



Ứng dụng trí tuệ nhân tạo trong phân tích chất lượng và ước tính khối lượng các loại hạt cà phê nhân

PHẠM MINH KHAN^a, LÊ HOÀNH SỬ^{b,*}

^a Trường Đại học Công nghệ Thông tin, Đại học Quốc gia TP. Hồ Chí Minh

^b Trường Đại học Kinh tế - Luật, Đại học Quốc gia TP. Hồ Chí Minh

THÔNG TIN	TÓM TẮT
Ngày nhận: 31/01/2023 Ngày nhận lại: 26/04/2023 Duyệt đăng: 27/04/2023 Mã phân loại JEL: C38; C45; C55. Từ khóa: Phân loại cà phê; Ước lượng khối lượng; Mạng thần kinh tích chập; Xử lý hình ảnh; Trí tuệ nhân tạo. Keywords: Coffee classification; Weight estimates; CNN;	Việt Nam là nước đứng thứ hai thế giới về sản lượng xuất khẩu cà phê, để duy trì thành tựu này thì chất lượng hạt cà phê nhân đóng vai trò rất quan trọng. Doanh nghiệp thu mua và xuất khẩu cà phê cần kiểm soát chất lượng cà phê từ khâu thu mua từng lô hàng tại hộ dân đến sơ chế tại nhà máy và xuất khẩu. Tuy nhiên, việc kiểm soát chất lượng cà phê nhân đòi hỏi chuyên môn và đang được thực hiện thủ công nên tốn thời gian, công sức và chi phí cho doanh nghiệp. Do đó, mục tiêu của nghiên cứu này là xây dựng quy trình, thiết kế thuật toán ứng dụng trí tuệ nhân tạo vào phân tích chất lượng và ước lượng khối lượng hạt cà phê nhân một cách tự động, nhanh chóng và giảm chi phí. Nghiên cứu được thực hiện bằng cách thu thập, dán nhãn mẫu dữ liệu hạt cà phê theo 12 loại hạt, huấn luyện mô hình bằng YOLOv5 trong nhận diện đối tượng, công nghệ xử lý ảnh mã nguồn mở OpenCV, thuật toán CNN (Convolutional Neural Networks) trong xử lý ảnh để phân loại. Từ đó, ước lượng khối lượng từng loại hạt cà phê. Kết quả thử nghiệm cho thấy các mô hình thuật toán ứng dụng cho kết quả tin cậy cao, có thể ứng dụng vào thực tiễn. Abstract Vietnam is the second largest country in the world in terms of coffee export volume, in order to maintain this achievement, the quality of green coffee beans plays a very important role. Coffee exporters need to control the quality of coffee from the stage of purchasing each

* Tác giả liên hệ.

Email: khanpm.14@grad.uit.edu.vn (Phạm Minh Khan), sulh@uel.edu.vn (Lê Hoàng Sử)

Trích dẫn bài viết: Phạm Minh Khan, & Lê Hoàng Sử. (2023). Ứng dụng trí tuệ nhân tạo trong phân tích chất lượng và ước tính khối lượng các loại hạt cà phê nhân. *Tạp chí Nghiên cứu Kinh tế và Kinh doanh Châu Á*, 34(5), 58–80.

Image processing;
Artificial Intelligence.

consignment at the household to preliminarily processing at the factory and exporting. However, the quality control of green coffee requires expertise and is being done manually, so it takes time, effort and cost for the exporters. Therefore, the objective of this study is to build a process, design an algorithm to apply artificial intelligence (AI) to analyze the quality and estimate the weight (ULKL) of green coffee beans automatically and quickly and reduce costs. The study was carried out by collecting and labeling samples of coffee bean data according to 12 types of beans, training the model with Yolov5 in object recognition, OpenCV open-source image processing technology, CNN (Convolutional) algorithm Neural Networks) in image processing for classification and weight estimation of each type of coffee bean. The research results show that the applied algorithm models give highly reliable results, which can be applied in practice.

1. Giới thiệu

Theo Tổ chức Cà phê Thế giới (International Coffee Organization), năm 2020 tổng sản lượng cà phê toàn cầu là 169,6 triệu bao (60 kg/bao); trong đó, 75% tổng sản lượng cà phê toàn cầu đến từ năm quốc gia, bao gồm: Brazil, Việt Nam, Colombia, Indonesia, và Ethiopia. Việt Nam xếp thứ hai thế giới về sản lượng cà phê với 29 triệu bao, đồng thời cũng là quốc gia có năng suất cà phê cao nhất thế giới với 2,4 tấn trên mỗi hecta, gần gấp đôi quốc gia đứng thứ 2 – Brazil (1,4 tấn/hecta) (Đức Anh, 2021).

Để giữ vững được vị thế, các doanh nghiệp xuất khẩu cà phê Việt Nam không chỉ đảm bảo số lượng mà còn phải đáp ứng về chất lượng hạt cà phê thông qua các tiêu chuẩn, quy định về chất lượng. Hiện nay, các tác vụ phân loại, đánh giá chất lượng và định giá thu mua được thực hiện thủ công gây tốn thời gian, nhân lực, tăng chi phí cho doanh nghiệp và ảnh hưởng chất lượng hạt khi thời gian lưu kho kéo dài. Không những thế, các việc này đôi khi cũng không chuẩn xác dẫn đến khó khăn trong việc xác định giá thu mua giữa doanh nghiệp và người nông dân. Rõ ràng, nhu cầu thực tế về một công cụ tự động hóa giúp hỗ trợ đánh giá chất lượng hạt cà phê nói chung và các loại nông sản khác nói riêng tại Việt Nam hiện nay là rất lớn.

Xuất phát từ nhu cầu đó, công nghệ trí tuệ nhân tạo (Artificial Intelligence – AI) với nhiều tiềm năng có thể được áp dụng để giải quyết các vấn đề về phân loại và ước tính khối lượng của nông sản. Việc ứng dụng các giải pháp AI với đặc điểm có thể học từ việc huấn luyện một tập dữ liệu (TDL) đã dán nhãn và có thể tự học để tiếp tục hoàn thiện, giải pháp AI có thể dùng để dự đoán, dự báo các đại lượng trên một TDL mới có những đặc trưng tương tự tập đã huấn luyện. Giải pháp AI với độ tin cậy cao, có thể tương tác với con người và các hệ thống khác giúp lưu trữ dữ liệu và lịch sử hoạt động để có thể được đánh giá và phân tích nhằm cải tiến hệ thống và hành vi người dùng. Theo tiêu chuẩn của Việt Nam, hạt cà phê hiện nay được phân thành 12 loại hạt (mô tả trong Bảng 1). AI có thể được áp dụng để nhận biết và phân loại hạt cà phê thành 12 loại hạt đồng thời ước lượng khối lượng (ULKL) từng loại hạt cà phê một cách tự động thông qua xử lý hình ảnh. Nghiên cứu này sẽ giải quyết ba mục tiêu tương ứng với ba bài toán là nhận diện, phân loại và ULKL hạt cà phê thông qua hình ảnh bằng cách sử dụng những mô hình thuật toán như CNN, Yolo (You Only Look Once).

