

Design and Development of Simple Inexpensive Disposable Colorimetric Sensor

This chapter is divided into two parts. The first part describes the design and development of a simple and inexpensive leak-proof paper pad, selection of paper as a base substrate, and optimization of samples and analytes etc. The second part deals with the accuracy of the result interpretation in a colorimetric based biosensor in this context. A combination of a smartphone with a range of lighting settings was constructed and optimized with the help of machine-learning approaches to determine glucose level using colorimetric signals.

3.1 CHEMOMETRIC-ASSISTED COLORIMETRIC-BASED INEXPENSIVE PAPER BIOSENSOR FOR GLUCOSE DETECTION

This study reports a simple and inexpensive leak-proof paper pad with an initial selection of a paper substrate on the grounds of surface morphology and fluid absorption time. Herein, a drying method is used for glucose detection on a paper pad through colorimetric analysis, and the spot detection of glucose is analyzed by optimizing the HRP concentration and volume to obtain accurate results.

3.1.1 Introduction

The first thought in the anticipation and treatment of disease is a meticulous diagnosis, but diagnostic technologies which are efficient in the economically industrialized world are often hard to avail in countless inaccessible places. It is expensive for common people to afford the examinations. As per the world health organization (WHO), diagnostic devices for remote locations in developing and underdeveloped countries should be affordable, sensitive, specific, user-friendly, rapid, robust, equipment-free, and deliverable (ASSURED) to the end-users. Since glucose metabolism is essential to life, detecting the human body's glucose level is one of the most frequently performed tests in hospitals and laboratories (Rifai *et al.*). Hypoglycemia and diabetes mellitus are medical conditions related to irregular glucose metabolism. According to the international diabetes federation (IDF), diabetes will affect approximately 643 million people globally by 2030, rising to 783 million by 2045 (*IDF Diabetes Atlas | Tenth Edition, 2022*). Early diabetes symptoms are still not clear. It could lead to severe complications, such as neuropathy, cardiovascular disease, and end-stage renal disease, by missing the proper diabetes management time (Soni & Jha, 2017). Therefore, creating a sensor that allows an adult to diagnose his or her diabetes quickly is crucial to address this unmet healthcare need. Along with micro-electro-mechanical system (MEMS) based drug delivery devices, flexible sensors are also playing a crucial role in biomedical applications. Flexible sensors with wireless sensing modules are being developed in the biomedical field for the ease of end-user purposes (S. Jena *et al.*, 2019)(Gupta & Pal, 2019)(Gupta & Pal, 2018a).

Previously, Müller and Clegg reported the first kind of paper-based microfluidic device in 1949 (Müller & Clegg, 1949). However, Whitesides' group later explored paper-based microfluidics and opened new pathways in this field for the use of papers to develop portable, on-site detection in bio-sensing applications (Martinez, Phillips, Butte, & Whitesides, 2007b).

These paper-based analytical devices (PADs) are primarily paper substrates with hydrophilic channels and hydrophobic barriers all around, created via stamping, photolithography, wax printing, and dipping, among many other fabrication techniques (Liana et al., 2012). PADs combine capillary management for the transportation of fluid, limited sample sizes, and minimum usage of reagents with the benefits of a paper platform, such as ease of availability, cost-effectiveness, bio-compatibility, flexibility, a light-weight quality, fluid movement without external pumps, and hydrophilicity (Morbioli et al., 2017) (X. Li et al., 2012) (Martinez, Phillips, Whitesides, et al., 2010). Bio-sensing can be performed via a dipstick model (Gupta et al., 2019), spot-based, and lateral flow analysis. Furthermore, electrochemical and colorimetric methods are the most frequently used detection techniques for the analysis of paper devices due to their portable, low-cost, simple, and easy-to-use instrumentation. Additionally, for more specific and enhanced colorimetric and electrochemical analyses, μ PADs are reported to be modified through nano-functionalized materials (Antibody Immobilization for ZnO Nanowire Based Biosensor Application, 2014) (Gupta et al., 2021) (Gupta et al., 2017) and biopolymers viz. chitosan (Atwe et al., 2014). The development of an advanced point-of-care testing device, which provides fast, accurate, and specific data, needs automation, structural data collection, planning, and monitoring which could further be acquired through the internet of things (IoT) (Kishnani, Park, et al., 2021). Paper-based sensors may capture color information in a variety of color schemes, including RGB (red, green, and blue), HSV (hue, saturation, and value), and $L^*a^*b^*$ (lightness, green-red, and blue-yellow) (Mutlu et al., 2017) (Jia et al., 2015) (Morsy et al., 2016) (Lopez-Ruiz et al., 2014). A number of articles have reported methods for applying different color spaces for the detection of various chemicals. For instance, the detection of textile dye (BR9) and alcohol content in saliva were both performed using the HSV color space, which was transformed from RGB (Bae et al., 2015) (Kuşçuoğlu et al., 2019). In comparison, sensitive pH measurements between 1 and 12 were performed using the $L^*a^*b^*$ color space (Shen et al., 2012a). For the purpose of detecting chlorine in water (Sumriddetchkajorn et al., 2014) and estimating the freshness of fruits (Intaravanne & Sumriddetchkajorn, 2015), the RGB color space was used directly. The colorimetric analysis was carried out using an analytical expression-based method to investigate the properties of the color space. Unfortunately, light sources and camera optics have a significant impact on colorimetric analysis. To eliminate the effects of light and increase the detection threshold, researchers employed 3D-printed light boxes in the field of colorimetric sensors (Heidari-Bafroui et al., 2021). In the quantitative evaluation process, sophisticated algorithms such as machine learning were suggested as a solution to this problem (Solmaz et al., 2018). Machine learning, which has useful features such as automated decision-making and self-learning from data, has become popular in the field of data analysis. The advantages of such tools in biosensors includes the opportunity of attaining realistic methodical outcomes from noisy, undesirable, and low-resolution sensing statistics that could comprehensively coincide with each other. Additionally, the proper deployment of machine-learning (ML) methods could realize unknown relations between sample constraints and sensing signals through data visualization. In particular, ML could be applied in a variety of ways to study the raw sensing data from a biosensor. Numerous systems executing complex algorithms have been created that conduct colorimetric analyses for trustworthy qualitative and quantitative assessments of colorimetric assays, thanks to recent advancements in smartphone technology. As an example, to detect nitrite in food, a mobile phone app was reported that combined a polyacrylonitrile nanosheets (PAN)-NSS color sensor and a deep convolutional neural network (DCNN) for fast sampling, transmission, and data processing (Z.-J. J. Huang et al., 2022). A smartphone-based lateral flow imaging system was developed by using ML classifiers for the detection of *salmonella* spp., a foodborne pathogen. Multiple models were used for the training of the data set, and it was found that support vector machines (SVMs) and k-nearest neighbors (KNNs) produced the highest accuracies, up to 95.56% (Min et al., 2021). To detect glucose in artificial saliva, a machine learning-based smartphone app was developed. This platform was classified on the basis of a color change under different illumination conditions and different smartphones (Mercan et al., 2021). In another work for the detection of pH, glucose, and protein in the urine, a colorimetric reader application was developed with inter-phone repeatability (Yetisen et al., 2014). In spite of

increasing admiration, smartphone-based colorimetric analyses have some issues, including simplicity, reliability, and accessories. It is necessary that the obtained results from smartphones should be accurate in any situation. The obtained results could deviate under different environmental conditions and accessories, viz., illumination, different phones, etc. The results are also sensitive towards the camera optics and smartphone (Berg et al., 2015) (Helfer et al., 2017) (Mercan et al., 2021) (Yetisen et al., 2014).

The expected range for normal fasting blood glucose levels is 3.9 mM to 5.6 mM or somewhere in between. Changes in lifestyles and the monitoring of glycemia are advised when fasting blood glucose levels are between 5.6 and 6.9 mM. The normal level of glucose is found within a limit of 2.5–5.3 mM in serum and 0.1–0.8 mM in urine. In this work, we demonstrated the colorimetric analysis of glucose on a paper pad, and machine learning was used for the optimization of the color variation.

3.1.2 Experimental procedure

(a) Material and Instrumentation

The following were utilized for experimentation: Whatman no. 1 and no. 4 filter paper (125 mm diameter) from GE Healthcare Life Sciences, permanent marker (Camlin, local vendor), Ultra Gold reagent grade Triple Deionized water (DIW) from Organo Laboratories Chemical Pvt Ltd., Delhi, India (specific conductivity < 1), Glycerol from Sigma-Aldrich (St. Louis, MO, USA) ($\geq 99.0\%$), Methylene Blue (MB) dye from Fisher Scientific (Waltham, MA, USA) (Sp. Absorptivity $A_{1\%}/1\text{ cm}$ max at 660–665 nm), glucose oxidase (GOx) from *Aspergillus niger* Type X-S, lyophilized powder, 100,000–250,000 units/g solid, Potassium Iodide (KI) from Sigma-Aldrich (ACS reagent $\geq 99.0\%$), Horseradish Peroxidase (HRP) from Sigma Aldrich, Type VI, (≥ 250 units/mg) (solubility: 0.1 M PBS soluble 10 mg/mL), Trehalose from Molychem, Mumbai, India (minimum assay- 99%), D-Glucose from Thermo Fisher Scientific Private Limited, India, and phosphate-buffered saline (PBS) from Sigma Aldrich, Water Soluble, pH: 7.2–7.6 (1 tablet/200 mL), HCl purchased from Sigma Aldrich India, Smartphone (Realme 8), scanning electron microscopy (SEM-Carl Zeiss-EVO 18, EDS-Oxford instruments-51-ADDD-0048, Oxford Instruments, Cambridge, UK), Atomic force microscopy (Park System-XE70, Park System, Seoul, Korea), contact angle goniometer (KRUS GmbH-DSA25B, KRUS, Hamburg, Germany) equipped with ADVANCE Software, printer (HP LaserJet M1005, HP, Delhi, India), and a commercially available laminator. Unless otherwise noted, all chemicals were acquired at the highest analytical grade and used without any purification. For better presentation of the images in the document some images are masked with black concentric circles.

