

Phát hiện và phân loại một số bệnh trên trái cam dựa trên kỹ thuật deep learning

Detection And Classification Of Some Orange Fruit Diseases Based On Deep Learning Techniques

Đặng Thị Xuân Tiên^{1,*}

¹ Khoa Công Nghệ, Trường Đại học Xây dựng Miền Tây.

*Tác giả liên hệ: dangthixuantien@mtu.edu.vn

■Nhận bài: 01/09/2024 ■Sửa bài: 18/10/2024 ■Duyệt đăng: 15/11/2024

TÓM TẮT

Phát hiện bệnh trên trái cây có múi nói riêng và trên cây nông nghiệp nói chung đã và đang được rất nhiều sự quan tâm nghiên cứu do tầm quan trọng của nó trong việc xây dựng một nền nông nghiệp bền vững và hiện đại. Các thành viên trong nhóm đã thu thập bộ dữ liệu hình ảnh một số loại bệnh gây hại cho cam trên Kaggle để phân tích các loại bệnh và đề xuất các biện pháp ngăn ngừa cũng như phương pháp trị bệnh tương ứng. Nhóm tác giả đã sử dụng một số kỹ thuật học sâu (Deep Learning) và thị giác máy tính (Computer Vision) qua đó mô hình học máy được huấn luyện trên tập dữ liệu lớn của các hình ảnh trái cam bị bệnh và không bệnh, từ đó học cách phân biệt giữa các trạng thái khác nhau dựa trên kiến trúc mạng YoloV8 và mạng MobileNetV2. Bước đầu qua đánh giá thử nghiệm cho thấy kết quả đạt được rất khả thi để ứng dụng vào thực tế nhằm dự đoán các bệnh trên trái cam.

Từ khóa: Phân loại bệnh trên cam, học sâu, thị giác máy tính, YoloV8, MobileNetV2.

ABSTRACT

Detecting diseases on citrus trees in particular and on agricultural crops in general is receiving much research attention due to its importance in building modern, sustainable agriculture. Team members have collected image data sets of several diseases that damage oranges on Kaggle to analyze the diseases and propose corresponding prevention and treatment methods. The authors used many of deep learning and computer vision techniques, through which the machine learning model was trained on a large data set of images, sick and disease-free oranges, thereby learning to distinguish different states based on the YoloV8 and Inception-V3 network architectures. Initial experimental evaluation shows that the achieved results are very feasible for practical application to predict diseases on oranges.

Keywords: Orange disease classification, Deep learning, Computer Vision, YoloV8, MobileNetV2.

1. GIỚI THIỆU

Nước Việt Nam với khí hậu nhiệt đới ẩm gió mùa vô cùng thích hợp cho sự phát triển của ngành nông nghiệp cây ăn quả. Trong đó cam là một trong những loại cây ăn quả được trồng phổ biến ở khắp Việt Nam, điển hình như ở các tỉnh: Hà Giang, Hà Tĩnh, Tiền Giang, Cần Thơ, Vĩnh Long [1, 2, 4]. Thế nhưng, khí hậu nhiệt đới ẩm

cũng là điều kiện thuận lợi cho vi khuẩn, nấm mốc phát triển và lây lan nhanh chóng trên cây ăn trái.

Một số bệnh thường gặp trên cam như bệnh ghẻ, vàng lá gân xanh và đốm đen. Đây là các thường gặp nhất trên cam, gây mất thẩm mỹ và giá trị của trái cam, bệnh nặng thì ảnh hưởng nghiêm trọng đến năng suất cây trồng.

Hiện nay có một số rất ít hệ thống nhận dạng bệnh học trên cây trồng của một số nhà nghiên cứu trong và ngoài nước như: Nhận dạng bệnh trên cây lúa bằng phương pháp học chuyển giao bởi Nguyễn Thái Nghe [5] hay Creative Common [8] đã giới thiệu về một phương pháp sử dụng mô hình mạng DNN kết hợp mạng VGGNet16 để phát hiện mức độ nghiêm trọng của bệnh trên trái cây có múi với tập dữ liệu được lấy trực tuyến từ PlantVillage và Kaggle. Độ chính xác của mô hình này qua kiểm tra trên các hình ảnh được lựa chọn ngẫu nhiên đối với tình trạng trái cây khỏe mạnh, mức độ bệnh thấp, mức độ bệnh cao và mức độ bệnh trung bình là rất cao.

