

HỌC KHÔNG GIÁM SÁT TRONG KIỂM TRA NGOẠI QUAN BỀ MẶT NẤP CAO SU

UNSUPERVISED LEARNING IN RUBBER STOPPER SURFACE ANOMALY DETECTION

PGS, TS. Lê Giang Nam, Lê Xuân Trung
Trường Cơ khí, Đại học Bách Khoa Hà Nội

TÓM TẮT

Phát hiện sự dị thường trên bề mặt kết cấu là một vấn đề quan trọng và cần thiết. Trong ngành y tế, thuốc và dung dịch y tế được sản xuất trong phòng sạch. Những dị thường trên bề mặt nắp cao su mang một tiềm ẩn ảnh hưởng trực tiếp đến chất lượng thuốc dẫn đến ảnh hưởng sức khỏe của con người hoặc động vật khi sử dụng. Một sự dị thường trên bề mặt nắp cao su cần phải được phát hiện để có thể loại bỏ sớm. Thách thức chính của nhiệm vụ này đó là phát hiện các vùng có sự thay đổi là dị thường nhỏ, nhiễu so với mức bình thường, thiếu dữ liệu nắp cao su dị thường được gán nhãn. Trong nghiên cứu này, chúng tôi chỉ ra một phương pháp tiếp cận kết hợp những từ mô hình ImageNet cho phát hiện phân loại những ảnh dị thường và bình thường của nắp cao su. Chúng tôi đã mở rộng và đề xuất mô hình PatchCore với bộ siêu tham số mới để đạt hiệu suất phát hiện các điểm dị thường. Đầu vào hình ảnh 1 kênh, đầu ra mặt nạ phân đoạn hình ảnh với các vùng dị thường trên bề mặt ảnh nắp cao su thử nghiệm. Chúng tôi tiếp tục báo cáo kết quả bộ dữ liệu bổ sung và cũng tìm thấy kết quả khả quan trong chế độ ít mẫu với chỉ số F1 và AUC lần lượt là 0,95 và 0,84.

Từ khóa: Xử lý ảnh; Phát hiện dị thường; Kiểm tra quang học; Học không giám sát; Nắp cao su.

ABSTRACT

Anomaly detection on the surface of a structure is an important and necessary issue. In the medical industry, medicines and medical solutions are produced in clean rooms. The anomalies on the surface of the rubber cap carry a potential that directly affects the quality of the drug leading to health effects in humans or animals when used. An anomaly on the surface of the rubber cap needs to be detected so that it can be removed early. The main challenge of this task is to detect areas where the variation is a small anomaly, noise from the normal level, and the lack of labeled anomaly rubber stopper data. In this study, we show an embedded combined approach from the ImageNet model for classification detection of normal and anomaly images of rubber caps. We have extended and proposed the PatchCore model with a new set of hyperparameters for anomaly detection performance. Image input is 1-channel, image segmentation mask output with an anomaly on test rubber cap image surface. We continue to report additional dataset results and find positive results in the few samples mode with F1-score and AUC of 0.95 and 0.84.

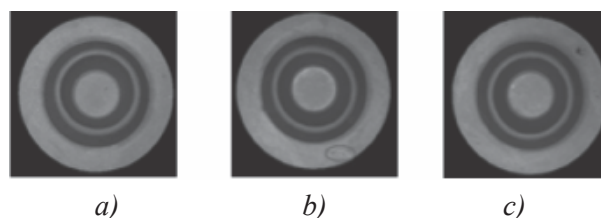
Keywords: Image processing; Anomaly detection; Visual inspection; Unsupervised learning; Rubber stopper.