(b) Selection of paper as a base substrate

Paper, in general, is composed of cellulose fibers, and basic cellulose is composed of lignin and hemicellulose, as shown in Figure 3.1. However, due to variations in material compositions and production methods, each variety of paper has unique mechanical characteristics. Given that different forms of paper have distinct components, it is challenging to generalize a set of attributes for paper. Based on its content and structure, each variety of paper performs a certain purpose differently. Cellulose fiber in printing paper is combined with a sizable quantity of filler material. Natural minerals like clay, talc, and limestone, as well as synthetic substitutes, viz., gypsum, precipitated calcium carbonate, and titanium dioxide, could be used as filler material. The composition, quality, and type of the filler ingredients determine the paper's appearance, thickness, and structure. The filler determines the paper's production costs, strength, brightness, and refractive index, as well as the energy needed for drying and friction, the burn rate, and the pore size. Cellulose is found in complex lignocellulose composites that are primarily composed of lignin and hemicellulose. Fillers could have a detrimental impact on the sheet's two-sidedness, durability, endurance, and abrasion resistance.

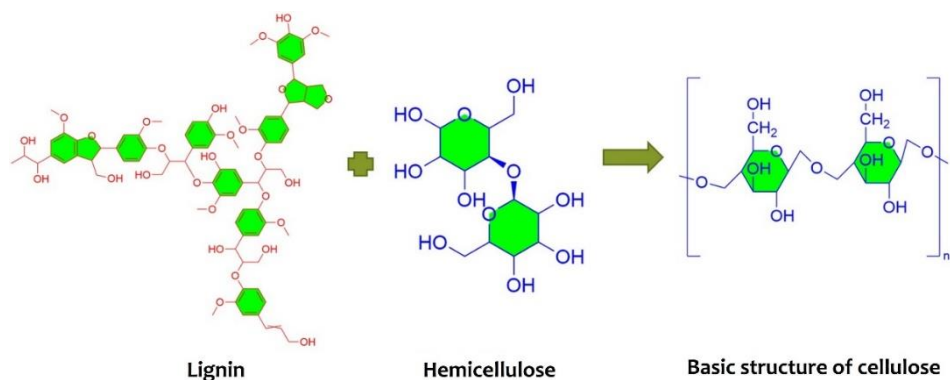


Figure 3.1 : Schematic of the basic structural composition of cellulose

For the selection of the right paper-substrate to develop paper-based biosensor, thirteen different paper substrates were collected and investigated using the drop analysis technique. A 0.5% MB dye solution was used for the drop analysis. The thirteen different substrates that were investigated were: Whatman no. 1 paper, bond paper, craft paper (200GSM), handmade blotting paper, tissue paper (2 ply), glossy paper, card sheet paper, practical sheet paper, ivory sheet paper, sketch sheet paper, drawing sheet paper, A4 paper, and Whatman no. 4 paper. A drop of 10 μ l of methylene blue (MB) dye solution was dropped onto the paper surface through a micropipette, and the time taken for the dye to spread and absorb on the paper surface was recorded using a stopwatch. Different paper substrates have different surface morphologies and pore sizes, which are responsible for the different absorption times.

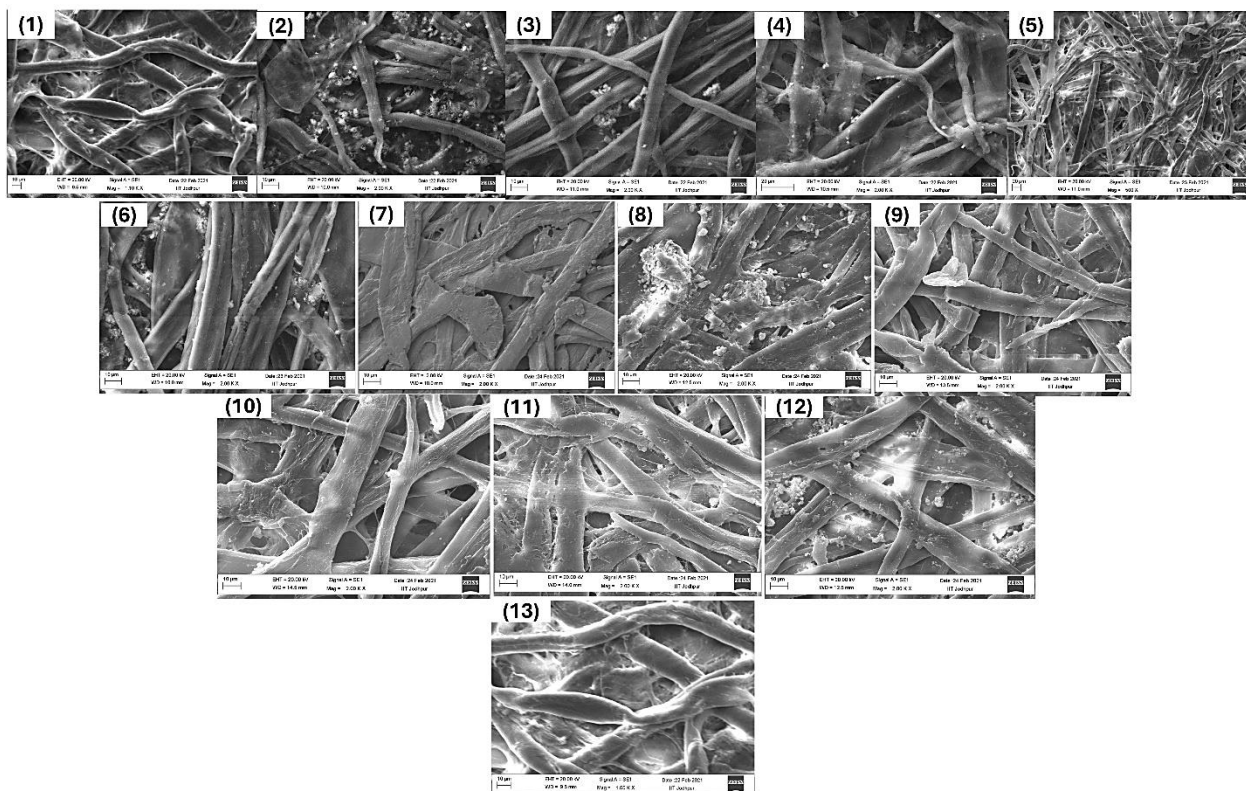


Figure 3.2 : SEM images of different types of papers available: (1) Whatman no. 1, (2) bond paper, (3) 200 GSM card sheet, (4) blotting paper, (5) paper towel (2 ply), (6) glossy paper, (7) 300 GSM card sheet, (8) practical sheet paper, (9) ivory sheet, (10) sketch sheet, (11) drawing sheet, (12) A 4 paper (67 GSM), and (13) Whatman no.

4

Figure 3.2 shows the SEM images of the thirteen commercial papers available on the market. It was found that most of the commercial paper contained foreign metals, viz., Si, Al, Ca,

Mg, and Cl, apart from cellulose as shown in Figure 3.3. These foreign materials are added to the paper pulp as a filler material in alkaline papers to increase the brightness and opacity. It is hard to create offset lithography in papers due to the availability of calcium carbonate. The papers could also include trace levels of magnesium carbonate, silicon dioxide, aluminum oxide, and iron oxide in addition to calcium carbonate. These filler materials also clog the pores of the paper substrate, which affect the porosity, permeability, and wicking properties of the paper.

Whatman filter paper was selected for further analysis because it showed good holding capacity and the lowest absorption time compared to other papers, as shown in Figure 3.4. The commercial A4 paper available on the market contained some foreign materials, as shown in Figure 3.2 (12), which was confirmed with the EDX analysis, through which it was determined that it contained C, O, Ca, and Cl Figure 3.3. These foreign materials blocked the pores of the A4 paper, which resulted in absorption time enhancement, as shown in Figure 3.4.