Bài báo này đề xuất một tiếp cận phát hiện và phân loại một số bệnh trên trái cam với tập dữ liệu được lấy từ Kaggle kết hợp sử dụng một số kỹ thuật học sâu và thị giác máy tính dựa trên kiến trúc mạng YoloV8 và mạng MobileNetV2.

2. NỘI DUNG

2.1 Phương pháp nghiên cứu

- Các phương pháp nghiên cứu:

Nghiên cứu tìm hiểu về kiến trúc mạng học sâu, thị giác máy tính, mô hình mạng tích chập Convolutional Neural Network (CNN) [7, 8]. Nghiên cứu về cách thức hoạt động của các mô hình YoloV8, MobileNetV2. Tìm hiểu

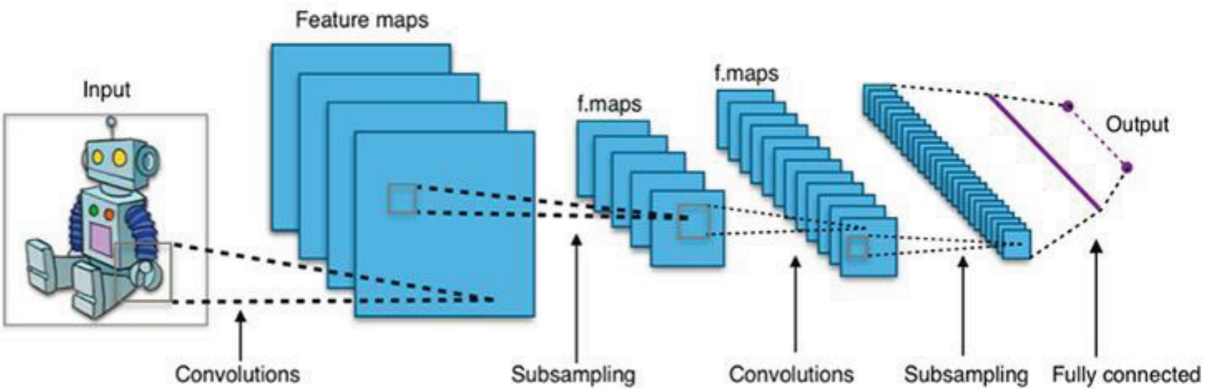
các tài liệu, các công cụ, thư viện và thực hiện chương trình.

Tiến hành thu thập bộ dữ liệu đồng thời xử lý các dữ liệu đã thu thập được, xây dựng mô hình. Phân chia dữ liệu, sử dụng mô hình CNN để trích xuất đặc trưng, tiến hành phân lớp và huấn luyện mô hình. Sau đó tiến hành phân tích, so sánh kết quả và đánh giá các phương pháp đã thực hiện. Từ đó đánh giá và đưa ra kết luận cho mô hình, tập dữ liệu và các phương án cần cải thiện cho mô hình đã thực hiện.

- Mạng nơ-ron tích chập

Mạng nơ-ron tích chập CNN là một mô hình học sâu có thể xây dựng được các hệ thống phân loại có khả năng nhận dạng và phân loại hình ảnh với độ chính xác cao. Mạng CNN cho phép người dùng xây dựng những hệ thống phân loại và dự đoán với độ chính xác cực cao. Hiện nay, mạng CNN được ứng dụng nhiều hơn trong xử lý, nhận diện và phân loại hình ảnh. Mạng CNN bao gồm các lớp cơ bản sau: lớp tích chập (Convolutional), Lớp phi tuyến Relu (Rectified Linear Unit), lớp Pooling và lớp kết nối đầy đủ (Fully Connected Layer).

Mô hình mạng CNN gồm: Input image => Convolutional layer (Conv) + Pooling layer (Pool) => Fully connected layer (FC) => Output.



Hình 1. Mô hình mạng CNN [10]

- Mạng YoLoV8

Yolo (You Only Look Once) [9] , một mô hình phát hiện đối tượng và phân đoạn hình

ảnh phổ biến dựa trên mạng CNN, được phát triển bởi Joseph Redmon và Ali Farhadi tại Đại học Washington. Ra mắt vào năm 2015,

YOLO nhanh chóng trở nên phổ biến vì tốc độ và độ chính xác cao. Các mô hình YOLO được huấn luyện trước trên các bộ dữ liệu lớn như COCO và ImageNet. Điều này cho phép chúng vừa có khả năng cung cấp dự đoán cực kỳ chính xác với các lớp đã được huấn luyện, vừa có thể học các lớp mới một cách tương đối dễ dàng.

