



THUYẾT MINH ĐỀ TÀI KHOA HỌC VÀ CÔNG NGHỆ CẤP SINH VIÊN 2024

A. THÔNG TIN CHUNG

A1. Tên đề tài

- Tên tiếng Việt (IN HOA): ĐO GIÁN TIẾP ĐỘ PH TRONG NƯỚC SỬ DỤNG CẢM BIẾN QUANG VÀ MÔ HÌNH HỌC MÁY
- Tên tiếng Anh (IN HOA): INDIRECT PH MEASUREMENT USING INFRARED SPECTROSCOPY AND MACHINE LEARNING

A2. Loại hình nghiên cứu

☒ Nghiên cứu cơ bản

☐ Nghiên cứu ứng dụng

☐ Nghiên cứu triển khai

A3. Thời gian thực hiện

..06.. tháng (kể từ khi được duyệt).

A4. Tổng kinh phí

(Lưu ý tính nhất quán giữa mục này và mục B8. Tổng hợp kinh phí đề nghị cấp)

Tổng kinh phí: ...6.. triệu đồng, gồm

- Kinh phí từ Trường Đại học Công nghệ Thông tin: ..6.. triệu đồng

A5. Chủ nhiệm

Họ và tên: **Nguyễn Duy Khanh**

Ngày, tháng, năm sinh: 12/02/2003

. Giới tính (Nam/Nữ): Nam

Số CMND: 068203000640

; Ngày cấp: 10/04/2021 ; Nơi cấp: Cục Cảnh sát Lâm Đồng.

Mã số sinh viên: 21520979

Số điện thoại liên lạc: 0915261749

Đơn vị (Khoa): Kỹ thuật máy tính

Số tài khoản: 6410974611

Ngân hàng: BIDV

A6. Thành viên đề tài

TT	Họ tên	MSSV	Khoa
1	Hà Việt Hoàng	21520240	Kỹ thuật máy tính
2	Nguyễn Duy Khanh	21520979	Kỹ thuật máy tính

B. MÔ TẢ NGHIÊN CỨU

B1. Giới thiệu về đề tài

Đo độ pH đóng vai trò quan trọng trong nhiều lĩnh vực, từ khoa học và kỹ thuật đến đời sống sinh hoạt hàng ngày. Trong lĩnh vực y tế, đo pH đóng vai trò quan trọng trong quá trình chẩn đoán và điều trị nhiều bệnh [1]. Trong lĩnh vực môi trường, đo độ pH giúp đánh giá mức độ ô nhiễm nguồn nước và quản lý xử lý nước thải. Ngoài ra, đo độ pH còn được ứng dụng trong nhiều lĩnh vực khác như: công nghiệp dệt may, sản xuất giấy, luyện kim. Chính vì vậy việc đo pH giúp chúng ta giám sát, kiểm soát, và điều chỉnh các yếu tố quan trọng trong nhiều lĩnh vực, từ đó đảm bảo chất lượng sản phẩm, an toàn sức khỏe cho con người và bảo vệ môi trường. Hiện nay, các phương pháp đo độ pH phổ biến bằng phương pháp trực tiếp như việc sử dụng bút đo pH cũng như cảm biến đo pH nhúng trực tiếp vào môi trường cần đo để xác định độ pH trong các dung dịch. Ví dụ với phương pháp đo pH bằng quỳ tím, ta chỉ sử dụng được 1 mẫu giấy cho 1 lần đo và cần phải có thêm một bước so sánh với bảng thang đo mới xác định được khoảng pH ước lượng dẫn tới nhiều sai số. Với phương pháp đo bằng bút đo pH, đầu bút thường bị hư hỏng khi liên tục tiếp xúc với nước, dẫn đến chi phí thay thế cao và ảnh hưởng đến độ tin cậy của kết quả đo. Hiện nay phương pháp đo độ pH gián tiếp chưa được phổ biến tại trong nước.

Ở các nước trên thế giới, việc đo pH gián tiếp với nhiều phương pháp khác nhau đã được nghiên cứu và ứng dụng vào thực tế như sử dụng đầu dò mô phỏng, laser,... [2] Tuy nhiên, với tiêu chí đo pH với giá thành hợp lý và độ chính xác cao, các phương pháp trên vẫn chưa thể đưa vào sử dụng phổ thông do thiếu nhiều yếu tố về giá thành, điều kiện sản xuất.

Với các vấn đề đã đặt ra, đề tài sẽ tập trung nghiên cứu, tìm hiểu và thiết kế bộ thí nghiệm đo pH trong nước bằng quang phổ hồng ngoại gần và áp dụng các mô hình máy học để đưa ra các dự đoán về độ pH với độ chính xác cao với giao diện phần mềm tích hợp dễ sử dụng.