Trên thế giới có nhiều công trình nghiên cứu về nhận diện những đối tượng nhỏ, như Keles và cộng sự (2022) đã sử dụng Yolov5 (You Only Look Once Version 5) để phát hiện các đối tượng nhỏ, TDL của họ là những bức ảnh để nhận diện các đối tượng rất nhỏ (xe hơi, xe tải, xe buýt...) trên bức ảnh được chụp từ trên cao và họ hướng tới ứng dụng vào thiết bị bay không người lái. Điều này cho thấy rằng việc ứng dụng Yolov5 trong nhận diện và phân loại đối tượng nhỏ như hạt cà phê là tiềm năng (Luis và cộng sự, 2022). Ngoài ra, nhiều bài báo khác cũng đã ứng dụng máy học, học sâu, họ sử dụng Gray-Level Cooccurrence Matrix (GLCM) và K-Nearest Neighbor (K-NN) để phát hiện các hạt cà phê bị lỗi (Jumarlis và cộng sự, 2022); dùng CNN để phân loại các hạt cà phê bị lỗi (Pinto và cộng sự, 2017). Trong bài toán dự báo ULKL thì Piao và cộng sự (2019) đã sử dụng CNN để dự báo giá nhà, còn Zhang và cộng sự (2020) thì đã ULKL và trọng lượng của quả táo.

Không chỉ dừng lại ở việc nhận diện và phân loại vài loại hạt cà phê, nghiên cứu này thực hiện nhận diện, phân loại tới 12 loại hạt. Nhóm tác giả thực hiện trên mẫu TDL được thu thập và phân loại theo 12 loại, sử dụng mô hình Yolov5 trong nhận diện đối tượng, công nghệ xử lý ảnh và mã nguồn mở OpenCV, thuật toán CNN trong xử lý ảnh để phân loại và ULKL. Cấu trúc của nghiên cứu bắt nguồn bằng việc xem xét bài toán doanh nghiệp trong việc phân tích, đánh giá chất lượng hạt cà phê và ước tính giá cả của nghiệp vụ thu mua hạt cà phê.

2. Bài toán và tổng quan các công trình liên quan

AI được ứng dụng để nhận biết và phân loại hạt cà phê đúng theo yêu cầu. Từ việc phân loại, ta có thể ước tính được khối lượng của chúng dựa trên các nghiên cứu và ứng dụng đã được thực hiện trên thế giới. Việc phát hiện ra chính xác các đối tượng trong một bức ảnh là điều không đơn giản. Đặc biệt là một đối tượng nhỏ như hạt cà phê thì việc chọn lựa phương pháp, kỹ thuật, mô hình giúp cho hệ thống có thể phát hiện ra mang lại nhiều thách thức cho các nhà nghiên cứu.

Keles và cộng sự (2022) trong nghiên cứu đánh giá các mô hình YOLO với Sliced Inference (thực hiện suy luận trên các lát nhỏ hơn của ảnh gốc và sau đó hợp các dự đoán trên các lát đó lại vào ảnh gốc) để phát hiện đối tượng nhỏ đã kiểm tra hiệu suất của các mô hình nhận diện đối tượng dựa trên Yolo hiện đại cho nhiệm vụ phát hiện đối tượng nhỏ.

Với phân loại hạt cà phê, Luis và cộng sự (2022) đã thực hiện phân loại hạt lỗi trên hạt cà phê nhân Robusta bằng cách sử dụng YOLO. Luis và cộng sự (2022) sử dụng Yolov5 để xác định các loại hạt, như: Hạt hồng, hạt đen và hạt bình thường; kết quả độ chính xác tổng thể là 95,11% trên cả ba loại hạt. Ở nghiên cứu khác, Jumarlis và cộng sự (2022) đã phân loại các lỗi của hạt cà phê bằng cách sử dụng GLCM và K-NN có độ chính xác khi đánh giá lỗi hạt cà phê là 90%. Theo đó, Jumarlis và cộng sự (2022) dùng GLCM để trích xuất các đặc trưng của ảnh cà phê và sử dụng K-NN để phân loại mức độ lỗi của hạt cà phê. Trong nghiên cứu phân loại hình ảnh hạt cà phê xanh dựa trên các loại lỗi sử dụng mạng CNN của Pinto và cộng sự (2017), CNN được sử dụng để phân loại các loại lỗi của hạt cà phê gồm 6 loại hạt như: Đen (Black), chua (Sour), phai (Fade), cà phê Culi (Peaberry), hư hại (Damage), và bình thường. Kết quả của Pinto và cộng sự (2017) cho thấy một số lớp có độ chính xác phân loại tốt (trên 90%) và một số lớp khác có độ chính xác thấp hơn (72%) đối với ảnh màu. Nghiên cứu của Lee và Jeong (2022) cũng sử dụng CNN trong dự đoán hạt cà phê bị lỗi, kết quả cho độ chính xác của hạt bình thường là 93% và độ chính xác hạt lỗi là 81%.

Ước tính khối lượng của các vật thể đã được thực hiện trong một số nghiên cứu tương tự trước đây, như: Piao và cộng sự (2019) sử dụng CNN trong dự báo giá nhà, Zhang và cộng sự (2020) xử lý hình ảnh và hồi quy qua thị giác máy tính để ước tính khối lượng và trọng lượng của quả táo bằng cách sử dụng phương pháp đo lường không tiếp xúc và tái tạo 3D.

