

# Kiểm tra khuyết tật dựa trên U-net trong hình ảnh phát quang điện quang điện

Trưởng Muhammad Rameez Ur Rahman về  
trí tuệ nhân tạo  
Đại học công nghệ Hà Bắc Thiên Tân,  
Trung Quốc  
rameezrehman83@gmail.com

Trưởng Trí tuệ  
nhân tạo Haiyong Chen  
Đại học công nghệ Hà Bắc Thiên Tân,  
Trung Quốc  
haiyong.chen@hebut.edu.cn

Trưởng  
trí tuệ nhân tạo Wenxi  
Đại học công nghệ Hà Bắc Thiên Tân,  
Trung Quốc  
15022327520@163.com

Tóm tắt-Phân đoạn khuyết tật hiệu quả từ hình ảnh phát quang điện (EL) quang điện (PV) là một quá trình quan trọng do nền không đồng nhất ngẫu nhiên và phân bố điểm ảnh nứt không nứt không cân bằng. Việc kiểm tra khuyết tật tự động của các tế bào năng lượng mặt trời ảnh hưởng rất lớn đến chất lượng của các tế bào quang điện, do đó cần phải kiểm tra các khuyết tật một cách hiệu quả và chính xác. Trong bài báo này, chúng tôi đề xuất một kiến trúc dựa trên học sâu đầu cuối mới để phân đoạn khuyết tật. Trong kiến trúc được đề xuất, chúng tôi giới thiệu một sự chú ý toàn cầu mới để trích xuất thông tin ngữ cảnh phong phú. Hơn nữa, chúng tôi đã sửa đổi U-net bằng cách thêm tích chập giãn nở ở cả phía bộ mã hóa và giải mã với các kết nối bỏ qua từ các lớp đầu đến các lớp sau ở phía bộ mã hóa.

Sau đó, sự chú ý toàn cầu được đề xuất được đưa vào U-net đã sửa đổi. Mô hình được đào tạo và thử nghiệm trên tập dữ liệu hình ảnh phát quang điện từ quang điện 512x512 và kết quả được ghi lại bằng cách sử dụng Giao điểm trung bình trên hợp nhất (IOU). Trong các thí nghiệm, chúng tôi đã báo cáo kết quả và so sánh giữa mô hình được đề xuất và các phương pháp tiên tiến khác.

IOU trung bình của phương pháp đề xuất là 0,6477 với độ chính xác pixel là 0,9738, tốt hơn các phương pháp hiện đại. Chúng tôi chứng minh rằng phương pháp đề xuất có thể mang lại kết quả hiệu quả với tập dữ liệu nhỏ hơn và hiệu quả về mặt tính toán.

Từ khóa-U-net, Phát hiện khuyết tật của pin mặt trời, phát hiện vết nứt, hình ảnh phát quang điện

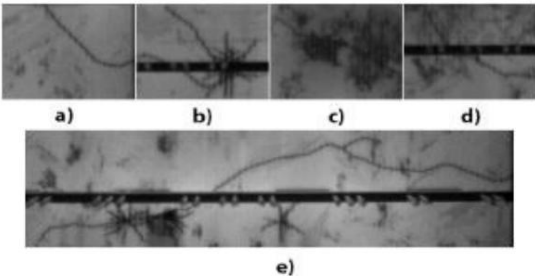
## I. GIỚI THIỆU

Trong quá trình sản xuất Pin mặt trời, kiểm tra khuyết tật trực quan được thực hiện để đảm bảo chất lượng sản phẩm. Pin mặt trời có thể được chia thành đơn tinh thể và đa tinh thể. Pin mặt trời đơn tinh thể có nền đồng nhất trong khi pin mặt trời đa tinh thể có nền kết cấu phức tạp phân bố ngẫu nhiên. Các khuyết tật bề mặt pin mặt trời có thể được phân loại thành các vết nứt, gián đoạn ngón tay và vỡ pin, v.v. Các khuyết tật này có nhiều kích thước và hình dạng khác nhau. Sự xuất hiện của các vết nứt nhỏ trong hình ảnh có độ phân giải cao đòi hỏi một phương pháp hiệu quả cao để phân đoạn các vết nứt chính xác. Các ngành công nghiệp sử dụng hệ thống kiểm tra khuyết tật thủ công. Các phương pháp kiểm tra khuyết tật thủ công không có khả năng phát hiện khuyết tật hiệu quả và tốn kém. Do những phát triển gần đây trong thị giác máy tính, kiểm tra khuyết tật thủ công được thay thế bằng các phương pháp thị giác máy tự động. Các phương pháp kiểm tra khuyết tật tự động sử dụng các thuật toán xử lý hình ảnh và trí tuệ nhân tạo có tốc độ cao, độ chính xác tốt hơn và chi phí thấp. Bài báo này đề xuất một phương pháp phân đoạn khuyết tật mạnh mẽ và hiệu quả dựa trên học sâu có thể phân đoạn các khuyết tật tinh vi trong hình ảnh EL.

Trong những năm gần đây, các thuật toán xử lý hình ảnh đã được áp dụng thành công trên hình ảnh EL quang điện để phát hiện các khuyết tật [1].