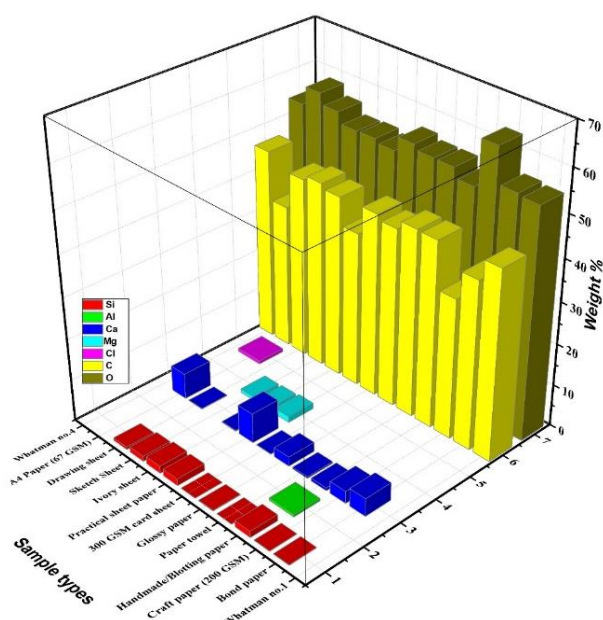


Figure 3.3 : Schematic image of elemental analysis of different type of papers

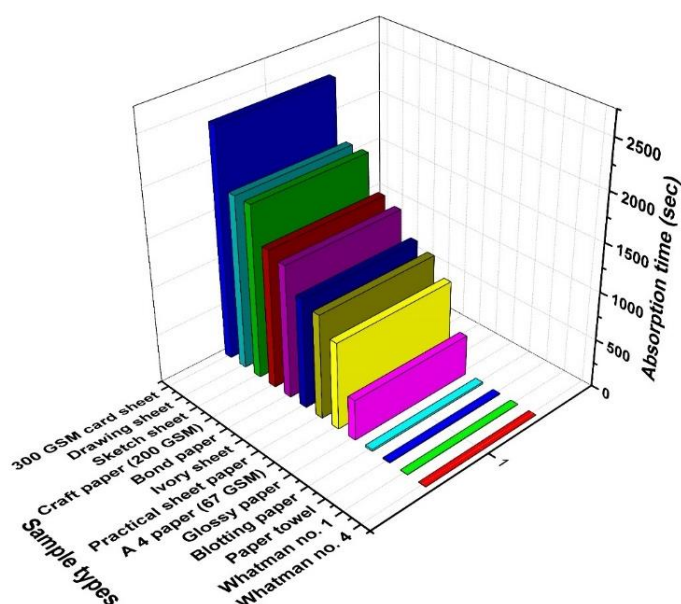


Figure 3.4 : Graph representing the absorption times for a fixed volume of dye solution with different papers

(c) Characterization Techniques

The surface morphology of the permanent ink-modified surface and the unmodified surface of the paper pad was analyzed through atomic force microscopy (AFM) using an advanced scanning probe microscope (Park Systems XE-70, Suwon, Korea). The data gain was set at -144.74×10^{-6} ($\mu\text{m}/\text{step}$), I Gain-1, and P Gain-1. Scanning electron microscope (SEM) (ZEISS and OXFORD instrument) images were obtained. Preceding the SEM readings, the samples were treated with a thin layer of gold to avoid and diminish charging on the surface. All contact angle (CA) measurements were performed with a goniometer (KRUS GmbH model no. DSA25B equipped with ADVANCE Software). For the static water CA measurements, DI water was dropped onto the sample surface, and the CA was measured at a relative humidity of $40 \pm 5\%$ and a temperature of 20°C . A piece of filter paper with a dimension of $10 \times 10 \text{ mm}^2$ modified with hydrophobic ink was taken for the contact angle measurement. There was no backing of tape or any other material to determine the exact contact angle generated with the hydrophobic ink. A droplet of a 0.47 mm mean diameter and a $0.045 \mu\text{l}$ mean volume (automatically calculated by the instrument software) was dropped over the modified surface. The contact angle was calculated after the generation of the static angle and was determined up to $108.69 \pm 0.25^\circ$.

(d) Fabrication of the paper-based sensor and colorimetric detection

Design patterns were drawn with SOLID WORKS (2020). The design of the microfluidic paper pad consisted of two concentric circles having sample zones of 4 mm . The area between the outer periphery of the inner circle and the inner periphery of the outer circle was used as a boundary to create a hydrophobic barrier, as shown in Figure 3.5. The designs were printed on Whatman paper with an HP printer using a black toner cartridge from HP LaserJet M1005. The back side of the sensing zone was covered with laminating film. Then, the area between the two concentric circles became hydrophobic, while the center area remained hydrophilic as a testing zone. The patterned paper was then placed in a vacuum desiccator for drying the paper pad, as well as for the prevention of atmospheric contact for further use.

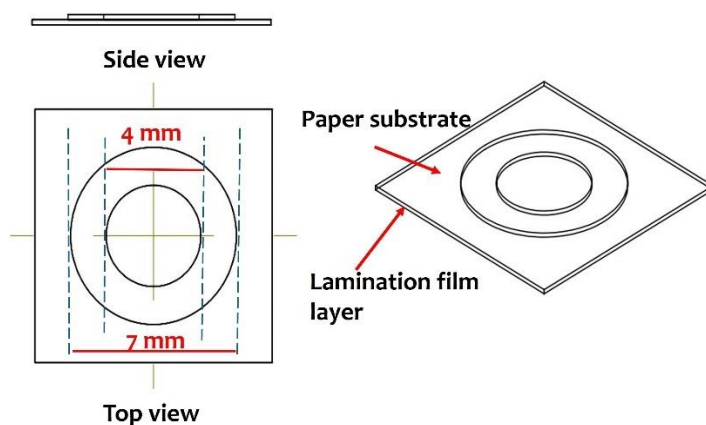


Figure 3.5 : Schematic representation of the dimensions and design of the paper pad

(e) Reagent Preparation

For the glucose assay, two different concentration ranges were chosen; the first was $0.1\text{--}0.5 \text{ M}$ with a difference of 0.1 M for the color development, and the other was $0.5\text{--}20 \text{ mM}$. The reason for selecting these values was that normal blood glucose levels lie in this range. To obtain both concentration ranges, glucose was dissolved in PBS (0.1 M , pH 6.0). KI, glucose oxidase, and HRP were dissolved in PBS (0.1 M , pH 6.0) to obtain 0.6 M , 120 U/mL , and 100 U/mL of stock solution. The prepared stock solution was kept in a refrigerator at 4°C , ceasing from the light, while the KI was stored at room temperature, wrapped with aluminum foil to protect from light for further use.

The colorimetric detection of glucose is based on oxidase enzyme activity. Glucose oxidase is specific for glucose molecules; it binds and converts it into gluconic acid, and thereafter, hydrogen peroxide is released. The hydrogen peroxide then combines with an indicator, here KI, and converts the iodide from potassium iodide to triiodide in the presence of horseradish peroxidase, resulting in the formation of a colored product, as shown in Figure 3.6.

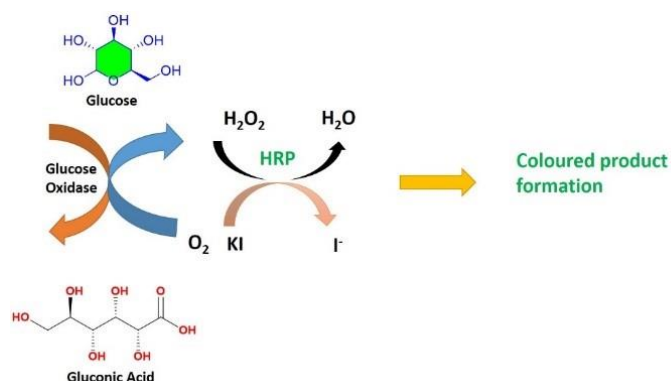


Figure 3.6 : Schematic representation of the reaction mechanism for the colored product formation using oxidase enzymes

(f) Optimization of the volume and concentration

The sample and reagent volume optimization were performed using a solution of MB dye and glycerol on the spots on Whatman no. 1 and no. 4 filter papers to determine the minimum volume of the sample and reagents required to fill the zone completely. For the optimization of HRP, different concentrations, i.e., 10 U/mL, 20 U/mL, 30 U/mL, 40 U/mL, 50 U/mL, 60 U/mL, 70 U/mL, 80 U/mL, 90 U/mL, and 100 U/mL were prepared. The steps for the optimization of the different concentrations of HRP for the detection of glucose are shown in Figure 3.7.

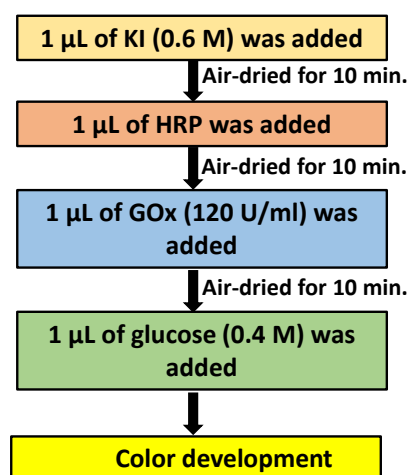


Figure 3.7 : Schematic showing the reaction mechanism for color development using different concentrations of HRP

For the detection of glucose, the indicator (potassium iodide) was first added to the detection pad, and then it was air-dried for 10 min, followed by the addition of horseradish peroxidase, after which it was air-dried for 10 min. After the indicator and HRP were completely dried at room temperature, glucose oxidase was added and air-dried for 10 min. After completing the process, the detection pad was ready for testing and was stored in a refrigerator at 4 °C for further use. The final test was conducted by adding different concentrations of glucose to the different spots.

(g) Method of the digitalization of the obtained result

The raw images obtained from the smartphones (Realme 8) were then pre-processed to progress to further steps, which included feature extraction and machine learning modeling. The images were pre-processed to obtain the actual region of interest (ROI) where the main color change took place. Various noise reduction techniques were used to obtain the ROI. After that, a mask was created, which was placed over the actual resized image to obtain the ROI. The images were converted to various color spaces or gamuts available in MATLAB to obtain the maximum possible derived feature vectors.

The color spaces used included HSV, L*a*b*, XYZ, grayscale (the scale varied from 0–255, in which 0 stood for white and 255 stood for black, and the in-between gray shades were obtained), and NTSC. The NTSC images had three attributes, as shown in Table 3.1.

Table 3.1 : The description of the attributes of the images

Attribute	Description
Y	Luma, also known as picture brightness. Values fall between [0, 1], where [0] designates black and [1] designates white. As Y grows, colors get brighter.
I	The proportion of blue or orange tones in the picture is known as in-phase. The values of “I” fall between [−0.5959, 0.5959], where a negative number represents a blue tone, and a positive number represents an orange tone. The intensity of the color grows as I’s magnitude rises.
Q	The proportion of green or purple tones in the picture is known as quadrature. Q contains a value of between (−0.5229) and (0.5229), where (−) denotes a green tone and (+) denotes a purple tone. The intensity of the color grows as Q’s magnitude rises.