YOLOV8 là phiên bản nâng cấp mới nhất, với khả năng nhận diện đối tượng nhanh hơn và chính xác hơn. Điều này được đạt được thông qua một số cải tiến, bao gồm mạng kim tự tháp đặc trưng, các mô-đun chú ý không gian và các kỹ thuật tăng cường dữ liệu tiên tiến.

- Mạng MobileNetV2

MobileNet-V2 [12] là một mô hình phân loại do Google phát triển. Nó cung cấp khả năng phân loại theo thời gian thực dưới các ràng buộc về tính toán trong các thiết bị thông minh. Việc triển khai này thúc đẩy việc chuyển giao việc học từ ImageNet sang tập dữ liệu. MobileNet-V2 được xây dựng dựa trên ý tưởng từ MobileNet-V1, sử dụng tích chập có thể phân tách theo chiều sâu làm các khối xây dựng hiệu quả. Tuy nhiên, MobileNet-V2 giới thiệu hai tính năng mới cho kiến trúc: tắc nghẽn tuyến tính giữa các lớp và kết nối lồi tắt giữa các tắc nghẽn.

- Các thước đo đánh giá mô hình:

Confusion Matrix (CM) là ma trận nhầm lẫn, bao gồm một tập hợp các tiêu chí nhằm đánh giá hiệu quả của một mô hình phân loại. Xét các giá trị trong ma trận nhầm lẫn thu được từ kết quả phân loại, cách sử dụng các chỉ số sau: True Positive (TP), True Negative (TN), False Positive (FP) và False Negative (FN).

- TP: Số lượng hình ảnh bị nhiễm bệnh mà hệ thống đã phát hiện chính xác.

- TN: Số lượng hình ảnh khỏe mạnh mà hệ thống đã phát hiện chính xác.

- FN: Số lượng hình ảnh khỏe mạnh được hệ thống xác định là trái cây bị nhiễm bệnh.

- FP: Số lượng hình ảnh bị nhiễm mà hệ thống đã phát hiện là khỏe mạnh.

Độ chính xác (Precision) là tỉ lệ chính xác khi nhận diện, thang đo độ chính xác của dự đoán. Precision càng cao đồng nghĩa mô hình càng tốt, ít đưa ra các dự đoán Positive sai.

Công thức tính Precision:

$$\text{Precision} = \frac{TP}{TP + FP} \quad (1)$$

Recall dùng để đo lường tỷ lệ dự báo chính xác các trường hợp positive trên toàn bộ các mẫu thuộc nhóm positive. Recall càng cao, mô hình càng tốt, ít bỏ sót trong việc xác định được tất cả các mẫu positive.

Công thức tính Recall:

$$\text{Recall} = \frac{TP}{TP + FN} \quad (2)$$

- F1 Score là trung bình điều hòa giữa precision và recall. Do đó nó đại diện hơn trong việc đánh giá độ chính xác trên đồng thời precision và recall.

- F1-Score càng cao, mô hình càng giỏi cả về chính xác và độ nhạy.

Công thức tính F1 Score:

$$\text{F1 Score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (3)$$

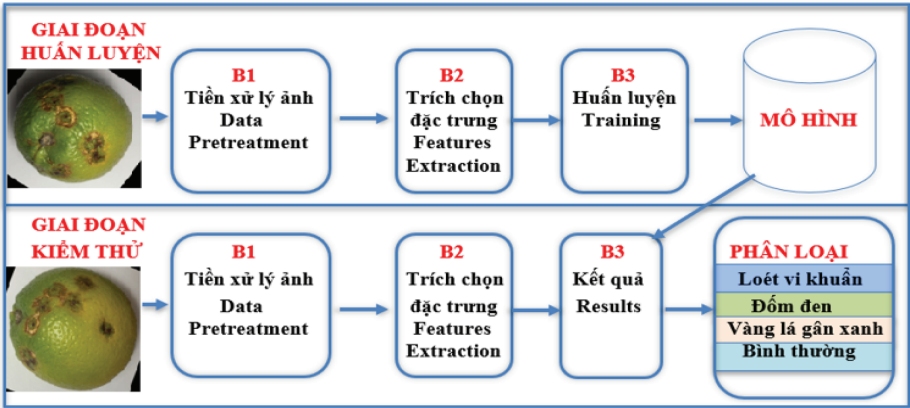
Độ chính xác-Accuracy giúp ta đánh giá hiệu quả dự báo của mô hình. Độ chính xác càng cao thì mô hình càng tốt. Độ chính xác được tính bằng tổng số các trường hợp được dự báo đúng chia cho tổng số các trường hợp.