Phương pháp chụp ảnh phát quang điện phát hiện thành công các khuyết tật nội tại (ví dụ như trật khớp, shunt) và các khuyết tật bên ngoài (ví dụ như vết nứt, tiếp xúc bị gián đoạn) trong hình ảnh EL có độ phân giải cao [2]. Trong [3] các vết nứt nhỏ và khuyết tật gián đoạn ngón tay được phát hiện và phân đoạn bằng cách sử dụng Fourier

biến đổi. Áp dụng tái tạo phân tích thành phần độc lập để phát hiện các khiếm khuyết trong các mô đun năng lượng mặt trời [4] [5]. Các lỗi ngón tay tự động được phát hiện bằng cách đầu tiên chọn vùng quan tâm và sau đó áp dụng cụm phổ [6]. Các vết nứt nhỏ của tế bào năng lượng mặt trời được phát hiện bằng cách đầu tiên trích xuất các đặc điểm thông qua ma trận đồng hiện mức xám, bộ phân loại máy vectơ hỗ trợ (SVM) được sử dụng để phân loại các đặc điểm này [7].



Hình 1. Các loại khuyết tật: a) khuyết tật đường, b) khuyết tật hình sao, c) Khuyết tật ben, trong cùng một mẫu cơ sở độ ánh sáng d) khuyết tật chia chấu đầu dò tiếp xúc, e) khuyết tật có tất cả các đặc tính (a) (b) (c) (d)

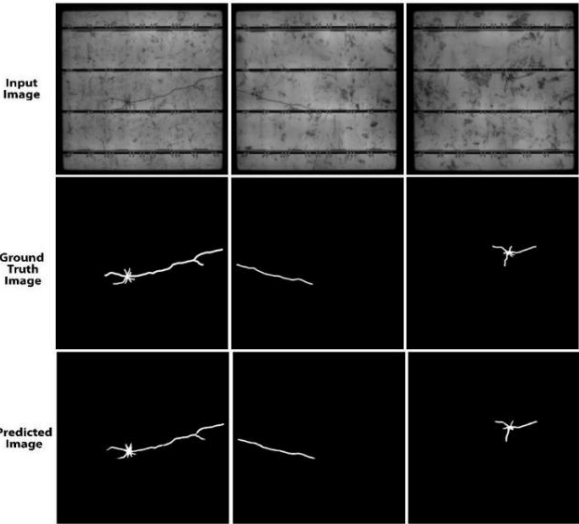
Các thuật toán xử lý hình ảnh truyền thống yêu cầu các đặc điểm thủ công được trích xuất bởi các chuyên gia, ảnh hưởng trực tiếp đến tỷ lệ phát hiện, do đó chúng không có khả năng ứng dụng tốt. Những tiến bộ gần đây trong mạng nơ-ron tích chập sâu (DCNN) đã được chứng minh là hiệu quả và nhanh hơn đối với nhiệm vụ phát hiện và phân đoạn đối tượng và chúng không phụ thuộc vào các đặc điểm thủ công phụ thuộc của chuyên gia. Do đó, việc áp dụng DCNN để phân đoạn các khuyết tật quang điện mang lại lợi ích đáng kể.

Tuy nhiên, nhiều phương pháp tiên tiến được đánh giá trên phân đoạn cảnh trong hình ảnh tự nhiên chứa thông tin ngữ nghĩa cấp cao và không có nhiều sự khác biệt giữa tỷ lệ nền và đối tượng. Một số lượng lớn các tập dữ liệu công khai có sẵn để các nhà nghiên cứu đánh giá các mô hình DCNN lớn. Đối với nhiệm vụ môi, các nhà nghiên cứu sử dụng các mô hình được đào tạo trước và tình hình chúng như mạng chiến lược này không hữu ích lắm đối với việc kiểm tra khuyết tật của hình ảnh phát quang điện quang điện. Bên cạnh đó, việc thu thập các mẫu hình ảnh EL quang điện bị lỗi từ môi trường công nghiệp là một nhiệm vụ tốn kém. Hình ảnh EL quang điện là hình ảnh thang độ xám với sự phân bố điểm ảnh không cân bằng của các khuyết tật và nền.

Các khuyết tật mặt trời se rất quan trọng do sự hiện diện của các hạt tinh thể phân bố ngẫu nhiên trên bề mặt tế bào quang điện. Để phân tích các khuyết tật mặt trời, chúng tôi đã phân loại các khuyết tật thành ba loại theo hình dạng vết nứt và nền kết cấu. Tất cả các vết nứt đều có kích thước khác nhau. Các ví dụ về các khuyết tật được phóng to được hiển thị trong hình 1. Khu vực không bị khuyết tật trong hình ảnh chứa các hạt tinh thể tạo nên các mẫu ngẫu nhiên. Các mẫu này là duy nhất cho mỗi hình ảnh và có thể có cùng giá trị cường độ điểm ảnh như khuyết tật khiến chúng trở thành lý do chính gây ra phân đoạn sai.

Các vết nứt đường thẳng có phân bố điểm ảnh đồng đều dễ dàng

phân đoạn hơn những phân đoạn khác. Một số vết nứt đường nằm bên trong mẫu hạt tinh thể ngẫu nhiên tối màu thể hiện cùng giá trị cường độ điểm ảnh như khuyết tật khiến khuyết tật khó phân đoạn. Các vết nứt giống hình sao thể hiện một cấu trúc phức tạp. Các vết nứt cũng được chia ra bởi đầu dò tiếp xúc tối có cùng giá trị cường độ điểm ảnh như vết nứt và có thể trở thành dự đoán tính giả mạo. Một số vết nứt thể hiện tất cả các đặc điểm đã đề cập ở trên. Dựa trên thảo luận ở trên, cần có một kiến trúc mới để khắc phục những thiếu sót của các phương pháp trước đây để phân đoạn khuyết tật trong hình ảnh phát quang điện quang điện.



