

Ứng dụng module internet vạn vật và thuật toán nhận diện ký tự quang học để trích xuất tự động số đọc của máy đo phóng xạ

Phạm Nguyên Phát, Nguyễn Tấn Được*

Ngành Vật lý Y sinh – Khoa Y, Trường Đại học Nguyễn Tất Thành, Thành phố Hồ Chí Minh

*ntduoc@ntt.edu.vn

Tóm tắt

Với sự phát triển của công nghệ kết nối internet vạn vật, các thiết bị đo đặc bức xạ ngày nay thường được thiết kế với khả năng kết nối mạng không dây. Tuy nhiên, trong thực tế nhiều thiết bị lại thiếu tính năng này, gây khó khăn trong việc trích xuất dữ liệu từ xa cũng như tích hợp vào chung một hệ thống quan trắc phóng xạ đa điểm. Nghiên cứu này nhằm thiết kế một hệ thống có tên RadOCR giúp trích xuất từ xa số đọc được hiển thị trên màn hình LCD của máy đo phóng xạ Radiation Alert Ranger® theo thời gian thực. Hệ thống bao gồm một thiết bị gắn kèm với máy đo và một phần mềm trên máy tính. Thiết bị RadOCR bao gồm một module ESP32-CAM giúp ghi nhận hình ảnh số đọc trên LCD của máy đo phóng xạ và gửi dữ liệu đến phần mềm thông qua kết nối wifi trên mạng nội bộ. Phần mềm sử dụng các thuật toán xử lý ảnh trong thư viện OpenCV và mô hình nhận diện ký tự quang học Tesseract OCR để trích xuất, hiển thị kết quả này trên một giao diện đồ họa và lưu trữ theo các mốc thời gian trong file csv. Kết quả thử nghiệm cho thấy hệ thống hoạt động ổn định với độ chính xác $(97,365 \pm 0,082) \%$.

© 2025 Journal of Science and Technology - NTTU

Nhận 28/07/2025

Được duyệt 16/09/2025

Công bố 28/10/2025

Từ khóa

internet vạn vật, nhận diện ký tự quang học, quan trắc phóng xạ

1 Đặt vấn đề

Ứng dụng năng lượng nguyên tử để phát triển kinh tế xã hội là một chiến lược lớn được đề ra trong Quy hoạch phát triển, ứng dụng năng lượng nguyên tử thời kỳ đến năm 2030, tầm nhìn đến năm 2050 của nước ta. Trong đó, năng lượng nguyên tử được kỳ vọng ngày càng có mặt rộng rãi trong các lĩnh vực từ nông nghiệp, công nghiệp, tài nguyên môi trường, năng lượng cho đến y tế [1]. Trong bối cảnh đó, quan trắc và đảm bảo an toàn bức xạ là công tác có ý nghĩa vô cùng quan trọng. Các thiết bị chủ yếu được sử dụng cho công tác này bao gồm liều kế và máy đo khảo sát phóng xạ (radiation survey meter). Liều kế thường được trang bị cho nhân viên làm công việc liên quan đến bức xạ, nhằm ghi lại liều lượng mà người đó đã phơi nhiễm trong quá trình làm việc. Mức liều lượng này phải thấp

hơn mức giới hạn an toàn theo quy định. Máy đo khảo sát bức xạ thường được sử dụng để quan trắc suất liều môi trường. Các máy đo này thường có cấu tạo gồm đầu dò bức xạ dạng chứa khí hoặc nhấp nháy rắn, hệ mạch điện tử xử lý tín hiệu, và bộ hiển thị kết quả đo (số đọc – readout) dạng đồng hồ analog hoặc màn hình LCD. Hiện nay, một số máy còn có khả năng truyền kết quả về máy tính thông qua kết nối bằng dây cáp, bluetooth, hoặc mạng wifi. Tuy nhiên, chi phí giấy phép (license) cho các tính năng này còn cao. Thêm vào đó, việc kết nối liên thông các thiết bị khác nhau vào cùng một hệ thống cũng là vấn đề nan giải. Ngoài ra, tại nhiều cơ sở vẫn đang sử dụng các máy đo khảo sát bức xạ thế hệ cũ, thiếu khả năng truyền tín hiệu để đo đặc từ xa hoặc liên kết thành mạng lưới quan trắc đa điểm. Để “internet hóa” các thiết bị dạng này, cần can thiệp vào phần cứng để gắn thêm module trích xuất tín

hiệu và kết nối mạng. Cách làm này sẽ ảnh hưởng đến độ chính xác, tính ổn định và an toàn của thiết bị. Một cách tiếp cận khác là sử dụng các kỹ thuật nhận diện ký tự quang học (optical character recognition – OCR) để trích xuất số đọc của thiết bị, sau đó sử dụng module internet vạn vật (Internet of things – IoT) để truyền kết quả này lên mạng chung.