Công thức tính Accuracy:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (4)$$

Kết quả nghiên cứu:

Để giải quyết bài toán trên nhóm tác giả đề xuất sử dụng mô hình tổng quát như Hình 2. Mô hình tổng quát gồm 2 giai đoạn. Giai đoạn huấn luyện và giai đoạn kiểm thử được mô tả cụ thể:



Hình 2. Mô hình đề xuất phát hiện một số bệnh trên cam

Giai đoạn huấn luyện:

Ở giai đoạn tiền xử lý ảnh. Tuy bộ dữ liệu đã được phân lớp sẵn, nhưng đối với mô hình mạng YoLoV8 buộc phải thực hiện khoanh vùng đối tượng và gán nhãn trên bộ dữ liệu Orange Diseases Dataset đã đề xuất bằng công cụ trực tuyến makesense.ai. Song song đó là loại bỏ các ảnh mờ không đạt tiêu chuẩn, sửa tên, tiến hành gán nhãn và phân loại dữ liệu. Kết quả thu được là các tập tin cần thiết cho quá trình huấn luyện.

Sau đó tiến hành giai đoạn tiếp theo là rút trích đặc trưng với hai mô hình mạng là YoloV8 và MobileNetV2 như đã đề cập ở trên, tiếp đó là huấn luyện mô hình.

Giai đoạn kiểm thử:

Giai đoạn kiểm thử cũng tiến hành tương tự như giai đoạn huấn luyện. Sau khi tiền xử lý dữ liệu. Tiếp theo chuẩn hoá hình ảnh về kích thước 224x224 đối với mô hình mạng MobileNetV2 và kích thước 640x640 đối với mô hình mạng YoloV8 để làm dữ liệu đầu vào. Ảnh sẽ được qua các mạng trích xuất đặc trưng để rút trích ra các đặc trưng tương ứng của từng loại bệnh

2.2. Kết quả thực nghiệm

Các kịch bản áp dụng:

Để tiến hành thực nghiệm cho mô hình đã đề xuất, nhóm tác giả thực hiện 2 kịch bản với tham số huấn luyện như sau:

Bảng 1: Các kịch bản được đề xuất và các tham số huấn luyện

Kịch bản	Mạng huấn luyện	Mạng trích xuất đặc Trưng	Tỉ lệ học	Kích thước mẫu	Số lớp
1	MobileNetV2	MobileNetV2	0,01	32	4
2	YoloV8	YoloV8	0,01	16	4

Môi trường cài đặt và tập dữ liệu thực nghiệm:

Môi trường cài đặt: Hệ thống được cài đặt bằng ngôn ngữ Python và chạy trên cùng một môi trường Google Colab. Thư viện hỗ trợ đào tạo mô hình mạng sử dụng là Tensorflow và Keras.

Tập dữ liệu thực nghiệm: Dữ liệu được nhóm tác giả thu thập là bộ dữ liệu Orange Diseases Dataset được phát triển để xây dựng

các thuật toán học máy và học sâu nhằm thực hiện việc phân loại bệnh trên cam. Trong tập dữ liệu này có lớp cam tươi và ba loại bệnh khác, bệnh loét vi khuẩn, đốm đen và bệnh vàng lá trên cây có múi, bao gồm: 201 hình ảnh về bệnh loét vi khuẩn, 206 hình ảnh về bệnh đốm đen, 369 hình ảnh về bệnh vàng lá gân xanh và 388 hình ảnh về những quả cam khỏe mạnh.