Hình 2. Phát hiện lỗi dựa trên phương pháp đề xuất

II. CÔNG TRÌNH LIÊN QUAN

Những tiến bộ gần đây trong mạng nơ-ron tích chập sâu (DCNN) đã được chứng minh là hiệu quả và nhanh hơn cho các nhiệm vụ phát hiện và phân đoạn đối tượng và chúng không dựa vào các tính năng thủ công phụ thuộc vào chuyên gia. Mạng niềm tin sâu (DBN) với học chuyển giao được sử dụng để phát hiện các khuyết tật mặt trời và năng lượng nó chỉ có thể phát hiện các khuyết tật ở cùng một vị trí [8]. Trong [9][10] Mạng nơ-ron tích chập (CNN) kết hợp với cửa sổ trượt được áp dụng để phát hiện các vết nứt và sáu loại khuyết tật khác với độ chính xác 99,8% nhưng kỹ thuật của sổ trượt làm cho quá trình này chậm lại. Một phương pháp ba giai đoạn dựa trên thị giác máy tính và máy học được áp dụng để phát hiện hình ảnh kính hiển vi điện tử [11]. Trước tiên, các khuyết tật được phát hiện bằng cách sử dụng máy dò đối tượng nổi tiếp và phương pháp sàng lọc dựa trên CNN được áp dụng để loại bỏ các kết quả dự đoán tính giả, sau đó thuật toán lưu trữ vệt [12] được áp dụng để trích xuất hình dạng và kích thước của các khuyết tật. Nhiều khuyết tật bề mặt thép được phân loại [13] bằng GooLeNet [14] với ảnh xạ danh tính. Trong [15], mạng hoàn toàn tích chập [16] và các mạng được giám sát sâu [17] được sử dụng để phân đoạn các vết nứt và sau đó áp dụng lọc có hướng dẫn để tinh chỉnh kết quả cuối cùng. Kiến trúc hình chữ U [18] bao gồm đường dẫn co lại và mở rộng, phụ thuộc nhiều vào việc tăng cường dữ liệu được áp dụng thành công trên nhiều tác vụ phân đoạn hình ảnh y tế khác nhau [19] [20] [26].

Kiến trúc này có các kết nối bỏ qua đối xứng kết hợp các tính năng cấp thấp với các tính năng cấp cao để cải thiện hiệu suất và định vị đối tượng. Kiến trúc hình chữ U có thể gặp phải một vấn đề là nó hợp nhất các bản đồ tính năng khác nhau dựa trên trọng số cố định. Một mạng đề xuất vùng kết hợp với U-net được sử dụng để phát hiện các khiếm khuyết của tế bào năng lượng mặt trời đa tinh thể nhưng vẫn đề với điều này

phương pháp là các hộp giới hạn có cùng kích thước và một lỗi có thể chứa nhiều hộp giới hạn [28].

Bài báo này đề xuất một kiến trúc dựa trên mạng nơ-ron tích chập hoàn toàn sâu để phát hiện khuyết tật dựa trên phân đoạn trong hình ảnh EL quang điện. Chúng tôi chọn mạng tích chập hoàn toàn hình chữ U vì nó hiệu quả, có thể đưa ra dự đoán dày đặc và quan trọng nhất là chấp nhận các đầu vào có bất kỳ kích thước nào. U-net kết hợp các đặc điểm cấp thấp với các đặc điểm cấp cao dựa trên trọng số cố định, điều này có thể làm giảm sức mạnh biểu diễn của nó. Do đó, chúng tôi giới thiệu mạng chú ý toàn cầu để cân nhắc các bản đồ đặc điểm theo mức độ liên quan của chúng. Sau đó, mạng chú ý toàn cầu được kết hợp vào U-net để tăng cường các đặc điểm khuyết tật cho phép U-net học các đặc điểm khuyết tật tốt hơn. Các ví dụ về phát hiện khuyết tật dựa trên phương pháp đề xuất được thể hiện trong Hình 2. Những đóng góp chính của bài báo này là

Chúng tôi giới thiệu mạng chú ý mới và kết hợp nó vào kiến trúc U-net để tăng sức mạnh biểu diễn tính năng của nó. Mô hình đề xuất được đánh giá trên tập dữ liệu hình ảnh EL quang điện thực để tìm lỗi.

Trong kiến trúc được đề xuất, các phép toán tích chập giãn nở được thêm vào U-net để mở rộng trường tiếp nhận mà không làm tăng số lượng tham số. Các kết nối bỏ qua còn lại với nhóm trung bình được giới thiệu ở phía bộ mã hóa giúp quá trình đào tạo dễ dàng hơn và học các tính năng tốt hơn.