B2. Mục tiêu, nội dung, kế hoạch nghiên cứu

B2.1 Mục tiêu

Mục tiêu tổng quan: Đề tài tập trung nghiên cứu và thiết kế bộ thí nghiệm đo độ pH trong nước áp dụng những ảnh hưởng của quang phổ cận hồng ngoại đến phân tử từ đó kết hợp với các phương pháp học máy để đưa ra mô hình dự đoán độ pH có độ chính xác hơn 80% so với các thiết bị thương mại có sẵn trên thị trường. Thiết kế và xây dựng bộ thí nghiệm mang tính chất thử nghiệm có khả năng đo pH.

Mục tiêu cụ thể:

- Tìm hiểu, nghiên cứu các phương pháp đo pH hiện nay.
- Xác định các bước sóng có ảnh hưởng mạnh nhất với phân tử nước.
- Xây dựng bộ dữ liệu với 200 mẫu, từ đó áp dụng các phương pháp tiền xử lý dữ liệu, mô hình máy học để đưa ra mối tương quan giữa các bước sóng thu được khi đã đi qua nước và tán xạ vào cảm biến với độ pH thực tế.
- Nghiên cứu các cảm biến, nguồn thu và nguồn phát phù hợp với đề tài.
- Thiết kế phần cứng với giá thành rẻ.
- So sánh và đưa ra các điều chỉnh phù hợp với độ sai số so với độ pH 9.2%.

[3]

B2.2 Nội dung và phương pháp nghiên cứu

2.2.1. Nội dung thực hiện

- Nội dung 1: Tìm hiểu tổng quan đề tài.
- Nội dung 2: Thiết kế cấu trúc phần cứng của hệ thống.
- Nội dung 3: Lập trình vi điều khiển, lập trình ứng dụng.
- Nội dung 4: Thu thập bộ dữ liệu và xây dựng mô hình máy học. Chọn ra bước sóng và mô hình có độ chính xác cao.
- Nội dung 5: Đóng gói, thử nghiệm, đánh giá sản phẩm.

2.2.2. Phương pháp nghiên cứu

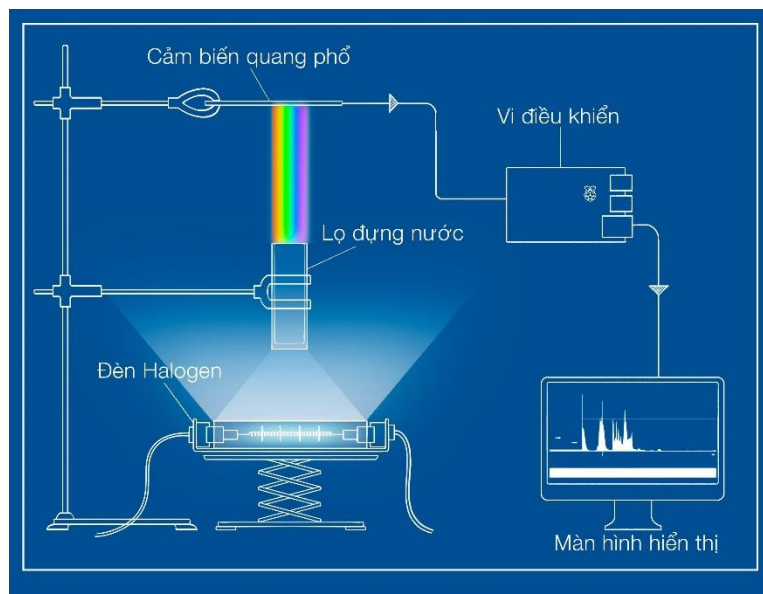
Đề tài sẽ hướng tới việc xây dựng một thiết bị đo pH nước ứng dụng các tính chất của quang phổ hồng ngoại gần và trí tuệ nhân tạo. Các phần cứng phục vụ cho thiết bị phải đáp ứng được các yêu cầu đặt ra. Vi điều khiển có thể chạy, lưu trữ các mô hình máy học,

các phương pháp tiền xử lý. Cảm biến quang phổ phải thu được các bước sóng trong khoảng NIR và giá thành hợp lí. Nguồn phát hay đèn có thể phát ra ánh sáng có những bước sóng phù hợp và bao phủ vùng NIR.