3.1.3 Results and Discussion

(a) Sample and reagent volume optimization

It is known that normal blood viscosity lies within the range of 3.5 cP to 5.5 cP. Therefore, in order to mimic blood spreading in the paper-based spot detection, 40% glycerol was used, which consisted of a viscosity of 3.7 cP, lying well within the range of normal blood viscosity. A 0.2% methylene blue dye solution was prepared in 40% glycerol. The reasons for choosing methylene blue were its properties, viz., antioxidant, cationic, and organic nature, ease of solubility in water and glycerol, etc., and its bright color for ease of detection. The optimized volume of the dye solution, which covered the entire testing zone, was found to be 1 μ l and 1.5 μ l for Whatman no. 4 and Whatman no. 1 filter paper, respectively Figure 3.8.

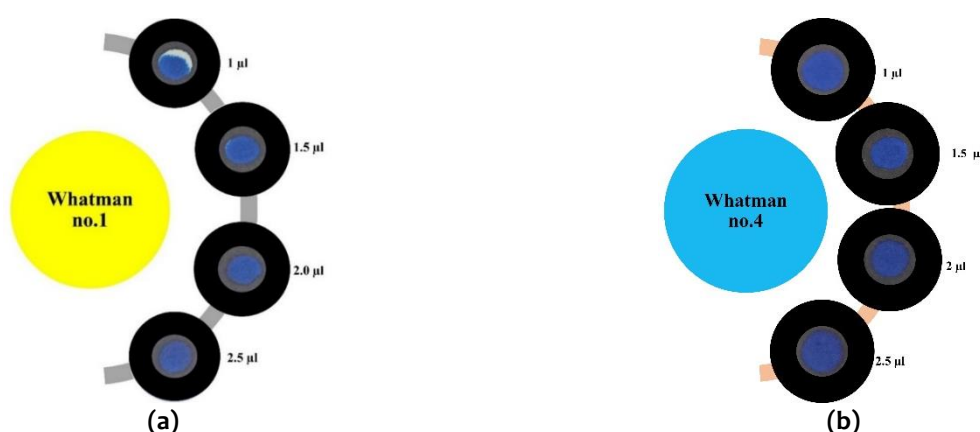


Figure 3.8 : Sample and reagent volume optimization using MB dye solution

(b) Characterization of the detection pad

Atomic force microscopy images of the permanent ink-modified surface and the unmodified surface of the detection pad revealed that the roughness of the modified surface increased slightly in comparison to the unmodified surface. However, the surface of the modified paper appeared to be smoother in comparison to the unmodified surface, as shown in Figure 3.9. The root mean square roughness factor for the unmodified and modified surface was calculated to be 35.160 nm and 75.449 nm, respectively. The unmodified surface of the paper showed a rough contour of the fibrous structure, while the ink-modified surface appeared smooth, as the ink absorbed over the fibrous surface and filled the pores of the paper surface. If the surface were rough instead of flat, the actual wet surface area would be greater in comparison to the geometric area. As a result, a net increase in the surface energy means that the surface would be more hydrophobic. Cross-sectional images of Whatman paper no. 4 are shown in Figure 3.9.

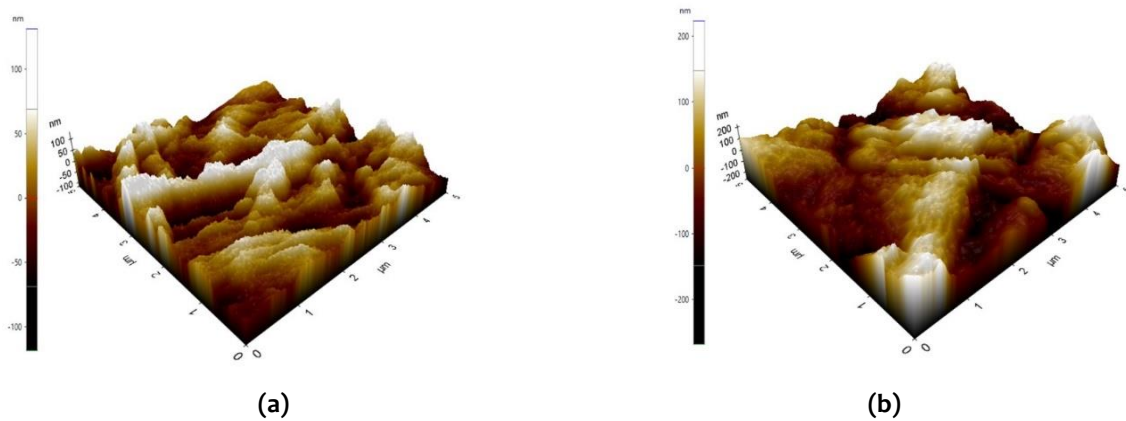


Figure 3.9 : AFM images of the (a) unmodified surface and (b) ink-modified surface of Whatman no. 4 paper

For the analysis of hydrophobicity, the paper was coated with a single-stroke and double-stroke of hydrophobic ink. To determine the contact angle over the modified paper surface, a sessile drop test was performed with a contact angle goniometer (DSA25B) at a mean temperature of 20.0 °C for a mean volume of 0.045 μ l of DI water. It was observed that a single stroke of hydrophobic ink was not sufficient for holding the water droplet because it decreased continuously. In comparison, using a double stroke of hydrophobic ink held the water properly for a longer time, as shown in Figure 3.10.

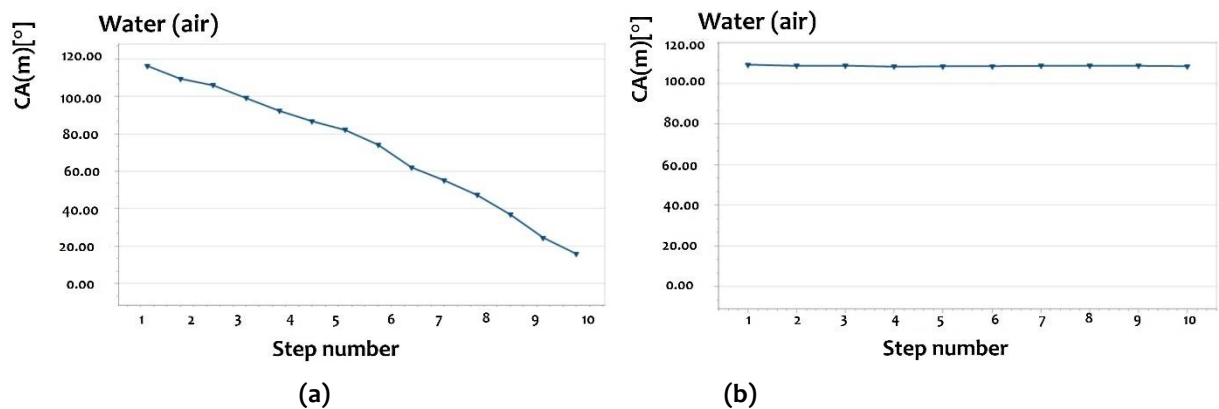


Figure 3.10 : Contact angle retention after an (a) single-stroke or (b) double-stroke

Figure 3.11 (a) shows the microscopic image of the ink-modified surface of Whatman no. 4 paper, and the red dotted line shows the boundary between the modified and unmodified surfaces. The mean contact angle of the double-stroke coating was observed up to $108.69 \pm 0.25^\circ$ (bottom right-most corner of Figure 3.11 (a)).

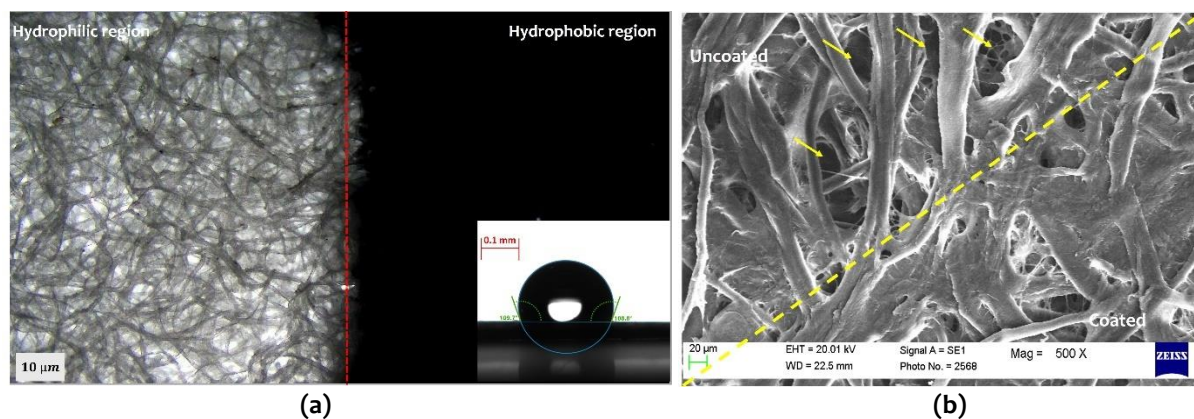


Figure 3.11 : (a) Microscopic image of the junction of the unmodified and modified paper (inset contact angle of the modified paper) (b) SEM image of the junction of the unmodified and modified paper

Figure 3.11 (b) shows the SEM image of the double-stroke hydrophobic ink-coated paper substrate. It was observed that the hydrophobic ink partially blocked the pores of the paper substrate, while the unmodified surface had deep and wide pores in comparison to the modified paper.