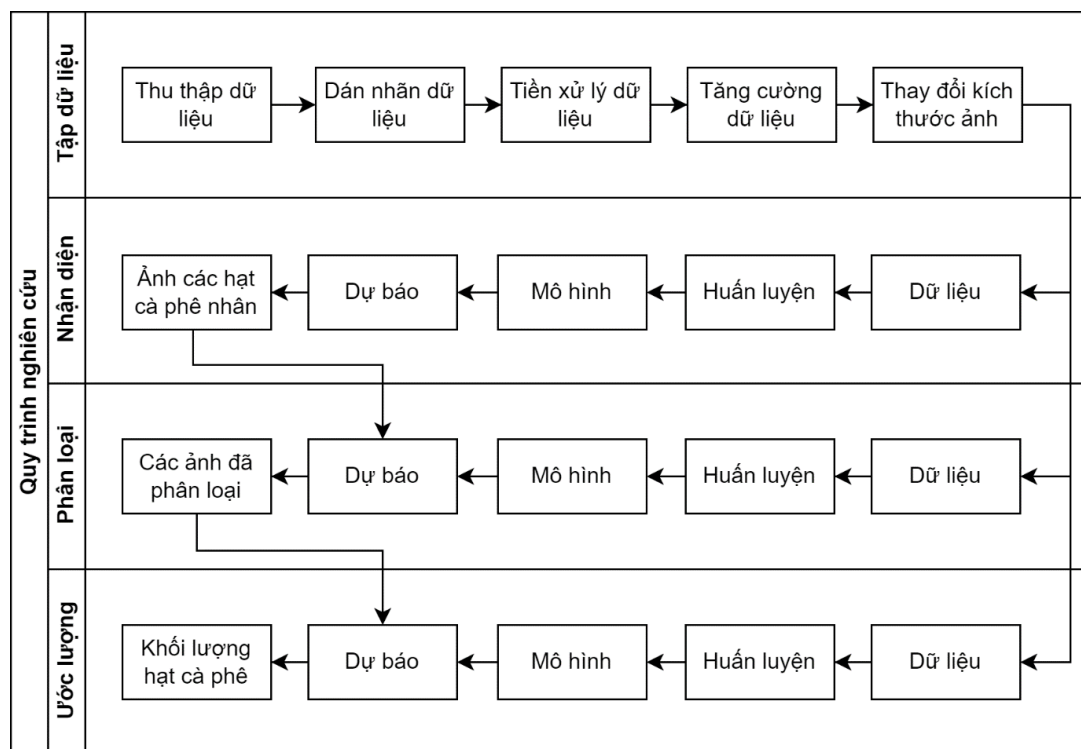
Nghiên cứu này ứng dụng nhiều thuật toán, mô hình khác nhau để giải quyết từng mục tiêu bài toán. Cụ thể, mục tiêu bài toán đầu tiên là nhận diện sẽ được sử dụng bằng Yolov5 (Xu và cộng sự, 2021) và Faster R-CNN (Ren và cộng sự, 2015) để nhận diện đối tượng là hạt cà phê và tờ A4 trong một ảnh. Những kết quả sẽ được so sánh với nhau để chọn mô hình tốt nhất. Sau đó là bài toán phân loại, nhóm tác giả sử dụng Yolov5, Faster R-CNN và EfficientNet (Tan & Le, 2019) để phân loại 12 loại hạt cà phê từ kết quả chọn ra mô hình tốt nhất. Cuối cùng là bài toán ULKL, nhóm tác giả sẽ sử dụng CNN để ULKL 12 loại hạt trên một ảnh.

Như vậy, dựa vào kết quả liên quan từ các công trình nghiên cứu đã thực hiện trên thế giới, nghiên cứu này sẽ ứng dụng các mô hình thuật toán vào từng bài toán như: Nhận diện, phân loại, ULKL. Để đảm bảo được các tiêu chí và giả định như: Mô hình hiện đại, độ chính xác cao, tốc độ xử lý nhanh, ít tốn tài nguyên, và đặc biệt là xử lý được các đối tượng nhỏ như hạt cà phê thì Yolov5, Faster R-CNN và EfficientNet đáp ứng yêu cầu. Trong phần lấy mẫu, tác giả giả định rằng tất cả các hạt cà phê được chọn ngẫu nhiên theo từng loại hạt và sẽ được đặt trên mẫu giấy A4 để chụp ảnh và gán nhãn, riêng TDL cho bài toán phân loại, nhóm tác giả sẽ thực hiện chọn lọc từng hạt (ảnh), để đảm bảo chúng mang những đặc trưng riêng của từng loại hạt và không có sự nhầm lẫn, số lượng ảnh trên mỗi lớp cũng phải đảm bảo theo kiến nghị của mô hình (như Yolov5 là 1.500 ảnh trên một lớp). Các phương pháp đo để đánh giá mô hình cho từng bài toán như sau: Bài toán nhận diện sẽ dùng *mAP* (*mean Average Precision*); bài toán phân loại dùng *Precision*, *Recall*, *F1*; và bài toán ULKL là MAE (Mean Absolute Error), MSE (Mean Squared Error), RMSE (Root Mean Squared Error), MAPE (Mean Absolute Percentage Error); R2 (Coefficient of Determination) để đánh giá và lựa chọn mô hình.

3. Dữ liệu và phương pháp nghiên cứu

3.1. Quy trình nghiên cứu

Trong nghiên cứu này, việc thu thập, gán nhãn, xử lý dữ liệu, tăng cường dữ liệu, thay đổi kích thước hình ảnh để phát hiện, phân loại và ước lượng hạt cà phê trải qua các bước được minh họa trong Hình 1, gồm: (1) Bước 1 – Dataset: Chuẩn bị TDL; (2) Bước 2 – Detection: Nhận diện hạt cà phê; (3) Bước 3 – Classifier: Phân loại cà phê; (4) Bước 4 – Estimate: Ước tính khối lượng hạt.

**Hình 1.** Quy trình nghiên cứu

3.2. Dữ liệu

3.2.1. Thu thập dữ liệu

Một số lượng lớn các hạt được quan sát và chọn lọc bằng mắt thường dựa theo các tiêu chí của tài liệu đặc tả, từ đó thu được TDL hạt chính xác.

Nhóm tác giả chọn lọc 12 loại hạt do chuyên gia trong ngành biên soạn, các mẫu hạt đó được chụp trên mẫu giấy A4 theo từng loại hạt riêng biệt ở định dạng mô hình màu RGB. Để tạo một bộ dữ liệu có thể huấn luyện tốt thì số lượng hạt cho từng loại hạt phải rất lớn. Để tiết kiệm thời gian và công sức, cũng như tránh nhầm lẫn trong việc chụp và cho bước gắn nhãn, nhóm thực hiện phân từng loại hạt và chụp nhiều hạt trong một ảnh. Tổng số 1.528 ảnh thuộc 12 lớp tương ứng 12 loại hạt như trong Bảng 1 dưới đây.

Bảng 1.

Tên gọi và hình ảnh về từng loại hạt cà phê

Tên hạt	Tên nhãn	Hình ảnh	Tên hạt	Tên nhãn	Hình ảnh
Hạt tốt	GOOD		Hạt đen từng phần	PARTLYBLACK	
Tạp nặng	HEAVYFM		Hạt vỡ	BROKEN	
Tạp nhẹ	LIGHTFM		Hạt vỡ lớn	BIGBROKEN	
Cherry	CHERRY		Hạt nâu	BROWN	
Hạt mốc	MOLD		Hạt sâu	INSECT	
Hạt đen	BLACK		Hạt non	IMMATURE	

3.2.2. Tiền xử lý dữ liệu

TDL được tiền xử lý với các phương pháp gồm: Tăng cường dữ liệu, thay đổi kích thước hình ảnh, và chuẩn hóa dữ liệu. Tăng cường dữ liệu sẽ giúp mô hình không bị quá khớp và cải thiện hiệu suất của mô hình. TDL không cân bằng này sẽ được thực hiện một số kỹ thuật nâng cao như đảo ngược hình ảnh sang trái và phải, xoay hình ảnh, thu phóng ảnh và cắt xén ảnh. Tất cả các ảnh khi đưa vào huấn luyện mô hình để dự báo khối lượng đều được điều chỉnh theo kích thước A4 (1.240 x 1.754) trước khi cắt các hạt ra khỏi bức ảnh. Đồng thời, kích thước mỗi hạt sẽ được đưa về 64 x 64 để làm đầu vào cho mô hình huấn luyện. Kích thước đầu vào của mô hình EfficientNet là 300 x 300 ở mỗi hạt.