OCR là một kỹ thuật của thị giác máy tính (computer vision) dùng để chuyển chuỗi ký tự dạng hình ảnh thành chuỗi ký tự có thể xử lý được trên máy tính. Một hệ thống OCR bao gồm các thành phần chính như: thiết bị quét hình ảnh quang học; các thuật toán định vị vùng phân đoạn, tiền xử lý, phân đoạn, tái thể hiện, trích xuất đặc trưng, huấn luyện và nhận diện, và hậu xử lý [2]. Bước đầu tiên trong tiến trình OCR là sử dụng thiết bị quét để số hóa hình ảnh. Vùng ảnh chứa chuỗi ký tự sẽ được lấy ra bằng các thuật toán định vị và phân đoạn. Thuật toán tiền xử lý giúp loại bỏ nhiễu trước khi ảnh được đưa vào bước trích xuất đặc trưng. Nội dung tương ứng với chuỗi ký tự trên mẫu ảnh sẽ được tạo ra bằng cách phân loại các đặc trưng trích xuất được dựa trên mô hình đã được huấn luyện với dữ liệu mẫu trước đó. Công đoạn khó khăn nhất trong tiến trình này là trích xuất đặc trưng (feature extraction) và phân loại (classification). Các mô hình OCR hiện đại trong những năm gần đây tập trung vào việc sử dụng mạng neural tích chập (convolution neural network – CNN) để trích xuất các thuộc tính bậc cao trong ảnh. Các thuộc tính này sau đó được phân loại bằng mạng neural nhân tạo (artificial neural network – ANN) [3]. Một số mô hình OCR sử dụng cách tiếp cận này có thể kể đến là Tesseract OCR, EasyOCR, Google Cloud Vision OCR, PaddleOCR, docTR. Ưu điểm của các mô hình sử dụng kỹ thuật học sâu là độ chính xác cao, áp dụng được cho nhiều loại ngôn ngữ, có thể được tinh chỉnh và huấn luyện chuyển tiếp (transfer training) để phù hợp cho từng trường hợp cụ thể. Ngoài ra, các mô hình đã trải qua huấn luyện với tập dữ liệu mẫu đủ lớn có thể được sử dụng để phát triển các ứng dụng như đọc biển số xe hoặc chỉ số đồng hồ điện, nước, gas [4-6]; trích xuất thông tin bìa hoặc gáy sách [7]; truy xuất nguồn gốc và truy vết đường đi của thực phẩm [8].

Tesseract OCR là mô hình nhận diện ký tự quang học được phát triển lần đầu tiên bởi Ray Smith và Hewlett-Packard vào năm 1985, được tài trợ và phát triển bởi

Google từ năm 2006 [9]. Từ phiên bản 4.0 trở lên, mô hình sử dụng mạng neural LSTM (long short-term memory neural network). Hiện nay, Tesseract OCR là mô hình mã nguồn mở được cung cấp miễn phí, không ngừng được bổ sung và phát triển bởi các chuyên gia trên khắp thế giới. Trong ngôn ngữ Python, việc giao tiếp với Tesseract OCR có thể được thực hiện thông qua trình đóng gói API (API wrapper) có tên là PyTesseract [10]. Người dùng có thể sử dụng mô hình đã được huấn luyện sẵn hoặc tái huấn luyện bằng tập dữ liệu của mình. Nhờ đó, Tesseract OCR có thể được sử dụng cho nhiều dự án khác nhau từ đơn giản như đọc biển số, chỉ số đồng hồ điện nước cho đến phức tạp như nhận diện chữ viết tay hoặc ký tự phi Latinh.



Hình 1 Máy đo phóng xạ Radiation Alert Ranger®

Máy đo Radiation Alert Ranger® (Hình 1) là thiết bị đo khảo sát phóng xạ môi trường. Máy được trang bị đầu dò bức xạ dạng buồng chứa khí Geiger-Mueller với độ nhạy cao, có thể đo đặc bức xạ alpha, beta, gamma, và tia X. Ngưỡng phát hiện của máy đối với tia alpha có năng lượng tối thiểu 2 MeV, tia beta tối thiểu 0,40 keV, tia gamma tối thiểu 10 keV. Số đọc được hiển thị trên màn hình LCD lớn với nhiều chế độ (mode) đo khác nhau tùy theo lựa chọn của người dùng. Tần số cập nhật số đọc trên màn hình là 1 lần mỗi giây. Máy đo được thiết kế nhỏ gọn, dạng cầm tay phù hợp cho các ứng dụng cần tính di động cao. Đây là một trong những thiết bị đo bức xạ được sử dụng rất phổ biến tại Việt Nam. Tuy nhiên, một số model cũ không được tích hợp tính năng truyền tín hiệu không dây và đo đặc từ xa. Do đó, máy đo Radiation Alert Ranger® được sử dụng trong nghiên cứu này như một đối tượng điển hình để xây dựng phần cứng và phần mềm phục vụ việc trích xuất số đọc từ xa qua kết nối không dây.