(c) Optimization of the concentration

Furthermore, the HRP concentration optimization was performed for the range of 10–100 U/ml. It was observed that initially, the mean gray value decreased until 30 U/mL, but after 30 U/mL, it rose sharply. The lowest mean gray value indicated that the darkest shade was obtained at 30 U/mL of HRP, while an increase in the mean gray value meant a lighter shade. This could be due to an increased concentration of enzyme for the fixed concentration over the substrate, as no additional substrate was available to bind to the enzyme. Figure 3.12 shows the optimization curve for HRP for a range of 10–100 U/ml.

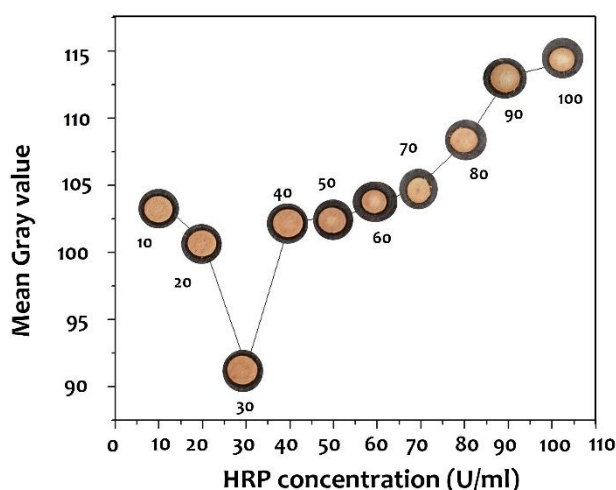


Figure 3.12 : Graph showing the mean gray value for different concentrations of HRP

(d) Glucose detection

Color developments were observed, and images were captured using the smartphone. Images of the detection pad were taken via a Realme 8 smartphone in the jpeg format with the following specifications: 64 MP, aperture: f/1.8, sensor size: 1/1.73", pixel size: 0.8 µm, focal length: 26 mm, and image stabilization: PDAF (phase detection auto-focus). The images were taken in an open ambient light without flashing homogeneously.

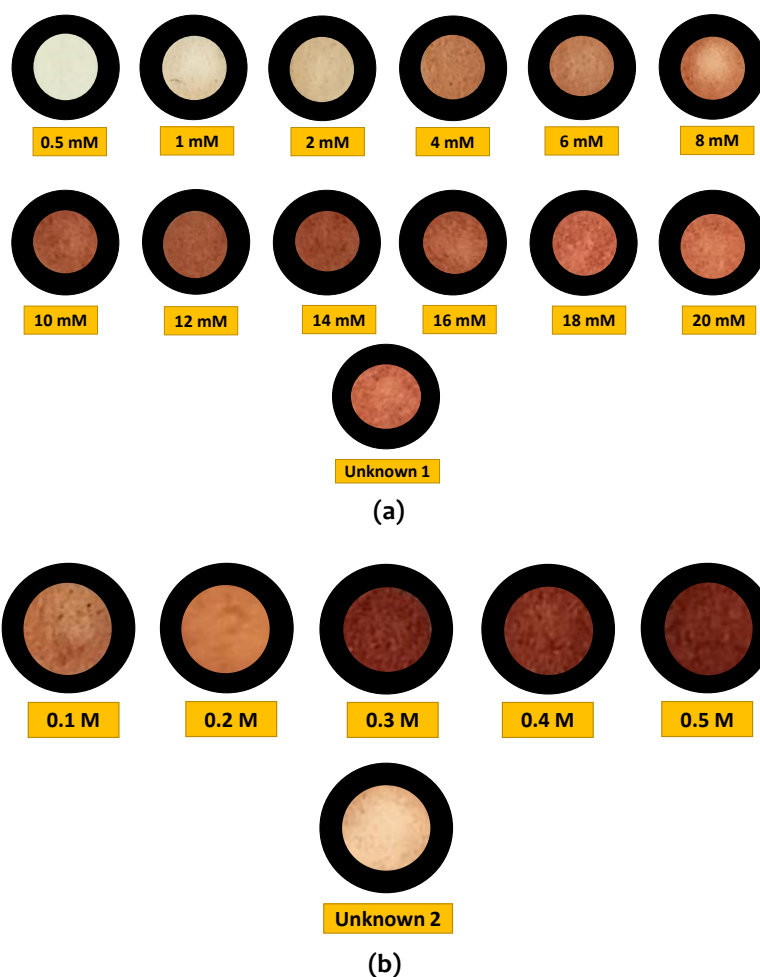


Figure 3.13 : Color development at different concentration of glucose solution (a) 0.5 mM-20 mM (b) 0.1 M-0.5 M

These images Figure 3.13 were taken for different concentrations, two of which were of unknown concentrations. The raw images were then pre-processed to progress to further processes, including feature extraction and machine learning modeling. Images were pre-processed to obtain the actual region of interest (ROI) where the main color change took place. Various noise reduction techniques were used to obtain the ROI. After that, a mask was created, which was placed over the actual resized image to obtain the ROI. A response curve of the glucose assay Figure 3.14 was constructed. It showed a non-linear curve with an $R^2=0.933$, and the image intensity was stable up to a 14 mM concentration. The image intensity was unstable above this range.

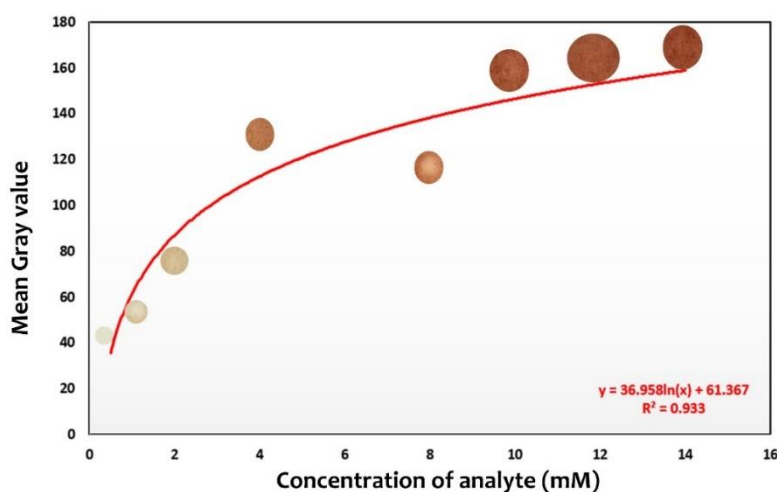
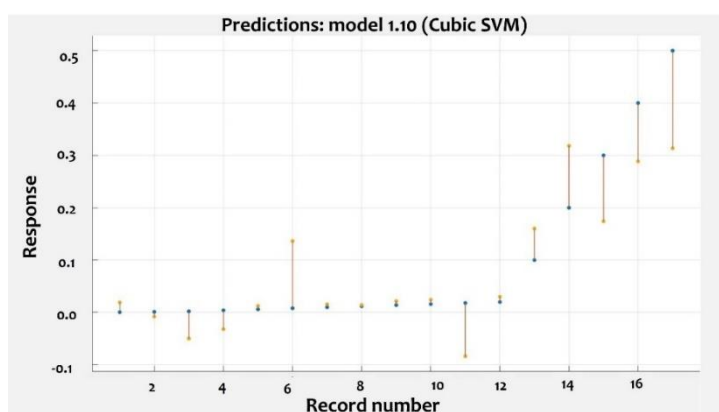


Figure 3.14 : Response curve of the spot based glucose assay with R^2 of 0.933

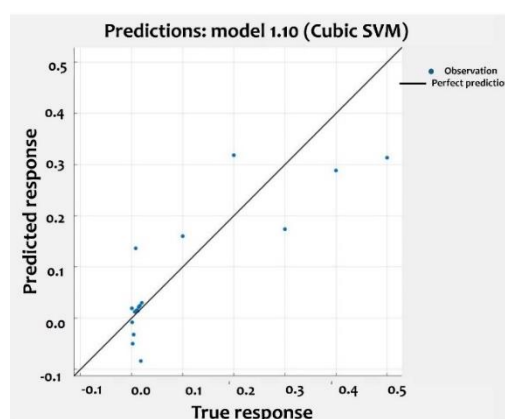
(e) Digitalization of the obtained results

The classification learners in MATLAB (R2021b) were used to assess the performance. The training data were analyzed by creating the training model using supervised learning. Support vector machines (SVMs) were used in MATLAB to discover the decision boundary that would best divide the classes and maximize the margin. SVMs are used to search for the hyperplane that maximizes the margin in the linearly separable scenario, provided that both classes are correctly identified. Soft Margin and Kernel Tricks are two principles that SVMs use to solve non-linearly separable scenarios. Kernel changes the dimensionality of the dataset if the dataset is not separable in the lower dimensions. Statistical feature vectors were obtained from all the color spaces, which included the mean, skewness (measurement of the asymmetry of the probability distribution), and kurtosis (its value defines the shape of the probability distribution), over all 45 features were created through this exercise. The gray-level co-occurrence matrix (GLCM) functions to characterize the texture of an image by calculating the frequency of the pairs of pixels with particular values and in a specified spatial relationship, creating a GLCM, and then retrieving the quantifiable metrics from this matrix, including contrast, correlation, energy, homogeneity, entropy, and intensity of the grayscale. All these features constituted a total of 51 features for each image.

Figure 3.15 (a) shows the error in the prediction with respect to the actual data value, where the blue dots represent the actual data, the orange dots represent the predicted data value, and the line between them shows the error in the data, while Figure 3.15 (b) shows the best-fit graph between the predicted and actual data. The R square value was 0.74. Further trained data were tested for two unknown samples, and the predicted concentrations were 7.18 mM and 0.35 M.