Cuối cùng là chuẩn hóa dữ liệu, giúp giới hạn dữ liệu trong phạm vi mong muốn để việc học nhanh hơn và giảm khả năng rơi vào tình trạng tối ưu hóa cục bộ (Lee & Jeong, 2022).

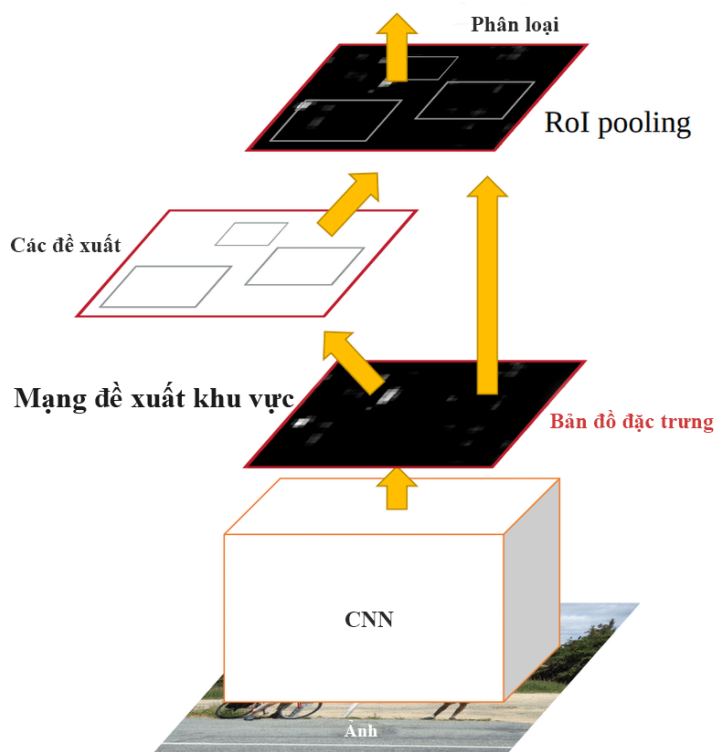
3.3. Phương pháp nghiên cứu

3.3.1. Nhận diện hạt cà phê

3.3.1.1. Mô hình nghiên cứu

Tác giả sử dụng Faster R-CNN (Ren và cộng sự, 2015), YOLOv5 và số lượng lớp Output là 2 (Cafe và A4).

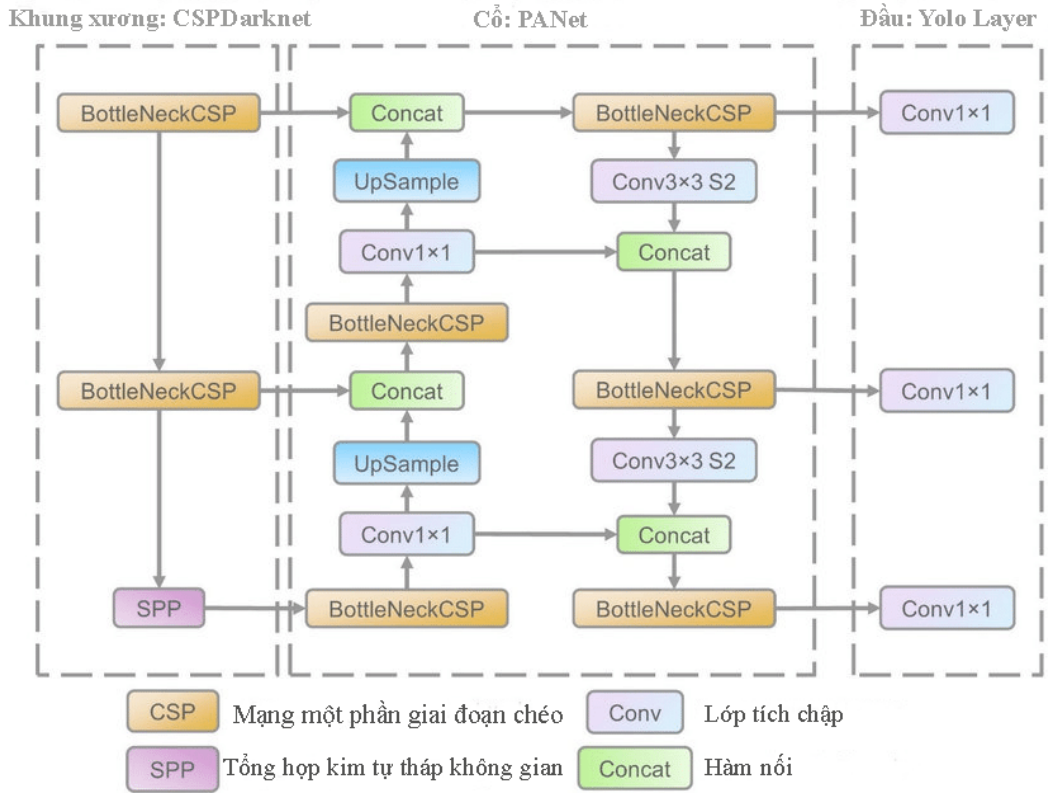
Faster R-CNN: Mô hình này bao gồm hai mô-đun. Mô-đun đầu tiên là một mạng tích hợp sâu đầy đủ để xuất các vùng, và mô-đun thứ hai là mô hình Fast R-CNN sử dụng các vùng được đề xuất. Toàn bộ hệ thống là một mạng thống nhất, duy nhất để phát hiện đối tượng như trình bày ở Hình 2 (Ren và cộng sự, 2015).



Hình 2. Mô hình Faster R-CNN để phát hiện đối tượng

Nguồn: Ren và cộng sự (2015).

YOLOv5: YOLO dùng để phát hiện đối tượng theo thời gian thực, tiên tiến nhất và YOLOv5 dựa trên YOLOv1-YOLOv4. Những cải tiến liên tục đã giúp nó đạt được hiệu suất cao nhất trên hai bộ dữ liệu phát hiện đối tượng chính thức: Pascal VOC (Everingham và cộng sự, 2015) và Microsoft COCO (Lin và cộng sự, 2014). Kiến trúc mạng của YOLOv5 bao gồm ba phần: (1) Khung xương (Backbone): CSPDarknet, (2) Cổ (Neck): PANet, và (3) Đầu (Head): YOLO Layer như ở Hình 3.

**Hình 3.** Kiến trúc mạng của YOLOv5

Nguồn: Xu và cộng sự (2021).

3.3.1.2. Tiêu chí đánh giá bài toán nhận diện

Để đánh giá cho bài toán này, nhóm tác giả dùng các độ đo chuẩn như: mAP theo chuẩn đánh giá của PASCAL VOC (Everingham và cộng sự, 2015) được trình bày theo công thức bên dưới:

Công thức tính độ chính xác trung bình:

$$mAP = \sum_{q=1}^n \frac{AP(q)}{n}$$

Trong đó:

n: Số lượng lớp đối tượng có trong TDL;

AP: Độ chính xác trung bình của từng lớp.