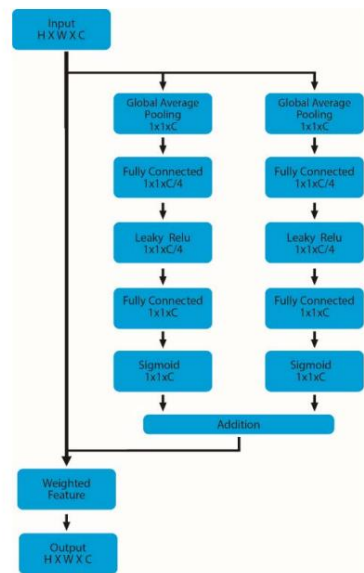
III. KIẾN TRÚC MẠNG

A. Khối chú ý toàn cầu

Trong quá trình trích xuất các tính năng, một số tính năng có thể hữu ích trong khi những tính năng khác có thể là nhiễu hoặc đối tượng nền. Để làm nổi bật các tính năng liên quan trọng nhất, mạng lưu ý chú ý toàn cầu được đề xuất. Khối Squeeze và excitation trong [25] được thiết kế nhằm mục đích phân loại và nó đóng vai trò quan trọng trong việc cải thiện hiệu suất của mạng bằng cách đặt nó vào từng khối của mạng. Để sử dụng nó trong phân đoạn các lỗi, chúng tôi đã sửa đổi khối SE ban đầu bằng cách nén các đầu vào bằng cách sử dụng nhóm trung bình toàn cầu và nhóm tối đa toàn cầu để có được các tính năng chất lượng cao có trọng số. Trong [29] và [30], tầm quan trọng của nhóm tối đa toàn cầu và nhóm trung bình toàn cầu được thể hiện đối với nhận dạng các đối tượng riêng biệt.

Như thể hiện trong hình 3, khối chú ý toàn cầu bao gồm hai nhánh. Sự khác biệt duy nhất giữa hai nhánh là, trong một nhánh, phép gộp trung bình toàn cục đầu tiên được áp dụng trong khi ở nhánh thứ hai, phép gộp tối đa toàn cục được áp dụng. Tư duy tự như khối SE, mỗi kênh đầu vào  $\{x_1, x_2, \dots, x_n\}$  được nén thành một giá trị số duy nhất bằng cách sử dụng phép gộp trung bình toàn cục. Điều này cung cấp đầu ra có chiều  $1 \times 1 \times \frac{C}{r}$ ,  $\{y_1, y_2, \dots, y_n\}$  trong đó  $n$  là số kênh.  $\{y_1, y_2, \dots, y_n\}$  = Sau đó, được đưa vào lớp được  $= \{z_1, z_2, \dots, z_n\}$  kết nối đầy đủ theo sau là hàm kích hoạt LeakyRelu tạo ra đầu ra

$\{z_1, z_2, \dots, z_n\}$  của chiều  $1 \times 1 \times \frac{C}{r}$ . Lớp kết nối đầy đủ cuối cùng theo sau là lớp sigmoid cho kết quả đầu ra  $\{s_1, s_2, \dots, s_n\}$ . Quá trình tư duy tự được áp dụng trong nhánh gộp toàn cục tối đa cho kết quả đầu ra  $\{t_1, t_2, \dots, t_n\}$ . Cuối cùng, đầu ra của cả hai nhánh được cộng lại với nhau  $\{s_1, s_2, \dots, s_n\} + \{t_1, t_2, \dots, t_n\}$ . Cân nhắc các kênh dựa trên việc đưa ra  $=$  dựa trên tầm quan trọng của chúng và phân bổ trọng số cao cho các tính năng quan trọng nhất.



Hình 3. Khối chú ý toàn cầu (GAB)

B. Những giác mơ

Kiến trúc hình chữ U ban đầu bao gồm đường dẫn lấy mẫu xuống để nắm bắt thông tin theo ngữ cảnh và đường dẫn lấy mẫu lên để định vị đối tượng chính xác. Đường dẫn lấy mẫu xuống bao gồm hai hình ảnh 3x3 liên tiếp được lặp lại các lớp tích chập không đệm để trích xuất các đặc điểm, tiếp theo là các lớp gộp tối đa 2x2 để lấy mẫu xuống. Số lượng kênh đặc trưng tăng gấp đôi sau mỗi bước giúp trích xuất các đặc trưng nâng cao. Kích thước của hình ảnh giảm khi độ sâu của mạng tăng. Đường dẫn lấy mẫu lên bao gồm lấy mẫu lên của các bản đồ đặc trưng theo sau là tích chập lên 2x2 để khôi phục kích thước của bản đồ phân đoạn. Các bản đồ đặc trưng lấy mẫu lên được nối với các bản đồ đặc trưng của đường dẫn lấy mẫu xuống ở cùng một cấp độ. Cuối cùng có 1x1

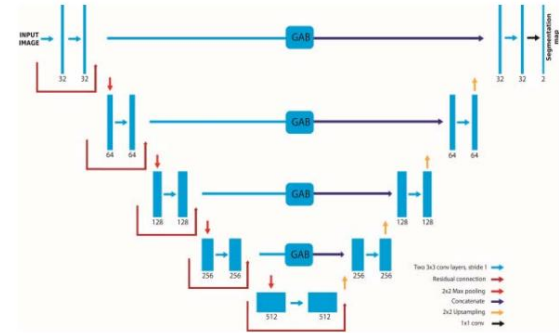
tích chập với 2 bản đồ đặc điểm vì chỉ có hai lớp.

C. Kiến trúc U-net được đề

xuất Trong phần này, một kiến trúc mới được đề xuất dựa trên khối chú ý toàn cầu. Hình 4 cho thấy kiến trúc của mạng được đề xuất. Lưu ý rằng khuôn khổ kiến trúc được đề xuất không giống với U-net ban đầu được đề xuất trong (Ronneberger, Fischer và Brox 2015). Chúng tôi đã sửa đổi U-net ban đầu và thêm các tính năng mới được chứng minh là hữu ích trong nhiệm vụ của chúng tôi. Kiến trúc được đề xuất bao gồm U-net 9 cấp kết hợp khối chú ý toàn cục, kết nối bỏ qua với nhóm trung bình và tích chập giãn nở. Khối chú ý toàn cục được đặt ở các kết nối giữa bộ mã hóa và bộ giải mã để cải thiện sức mạnh biểu diễn của kiến trúc U-net. Khối chú ý toàn cục học cách cân nhắc các kênh bản đồ tính năng đến từ đường dẫn bộ mã hóa đến hoàn thành cơ chế chú ý. Đầu ra của khối chú ý được nối với các tính năng được lấy mẫu lại của đường dẫn bộ giải mã ở cùng một cấp độ. Ở phía bộ mã hóa và giải mã, chúng tôi sử dụng tỷ lệ giãn nở là 1, 1, 2, 2, 3 từ cấp trên xuống cấp dưới. Sự giãn nở [23] tùy ý mở rộng trường nhìn của bộ lọc để không bỏ sót bất kỳ khuyết tật nhỏ nào. Ở phía Bộ mã hóa, lớp Maxpooling 2x2 được sử dụng để giảm mẫu bản đồ đặc điểm xuống