Các kết quả thực nghiệm:

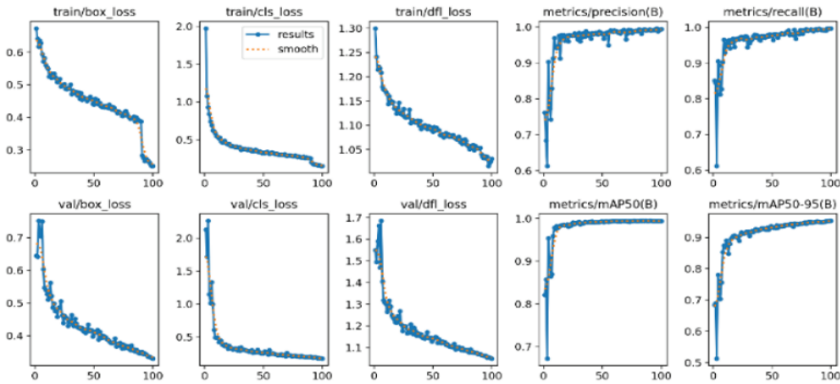
Bảng 2 : Thể hiện các giá trị đánh giá mô hình YoloV8

	Instances	Box(p)	R	mAP50
Fresh	399	0,997	0,997	0,995
Canker	200	0,991	0,99	0,993
Blackspot	207	0,995	0,986	0,995
Greening	369	0,999	1	0,995
Accuracy				99,4

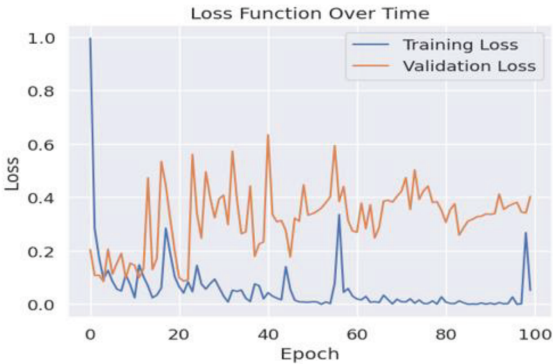
Bảng 3: Thể hiện các giá trị đánh giá mô hình MobileNet-V2

	Precision	Recall	F1-Score	Support
Fresh	1.00	1.00	1.00	78
Canker	0.92	0.97	0.95	36
Blackspot	0.98	0.94	0.96	48
Greening	1.00	1.00	1.00	71
Accuracy			0.98	233
Macro avg	0.97	0.98	0.98	233
Weighted avg	0.98	0.98	0.98	233

Độ đo Loss và Accuracy: Hình 3 thể hiện giá trị Loss và Accuracy của các kịch bản 1 và kịch bản 2. Độ chính xác của các kịch bản lần lượt là 99,4% và 98%. Kết quả thực nghiệm cho thấy kịch bản 1 độ chính xác cao hơn trên tập dữ liệu đề xuất, phù hợp cho nhận dạng bệnh trên cam trong thực tế.



a) Kết quả huấn luyện trên mô hình YoloV8



b) Mô hình MobileNet-V2

Hình 3. Kết quả huấn luyện trên các mô hình

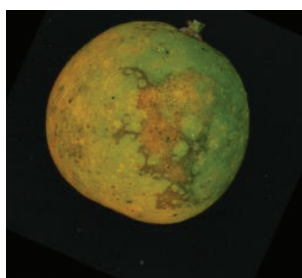
Một số kết quả thực nghiệm của quá trình nhận dạng bệnh trên trái cam dựa trên mô hình mạng đã đề xuất được trình bày như Hình 4.

Mô hình MobileNetV2



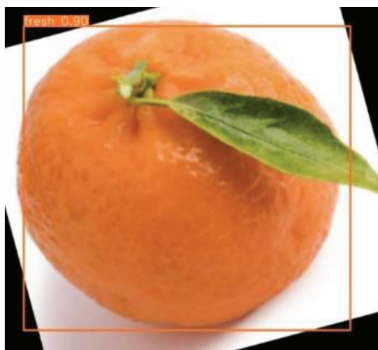
1/1 [=====] - 0s 29ms/step
[[9.9999976e-01 1.8295967e-07 1.4008230e-09 5.1890878e-20]]
Giá trị dự đoán là: BỆNH ĐÓM ĐEN