Trong nghiên cứu này, các tác giả ứng dụng mô hình Tesseract OCR đã được huấn luyện sẵn để trích xuất tự động số đọc hiển thị trên màn hình LCD của máy đo phóng xạ Radiation Alert Ranger®. Module IoT

ESP32-CAM được sử dụng để chụp ảnh màn hình LCD của máy đo, đồng thời truyền luồng ảnh ghi được về phần mềm xử lý (được gọi là RadOCR) trên một máy tính thông qua kết nối mạng wifi. Trước khi đưa vào Tesseract OCR để trích xuất số đọc, một loạt thao tác xử lý ảnh được thực hiện để loại bỏ các chi tiết gây nhiễu nhằm tăng độ chính xác của kết quả được trích xuất. Phần mềm RadOCR còn được thiết kế với giao diện đồ họa (graphical user interface – GUI) để thuận tiện cho việc hiển thị số đọc theo thời gian thực hoặc thống kê kết quả dưới dạng file csv. Thử nghiệm đánh giá thiết bị và phần mềm RadOCR cũng được tiến hành trên 2000 ảnh. Kết quả cho thấy độ chính xác đạt $(97,365 \pm 0,082) \%$.

2 Phương pháp

2.1 Thiết kế và gia công phần cứng

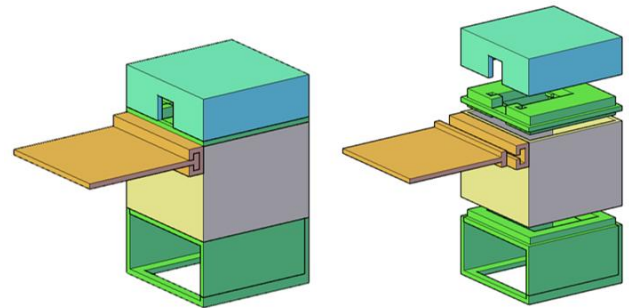
Phần cứng của thiết bị RadOCR bao gồm một module ESP32-CAM có tích hợp camera OV2640, nguồn điện 5 V, và một bộ giá đỡ để cố định vị trí của camera và máy đo. Module ESP32-CAM là một bản mạch IoT được sử dụng rất phổ biến với giá thành thấp. ESP32-CAM có kích thước chỉ $(27,0 \times 40,5 \times 4,5)$ mm, rất nhỏ gọn phù hợp cho các ứng dụng tích hợp IoT.



Hình 2 Minh họa cách kết nối và truyền thông tin giữa ESP32-CAM và phần mềm RadOCR
Module này cho phép giao tiếp thông qua kết nối không dây wifi, bluetooth; hoặc kết nối nối tiếp (serial) qua

cổng micro USB. Camera OV2640 có độ phân giải tối đa 2 MP cho phép ghi hình rõ trong khoảng cách gần. Trong nghiên cứu này, module ESP32-CAM được sử dụng như một máy ảnh IoT để ghi hình số đọc hiển thị trên LCD của máy đo Radiation Alert Ranger®. Các hình ảnh này sẽ được truyền về máy tính có cài phần mềm xử lý thông qua kết nối wifi chung đến một mạng nội bộ bất kỳ như minh họa trong Hình 2.

Bộ giá đỡ có tác dụng cố định vị trí giữa ESP32-CAM và máy đo phóng xạ Radiation Alert Ranger®. Giá đỡ được thiết kế trên phần mềm AutoCAD dưới dạng module để dễ dàng lắp ghép lên máy đo. Hình 3 là mô phỏng 3D của giá đỡ này. Nó được thiết kế thành 4 tầng riêng biệt, được kết nối với nhau thông qua các khớp nối ma sát. Tầng 1 có thể trượt vào khu vực màn hình của máy đo phóng xạ, giúp cố định thiết bị chắc chắn lên máy đo. Tầng 2 là bộ phận trung gian kết nối tầng 1 và tầng 3, giúp giữ khoảng cách cố định từ camera tới LCD của máy đo để hình ảnh thu được rõ nét. Bên ngoài tầng 2 được gắn với một chân chống giúp thiết bị đứng vững khi gắn vào máy đo. Tầng 3 là khoang chứa ESP32-CAM và dải đèn LED cung cấp ánh sáng. Tầng 4 là nắp đậy giúp cố định cũng như bảo vệ camera. Các bộ phận của bộ giá đỡ được gia công bằng phương pháp in 3D trên chất liệu nhựa nhiệt dẻo acrylonitrile butadiene styrene (ABS).