(a)



(b)

Figure 3.15 : (a) Errors in prediction with respect to the actual data, (b) the predicted vs. actual data

(f) Shelf-life testing

Shelf-life testing was also performed after the immobilization of the enzymes (HRP and glucose oxidase) and the reagents on the paper. Trehalose was added in one set that was kept at room temperature and one kept at 4 °C. A comparison was performed between the detection pad kept at room temperature and 4 °C for 14 days. Shelf-life testing was conducted to determine the performance and lifetime of the pad. The color appeared to be darker for the spots in which trehalose was added, as it maintained the enzyme stability Figure 3.16. Different spots were tested at different intervals of days using a glucose solution of a fixed concentration. The shade or darkness of the color decreased as the number of days increased. Images of the color were developed after the addition of the glucose solution at different numbers of days and were captured using a smartphone for all four sets, and the mean gray value was analyzed using MATLAB.

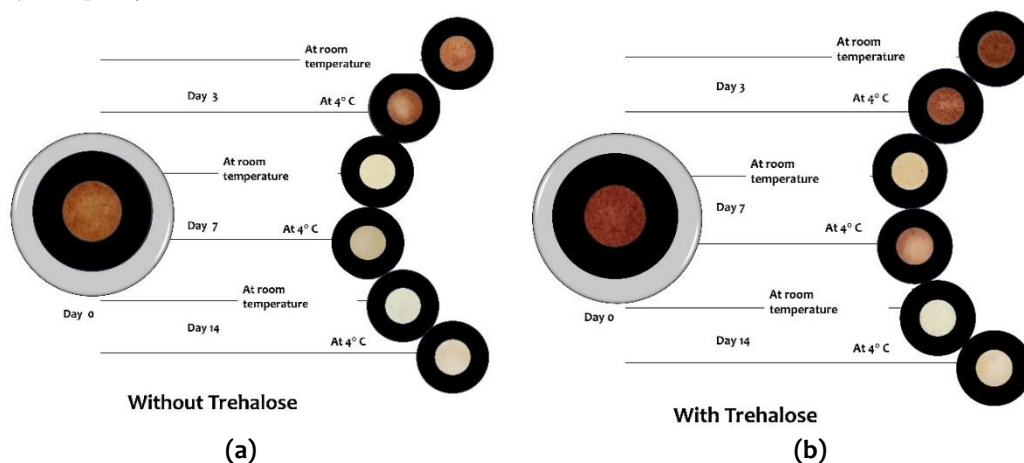
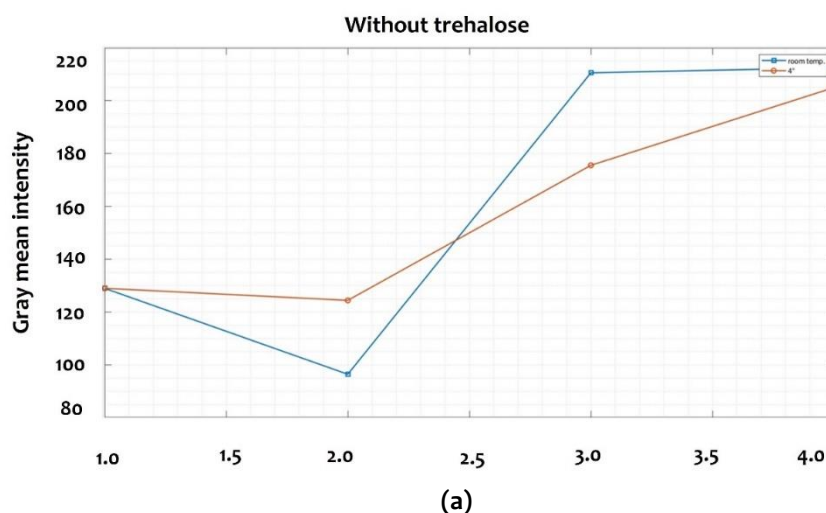
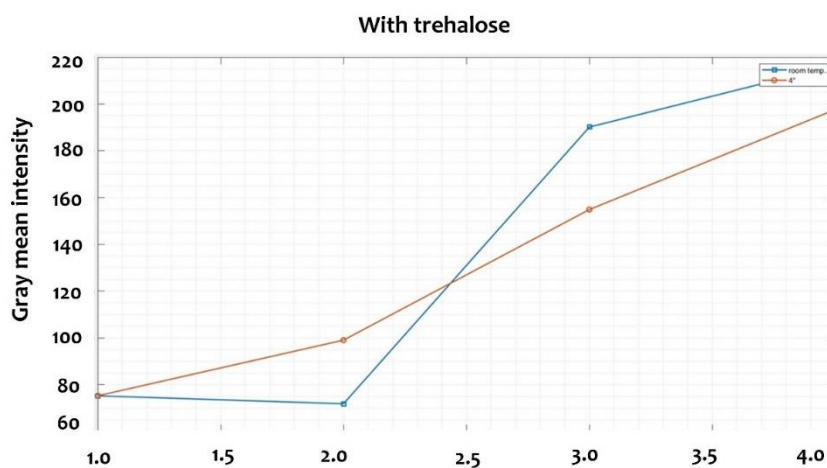


Figure 3.16 : Shelf-life testing of spots kept at room temperature and at 4 °C for 14 days (a) without trehalose, (b) with trehalose

Figure 3.17 shows the mean gray value, which first decreases and then increases with the number of days, which indicates a lighter shade of color developed. The mean gray value on the 14th day was lower for the one which was stored at 4 °C and had trehalose immobilized. However, the color developed for this spot was also lighter as compared to the first day and showed approximately 62% decreased color intensity. It could be stored at 4 °C, but the result would not be accurate if a surface modification were not conducted.





(b)

Figure 3.17: Graph showing the mean gray values for the set (a) in which trehalose was added (b) in which trehalose was not added and kept at room temperature and 4 °C.

3.1.4 Summary

In this work, paper as an inexpensive and potential substrate material for point-of-care testing was investigated under ASSURED criteria. Herein, colorimetric analyses were performed through machine learning techniques in a controlled environment for the detection of glucose on a leak-proof pad, which was fabricated by creating a hydrophobic zone through the proper penetration of ink into the pores of the paper. A detailed analysis was carried out for the ink-coated paper zone through the surface roughness (75.449 nm) analysis and contact angle ($108.69 \pm 0.25^\circ$) measurement. Further, the fabricated detection pad was tested for the optimization of samples, and the determination of glucose for different concentrations with a LOD of 2.92 mmol L⁻¹. Furthermore, machine learning classifiers were implemented for the color-change analysis in μ PAD for the detection of glucose levels in the samples. These classifiers were trained using images captured through smartphones under ambient lighting conditions. Trained data were tested for the two unknown concentrations. The obtained R square value for the trained data was 0.74, and the predicted concentrations were 7.18 mM and 0.35 M. To further improve accuracy, work was also performed on increasing the dataset for the determination of closer concentration levels. It could be concluded that the fabricated biosensor platform with the explored chemometrics is economical, robust, and could be fabricated easily without the use of any heavy instrumentations, and the generated machine learning platform could be helpful in the statistical analysis for the detection of biological samples.

3.2 COLORIMETRIC BASED GLUCOSE DETECTION WITH MACHINE LEARNING APPROACHES

The accuracy of the result interpretation in a colorimetric based biosensor is a major challenge. In this context, a combination of a smartphone with a range of lighting settings was constructed using machine-learning approaches to determine glucose level using colorimetric signals. Herein, five different smartphones have been used to capture images of the reaction zone with varying colour intensities under different lighting atmosphere. The images were initially processed for region of interest extraction using noise reduction logics, then for feature extraction, and finally for machine learning classifier training, resulting in a more vigorous and adaptive platform that is more resistant to lighting variations and camera optics. For each concentration and lighting situation, various machine-learning (ML) classifiers were tried, and the optimal ML classifier was determined.

3.2.1 Introduction

Traditional colorimetric tests, such as pregnancy tests, rely on true or false results determined by the naked eye and are incapable of providing quantifiable values due to their low accuracy and sensitivity to environmental factors. The widespread use of mobile phones and the quickly expanding global telecommunications infrastructure give an unparalleled potential for reading and transmitting point-of-care diagnostic data. Smartphone cameras struggle with standardisation in optical analysis of colorimetric tests; their built-in colour balancing features are optimised for photography in bright environments. Additionally, its unidentified built-in automated image correcting procedures like as automatic brightness management and white balancing preclude a smartphone camera from getting steady colour information from colorimetric sensors (Kishnani, Park, et al., 2021) (Shen et al., 2012b).

It is difficult to identify minor colour variations with the naked eyes, visual colorimetry has a lower measurement precision than photoelectric colorimetry does. The importance of having access to and being able to share analytical data quickly on-site is growing exponentially in modern culture. Colorimetric analysis of pictures taken via digital cameras, mobile phones, scanners, webcams, and so forth may be done using the pixelated digital image colorimetry (PDIC) technique (Firdaus et al., 2014b). Smartphones and digital cameras are used in PDIC significantly more frequently than webcams and scanners because of their mobility and small weight. As a result of their quick uptake, rapid advancements in camera technology, and the broad adoption of mobile apps (APPs), cell phones are increasingly being employed in PDIC as image capture tools (Coskun et al., 2013b). Image capture by mobile phone and colour assessment using image processing software (MATLAB, Pantone Studio, Image J, Adobe Photoshop, etc.) in an appropriate colour space are two required phases in the PDIC approach. Analyte concentration may then be correlated with quantified pictures data and findings read out utilising bespoke mobile apps.