1. ĐẶT VẤN ĐỀ

Kiểm tra quang học tự động ngày nay là một vấn đề được quan tâm, với sự phát triển nhanh chóng của học sâu ngày càng được triển khai phổ biến rộng rãi hơn trong lĩnh vực phát hiện khiếm khuyết. Trong lĩnh vực y tế, dung dịch y tế hoặc thuốc được sản xuất trong phòng sạch, được đựng trong chai thủy tinh và đậy nắp cao su kín. Hình 1a thể hiện một nắp cao su được định nghĩa là bình thường, không có bất kỳ lỗi nào được xuất hiện trên bề mặt (top view). Vì vậy, điều quan trọng là cần phải phát hiện sớm các khuyết tật bề mặt nắp cao su, đặc biệt là mặt tiếp xúc trực tiếp với dung dịch y tế hay thuốc. Các khuyết tật này gọi là vùng dị thường sẽ ảnh hưởng tới chất lượng thuốc và dung dịch, gây ảnh hưởng tới sức khỏe con người hoặc động vật khi sử dụng. Các phương pháp phát hiện được các vùng dị thường trên bề mặt sản phẩm đã có nhiều nghiên cứu nói tới. Tuy nhiên, phát hiện dị thường trên nắp cao su vẫn còn là một vấn đề mở mà vẫn chưa có nhiều nghiên cứu thực hiện. Pithadiya và Kunal [1] đề xuất một phương pháp phát hiện nắp cao su bị lỗi, được gọi là Bộ lọc hàm mũ đối xứng vô hạn (ISEF). Từ khoảng cách cho các khu vực khác nhau để tìm sự hình thành nắp cao su và bán kính của nắp cao su, phương pháp dựa trên phát hiện cạnh tối ưu để tìm cung tròn từ kỹ thuật Laplace nhị phân, sau đó phát hiện đối tượng với mục đích chọn ra ngưỡng tốt nhất trong số bộ ngưỡng (không triệt tiêu tối đa). Tuy nhiên, phương pháp này không mang lại hiệu quả cho phát hiện dị thường bề mặt.

Các thách thức mà hình ảnh trong công nghiệp mang lại là hình ảnh dị thường trong thực tế rất khó để có thể thu thập đủ các mẫu cho quá trình huấn luyện để có thể thử nghiệm. Dữ liệu hình ảnh dị thường sẽ ít phổ biến hơn,

công việc thu thập dữ liệu và gán nhãn đòi hỏi công sức và chi phí lớn. Hơn nữa, sự dị thường có thể là mô hình không mong muốn và không được dự báo trước. Hình 1 (b và c) bao gồm một số hình ảnh nắp cao su che cổ phần trên của chai thủy tinh có dị thường trên bề mặt. Các lỗi thường xuất hiện ở nắp cao su như vết xước bề mặt, rỗ khí, nứt tế vi,... những lỗi này được định nghĩa là vùng dị thường. Đã có nhiều nghiên cứu giải quyết vấn đề phát hiện dị thường. Các phương pháp huấn luyện dựa trên tập dữ liệu tự nhiên ImageNet [2] để trích xuất các đặc trưng mà không có bất kỳ sự tương đồng nào. Mô hình huấn luyện dựa trên đa số dữ liệu là hình ảnh bình thường và dữ liệu ảnh dị thường được gán nhãn thiểu số như mô hình PatchCore [3], PaDim [4], và SPADE [5] thể hiện một hiệu suất cao.



Hình 1. (a) Nắp cao su bình thường (top view); (b), (c) Nắp cao su dị thường bề mặt