AP được tính bằng công thức:

$$AP = \frac{1}{11} \sum_{r \in \{0,0.1,\dots,1\}} P_{inter} p(r)$$

$P_{inter}p(r)$ được tính bằng công thức:

$$P_{inter}p(r) = \max_{\tilde{r}: \tilde{r} \geq r} p(\tilde{r})$$

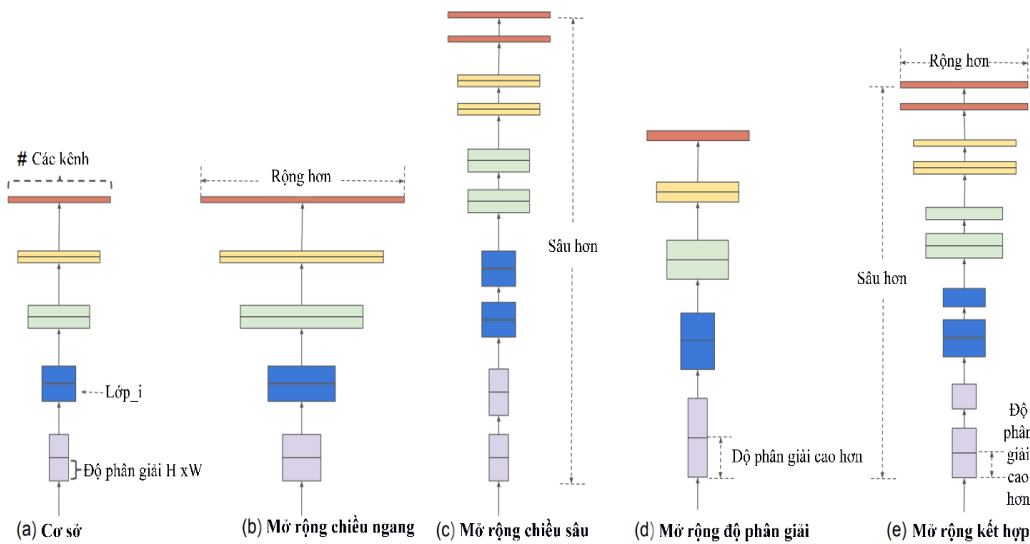
Và $p(\tilde{r})$ là độ chính xác (Precision) được đo tại độ phủ (Recall) $\tilde{r}$.

3.3.2. Phân loại hạt cà phê

3.2.2.1. Mô hình nghiên cứu

Nhóm tác giả sử dụng FasterR-CNN, Yolov5 và EfficientNet (Tan & Le, 2019) để phân loại 12 lớp cà phê như Bảng 1. Mô hình Yolov5 và Faster R-CNN đã được mô tả ở trên.

Tan và Le (2019) nghiên cứu họ CNN mới và gọi CNN mới là EfficientNet. Mô hình này làm tăng độ chính xác và cải thiện hiệu suất bằng cách giảm số tham số và thước đo hiệu suất máy tính FLOPS (Floating Point Operations Per Second) so với các mô hình hiện đại như Gpipe (Huang và cộng sự, 2018). Tan và Le (2019) cũng giới thiệu cách thiết kế kiến trúc cơ sở đơn giản kích thước di động EfficientNet-B0 và cung cấp phương pháp mở rộng kết hợp một cách hiệu quả, giúp tăng kích thước mô hình để đạt được độ chính xác cao nhất như Hình 4.



Hình 4. Mô hình mở rộng

Nguồn: Tan và Le (2019).

Mạng CNN này có thể được mở rộng theo ba chiều: Độ sâu (Depth), độ rộng (Width), độ phân giải (Resolution). Độ sâu của mạng nơ-ron thần kinh tương ứng với số lớp trong mạng, độ rộng liên quan đến số nơ-ron trong mỗi lớp hoặc số Filters trong mỗi lớp Conv (số các kênh của đầu ra; độ phân giải đơn giản là chiều cao và chiều rộng của ảnh đầu vào).

3.2.2.2. Tiêu chí đánh giá cho bài toán phân loại

Theo Trần Thanh Điện và cộng sự (2019), trong bài toán phân loại thì lớp dữ liệu quan trọng là lớp dương, lớp còn lại là lớp âm. Để đánh giá cho bài toán phân loại các loại hạt cà phê, nhóm sử dụng độ đo là Precision (tỷ lệ số điểm **đương đúng** trong số những điểm **được phân loại là dương**), Recall (tỷ lệ số điểm dương đúng trong số những điểm thực sự là dương), F1 (là sự kết hợp của Precision và Recall) với các công thức như sau:

$$Precision = \frac{True\ Positive}{True\ Positive + False\ Positive}$$

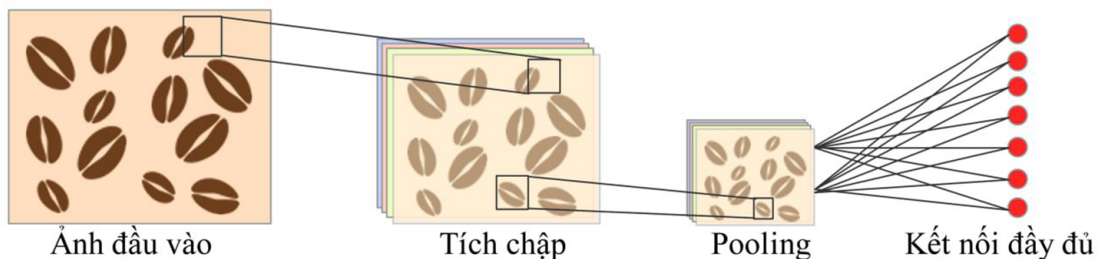
$$Recall = \frac{True\ Positive}{True\ Positive + False\ Negative}$$

$$F1 = \frac{2 \times Precision \times Recall}{Precision + Recall}$$

3.3.3. Ước lượng khối lượng

3.3.3.1. Mô hình nghiên cứu

Nhóm tác giả sử dụng mô hình CNN, kiến trúc truyền thống của CNN là một mạng nơ-ron tích chập, là một dạng mạng nơ-ron được cấu thành bởi các tầng như Hình 5.



Hình 5. Kiến trúc truyền thống của một mạng CNN

CNN chủ yếu bao gồm ba loại cấu trúc phân cấp, đó là: Lớp tích chập, lớp tổng hợp, và lớp kết nối đầy đủ.

Nhóm tác giả chọn CNN bởi đây là một mô hình học sâu được sử dụng phổ biến để xử lý các công việc liên quan đến hình ảnh và video. CNN được thiết kế để tự động học và trích xuất các đặc trưng từ dữ liệu đầu vào (như: Hình ảnh) bằng cách sử dụng tầng tích chập và tầng Pooling. Tuy nhiên, để các mô hình CNN hoạt động hiệu quả, một số giả định nhất định về dữ liệu đầu vào cần được xem xét. Giả định thứ nhất là dữ liệu yêu cầu đầu vào có kích thước hoặc và độ phân giải cố định. Giả định này được sử dụng trong kiến trúc và thiết kế của CNN, trong đó, các kích thước của dữ liệu đầu vào và các bộ lọc trong các lớp tích chập được xác định trước và bất kỳ dữ liệu đầu vào nào có các kích thước khác nhau cần được xử lý trước để phù hợp với kích thước cố định này. Đây là lý do trong nghiên cứu này, quy trình thu thập và tiền xử lý dữ liệu được thực hiện bao chuẩn hóa kích thước và độ phân giải của hình ảnh hạt cà phê trên một tờ giấy A4 là kích thước đã biết trước, sau đó, các ảnh mẫu được điều chỉnh kích thước về kích thước đã định trước. Ngoài ra, các mô hình CNN giả định về dữ liệu có tính dừng (Stationary), tính bất biến theo không gian (Partial Invariance). Nghiên cứu này thu thập mẫu và xử lý hình ảnh mẫu hạt cà phê với nhiều loại, nhiều thời điểm và điều kiện môi trường khác nhau nhằm tối đa hóa đáp ứng các giả định này nhiều nhất có thể.