2.2.2.1. Xây dựng hệ thống phần cứng

Phần cứng cơ bản sẽ gồm 4 thành phần chính:

- Vi điều khiển: là bộ điều khiển trung tâm, nơi tiếp nhận xử lý các tín hiệu từ cảm biến quang phổ từ đó xử lý dữ liệu, và áp dụng các mô hình máy học, để đưa ra dự đoán.
- Cảm biến quang phổ: dùng để thu thập các bước sóng được phát ra từ nguồn phát sau khi đi qua nước và tán xạ vào cảm biến; cảm biến quang phổ phải thu thập các bước sóng nằm trong dải NIR với giá thành dễ tiếp cận.
- Nguồn sáng: đèn Halogen: loại đèn phổ biến nhất trên thị trường hiện hoặc đèn Leds với công suất tiêu thụ điện thấp và tuổi thọ sử dụng lâu dài. Tuy nhiên, đèn Leds với dải bước sóng rộng chưa được phổ biến và giá thành cao, và loại Leds có thể dễ dàng tiếp cận thì chỉ có thể phát ra ánh sáng với 1 bước sóng cố định, vì thế việc sử dụng nhiều đèn Leds sẽ cồng kềnh và không phù hợp với đề tài. Đèn Halogen có khả năng phát ra ánh sáng với dải bước sóng trải dài từ 350 nm – 1000 nm và có tuổi thọ tương đối cao, vì thế sẽ phù hợp với đề tài hơn.

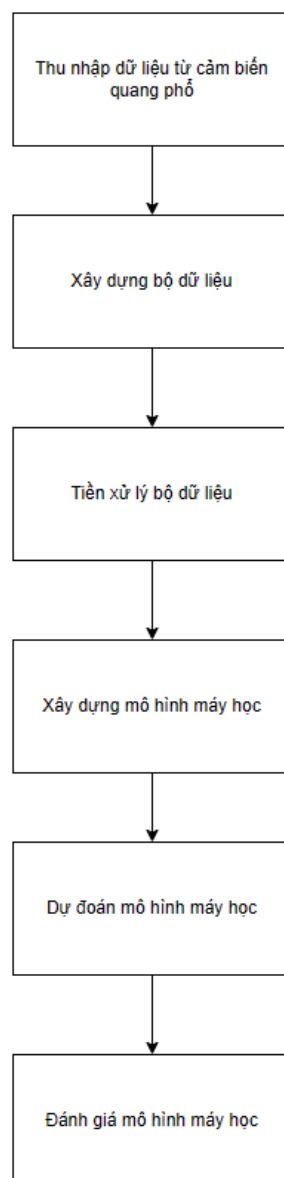


Hình 1: Bộ thí nghiệm với những phần cứng chính

2.2.2.2. Xây dựng hệ thống phần mềm

Hình 2 mô tả mô hình xây dựng dự kiến của hệ thống, thiết bị sẽ thực hiện việc đo lấy dữ liệu từ cảm biến, sau đó được đưa vào các phương pháp tiền xử lý để tối ưu hóa dữ liệu.

Dữ liệu này sẽ được phục vụ cho quá trình huấn luyện các mô hình máy học. Sau khi hoàn tất huấn luyện, dựa vào các phương pháp đánh giá sẽ chọn ra một mô hình máy học phù hợp nhất. Từ đó thực hiện việc đo và dự đoán độ pH, cuối cùng sẽ kết quả dự đoán sẽ hiển thị trên màn hình. Vi điều khiển đóng vai trò quan trọng trong việc chuyển đổi quang phổ dữ liệu thành giá trị pH và điều khiển các cảm biến, đèn Halogen. Khi khởi động, vi điều khiển sẽ truyền tín hiệu để khởi tạo các cảm biến sau đó sẽ khởi động nguồn phát. Khi ánh sáng của đèn Halogen được chiếu xuyên qua nước thì cảm biến sẽ thu được giá trị của các bước sóng. Khi đã có dữ liệu từ cảm biến, vi điều khiển sẽ xử lý dữ liệu và áp dụng các mô hình máy học để huấn luyện dữ liệu.



Hình 2: Mô hình xây dựng của hệ thống

2.2.2.3. Phương pháp chọn bước sóng

Hệ thống sẽ ứng dụng những tính chất của quang phổ hồng ngoại gần cụ thể là ở bước sóng nằm trong khoảng $750\text{ nm} - 970\text{ nm}$ (SW-NIR) bởi sự hấp thụ của gốc -OH trong phân tử nước đến dải bước sóng này, bước sóng 970 nm sẽ được quan tâm nhiều để nghiên cứu và thử nghiệm vì nó đã được sử dụng ở các nghiên cứu khác nhau [3][4]. Đèn halogen chính là sự lựa chọn phù hợp để làm nguồn phát, từ đó ta sẽ phát hiện được chính xác các bước sóng tác động mạnh đến phân tử nước và tối ưu hóa số lượng bước sóng để có thể tăng tốc độ xử lý để đưa ra kết quả một cách nhanh nhất nhưng vẫn đảm bảo về độ chính xác.