Hình 3 Thiết kế của bộ giá đỡ



Hình 4 Các module của phần mềm

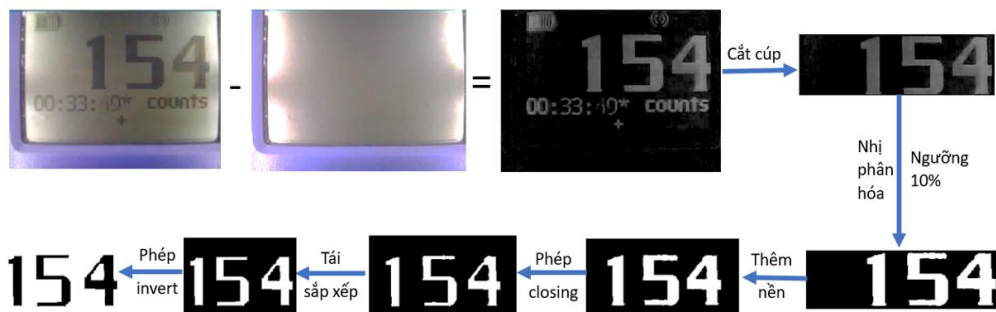
2.2 Cấu trúc phần mềm

Hình 4 trình bày cấu trúc của phần mềm. Trên ESP32-CAM, một firmware được cài đặt để liên tục chụp ảnh

màn hình máy đo phóng xạ, gửi luồng ảnh này đến phần mềm xử lý thông qua một socket mạng. Module ESP32-CAM sẽ được định danh trên mạng thông qua

địa chỉ IP và port. Phần mềm xử lý trên máy tính được lập trình bằng ngôn ngữ Python, bao gồm 3 module chính: module giao tiếp camera, module xử lý ảnh, và module giao diện đồ họa người dùng (GUI).

Module giao tiếp camera quản lý socket kết nối đến firmware để nhận luồng ảnh liên tục theo thời gian thực với tốc độ mặc định là 10 khung hình mỗi giây (frame per seconds - fps), kích thước mỗi ảnh là (320×240)



Hình 5 Các bước xử lý ảnh

Module xử lý ảnh là trái tim của cả hệ thống phần mềm. Nó chịu trách nhiệm chuyển đổi số đọc dạng hình ảnh thành số đọc dạng số bằng cách sử dụng thuật toán nhận diện ký tự quang học Tesseract OCR. Tuy nhiên, việc sử dụng trực tiếp các khung hình thô từ module giao tiếp camera để chạy thuật toán Tesseract OCR sẽ cho kết quả rất tệ vì ảnh đầu vào chứa nhiều chi tiết gây nhiễu. Do đó, ảnh thô cần phải được tiền xử lý (pre-processing) [1] qua nhiều bước như minh họa trên Hình 5. Các bước này được thực hiện chủ yếu dựa vào bộ thư viện OpenCV của Python. Trừ nền và cắt cúp là bước giúp loại bỏ vùng nền (background) chứa các chi tiết không liên quan và khu trú vào vùng chứa số đọc. Nhị phân hóa (với ngưỡng 10 %) sẽ giúp phân đoạn ảnh thành vùng chữ số và nền với độ tương phản tối đa. Phép biến đổi hình thái học đóng (morphology closing) được sử dụng để tiếp tục loại bỏ nhiễu và làm mịn đường biên của các chữ số. Bước phân tách và tái sắp xếp giúp khoảng cách giữa các chữ số trở nên đều nhau nhằm tăng hiệu suất nhận diện. Ảnh sau xử lý tái sắp xếp cần được nghịch đảo (invert) cường độ trước khi đưa vào Tesseract OCR theo khuyến cáo. Các thông số thiết lập cho mô hình Tesseract OCR bao gồm: loại ngôn ngữ (Lang), tùy chỉnh phân đoạn (Page Segmentation Modes - PSM), tùy chọn dữ liệu đầu ra (Output), và bộ tham số mô hình. Để phù hợp cho việc nhận diện số đọc, nghiên cứu này sử dụng một bộ tham

pixel, tương ứng với kích thước vật lý là $(2,0 \times 1,6)$ inch. Luồng ảnh này nếu không bị giới hạn lại thì sẽ làm quá tải các module phía sau. Do đó, module giao tiếp camera có nhiệm vụ tải lấy mẫu trên luồng ảnh với tốc độ 2 khung hình mỗi giây. Tốc độ tải lấy mẫu này được chọn dựa trên thực tế là tần số cập nhật số đọc lên màn hình máy đo Radiation Alert Ranger® là 1 lần mỗi giây.

