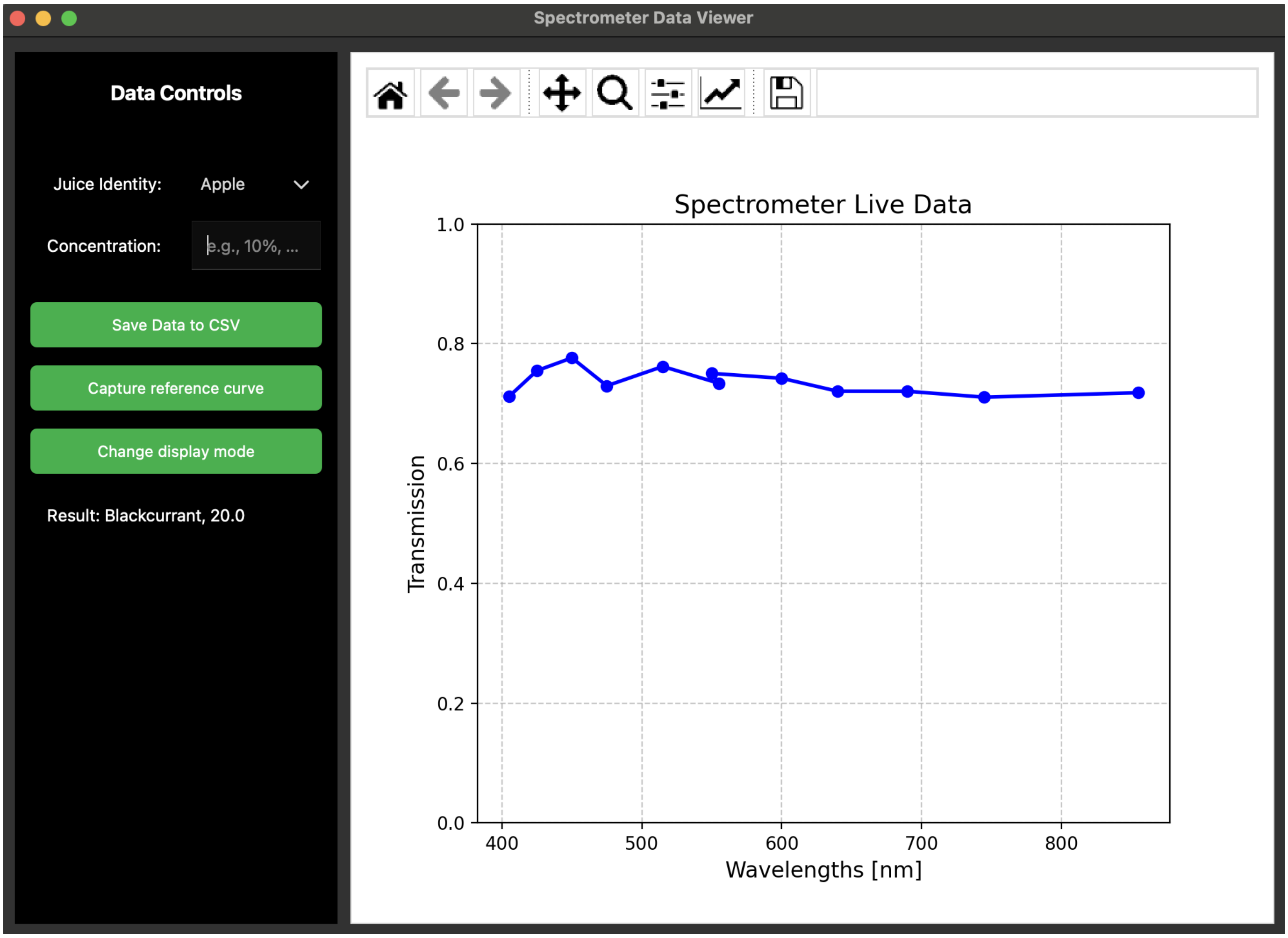

Fig. 10. Screenshot showing the PyQt data app. Left column: juice identity selector and concentration input box for saving to csv. Save data to CSV saves raw counts to a csv. Capture reference curve saves a blank reference into memory. Change display mode toggles from displaying transmission based on the saved blank spectrum to displaying raw counts. The result shows an instantaneous the classified juice and concentration based on the NN algorithm. The right hand side shows a live plot of the datapoints.