một nửa. Chúng tôi đã giới thiệu các kết nối bỏ qua với nhóm trung bình ở phía bộ mã hóa để lấy mẫu xuống các bản đồ đặc điểm. Việc sử dụng tối đa các hoạt động gộp nhóm và gộp nhóm trung bình cùng nhau dẫn đến việc học các đặc điểm phân biệt dẫn đến phát hiện tốt hơn các khuyết tật và cạnh. Các kết nối bỏ qua giúp dễ dàng đào tạo mạng bằng cách cung cấp một giải pháp thay thế cho gradient để truyền ngược và hỗ trợ duyệt thông tin trong mạng. Khác với U-net, chúng tôi đã thêm lớp kích hoạt Chuẩn hóa theo lô (BN) [21] và LeakyRelu [22] sau mỗi lớp tích chập.

Tích chập 1x1 với kích hoạt sigmoid được sử dụng ở cuối để xuất ra hình ảnh phân đoạn. Không giống như U-net gốc, hình ảnh không bị cắt xén, kích thước đầu vào và đầu ra được giữ nguyên để tránh mất cạnh của hình ảnh.



Hình 4. Kiến trúc U-net được đề xuất

D. Hàm mất mát Trong

quá trình đào tạo mạng, cần phải ước tính trọng số để tăng độ mạnh mẽ và độ chính xác của mạng. Điều này đạt được bằng cách sử dụng hàm mất mát thích hợp sẽ được giảm thiểu trong quá trình đào tạo. Vì tập dữ liệu của chúng tôi có phân phối pixel không cân bằng, khoảng 90% nền và 10% lỗi. Vì vậy, trong công việc này, chúng tôi sử dụng Dice loss [24] được thiết kế đặc biệt cho phân phối dữ liệu không cân bằng và được định nghĩa là

$$\dots \frac{\cap}{\dots}$$

Trong đó  $|X_g|$  là hình ảnh thực tế trong khi  $|X_p|$  là hình ảnh dự đoán. Đường cong mất mát và độ chính xác khi đào tạo được hiển thị trong hình 5.

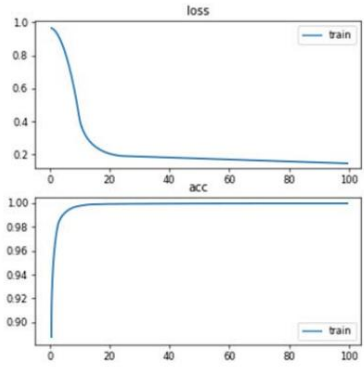
E. Bộ dữ liệu

Tất cả các thí nghiệm được thực hiện trên hình ảnh EL 512x512. Các hình ảnh được lấy từ môi trường công nghiệp thực tế. Bộ dữ liệu chứa 400 hình ảnh lỗi, trong đó 350 hình ảnh được sử dụng để đào tạo và 50 hình ảnh để thử nghiệm mạng. Tất cả các hình ảnh đều là thang độ xám 512x512x1. Tất cả các hình ảnh đào tạo và dữ liệu thực tế đều được dán nhãn cẩn thận bằng công cụ chú thích LabelME. Mỗi pixel trong hình ảnh được dán nhãn là lỗi hoặc nền. Để khắc phục vấn đề về tập dữ liệu nhỏ và có lợi cho việc đào tạo mô hình học sâu; hoặc là tăng tốc độ hội tụ hoặc hoạt động như bộ điều chỉnh, một phép tăng cường dữ liệu mở rộng được sử dụng. Một số tham số được sử dụng để tăng cường hình ảnh bao gồm, xoay = 40, dịch chuyển chiều rộng = 0,05, dịch chuyển chiều cao = 0,05, phạm vi cắt = 0,05, lật ngang, lật dọc và cân bằng histogram thích ứng (AHE). Trong cân bằng histogram thích ứng, một số histogram khác nhau được tính toán cho

mỗi phần khác nhau của hình ảnh và được sử dụng để điều chỉnh giá trị độ sáng của từng phần hình ảnh. AHE cải thiện độ tương phản cục bộ và tăng cường định nghĩa các cạnh trong mỗi vùng hình ảnh. Tăng cường dữ liệu làm tăng khả năng khái quát hóa và độ chính xác của mô hình.

IV. KẾT QUẢ THÍ NGHIỆM VÀ PHÂN TÍCH

Chúng tôi đánh giá hiệu suất của mạng được đề xuất trên bộ dữ liệu hình ảnh phát quang điện được thu thập từ môi trường công nghiệp thực tế nhằm mục đích phân đoạn các khuyết trong hình ảnh. Cụ thể, chúng tôi so sánh kết quả của bộ thử nghiệm với hai phương pháp tiên tiến nhất: U-net [18] và R2-Unet [26]. Chúng tôi sử dụng tất cả các thông số giống nhau cho các thí nghiệm so sánh.