số đã được Shreeshrii huấn luyện trên 8704 ảnh, dành riêng cho ký tự dạng chữ số [12]. Tesseract OCR sẽ trả về kết quả bao gồm giá trị nhận diện được (OCR value) và độ tự tin của thuật toán (confident factor). Chỉ những giá trị được nhận diện với độ tự tin tối thiểu 90 % mới được chấp nhận.

2.3 Phương pháp chạy thử và đánh giá

Do mục tiêu chính của đề tài là phát triển thiết bị và phần mềm giúp trích xuất số đọc từ máy đo phóng xạ Radiation Alert Ranger® nên phần này chỉ trình bày việc kiểm thử và đánh giá độ chính xác của quá trình trích xuất số đọc, không liên quan đến độ chính xác đo đạc của máy đo. Hình 6 mô tả bố trí hình học của phép đo. Một nguồn phóng xạ Cs-137, hoạt độ 180,61 μCi (hiệu chỉnh suy giảm đến 01/02/2024) được sử dụng để tạo tín hiệu cho máy đo. Để thu được nhiều kết quả đo khác nhau, cần giảm hoạt độ nguồn để không gây bão hòa kết quả đo của máy. Điều này có thể đạt được bằng cách giữ nguyên nguồn phóng xạ trong vỏ chì và giữ cho khoảng cách từ nguồn đến máy là 26 cm. Để thu được các kết quả trải đều trên toàn bộ thang đo của máy (0 đến 9999), cần thực hiện chạy thử trong thời gian 10 phút mỗi lần

Cỡ mẫu tối thiểu cần lấy được tính toán bằng công thức (1) [13]

$$n = \frac{t_{\alpha}^2 \times p(1-p) \times N}{(N-1) \times \alpha^2 + t_{\alpha}^2 \times p(1-p)} \quad (1)$$

Trong đó, N là số giá trị khả dĩ của số đọc (kích thước đám đông), p là tỉ lệ chính xác, α là mức nghĩa thống kê, t_{α} là giá trị của phân bố chuẩn tương ứng với khoảng tự tin (confident interval, $CI = 1 - \alpha$). Trong khảo sát này, mức ý nghĩa thống kê được chọn là $\alpha = 0,05$. Khi đó, $t_{0,05} = 1,96$. LCD của máy đo Radiation Alert Ranger® có thể hiển thị 4 chữ số từ 0 đến 9. Do đó, số giá trị khả dĩ của số đọc là $N = 10\,000$. Do tỉ lệ chính xác chưa biết nên chọn giá trị điển hình là $p = 0,5$. Khi đó, cỡ mẫu tính được là $n = 380$. Đây là kích thước mẫu tối thiểu cần khảo sát để đạt được mức ý nghĩa thống kê 0,05. Để đạt được cỡ mẫu này, đồng thời đảm bảo tính ngẫu nhiên độc lập và phân phối đồng nhất (independent and identical distribution – IID) của mẫu được chọn, cần duy trì mỗi lần đo trong khoảng 10 phút.



Hình 6 Bố trí hình học của phép đo độ chính xác

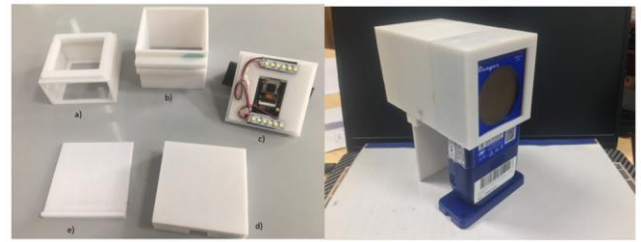
Ở mỗi lần đo, một loạt các số đọc kèm theo kết quả OCR sẽ được lưu lại. Các hình ảnh này sẽ được lọc để loại bỏ các số đọc trùng nhau. Thao tác kiểm tra thủ công sẽ được thực hiện trên từng ảnh để đếm số trường hợp cho kết quả OCR đúng so với số đọc được hiển thị. Khi đó, độ chính xác sẽ được tính theo công thức (2)

$$P (\%) = \frac{T}{S} \times 100 \quad (2)$$

Trong đó, T là số lượng kết quả đúng, S là cỡ mẫu thực tế.