According to the publications listed in (Oberholzer et al., 1996) (Hameed et al., 2020), the concept of digital image processing was also widely discussed. (Gölcez et al., 2020) uses a PAD (Paper based analytical device) combined with a smartphone-based glucose testing system. A custom-designed 3D printed smartphone cover was utilized to avoid ambient light interference while photographing. MATLAB was used to analyze the pictures in JPEG & RAW formats and ROI (Region of Interest) was extracted. The data obtained used to get calibration curves. The sensor showed a linear response for glucose ranging from 0-1 mM, with a calculated detection limit of 29.65 μM of glucose. An app 'Glucosense' was made to make process user-friendly. (Yetisen et al., 2014) showed an approach with inter-phone repeatability that enables a smartphone camera to quickly interpret semi-quantitative tests with minimum operator interaction. (S. D. Kim et al., 2017) used 3D-printed small light box, which blocks the ambient light effects and has two LED (Light emitting diode) backlights. Images captured in 'auto mode'. (Solmaz et al., 2018) developed a smartphone application that helps binary and multi-class

learning models to automatically classify peroxide content depending on the color of the test strip. Random Forest and LS-SVM (Least squares – Support Vector Machine) were given the mean RGB, HSV, and LAB values of the test strip picture under varied lighting circumstances, with and without the grey-world color constancy method. Training data with 10-fold cross validation yields above 95% accuracy, while mobile app user tests reveal a 90.33% accuracy rate for the primary group with six classes in the primary classification. Mercan et al. (Mercan et al., 2021) trained classifiers by using photographs taken in seven distinct lighting settings with four different cellphones, enhancing the platform's ability to withstand the effects of camera optics. So, no matter what brand of smartphone, when, or where an analysis is performed, the present platform can accurately determine a sample's glucose level. An image processing smartphone app (GlucoSensing) was built to interface with the distant server running several machine learning classifiers. In order to reduce human mistakes, GlucoSensing has a built-in image processing tool that automatically locates the ROI.

The image processing and analysis technique has a broad range of applications, including the examination of specific molecules, physical, chemical, and other processes, and illness diagnosis. An enzymatic or other chemical interaction that results in a colored output or an affinity interaction that results in the capture of a fluorescent or colored trace are examples of bioassay methods (Pohanka, 2020) (Kishnani, Kumari, et al., 2022).

In this study, we aim to find the relation between concentration of glucose and change in color shade which would fit in every aspect whether it is the change light conditions or the prediction of any unknown data. Using different machine learning algorithms, the training of model plays a significant role in the accuracy of model. To make a robust model we must train the model with images captured with maximum possible illumination conditions and with different smartphones. Finally, we must make the process easy and user friendly to operate which can be done by making a smartphone application or an application which can be used by any wearable device.

3.2.2 Materials and Methods

Whatman No. 1 Filter Paper (GE Healthcare Life Sciences), Permanent Marker ink (Camlin), Deionized Water, Glucose oxidase (GOx), Potassium Iodide (KI), Horseradish Peroxidase (HRP), Glucose, and Phosphate Buffered Saline (PBS), were purchased from Sigma Aldrich India.

(a) Device fabrication and Reagent preparation

Spot detection of glucose concentration is performed on paper-based sensors by creating a confined zone of reaction through the hydrophobic ink, various range of concentration were tested in different atmospheric condition. Figure 3.18 shows 3D model of a container box, which was further used for different atmosphere.

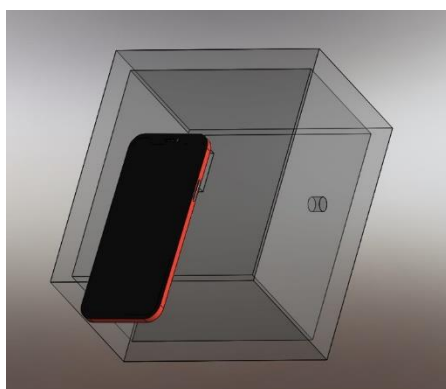


Figure 3.18: A 3D model of close container box for dark room image capture

Smartphones	Conditions	Concentration (mM)	10	20	30	40	50
iPhone	Room						
	Box Without Flash						
	Box With Flash						
	LED						
	Sunlight under shade						
	direct Sunlight						
One Plus	Room						
	Box Without Flash						
	Box With Flash						
	LED						
	Sunlight under shade						
	direct Sunlight						
Realme	Room						
	Box Without Flash						
	Box With Flash						
	LED						
	Sunlight under shade						
	direct Sunlight						
Samsung M31	Room						
	Box Without Flash						
	Box With Flash						
	LED						
	Sunlight under shade						
	direct Sunlight						
Smasung M51	Room						
	Box Without Flash						
	Box With Flash						
	LED						
	Sunlight under shade						
	direct Sunlight						

Figure 3.19 : Colorimetric data generated with variation of phones and lightening condition

The concern experiment was repeated for all different atmospheric conditions *viz.*, room, box without flesh, box with flesh, LED, sunlight under shade, direct sunlight for a wide range of concentrations 10 mM to 50 mM with a difference of 10 mM. The entire set of experiment was repeated for the five different smartphones *viz.*, iPhone, One Plus, Realme, Samsung M31, and Samsung M51 as shown in Figure 3.19

(b) Image pre-processing:

Raw Images are then preprocessed to be able to feed to further process which includes feature extraction and Machine learning modelling. Images are pre- processed to get actual Region of Interest (ROI) in which the main change of color takes place. Various noise reduction techniques are used to get ROI. After that a mask is created which is placed over the actual resized image to get ROI.

Step-1

Image captured by camera is resized to 270 x 270 pixels to get an equivalent distribution of pixels in all the images as shown in Figure 3.20.

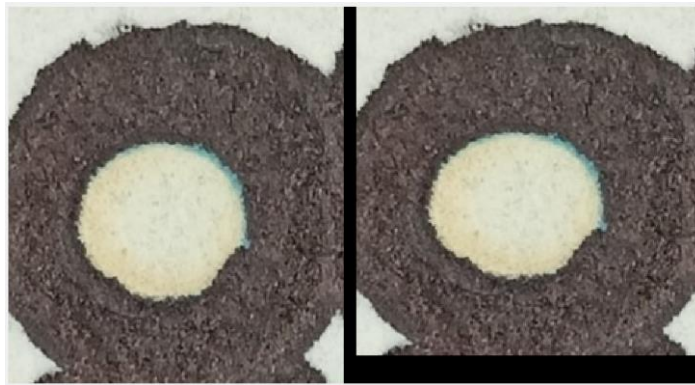


Figure 3.20: Image pre-processing Left: Original image, right: Resized Image

Step-2

The image is converted to grayscale which converts the image from RGB scale to grayscale. RGB scale has image in the form of 3 layers in which each red, green and blue layer has pixel range from 0 - 255 which is transformed to only gray pixels having range of 0-255 in which 0 stands for black and 255 stands for white and corresponding range in between have gray shades. An example of operation is shown in Figure 3.21

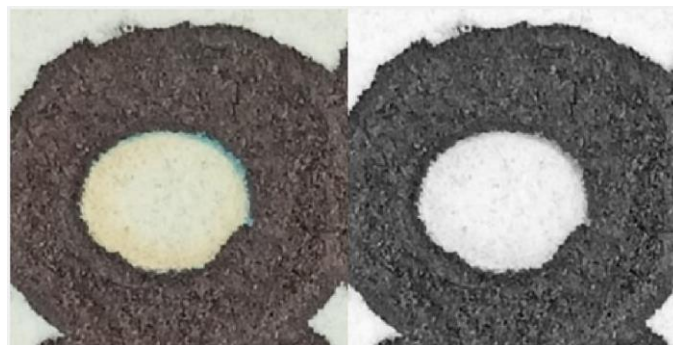


Figure 3.21: Gray scale conversion of the resized image Left: RGB Image, right: Grayscale Image

Step-3

Now Image threshold is minimized by the Otsu method. It chooses a threshold which minimizes the intraclass variance of the thresholded black and white pixels in the image.

Step-4

The result image from above operations is changed to binary image which has only two-pixel intensities 0 means black and 1 means white to create a mask but obtained image has noise so further processing is needed. An example of an operation is shown in Figure 3.22.

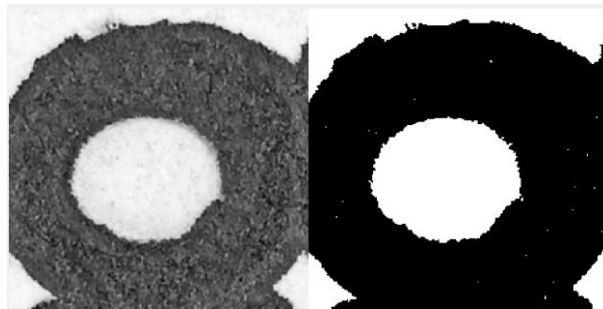


Figure 3.22 : Binary image conversion of the grayscale image Left: Grayscale Image, right: Binary Image with noise

Step-5

Noise reduction techniques are applied on the obtained image which include clearing the border which suppresses structures in image that are lighter than their surroundings and that are connected to the image border and then combination various filters & dilation of image results in the desired image as shown in Figure. 3.23.