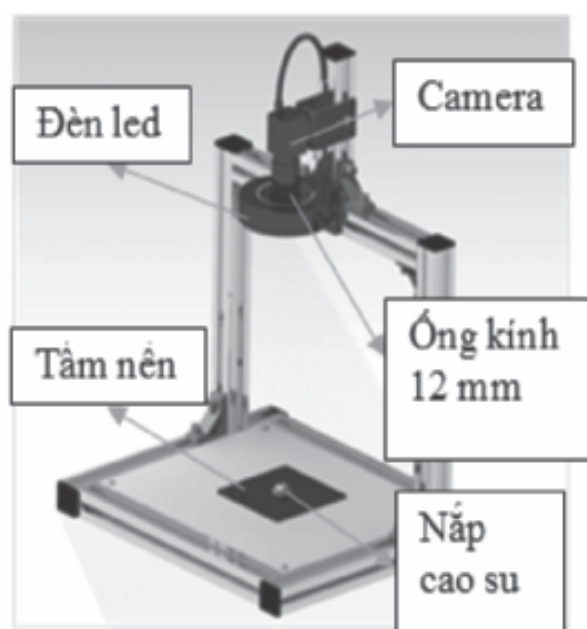
Cuối cùng, chúng tôi tiếp cận với hình ảnh nắp cao su thu thập bằng camera cùng với hệ chiếu sáng đầy đủ. Chúng tôi sử dụng mô hình PatchCore cho khả năng phát hiện hình ảnh và phân đoạn hình ảnh xuất hiện vùng dị thường. Chúng tôi thử nghiệm bộ siêu tham số do chúng tôi đề xuất với dữ liệu ảnh bình thường nhiều hơn. Mạng đào tạo trước Wide-ResNet [6] được sử dụng mang lại hiệu suất tốt, độc lập với dữ liệu đào tạo ở giai đoạn dự đoán. Kết quả cho thấy, hệ thống có thể nhận diện được vùng khuyết tật, cũng như có thể chỉ ra được vị trí trong hình ảnh.

2. VẬT LIỆU VÀ PHƯƠNG PHÁP NGHIÊN CỨU

2.1. Bộ dữ liệu

a. Hệ chiếu sáng

Hệ thống sử dụng 1280x1024 pixels CMOS camera monochrom có kích cỡ cảm biến $\frac{1}{2}$ " với ống kính 12 mm. Trường nhìn FOV là 50x50 mm, độ phân giải 4,8 μ m x 4,8 μ m và độ phóng đại là 0,063. Ánh sáng chiếu theo nguồn sáng tán xạ [7] đã được ứng dụng nhiều trong công nghiệp với đèn led vòng điều chỉnh được cường độ sáng được cung cấp với dòng điện 1A. Lựa chọn phương pháp chiếu sáng tán xạ và nguồn sáng để chiếu sáng tán xạ đều hết bề mặt có nắp cao su. Hệ thống chiếu sáng và thu nhận ảnh được mô tả như hình 2.

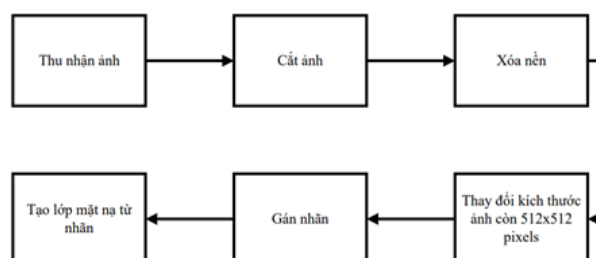


Hình 2. Hệ thống thu nhận ảnh

Bề mặt tấm nền được chọn là bề mặt nhẵn, có độ nhám thấp để tránh độ tương phản cao từ máy ảnh, gây nhiễu cho hình ảnh. Quá

trình này mang lại hiệu quả cao cho xử lý ảnh nhanh và ngắn gọn hơn, chuẩn bị dữ liệu cho quá trình huấn luyện có hiệu suất tốt hơn.