3 Kết quả và bàn luận

Hình 7 là các bộ phận và lắp ráp hoàn chỉnh của thiết bị RadOCR. Hình 8 là giao diện đồ họa (GUI) của phần mềm. Giao diện này có các khung để hiển thị theo thời gian thực ảnh đầu vào được truyền từ ESP32-CAM, ảnh chứa kết quả OCR; các nút chức năng để khởi động (start), ngừng (stop) quá trình đo, hoặc xuất (export) file thống kê csv.



Hình 7 Các thành phần của thiết bị RadOCR trước và sau khi được lắp ráp hoàn chỉnh vào máy đo



Hình 8 Giao diện phần mềm RadOCR

Bảng 1 trình bày kết quả đánh giá độ chính xác trong trích xuất số đọc của thiết bị và phần mềm RadOCR. Có tất cả 5 mẫu được chọn. Độ chính xác trung bình đạt $(97,365 \pm 0,832) \%$. Trong đó, độ chính xác thấp nhất là 96,190 %, cao nhất là 98,165 %. Kết quả cho thấy thiết bị và phần mềm RadOCR hoạt động ổn định và có độ chính xác cao.

Bảng 1 Kết quả đánh giá độ chính xác

Mẫu	Cỡ mẫu (S)	Số kết quả sai (S – T)	Số kết quả đúng (T)	Độ chính xác P (%)
Mẫu 1	409	9	400	97,799
Mẫu 2	420	16	404	96,190
Mẫu 3	414	13	401	96,860
Mẫu 4	411	9	402	97,810
Mẫu 5	436	8	428	98,165
Trung bình				97,365
Độ lệch chuẩn				0,832

Độ chính xác của RadOCR được phát triển trong nghiên cứu này nhất quán với các nghiên cứu khác cũng sử dụng Tesseract OCR cho nhận diện ký tự dạng số. Nghiên cứu [8] sử dụng mô hình này để trích xuất số serial trên bao bì thực phẩm với độ chính xác đạt được là 87,5 % trong trường hợp ảnh đầu vào được tiền xử

lý bằng các công cụ trong thư viện OpenCV. Độ chính xác giảm còn 62,0 % nếu ảnh đầu vào không trải qua công đoạn tiền xử lý. Nghiên cứu [7] cũng sử dụng phương pháp tương tự để trích xuất tự động mã số sách dựa vào hình ảnh gáy sách với độ chính xác đạt được hơn 90 %. Ngoài ra, tuy tận dụng mô hình Tesseract OCR được huấn luyện sẵn, độ chính xác trong nghiên cứu của chúng tôi vẫn đạt mức có thể so sánh được với các nghiên cứu sử dụng mạng neural tích chập được huấn luyện trên tập dữ liệu lớn. Mô hình CNN truyền thống được sử dụng trong nghiên cứu [6] và [14] đối với hình ảnh là chỉ số đồng hồ gas có độ chính xác được báo cáo lần lượt là 96 % và 85,71 %. Trong khi đó, các mô hình CNN phức tạp hơn giúp đạt độ chính xác cao hơn. Nghiên cứu [15] đã sử dụng mạng neural tích chập đa nhánh (multi-branch CNN) để nhận diện từng chữ số riêng lẻ trong dãy số đọc của công tơ điện với độ chính xác 96,3 %. Nghiên cứu [5] đề xuất một cách tiếp cận khác trong đó sử dụng Mask-RCNN để tự động trích xuất vùng ảnh chứa chữ số, sau đó sử dụng mô hình GoogleLeNet đã tinh chỉnh để trích xuất đặc trưng. Kết quả cho thấy độ chính xác đạt xấp xỉ 100 %.

Thiết bị và phần mềm RadOCR là một giải pháp giúp “internet hóa” máy đo phóng xạ mà không cần phải can thiệp vào phần cứng hoặc phần mềm của máy. Bằng việc tận dụng mạng wifi nội bộ để kết nối đến máy đo phóng xạ, giải pháp này không cần dùng hệ thống dây cáp truyền tín hiệu công kênh nhưng vẫn cho phép ghi lại số đọc của máy từ xa. Qua đó, nó giúp nhân viên đo đạc không bị phơi nhiễm bởi nguồn phóng xạ nguy hiểm. Giải pháp này có thể tiếp tục được cải tiến nhằm tích hợp cùng lúc nhiều máy đo phóng xạ thành một mạng lưới quan trắc đa điểm theo thời gian thực, có thể kết nối để hiển thị và lưu trữ trên một máy tính trung tâm. Ngoài ra, bộ giá đỡ của thiết bị RadOCR có thể được điều chỉnh để phù hợp với các loại máy đo phóng xạ khác.