2.2.2.4. Xây dựng bộ dữ liệu

Với mục tiêu đã đề ra, đề tài sẽ thu thập đa dạng dữ liệu hơn bằng cách tăng độ trải dài pH của các mẫu nước từ đó bộ dữ liệu sẽ trực quan và đáng tin cậy hơn. Bộ thí nghiệm sẽ áp dụng các mô hình máy học để tìm ra mối liên hệ giữa các bước sóng và độ pH thực tế, bút đo pH sẽ được dùng để thu thập dữ liệu về độ pH thực tế từ đó các mô hình máy học tìm ra được sự liên quan giữa bước sóng và pH thực tế.

2.2.2.5. Phương pháp máy học

Phân tích hồi quy là một phương pháp thống kê được sử dụng để nghiên cứu mối quan hệ giữa một biến phụ thuộc và một hoặc nhiều biến độc lập. Mục tiêu của phân tích hồi quy là xác định mô hình toán học tốt nhất để dự đoán hoặc giải thích biến phụ thuộc dựa trên các biến độc lập. Công trình sẽ ứng dụng các mô hình hồi quy để tìm ra mối tương quan giữa độ pH và các bước sóng được thu bởi cảm biến. Sau khi chọn ra được mô hình hồi quy phù hợp, mô hình sẽ được nạp lên vi điều khiển và tiến hành dự đoán độ pH của nước.

2.2.2.6. Phương pháp đánh giá kết quả

Đánh giá mô hình học máy bằng các thông số:

- R-square (Hệ số xác định): Đo lường phần trăm phương sai của biến phụ thuộc mà mô hình có thể giải thích. Hệ số xác định giữa 0 và 1. Giá trị càng cao, mô hình càng tốt. Nếu R-square là 1, mô hình dự đoán hoàn toàn giống với dữ liệu thực tế.
- RMSE (Root Mean Squared Error - Sai số trung bình bình phương căn bậc hai): Đo lường độ lớn của sai số giữa giá trị dự đoán và giá trị thực tế. RMSE tính giá trị trung bình

của sai số, giúp đánh giá sự chênh lệch trung bình giữa dự đoán và thực tế. Giá trị càng thấp, mô hình càng chính xác.

B2.3 Kế hoạch nghiên cứu

STT	Nội dung	Thời gian					
		Tháng thứ 1	Tháng thứ 2	Tháng thứ 3	Tháng thứ 4	Tháng thứ 5	Tháng thứ 6
1	Nội dung 1	X					
2	Nội dung 2		X	X			
3	Nội dung 3			X	X		
4	Nội dung 4				X	X	
5	Nội dung 5						X

B3. Kết quả dự kiến

Các kết quả cần đạt được đề làm tiêu chí đánh giá trong thời gian 6 tháng thực hiện đề tài như sau:

1. Phần cứng được thiết kế ổn định, chắc chắn, hạn chế bị nhiễu bởi điều kiện môi trường và sai số giữa các lần đo.
2. Xây dựng được dataset lớn hơn 200 mẫu với độ pH đa dạng của các mẫu nước.
3. Chọn ra 3 - 4 bước sóng ảnh hưởng cao nhất đến phân tử glucose nhằm tối ưu hóa bộ dữ liệu
4. Các phương pháp đánh giá mô hình máy học đạt hiệu quả tốt. Chọn ra một mô hình đạt hiệu quả cao nhất (R^2 phải đạt trên 70%), phù hợp với vi điều khiển.
5. Hệ thống hoạt động ổn định có độ chính xác cao. Đạt kết quả sai số trung bình giữa thiết bị và độ pH chuẩn không quá 9.2%.

B4. Tài liệu tham khảo

- [1] M. Frazzoni, N. De Bortoli, L. Frazzoni, S. Tolone, V. Savarino, and E. Savarino, “Impedance-PH monitoring for diagnosis of reflux Disease: New perspectives,” *Digestive Diseases and Sciences*, vol. 62, no. 8, pp. 1881–1889, May 2017.
- [2] M. C. O. Monteiro and M. T. M. Koper, “Measuring local pH in electrochemistry,” *Current Opinion in Electrochemistry*, vol. 25, p. 100649, Feb. 2021.
- [3] D. Li and L. Li, “Detection of water pH using visible Near-Infrared spectroscopy and One-Dimensional convolutional neural network,” *Sensors*, vol. 22, no. 15, p. 5809, Aug. 2022.
- [4] B. M. Saalidong, S. A. Aram, S. Otu, and P. O. Lartey, “Examining the dynamics of the relationship between water pH and other water quality parameters in ground and surface water systems,” *PloS One*, vol. 17, no. 1, p. e0262117, Jan. 2022.

Ngày __ tháng __ năm 20__
Giảng viên hướng dẫn
(Ký và ghi rõ họ tên)

TS. Phạm Quốc Hùng

Ngày 25 tháng 4 năm 2024
Chủ nhiệm đề tài
(Ký và ghi rõ họ tên)



Nguyễn Duy Khanh