Hình 5. Mất mát và độ chính xác khi đào tạo

A. Chi tiết thực hiện

Mạng được đề xuất và hai mạng cơ sở được triển khai trong khuôn khổ Keras 2.2.4. Chúng tôi sử dụng trình tối ưu hóa RmsProp [27] với tốc độ học là 0,0001 và rho = 0,9 để tối ưu hóa mô hình. Các tham số được khởi tạo bằng trình khởi tạo thống nhất Xavier. Chúng tôi đã đào tạo mô hình trên GPU GTX 1080. Kích thước lô được đặt thành 1 và mô hình được đào tạo trong 100 kỷ nguyên. Các tham số CPU chính được hiển thị trong Bảng I.

BẢNG I. THÔNG SỐ CPU	
Tên tham số	Giá trị tham số
Bộ vi xử lý	Intel® core™ i7-CPU 6700K @ 4.00 GHZ (8CPU)
Đáp	32GB
Đĩa cứng	2TB

B. Số liệu

Nhiều số liệu đánh giá được đề xuất và thường được sử dụng cho nhiệm vụ phân đoạn ngữ nghĩa. Để đánh giá hiệu suất của mô hình, kết quả dự đoán được so sánh với sự thật cơ bản bằng ba số liệu: Giao điểm trên hợp nhất (IOU), độ nhạy và độ đặc hiệu. Các số liệu này được thiết kế riêng cho nhiệm vụ phân đoạn ngữ nghĩa.

IOU tính toán sự chồng lấn không gian giữa các dự đoán và thực tế, được xác định giữa 0 và 1 trong khi 0 biểu thị

không chồng chéo và 1 biểu thị sự chồng chéo hoàn toàn. IOU được đưa ra bằng

= (2)

TP, FP, FN lần lượt là các mẫu dương tính thật, dương tính giả và âm tính giả.

Độ nhạy được sử dụng để tìm Tỷ lệ dương tính thực (TPR) của mạng được đề xuất. Độ nhạy là thước đo các điểm ảnh dương trong sự thật cơ bản cũng được xác định là dương theo phân đoạn đang được đánh giá. Nó được đưa ra theo công thức

= (3)

Độ đặc hiệu được sử dụng để tìm Tỷ lệ âm tính thực (TNR) của mạng được đề xuất. Độ đặc hiệu là thước đo các điểm ảnh âm trong sự thật cơ bản cũng được xác định là âm theo phân đoạn đang được đánh giá. Nó được đưa ra theo công thức

= (4)

Trong đó TP, FP, TN, FN lần lượt là dương tính thật, dương tính giả, âm tính thật và âm tính giả. Cả độ đặc hiệu và độ nhạy đều được báo cáo trong khoảng từ 0 đến 1.

C. Kết quả và thảo luận

Trong nghiên cứu này, chúng tôi trình bày một phương pháp tiếp cận dựa trên học sâu có khả năng phân đoạn hiệu quả các khuyết từ hình ảnh phát quang điện quang điện. Phương pháp được đề xuất có lợi thế của phép tích chập giãn nở, kết nối bỏ qua còn lại và cơ chế chú ý toàn cục. Sau khi đào tạo và đánh giá thành công trên hình ảnh EL, chúng tôi có thể liên tục đạt được kết quả phát hiện tốt. Do đó, chúng tôi cung cấp một khuôn khổ phân đoạn khuyết tật mạnh mẽ cho hình ảnh EL quang điện. Mặc dù các tham số của mạng lưới của chúng tôi chỉ là 8,3M, bằng ¼ U-net ban đầu, nhưng kết quả đầy hứa hẹn đã đạt được trên tập dữ liệu EL. Bảng II thể hiện sự so sánh về số lượng tham số.

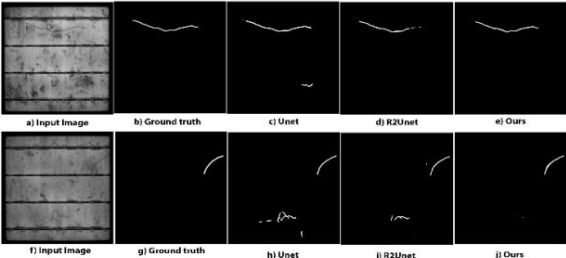
BẢNG II. SO SÁNH UNET ĐỀ XUẤT VỚI U-NET VÀ R2-UNET VỀ SỐ LƯỢNG THAM SỐ			
Phương pháp	Mạng U	R2-Giác mơ	Của chúng tôi
Tham số của số (triệu)	30,8	48,9	8.3

BẢNG III. SO SÁNH CÁC SỐ LIỆU CỦA UNET ĐỀ XUẤT VỚI U-NET VÀ R2-DREAMS			
Phương pháp	Mạng U	R2-Giác mơ	Của chúng tôi
Nghĩa là IOU	0,6296	0,6384	0,6477
Độ nhạy trung bình	0,8304	0,8313	0,8072
Độ đặc hiệu trung bình	0,9787	0,9790	0,9793
Độ chính xác của điểm ảnh	0,9736	0,9737	0,9738

Sau khi đào tạo và thử nghiệm, chúng tôi đã thu được kết quả của mạng được đề xuất bằng cách sử dụng bộ thử nghiệm gồm 50 hình ảnh 512x512. Ngoài việc chuẩn hóa, chúng tôi không sử dụng bất kỳ kỹ thuật tiền xử lý hoặc hậu xử lý nào. Kết quả đầu ra là

bản đồ phân đoạn nhị phân của các khuyết tật và có cùng kích thước với đầu vào. Chúng tôi đã tính toán IOU trung bình, độ nhạy trung bình và độ đặc hiệu trung bình để đánh giá kết quả phân đoạn và trình bày sự so sánh giữa mạng được đề xuất và các phương pháp tiên tiến. Kết quả của các số liệu đánh giá được trình bày trong Bảng III. Mạng được đề xuất đạt được IOU trung bình cao nhất là 0,6477, độ chính xác pixel là 0,9738 và độ đặc hiệu là 0,8072.