Một số hạn chế của nghiên cứu có thể kể đến là: phần mềm RadOCR chỉ mới trích xuất được các số đọc dạng số nguyên không chứa dấu chấm thập phân; kích thước bộ giá đỡ còn to và cồng kềnh; một số trường hợp kết quả trích xuất bị sai so với số đọc. Các tác giả đã thử nghiệm trích xuất đối với các thang đo mà số đọc có chứa dấu chấm thập phân. Kết quả cho thấy độ chính xác thấp, chỉ đạt khoảng 60 %. Nguyên nhân chủ yếu

là do dấu chấm thập phân bị dính vào chữ số lân cận như minh họa trong Hình 9. Vấn đề này có thể được giải quyết bằng cách tăng số lần lặp khi thực hiện phép co (erosion) trong bước tiền xử lý ảnh. Tuy nhiên, việc co quá mức có thể làm biến mất các chữ số. Một cách giải quyết khác là sử dụng Mask-CNN để xác định vùng ảnh chứa số đọc [5]. Do camera OV2640 có tiêu cự dài nên cần khoảng cách lớn từ camera đến LCD của máy đo, làm cho kích thước của bộ giá đỡ trở nên cồng kềnh. Để giảm kích thước, cần sử dụng camera với thấu kính có khả năng ghi hình gần (macro lens).



Hình 9 Một ví dụ có kết quả OCR sai khi số đọc có chứa dấu chấm thập phân

Các kết quả OCR sai được đã được phân tích, qua đó thấy được có hai nguyên nhân chính: số đọc đang trong quá trình cập nhật hoặc lỗi do thuật toán xử lý ảnh. Do LCD của máy đo Radiation Alert Ranger® có thời gian chuyển đổi tương đối dài, nên những ảnh được ghi nhận trong lúc số đọc đang cập nhật từ giá trị này sang giá trị khác thì sẽ chứa số đọc không rõ ràng như minh họa trong Hình 10a. Điều này làm cho kết quả trích xuất bị sai. Vấn đề này có thể được giải quyết bằng cách tận dụng các ảnh khác trong luồng mà có thời điểm ghi nhận xung quanh thời điểm cập nhật số đọc để hiệu chỉnh lẫn nhau. Nguyên nhân thứ hai gây ra kết quả sai là do lỗi của thuật toán. Một số hình ảnh hiển thị rõ nét nhưng kết quả trích xuất vẫn sai như minh họa trong Hình 10b. Nguyên nhân là do mô hình Tesseract OCR chưa được huấn luyện một cách tối ưu, nhất là đối với kiểu chữ số hiển thị trên màn hình máy đo Radiation Alert Ranger®. Hạn chế này có thể được khắc phục bằng cách huấn luyện lại hoặc huấn luyện chuyển tiếp (transfer learning) mô hình Tesseract OCR bằng một bộ ảnh phù hợp hơn.



Hình 10 (a) Kết quả OCR sai do LCD đang chuyển từ hiển thị số đọc này sang số đọc khác. (b) Kết quả OCR sai ngay cả khi số đọc hiển thị rõ ràng

4 Kết luận

Nghiên cứu này đã thiết kế và gia công thiết bị và phần mềm RadOCR để trích xuất từ xa, tự động số đọc của máy đo phóng xạ Radiation Alert Ranger®. Module IoT ESP32-CAM được sử dụng như một camera để ghi nhận số đọc hiển thị trên LCD của máy đo và truyền

luồng hình ảnh theo thời gian thực về phần mềm xử lý thông qua kết nối mạng wifi. Các công cụ trong thư viện OpenCV được sử dụng để tiền xử lý ảnh trước khi đưa vào trích xuất số đọc bằng Tesseract OCR. Kết quả đánh giá độ chính xác cho thấy thiết bị và phần mềm hoạt động ổn định với độ chính xác cao. Tuy nhiên cứu được thực hiện trên máy đo Radiation Alert Ranger® như một trường hợp điển hình, phương pháp đề ra có thể được hiệu chỉnh để phù hợp với nhiều loại máy đo khác nhau. Đây là giải pháp hữu ích cho phép đo đạc từ xa, giúp hạn chế phơi nhiễm do tiếp xúc gần nguồn phóng xạ. Ngoài ra, giải pháp này cũng có thể được phát triển để tích hợp nhiều máy đo khác nhau vào một hệ thống quan trắc phóng xạ đa điểm.