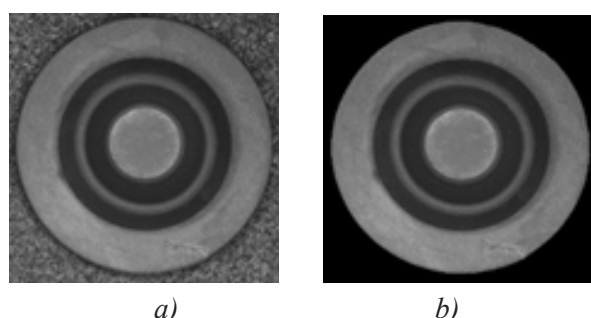
b. Xử lý ảnh



Hình 3. Xử lý ảnh và gán nhãn

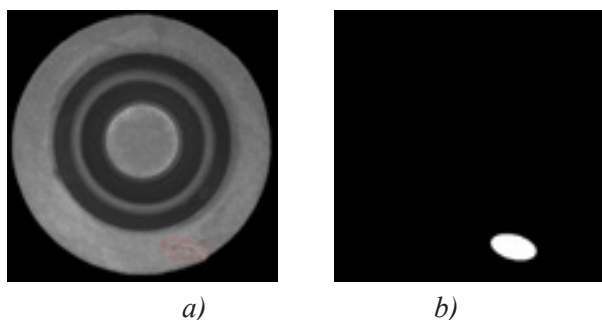
Để quá trình huấn luyện mô hình đạt hiệu suất tốt, chúng tôi đã thực hiện các bước xử lý ảnh như trên hình 3.

Hình ảnh sau khi chúng tôi thu nhận, được tiến hành cắt ảnh chỉ còn nắp cao su và một phần của tấm nền. Tiếp theo, chúng tôi thực hiện bước xóa nền, chỉ còn đối tượng là nắp cao su trong ảnh như hình 4 (a và b) nhằm mục đích giảm nhiễu trong ảnh để hiệu suất huấn luyện đạt cực đại.



Hình 4. (a) Ảnh trước khi xóa nền; (b) Ảnh đã xóa nền

Sau đó, chúng tôi tiến hành thay đổi kích thước ảnh về 512 x 512 pixels và tiến hành gán nhãn những vùng khuyết tật trên nắp cao su. Cuối cùng là phần tạo lớp mặt nạ pixel cho vùng khuyết tật như hình 5 (a và b).



Hình 5. (a) Vết xước bề mặt; (b) Mặt nạ chú thích gắn nhãn

2.2. Phương pháp

a. Khung học tập

Khung học tập PatchCore [3] đã được triển khai để thiết kế phân phối tính năng cho các hình ảnh đào tạo thông thường của chúng tôi. Mạng lưới có ảnh hưởng được đào tạo trước về các nhiệm vụ phân loại hình ảnh. Mô hình huấn luyện trước dựa trên tập dữ liệu tự nhiên ImageNet với khung xương ResNet [8] để trích xuất vector đặc trưng hình ảnh bình thường được đào tạo. Đặc điểm của hình ảnh đích được hiểu là bình thường hoặc dị thường bằng cách sử dụng tìm kiếm k lân cận (k-NN) được tính bằng hàm trung bình tổng hợp (AvgPooling2D). Cuối cùng thu được các đặc trưng đã trích xuất lưu vào ngân hàng bộ nhớ. Bản đồ dị thường được phóng lên bằng nội suy song tuyến và làm mịn bằng một bộ lọc Gauss có độ rộng nhân không đổi $\sigma = 4$.

Phương pháp của chúng tôi sử dụng ảnh đầu vào kích thước 512x512 pixels trên bộ dữ liệu tự xây dựng cho phép chúng tôi trích lọc và giữ lại các đặc trưng đại diện của các vùng dị thường nhỏ. Chúng tôi xem xét một quy trình tương tự như trên, thay đổi bộ tham số phù hợp và sử dụng hàm tổng hợp trung bình với bộ tham số kernel size = (3, 3), stride = 1, padding = 1 cài đặt AvgPooling2D cho tính năng nhận

diện cục bộ. Khoảng cách tính năng là một điểm dị thường ở mọi vị trí được tính toán bằng faiss thông qua k-NN [5]. Bộ siêu tham số được trình bày trong bảng 1.