Về độ nhạy, mạng được đề xuất có kết quả thấp hơn so với các phương pháp tiên tiến. Bản đồ dự đoán cuối cùng của hình ảnh đầu vào và hình ảnh thực tế được minh họa trong hình 6.



Hình 6. So sánh kiến trúc (a)(f) hình ảnh đầu vào, (b)(g) thực tế, (c)(h) với u-net, (d)(i) R2-Net, (e)(j) của chúng tôi

Rõ ràng là phương pháp đề xuất phân đoạn các khuyết tật một cách chính xác và có ít phát hiện sai hơn so với các phương pháp tiên tiến. Việc chèn mạng lưới chú ý toàn cầu làm tăng cường các đặc điểm khuyết tật và tập trung nhiều hơn vào các khuyết tật trong khi cải thiện độ chính xác của việc phát hiện vết nứt. Việc chèn giãn nở đóng vai trò quan trọng trong việc trích xuất thông tin ngữ nghĩa lớn. U-net gốc phát hiện nhiều đối tượng hơn là lỗi dẫn đến phân đoạn sai. R2-Net phát hiện lỗi tốt hơn U-net gốc nhưng nó coi kết cấu nền là lỗi.

V. KẾT LUẬN

Trong bài báo này, một khối chú ý toàn cục mới được đề xuất và thêm vào kiến trúc U-net để phân đoạn khuyết tật trong hình ảnh EL lấy từ môi trường công nghiệp thực tế. Kiến trúc được đề xuất bao gồm các phép tích chập giãn nở với các kết nối bỏ qua còn lại chứa nhóm trung bình nhằm mục đích học các đặc điểm phân biệt và đạt được kết quả có ý nghĩa. Mô hình U-net được cải tiến có 8,3M tham số. Nhược điểm duy nhất của mạng được đề xuất là độ nhạy kém hơn so với các phương pháp tiên tiến. Điều này có thể là do chúng tôi không sử dụng bất kỳ kỹ thuật tiền xử lý hoặc hậu xử lý nào. Chúng tôi sẽ xem xét nhược điểm này trong công việc tương lai của mình và tăng

hiệu suất của mạng với độ nhạy được cải thiện. Mô hình đề xuất được so sánh với các mô hình tiên tiến nhất, tức là U-net và R2-Net và kết quả tổng thể cho thấy có thể đạt được kết quả tốt hơn bằng cách áp dụng phương pháp đề xuất.

TÀI LIỆU THAM KHẢO

[1] A. Mansouri, M. Zettl, O. Mayer, M. Lynass, M. Bucher và O. Stern, "Phát hiện khuyết tật trong các mô-dun quang điện bằng cách sử dụng hình ảnh phát quang điện", Hội nghị Năng lượng Soli Photovoltaic lần thứ 27 của Châu Âu, tập 64617926, trang 3374-3378, 2012.

[2] Ebner, R., B. Kubicek và G. Ujvari, Các kỹ thuật không phá hủy để kiểm soát chất lượng của các mô-dun PV: Nhiệt ảnh hồng ngoại, hình ảnh phát quang điện và quang. 2013. 8104-8109.

[3] Tsai, D.-M., S.-C. Wu, và W.-C. Li, Phát hiện khuyết tật của pin mặt trời trong hình ảnh phát quang điện bằng cách sử dụng tái tạo hình ảnh Fourier. Vật liệu năng lượng mặt trời và pin mặt trời, 2012. 99: trang 250-262.

[4] Tsai, D., S. Wu, và W. Chiu, Phát hiện lỗi trong các mô-dun năng lượng mặt trời bằng cách sử dụng hình ảnh cơ sở ICA. Giao dịch IEEE về tin học công nghiệp, 2013. 9(1): trang 122-131

[5] Zhang, X., et al., Một phương pháp mới để phát hiện khuyết tật bề mặt của mô-dun quang điện dựa trên phân tích thành phần độc lập % Các vấn đề toán học trong kỹ thuật. 2013. 2013: trang 8.

[6] Tseng, D.-C., Y.-S. Liu, và C.-M. Chou, Phát hiện gián đoạn ngón tay tự động trong hình ảnh phát quang điện của tế bào năng lượng mặt trời đa tinh thể % Các vấn đề toán học trong kỹ thuật. 2015. 2015: trang 12.