3.3.3.2. Tiêu chí đánh giá cho bài toán ước lượng khối lượng

Để đánh giá cho bài toán có hồi quy này, nhóm tác giả sử dụng các chỉ số đánh giá mô hình như sau:

Trung bình sai biệt tuyệt đối (MAE):

$$MAE = \frac{1}{N} \sum_{i=1}^N |y_i - \hat{y}|$$

Trung bình bình phương sai số (MSE):

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2$$

Căn bậc 2 của trung bình bình phương sai số (RMSE):

$$RMSE = \sqrt{MSE} = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y})^2}$$

Phần trăm sai số trung bình tuyệt đối (MAPE):

$$MAPE = \frac{1}{N} \sum_{i=1}^N \left| \frac{(y_i - \hat{y})}{y_i} \right|$$

R^2 cho biết mô hình hợp với dữ liệu ở mức bao nhiêu %:

$$R^2 = 1 - \frac{\sum (y_i - \hat{y})^2}{\sum (y_i - \bar{y})^2}$$

Trong đó,

$\hat{y}$: Giá trị dự báo của y ;

$\bar{y}$: Giá trị trung bình của y .

3.3.4. Các kỹ thuật xử lý hiện tượng mô hình quá khớp

Để ngăn hiện tượng mô hình quá khớp, ngoài việc thu thập dữ liệu với kích thước mẫu phù hợp, thực hiện tăng cường dữ liệu như đã trình bày ở mục 3.2.2, nhiều kỹ thuật khác có thể được sử dụng, như: Kỹ thuật quy chuẩn hóa (Regularization), kỹ thuật tắt ngẫu nhiên một số nút trong mô hình (Dropout), kỹ thuật kiểm tra chéo (Cross-Validation)... Các kỹ thuật này thường đòi hỏi người dùng có một nền tảng kiến thức sâu về việc lập và xây dựng các mô hình học sâu (Hà Mạnh Hùng và cộng sự, 2020). Trong phạm vi nghiên cứu này, hai thuật toán có tính hiệu quả rất cao nhưng lại dễ hiểu là kỹ thuật dừng sớm (Early Stopping – ES) và kỹ thuật nhớ mô hình (Model Checkpoint – MC) được giới thiệu và ứng dụng.

3.4. Kỹ thuật dừng sớm

Trong nghiên cứu này, kỹ thuật dừng sớm được sử dụng để tránh xảy ra hiện tượng quá khớp khi số Epoch quá lớn sẽ dẫn đến độ chính xác của mô hình đối với tập kiểm tra bị giảm đi. Như vậy, một cách đơn giản nhất để xử lý vấn đề này là người nghiên cứu cần tìm ra một thời điểm thích hợp dừng việc huấn luyện lại để đảm bảo mô hình thu được cho sai số đối với tập kiểm tra là nhỏ nhất, đây gọi

là kỹ thuật dừng sớm. Để triển khai, TDL được chia thành hai tập là tập huấn luyện và tập kiểm định. Trong quá trình huấn luyện, độ chính xác của mô hình thu được sẽ liên tục được đánh giá dựa trên tập kiểm định nhằm tìm ra được thời điểm khi nào thì dừng cho hợp lý (Hà Mạnh Hùng và cộng sự, 2020). Quá trình huấn luyện dừng khi giá trị Loss của mô hình hiện tại trên tập kiểm định lớn hơn giá trị tương ứng ở vòng lặp trước đó.

3.5. Kỹ thuật nhớ mô hình

Kỹ thuật MC là khi kết thúc một bước lặp, mô hình tìm được sẽ được ghi nhớ lại và trong bước lặp tiếp theo, nếu mô hình mới tìm được tốt hơn mô hình trước thì mô hình mới được ghi lại thay thế mô hình cũ. Mô hình mới được đánh giá là tốt hơn mô hình cũ nếu giá trị Loss của tập kiểm định của mô hình mới nhỏ hơn mô hình cũ. Với cách thức này thì mô hình được ghi nhớ cuối cùng là mô hình có giá trị Loss đối với tập kiểm định là nhỏ nhất. So với kỹ thuật ES, kỹ thuật MC đảm bảo mô hình tìm được là mô hình tốt nhất, tuy nhiên, số lần lặp của kỹ thuật này lại lớn hơn rất nhiều, dẫn đến thời gian huấn luyện dài hơn. Điều này có thể trở thành vấn đề nghiêm trọng đối với các mô hình lớn mà ở đó, thời gian huấn luyện là khá dài. Để khắc phục điều này, có thể kết hợp hai kỹ thuật ES và MC, trong đó ES đóng vai trò dừng chương trình huấn luyện lại khi mô hình không có sự tiến bộ đáng kể nào và MC đóng vai trò ghi nhớ lại mô hình tốt nhất tìm được trước khi chương trình huấn luyện dừng lại. Rõ ràng, sự kết hợp này phát huy thế mạnh cũng như khắc phục nhược điểm khi sử dụng riêng từng kỹ thuật ES và MC (Hà Mạnh Hùng và cộng sự, 2020).

4. Thực nghiệm và kết quả

4.1. Môi trường nghiên cứu

Công việc huấn luyện mô hình được thực hiện trên Server với cấu hình phần cứng được mô tả ở Bảng 2 dưới đây.

Bảng 2.

Mô tả phần cứng server

Thành phần	Thông số
CPU	Intel(R) Xeon(R) CPU E5-2696 v3 @ 2.30GHz
RAM	128 GB
GPU	3 x Card Nvidia Quadro K6000, 12.207MiB

4.2. Mô tả dữ liệu

Với 1.528 ảnh được chụp, các hạt trên ảnh được đặt trên tờ giấy A4 để chuẩn hóa kích thước của các hạt, việc chuẩn hóa sẽ được sử dụng cho phần ULKL hạt. Các ảnh trên sẽ được gán nhãn thủ công với 84.732 hạt thuộc 1.528 ảnh và được đặt tên là Coffee84k. Chi tiết số hạt ở mỗi loại của tập Coffee84k được thể hiện trong Bảng 3.

Bảng 3.