a) Dự đoán trên bệnh đốm đen



1/1 [=====] - 0s 23ms/step
[[0. 0. 1.]]
Giá trị dự đoán là: Bệnh vàng lá

b) Dự đoán trên bệnh vàng lá gân xanh



Mô hình YoloV8

c) Dự đoán trái cam bình thường



Mô hình YoloV8

d) Dự đoán trái cam bệnh ghẻ

Hình 4: Kết quả dự đoán trên các loại bệnh

Kết quả trên cho thấy phát hiện và phân loại trên trái cam trên cả hai mô hình đạt độ chính xác cao trong nhận dạng 3 loại bệnh và nhận dạng cam bình thường, thời gian xử lý rất nhanh và chính xác với các loại ảnh cam bệnh và ảnh cam không bệnh trên tập ảnh Test.

3. KẾT LUẬN

Trong bài báo này nhóm tác giả đã tiến hành nghiên cứu lý thuyết về mạng học sâu, mạng CNN. Nghiên cứu các mạng YoloV8, mạng MobileNetV2 cho việc phát hiện và phân loại một số bệnh trên trái cam.

Ứng dụng phương pháp giúp phát hiện và phân loại một số bệnh trên trái cam dựa trên các mô hình mạng nêu trên với độ chính xác là 99,4% và 98%.

Kết quả cho thấy việc ứng dụng mô hình học chuyển giao trong nhận dạng bệnh hại trái cam hoàn toàn khả thi. Ứng dụng này có thể hỗ trợ người nông dân phát hiện bệnh hại, kịp thời điều trị đúng bệnh.

TÀI LIỆU THAM KHẢO

- [1]. sfesolution.vn S.S.C. LTD- Bệnh vàng lá thối rễ hoành hành vùng cây có múi Bắc Giang. CÔNG TY TNHH TM & SX NGỌC YẾN, <<http://ngocyen.com.vn/vi/benh-vang-la-thoi-re-hoanh-hanh-vung-cay-co-mui-bac-giang/n1579.html>>.
- [2]. Phòng trừ bệnh loét trên cây có múi trong mùa mưa. <<http://dost-bentre.gov.vn/tin-tuc/606/phong-tru-benh-loet-tren-cay-co-mui-trong-mua-mua>>, accessed: 28/08/2023.
- [3]. Tìm giải pháp khắc phục triệt để diện tích cam bị bệnh ở Hà Giang (2022). Báo Nhân Dân điện tử, <<https://nhandan.vn/post-710200.html>>, accessed.
- [4]. Vĩnh Long khắc phục bệnh vàng lá trên cây cam sành (2006). Báo Nhân Dân điện tử, <<https://nhandan.vn/post-591151.html>>, accessed: 27/08/2023.
- [5]. Thanh T.T.P. và Nghe N.T. (2022). Nhận dạng bệnh trên lá lúa bằng phương pháp học chuyển giao. Tạp Chí Khoa Học Đại Học Cần Thơ, 58(4), 1–7.

- [6]. Dhiman P., Kukreja V., Manoharan P. và cộng sự. (2022). A Novel Deep Learning Model for Detection of Severity Level of the Disease in Citrus Fruits. *Electronics*, 11(3), 495.
- [7]. krishna M., Neelima M., Mane H. và cộng sự. (2018). Image classification using Deep learning. *Int J Eng Technol*, 7, 614.
- [8]. Simonyan K. và Zisserman A. (2014). Very Deep Convolutional Networks for Large-Scale Image Recognition. *CoRR*.
- [9]. Neural Network là gì? Ứng dụng và phân loại (2022) <<https://plus.vtc.edu.vn/neural-network-ung-dung-va-phan-loai/>>, accessed: 29/08/2023.
- [10]. Những Thông Tin Về Cấu Trúc Mạng CNN Là Gì?<<https://itnavi.com.vn/blog/cnn-la-gi>>, accessed: 20/04/2024.
- [11]. xdevlabs (2023). VinBigData, <<https://vin-bigdata.com/cong-nghe-hinh-anh/yolov8-co-gi-nang-cap-so-voi-cac-phien-ban-truoc.html>>, accessed: 04/03/2024.
- [12]. MobileNetV2: The Next Generation of On-Device Computer Vision Networks (2018). <<https://blog.research.google/2018/04/mobilenetv2-next-generation-of-on.html>>, accessed: 08/03/2024.