Bảng 1. Bảng bộ siêu tham số mô hình

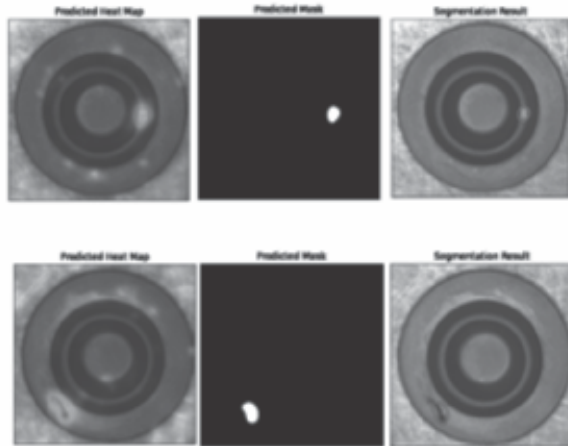
Mô hình đào tạo trước	Wide-ResNet50
σ	4
f_{agg} (Hàm tổng hợp)	AvgPooling2D
p (Kích cỡ khu lân cận)	2
Thuật toán k-NN	faiss. IndexIVFFlat

b. Quá trình huấn luyện

Thư viện mã nguồn mở được sử dụng: OpenCV, Pytorch. Mô hình huấn luyện dựa trên đào tạo trước ImageNet được triển khai với bộ khung xương Wide-ResNet50. Mô hình đào tạo trên bộ dữ liệu của chúng tôi với 1200 ảnh bình thường và 150 ảnh dị thường kèm theo phép tăng cường dữ liệu như: xoay ngẫu nhiên (-5° , $+5^\circ$), lật lên và lật xuống, chuẩn hóa ma trận ảnh với trung bình và độ lệch chuẩn. Trong đó, 20% hình ảnh bình thường được đánh giá dựa trên thời gian kiểm tra, với phần còn lại được sử dụng cho quá trình huấn luyện. Quá trình huấn luyện mô hình sử dụng Google Colab [9], môi trường phần mềm: Python 3.8, Pytorch.

3. KẾT QUẢ

Trong bài báo này, trải qua quá trình xử lý ảnh như tách nền khỏi nắp chai cao su, sau đó xử lý các bước tiếp theo như cắt ảnh, thay đổi kích thước còn 512x512 pixels và xoay ảnh, lật ảnh tăng cường dữ liệu huấn luyện. Hình 6a thể hiện kết quả của bản đồ nhiệt vùng dị thường, kết quả dự đoán mặt nạ pixel trên hình 6b và vùng dị thường trên ảnh được mô tả trên hình 6c.



Hình 6. Kết quả ví dụ trên tập ảnh thử nghiệm:
(a) Bản đồ nhiệt; (b) Mặt nạ pixel dự đoán;
(c) Vùng dị thường trên ảnh thử nghiệm

Trong nghiên cứu này, hiệu suất phát hiện các dị thường trên hình ảnh được đo bằng chỉ số diện tích đường cong (AUROC) [10] và F1-score [11], là chỉ số mạnh mẽ để đánh giá hiệu suất mô hình có tập dữ liệu không cân bằng đã được nhắc đến trong nhiều nghiên cứu từ trước. Hiệu suất mô hình đạt được F1-score = 0,95 và AUROC = 0,84.

Bảng 2. Kết quả so sánh

Nắp cao su	Giải pháp cải tiến	PatchCore 100%
Cấp độ hình ảnh	84%	80%
Cấp độ pixel	95%	95%

Bảng 2 cho thấy được trên cùng tập dữ liệu của chúng tôi, mô hình của chúng tôi sau khi thay đổi bộ tham số đạt độ chính xác dự đoán cao hơn so với mô hình PatchCore nguyên thủy. Những thay đổi của chúng tôi cho thấy một hiệu suất tốt của phát hiện dị thường trên nắp cao su. Kết quả này cho thấy phương pháp được đề xuất của chúng tôi đơn giản và đạt hiệu quả. Điều này rất tốt cho việc xây dựng mô hình thực tế trong công nghiệp vì một tập dữ liệu được chú thích nhãn chất lượng cao rất tốn thời gian và giá thành cao.