Mô tả số lượng từng loại hạt

Loại hạt	Số hạt	Loại hạt	Số hạt	Loại hạt	Số hạt
Bigbroken	6.924	Immature	3.438	Cherry	6.391
Black	15.707	Insect	6.555	HeavyFN	7.177
Broken	9.919	LightFM	5.188	PartlyBlack	5.601
Brown	4.867	Mold	10.045	Good	2.920
Tổng cộng			84.732		

Trong bài toán nhận diện giữa hạt cà phê và A4, nhóm có thêm nhãn A4 trên mỗi ảnh ngoài 12 loại hạt. Vì thế trong bài toán nhận diện nhóm gộp 12 nhãn của 12 loại hạt thành 1 loại có tên là cà phê và 1 nhãn A4 như Bảng 4.

Bảng 4.

Thống kê số lượng hạt cà phê

Nhãn	Tập huấn luyện	Tập kiểm định	Tổng
Cà phê	67.913	16.819	84.732
A4	1.528	252	1.276

- *Bài toán phân loại:* TDL rút trích còn 2.000 hạt (ảnh) cho mỗi loại hạt. Trong 2.000 hạt này đều có những đặc trưng riêng của từng hạt, khi nhìn sẽ thấy ngay sự khác biệt. Hạt số lượng nhỏ nhất là Immature (3.438 hạt), nếu chọn ít hơn 2.000 hạt thì mô hình không đủ độ tin cậy, nếu chọn nhiều hơn thì số lượng hạt trong TDL ban đầu không đáp ứng, vì thế chọn 2.000 hạt vừa đảm bảo được hiệu quả của mô hình, vừa đảm bảo được dữ liệu ban đầu đáp ứng đủ số lượng để chọn. Riêng hạt nâu chỉ có thể tìm thấy 1.228 ảnh có đặc trưng riêng nên bắt buộc phải sử dụng thêm kỹ thuật tăng cường dữ liệu để có thêm 772 ảnh từ 100 hạt được chọn lọc riêng ngẫu nhiên từ 1.228 ảnh trên. TDL này có tên là Coffee2k, Bảng 5 thể hiện số lượng hạt của từng loại cho từng TDL tập huấn luyện, tập kiểm định, tập kiểm tra.

Bảng 5.

Thống kê số lượng 2.000 hạt cà phê mỗi loại

Loại hạt	Số hạt	Tập huấn luyện	Tập kiểm định	Tập kiểm tra
Bigbroken	2.000	1.600	200	200
Black	2.000	1.600	200	200
Broken	2.000	1.600	200	200
Brown	2.000	1.600	200	200
Cherry	2.000	1.600	200	200

Loại hạt	Số hạt	Tập huấn luyện	Tập kiểm định	Tập kiểm tra
HeavyFM	2.000	1.600	200	200
Immature	2.000	1.600	200	200
Insect	2.000	1.600	200	200
LightFM	2.000	1.600	200	200
Mold	2.000	1.600	200	200
Good	2.000	1.600	200	200
Tổng cộng	24.000	19.200	2.400	2.400

• *Bài toán ước lượng khối lượng*: Mẫu hạt LightFM sẽ được loại bỏ các hạt có ảnh chụp không rõ, bị nhòe hoặc các ảnh có tổng khối lượng không đầy đủ. Do đó, số lượng hạt sẽ được loại bỏ bớt từ 5.188 xuống còn 1.297 hạt. Hạt Immature do thiếu tổng khối lượng các hạt ở mỗi hình, nên mẫu hạt Immature cũng không được sử dụng. Thay vào đó, mẫu Immature sẽ được làm lại mẫu mới, số ảnh chụp lại cho mẫu Immature là 3.513 hạt (ảnh), TDL mới cho bài toán này sẽ có tên là Coffee84kv2 và được chi tiết trong Bảng 6.

Bảng 6.

Thống kê số lượng hạt cà phê mỗi loại

Loại hạt	Số hạt	Loại hạt	Số hạt	Loại hạt	Số hạt
Bigbroken	6.924	Immature	3.513	Cherry	6.391
Black	15.707	Insect	6.555	HeavyFN	7.177
Broken	9.919	LightFM	1.297	PartlyBlack	5.601
Brown	4.867	Mold	10.045	Good	2.920

Để ULKL hạt chính xác thì ngoài ảnh của từng loại hạt cà phê, còn có thêm giá trị số Pixel và khối lượng của hạt đó.

4.3. Phát hiện đối tượng

Trong bài toán nhận diện thì số lớp để nhận diện bao gồm lớp cà phê và lớp A4, TDL 84.732 hạt gồm 1.528 bức ảnh. Tham số cài đặt ban đầu để huấn luyện được trình bày ở Bảng 7 và kết quả đánh giá nhận diện hạt cà phê được thể hiện ở Bảng 8.

Bảng 7.

Tham số để huấn luyện mô hình

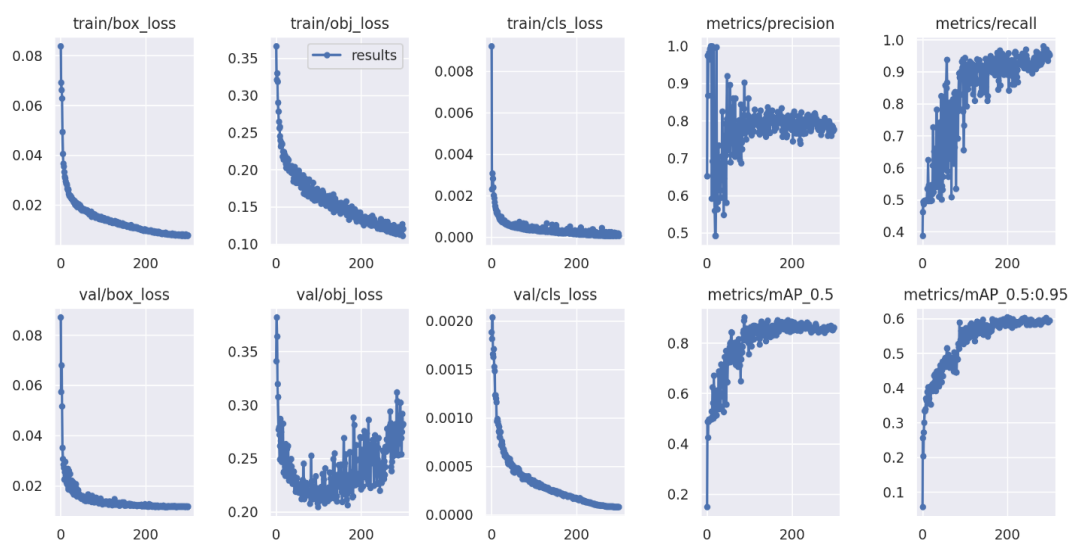
Tham số	Yolov5	Faster R-CNN
Epoch	300	300
Model	Yolov5m	fasterrcnn_resnet50
Số tham số	20.915.769	41.304.286
Bath size	12	8
Learning rate	0,01	0,001

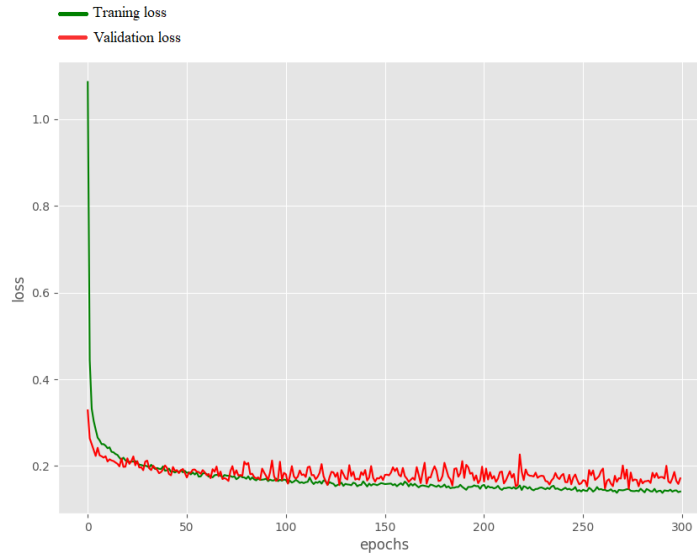
Bảng 8.