Lời cảm ơn

Chúng tôi xin cảm ơn Trường Đại học Nguyễn Tất Thành, Thành phố Hồ Chí Minh đã hỗ trợ cho nghiên cứu này.

Tài liệu tham khảo

1. Chính phủ Nước CHXHCN Việt Nam. (2025). *Quyết định số 245/QĐ-TTg của Thủ tướng Chính phủ: Về việc phê duyệt Quy hoạch phát triển, ứng dụng năng lượng nguyên tử thời kỳ đến năm 2030, tầm nhìn đến năm 2050*.
2. H. Wang, C. Pan, X. Guo, C. Ji, & K. Deng. (2021). From object detection to text detection and recognition: A brief evolution history of optical character recognition. *Wiley Interdisciplinary Reviews: Computational Statistics*, 13(5), 1547. <https://doi.org/10.1002/wics.1547>
3. N. Sarika, N. Sirisala, & M.S. Velpuru. (2021). *CNN based optical character recognition and applications*. Paper presented at the 2021 6th International conference on inventive computation technologies (ICICT).
4. A. Azeem, W. Riaz, A. Siddique, & U.A.K. Saifullah. (2020). *A robust automatic meter reading system based on mask-rcnn*. Paper presented at the 2020 IEEE International Conference on Advances in Electrical Engineering and Computer Applications (AEECA).
5. A. Shanthakumari, R. Kalpana, J. Jayashankari, B. Umamaheswari, & M. Siriya. (2022). *Mask RCNN and Tesseract OCR for vehicle plate character recognition*. Paper presented at the AIP Conference Proceeding, Krishnagiri, India.
6. A.K. Sharma, & K.K. Kim. (2021). Lightweight CNN based meter digit recognition. *Journal of Sensor Science and Technology*, 30(1), 15-19. <https://doi.org/10.46670/JSST.2021.30.1.15>
7. L. Cao, M. Liu, Z. Dong, & H. Yang. (2019). *Book spine recognition based on OpenCV and tesseract*. Paper presented at the 2019 11th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC).
8. S. Čakić, T. Popović, S. Šandi, S. Krčo, & A. Gazivoda. (2020). *The use of tesseract ocr number recognition for food tracking and tracing*. Paper presented at the 2020 24th International Conference on Information Technology (IT).
9. S. Ray, & Hewlett-Packard. (2021). Tesseract user manual. Retrieved from <https://tesseract-ocr.github.io/tessdoc/#introduction>

10. Madmaze. (2024). PyTesseract 0.3.13. Retrieved from <http://pypi.org/project/pytesseract/>
11. Tesseract OCR. (2021). Improving the quality of the output. Retrieved from <https://tesseract-ocr.github.io/tessdoc/ImproveQuality.html>
12. Shreeshrii. (2019). Tessdata Shreetest. Retrieved from https://github.com/Shreeshrii/tessdata_shreetest
13. M. Noordzij, F.W. Dekker, C. Zoccali, & K.J. Jager. (2011). Sample size calculations. *Nephron Clinical Practice*, 118(4), 319-323.
14. C. Son, S. Park, J. Lee, & J. Paik. (2019). Deep learning–based number detection and recognition for gas meter reading. *IEIE Trans. Smart Process. Comput.*, 8(5), 367-372.
15. H. Xiu, J. He, Y. Qi, X. Zhang, & L. Wang. (2020). *Numeric CNNs: CNNs for Reading Numeric Characters on Meters*. Paper presented at the International Conference on Cyberspace Data and Intelligence.

Application of Internet of Things module and optical character recognition algorithm to automatically extracting readouts of radiation survey meters

Pham Nguyen Phat, Nguyen Tan Duoc*

Biomedical Physics Department – Faculty of Medicine, Nguyen Tat Thanh University, Ho Chi Minh City

*ntduoc@ntt.edu.vn

Abstract With the development of Internet of Things, radiation measuring devices today are embedded with wireless network connections. However, many devices being used in facilities lack this feature, making it difficult to remotely access data as well as to integrate into a multi-point radioactive monitoring system. In this study, we designed a system called RadOCR that extracts the readouts displayed on the LCD screen of Radiation Alert Ranger® survey meter in real time. The system consisted of a device attached to the meter and a software deployed on a computer. The RadOCR device consisted of a ESP32-CAM module which took photo of the meter's LCD and sent to the software via a WiFi connection on a local network. The software utilized image processing algorithms from OpenCV and the optical character recognition Tesseract OCR model to extract the readouts, then showed them on a graphical interface, and saved them with timestamps to a csv file. The test results show that the system operates stably with an accuracy of $(97.365 \pm 0.082) \%$.

Keywords Internet of Things, optical character recognition, radiation measurement