4. THẢO LUẬN

Trong bài báo này, trên bộ dữ liệu tự xây dựng từ mô hình cung cấp đánh giá vùng dị thường thông qua học không giám sát. Ngoài khả năng phát hiện lỗi trên nắp cao su thì phương pháp này xác định được những nắp cao su bình thường. Chúng tôi thay đổi bộ siêu tham số để giảm thiểu quá khớp (overfitting) cho mô hình huấn luyện. Trong nghiên cứu này chúng tôi đã chỉ ra những định dạng lỗi đa dạng hơn trong thực tế được thu thập từ mô hình. Điều này có ích cho thực tế sản xuất vì những vùng dị thường xuất hiện mà không có thông tin và không được dự báo. Do đó, có thể mong đợi rằng mô hình đề xuất sẽ được áp dụng nhiều hơn trong bối cảnh công nghiệp thực tế nơi số lượng mẫu dị thường thường bị hạn chế. Mặc dù kết quả phân loại đạt hiệu suất tốt, tuy nhiên cần phải bổ sung dữ liệu ảnh dị thường nắp cao su để hiệu suất được nâng cao hơn và có thể kiểm tra khiếm khuyết ở bề mặt còn lại. Cuối cùng, trọng tâm công việc tiếp theo của chúng tôi là làm thế nào để loại bỏ những nắp chai cao su lỗi trong dây chuyền công nghiệp. ❖

Ngày nhận bài: 28/9/2022

Ngày phản biện: 11/10/2022



Tài liệu tham khảo:

- [1]. Pithadiya, Kunal J; "*Inspection of Rubber Cap Using ISEF (Infinite Symmetric Exponential Filter): An Optimal Edge Detection Method*". International Journal of Research in Engineering and Social Sciences, 2016.
- [2]. Jia. D, Wei. D, R. Socher, Li-Jia. L, Kai. L, Li Fei-Fei; "*ImageNet: A large-scale hierarchical image database*." IEEE Conference on Computer Vision and Pattern Recognition, 2009.
- [3]. Karsten Roth, Latha Pemula, J. Zepeda, B. Schölkopf; "*Towards Total Recall in Industrial Anomaly Detection*". IEEE Computer Vision and Pattern Recognition Conference (CVPR), 2022.
- [4]. Thomas Defard, Aleksandr Setkov; "*PaDiM: a Patch Distribution Modeling Framework for Anomaly Detection and Localization*". ICPR International Workshops and Challenges, 2021.
- [5]. Niv Cohen, Yedid Hoshen; "*Sub-Image Anomaly Detection with Deep Pyramid Correspondences*". ArXiv:2005.02357, 2020.
- [6]. Komodakis, S. Zagoruyko and N; "*Wide residual networks*". BMVC, 2016.
- [7]. Bruce G. Batchelor (Ed.); "*Machine Vision Handbook*". Library of Congress Control Number: 2011942208, Springer-Verlag London Limited 2012.
- [8]. Kaiming He, Xiangyu Zhang, Shaoqing Ren, Jian Sun; "*Deep Residual Learning for Image Recognition*". ArXiv.1512.03385, 2015.
- [9]. Zhe Ming Chng; "*Google Colab for Machine Learning Projects*". In Python for Machine Learning, April 28, 2022.
- [10]. Andrew P. Bradley; "*The use of the area under the ROC curve in the evaluation of machine learning algorithms*". Pattern Recognition, 1997.
- [11]. Damien Fourure, M. Usama Javaid; "*Anomaly Detection: How to Artificially Increase your F1-Score with a Biased Evaluation Protocol*". ECML-PKDD, 2021.