[7] Teow Wee Teo và Mohd Zaid Abdullah, "Phát hiện vết nứt nhỏ của tế bào năng lượng mặt trời bằng cách sử dụng phân tích kết cấu cục bộ", Tạp chí hình ảnh và đồ họa, Tập 6, Số 1, trang 54-58, tháng 6 năm 2018

[8] Ri-Xian, L., Y. Ming-Hai, và W. Xian-Bao, Phát hiện khuyết tật dựa trên học sâu và học chuyển giao. Metall Min Ind.Vol. 7. 312-321. 2015

[9] YJ Cha, W. Choi, O. Büyüköztürk, Phát hiện hư hỏng vết nứt dựa trên học sâu bằng cách sử dụng mạng nơ-ron tích chập, Comput. Aided Civ. Inf. Eng. 32 (5) (2017)

[10] T. Wang, Y. Chen, M. Qiao và H. Snoussi, "Mô hình phát hiện lỗi dựa trên mạng nơ-ron tích chập nhanh và mạnh mẽ trong kiểm soát chất lượng sản phẩm," Int. J. Adv. Manuf. Technol., tập 94, trang 3465-3471, tháng 2 năm 2018

[11] Li, W., KG Field, và D. Morgan, Phân tích khuyết tật tự động trong hình ảnh kính hiển vi điện tử. npj Computational Materials, 2018. 4(1): tr. 36.

[12] Vincent, L. và P. Soille, Lưu vực trong không gian kỹ thuật số: một thuật toán hiệu quả dựa trên mô phỏng nhúng. Giao dịch IEEE về Phân tích mẫu và Trí tuệ máy móc, 1991. 13(6): trang 583-598.

[13] Liu, Y., Geng, J., Su, Z., Zhang, W., Li, J.: Phân loại thời gian thực các khuyết tật bề mặt dãi thép dựa trên CNN sâu. Trong: Ghi chú bài giảng về Kỹ thuật điện, trang 257-266. Số 529 (2019)

[14] Szegegy, C., et al. Đi sâu hơn vào phép tích chập. trong Hội nghị IEEE năm 2015 về Tầm nhìn máy tính và Nhận dạng mẫu (CVPR). 2015.

[15] Zou, Q., et al., DeepCrack: Học các tính năng tích chập phân cấp để phát hiện vết nứt. Giao dịch IEEE về xử lý hình ảnh, 2019. 28(3): tr. 1498-1512.

[16] Long, J., E. Shelhamer và T. Darrell. Mạng hoàn toàn tích chập để phân đoạn ngữ nghĩa. trong Hội nghị IEEE năm 2015 về Tầm nhìn máy tính và Nhận dạng mẫu (CVPR). 2015.

[17] Lee, C.-Y., et al., Deeply-Supervised Nets, trong Biên bản Hội nghị quốc tế lần thứ mười tám về Trí tuệ nhân tạo và Thống kê, L. Guy và SVN Vishwanathan, Biên tập viên. 2015, PMLR % Biên bản nghiên cứu học máy: Biên bản nghiên cứu học máy. tr. 562-570.

[18] Ronneberger, O., P. Fischer và T. Brox. U-Net: Mạng tích chập để phân đoạn hình ảnh y sinh. 2015. Cham: Springer International Publishing.

[19] Lozej, J., et al., Phân đoạn mất từ đầu đến cuối bằng U-Net. 2018. 1-6.

[20] Liu, Z., et al., Phân đoạn trình tự CT gan dựa trên U-Net cải tiến và cắt đồ thị. Hệ thống chuyên gia có ứng dụng, 2019. 126: tr. 54-63.

[21] Ioffe, S. và C. Szegedy, Chuẩn hóa theo lô: Tăng tốc đào tạo mạng sâu bằng cách giảm sự dịch chuyển biến phụ thuộc nội bộ, trong Biên bản báo cáo Hội nghị quốc tế lần thứ 32 về học máy, B. Francis và B. David, Biên tập viên. 2015, PMLR % Biên bản báo cáo nghiên cứu học máy: Biên bản báo cáo nghiên cứu học máy. tr. 448-456.

[22] AL Maas, AY Hannun và AY Ng. Các phi tuyến tính của bộ chính lưu cải thiện các mô hình âm thanh của mạng nơ-ron. Trong Hội nghị quốc tế về học máy (ICML).2013

[23] F. Yu và V. Koltun, "Tổng hợp ngữ cảnh đa thang đo bằng cách giãn nở "sự tích chập", trong ICLR, 2016.

[24] Sudre, CH, Li, W., Vercauteren, T., Ourselin, S., Cardoso, MJ: Sự chống chéo Dice tổng quát như một hàm mất mát học sâu cho các phân đoạn không cân bằng cao. Trong: Học sâu trong phân tích hình ảnh y tế và học đa phương thức để hỗ trợ quyết định lâm sàng. trang 240-248 (2017)

- [25] J. Hu, L. Shen và G. Sun. Mạng nén và kích thích. Bản in trước arXiv arXiv:1709.01507, 2017.
- [26] MZ Alom, M. Hasan, C. Yakopcic, TM Taha và VK Asari, "Mạng nơ-ron tích chập dư hồi quy dựa trên u-net (R2U-Net) để phân đoạn hình ảnh y tế," tháng 2 năm 2018, arXiv:1802.06955
- [27] Tieleman, T., Hinton, G.: Bài giảng 6.5-rmsprop: chia độ dốc cho giá trị trung bình đang chạy của độ lớn gần đây của nó. Trong: COURSERA: Mạng nơ-ron cho máy học (2012)
- [28] Han, H., et al., Phân đoạn khuyết tật wafer silicon đa tinh thể dựa trên mạng nơ-ron tích chập sâu. Thư nhận dạng mẫu, 2018
- [29] M. Lin, Q. Chen, S. Yan, "Mạng trong mạng," trong International Biểu diễn học tập tại hội nghị. (ICLR), 2014, trang 1-10.
- [30] [20] Zhou, B., Khosla, A., Lapedriza, A., Oliva, A., Torralba, "Học các tính năng sâu để định vị phân biệt," trong Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), 2016, tr. 2921-2929.