Figure 3.23: Noise reduction of the binary image Left: Noisy Image, right: Noiseless Image with ROI mask

Step-6

Now bounding box is made on the ROI's area and masked out. This mask is then clubbed with the original image to get the ROI through the original image. An example of an operation is shown in Figure. 3.24

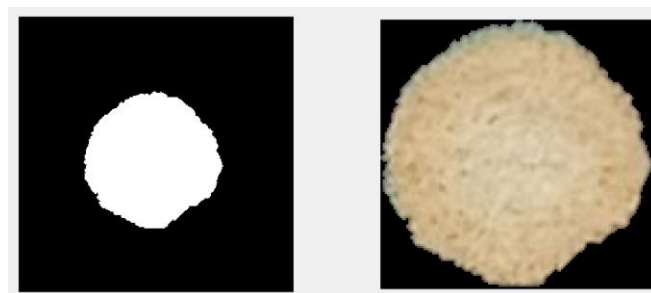


Figure 3.24 : Masking of the noiseless image Left: Noiseless Image with ROI mask, Right: Masked Image with ROI

(c) Feature Extraction

Image is converted to various color spaces or gamut available in MATLAB to get maximum possible derived feature vectors. Color spaces used includes HSV (Hue, Saturation, Value), L*a*b* (L* for perceptual lightness, and a* and b* for the colors), XYZ (X is a mix of RGB curves treated as non-negative, Y is the luminance, and Z is the quasi equal to blue), Grayscale (varies from 0 – 255), NTSC. Statistical feature vectors are obtained from all the color spaces which include mean, skewness (measurement of asymmetry of probability distribution), kurtosis (its value define the shape of probability distribution) makes 45 features. Image texture features are obtained by gray level co-occurrence matrix (GLCM) and then extracting statistical measures from this matrix which includes entropy which measures the randomness and intensity of gray scale pixels is also calculated. All of these features constitute a total of 51 features for each image.

3.2.3 Results and Discussion

(a) Classification Analysis

Many machine learning classifiers were developed and assessed on their capacity to properly predict concentration from the colorimetric rankings of different reagents in order to discover and recognize the association between color change and concentration. Smartphone-based technology was able to identify the best-performing classifiers for each combination from among the evaluated ones. Linear discriminant analysis (LDA) is the best of the classifiers that have been found so far. Herein, k for cross validation is set to 10 which was reported to be adequate to avoid issues caused by high bias and variance. In this case total 30 images are taken in each bin and concentrations are 10, 20, 30, 40, 50 mM respectively for each bin.

(b) Classification Model: Linear Discriminant Analysis (LDA)

A supervised classification technique called the LDA approach uses data to calculate the mean and variance for each class. If all classes have equal covariance, the Bayesian rule is used to calculate the maximum likelihood for the input data among all classes using the discriminant function to produce predictions. LDA occurs when we assume equal covariance between K classes. That is, rather of having a unique covariance matrix for each class, all classes share a common covariance matrix. Then we may derive the discriminant function as follows:

$$\delta_k(x) = x^T \Sigma^{-1} \mu_k - \frac{1}{2} \mu_k^T \Sigma^{-1} \mu_k + \log \pi_k \text{ ---- (2)}$$

Discriminant analysis classification essentially implies that we assign x to class k if x is in region k

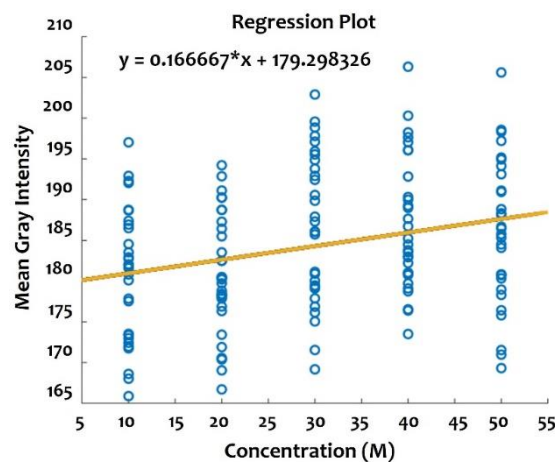


Figure 3.25: Variation of Gray mean intensity of image versus Concentration of glucose (10 – 50 mM)

Figure 3.25 shows the regression fit line for the glucose concentration value ranging from 10 – 50 mM of all the types of images i.e., 30 images in each pool of concentration with gray mean intensity of image i.e., mean of all pixel values of gray scale image of region of interest. It also

depicts that in spite of the variation of the smartphones and illumination conditions the regression fit remains same.

Table 3.2 : Training results of linear discriminant analysis

Accuracy (Validation)	72.7%
Speed of Prediction	~660 obs/sec
Training time	9.2782 sec

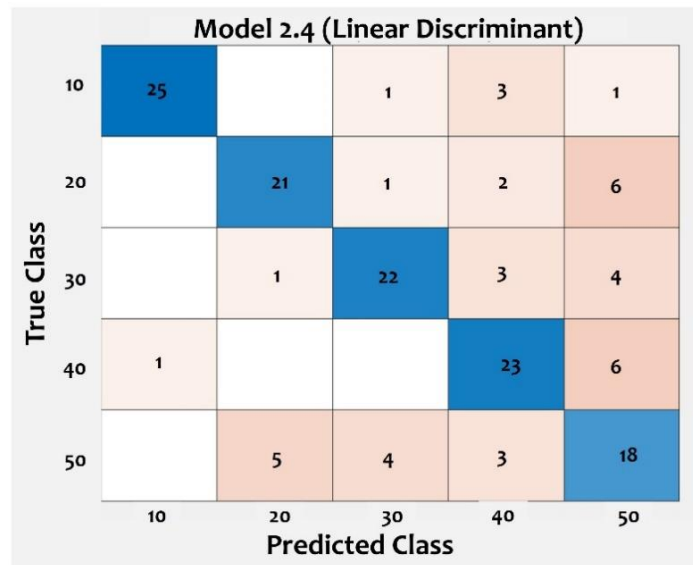


Figure 3.26: Confusion Matrix of the predicted and true class data

In Figure 3.26 the elements on diagonal show the true positives (Eq.3.2) and respective horizontal row with respect to that element are false positives (Eq.3.3) and respective vertical column depicts false negatives, and all left elements are true negatives. If we treat it as the binary classification problem, the area under curve in receiver operating characteristic curve can be found by the below formula. For multiclass classification accuracy (Eq.3.1) can be calculated as:

$$\text{Accuracy} = \frac{\text{Sum of element on diagonal}}{\text{Sum of element on matrix}} \quad (3.1)$$

$$= \frac{25 + 21 + 22 + 23 + 18}{150} = 0.7266$$

(c) Receiver operating characteristic (ROC) curve

A graph that shows how well a classification model performs at every level of categorization is called the receiver operating characteristic curve (ROC curve). It displays a plot of two parameters rate of True Positives and False Positives.

$$\text{True Positive Rate (TPR)} = \frac{\text{True Positive (TP)}}{\text{True Positive (TP)} + \text{False Negative (FN)}} \quad (3.2)$$

$$\text{False Positive Rate (FPR)} = \frac{\text{False Positive (FP)}}{\text{False Positive (FP)} + \text{True Negative (TN)}} \quad (3.3)$$

Figure 3.27 shows a ROC curve plots (TPR vs. FPR) at different classification thresholds. More items are classified as positive when the classification threshold is lowered, which raises the number of both False Positives and True Positives.

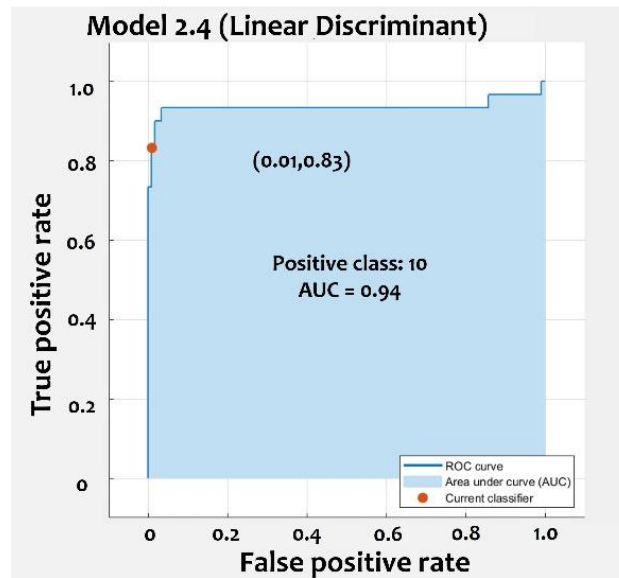


Figure 3.27: ROC curve demonstrating how well a classification model performs across all categorization thresholds

3.2.4 Conclusion

To categorize glucose concentration levels, machine learning regressors and classifiers were proposed in this study using a color change in μ PADs. In order to increase the platform's resistance to variations in illumination and camera optics, ML regressors and images taken with several cellphones in various lighting conditions were used to train classifiers. By this approach, it is possible to accurately determine the glucose level in a given sample irrespective of the brand, time, and lightning condition. If necessary, the platform can be taught to adjust to any other circumstance outside of its control. This work includes machine learning classifiers with paper-based detection technology for glucose measurement with an accuracy of (72.7%). The training dataset can be supplemented with images from additional smartphone brands taken using a variety of lighting conditions and ML techniques to further increase classification accuracy. The system's categorization sensitivity can also be enhanced by training it with higher concentration levels. Finally, performance could be enhanced by using deep learning and transfer learning techniques. Also, this research might be expanded to assist in the detection of blood, urine, etc. by modifying the algorithm as per the required conditions.

3.3 RESEARCH OUTCOMES

- Kishnani, V., Kumari, S. and Gupta, A., 2022. A chemometric-assisted colorimetric-based inexpensive paper biosensor for glucose detection. *Biosensors*, 12(11), p.1008, Doi: [10.3390/bios1211008](https://doi.org/10.3390/bios1211008)
- Kashyap, N., Kishnani, V. and Gupta, A., 2023. Colorimetric Based Glucose Detection with Machine Learning Approaches. *Recent Advances in Mechanical Engineering*, 2, p.466.