Kết quả nhận diện hạt cà phê bằng Yolov5 và Faster R-CNN

Tham số	Yolov5	Faster R-CNN
Precision	0,7973	0,2153
Recall	0,9251	0,4766
mAP_0,5	0,8793	0,1276
mAP_0,5:0,95	0,6035	0,1016
Best validation mAP	0,8793	0,2560

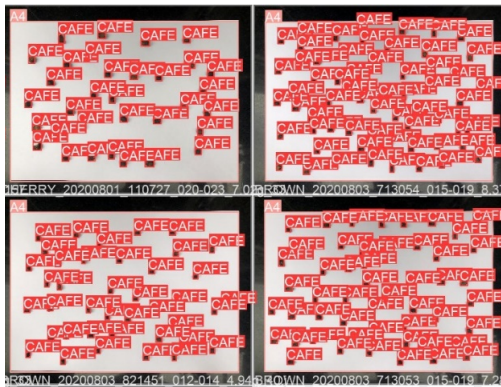
Biểu đồ kết quả của Yolov5 và Faster R-CNN được trình bày trong Hình 6 và Hình 7.

**Hình 6.** Kết quả biểu đồ khi huấn luyện nhận diện bằng Yolov5



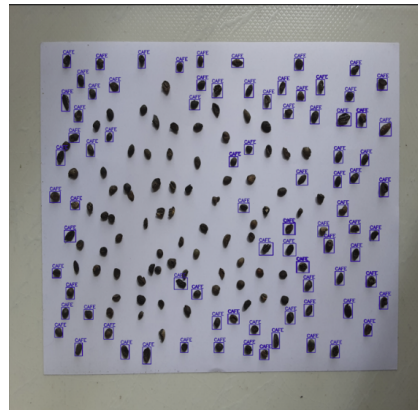
Hình 7. Kết quả biểu đồ Loss khi huấn luyện nhận diện bằng Faster R-CNN
 Kết quả khi mô hình chạy trên tập kiểm chứng (Validation) được minh họa trong Hình 8 và Hình 9.

Yolov5



Hình 8. Kết quả nhận diện bằng Yolov5

Faster R-CNN



Hình 9. Kết quả nhận diện bằng Faster R-CNN

4.4. Phân loại

Trong bài toán phân loại thì số lớp để nhận diện bao gồm 12 lớp ở Bảng 1, TDL 24.000 hạt (ảnh) bao gồm tập huấn luyện với 19.200 mẫu, tập kiểm chứng 2.400 mẫu, và tập thử nghiệm có 2.400 ảnh.

Tham số cài đặt ban đầu để huấn luyện được trình bày trong Bảng 9 và kết quả huấn luyện được trình bày trong Bảng 10.

Bảng 9.

Tham số để huấn luyện mô hình

Tham số	Yolov5	Faster R-CNN	EfficientNet
Epoch	300	300	300
Model	Yolov5m	fasterrcnn_resnet50	EfficientNetB3
Số tham số	20.915.769	41.304.286	12.000.000
Size ảnh	1.280	1.280	300
Bath size	12	8	8
Learning Rate	0,01	0,001	0,001

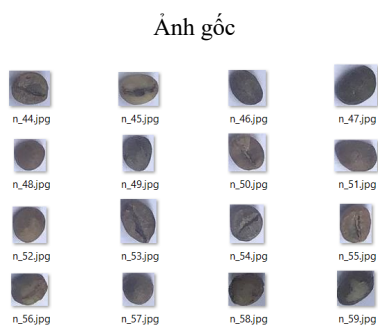
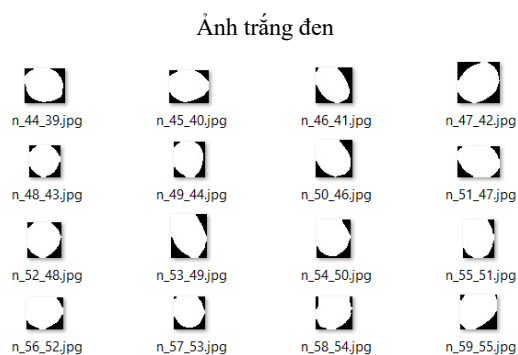
Bảng 10.

Kết quả huấn luyện mô hình của Yolov5, Faster R-CNN, EfficientNet

Tiêu chí	Yolov5	Faster R-CNN	EfficientNet
Precision	0,9460	0,4110	0,9762
Recall	0,9630	0,8350	0,9754
F1	0,9544	0,5508	0,9758

4.5. Ước lượng khối lượng

Ngoài các ảnh của từng hạt thì cần thêm dữ liệu đầu vào để huấn luyện mô hình, như: Số lượng Pixel và khối lượng của từng hạt. Cách thức để thực hiện như sau: Một ảnh mẫu có nhiều hạt trên đó và được xử lý bằng cách cắt các hạt trong ảnh mẫu ra từng hạt ứng với từng hình, từ mỗi hình sử dụng kỹ thuật ngưỡng (Threshold) trong openCV để chuyển ảnh về trắng đen, lấy Pixel phần trắng tương ứng là phần hạt cả phê trên bức ảnh. Từ đó, ta được số Pixel tương ứng là số Pixel của hạt đó như minh họa ở Hình 10 và Hình 11.

**Hình 10.** Ảnh gốc của hạt cà phê trước khi dùng kỹ thuật ngưỡng**Hình 11.** Ảnh trắng đen của hạt cà phê sau khi áp dụng kỹ thuật ngưỡng

Do khối lượng hạt quá nhỏ, không thể dùng cân để cân thực tế từng hạt nên sẽ được tính bằng các công thức dựa trên khối lượng tổng các hạt (giá trị này được tính bằng cách sử dụng cân tiểu ly để cân thủ công tổng số hạt có trên một ảnh lớn trước khi được cắt nhỏ ra từng hạt nhỏ riêng biệt để sử dụng cho bài toán phân loại và ULKL), số Pixel của từng hạt và số lượng hạt có trên một ảnh.

Trước tiên, cần tính khối lượng trung bình (KLTB), Pixel trung bình (PTB) của các hạt trên một ảnh của từng loại hạt theo công thức sau:

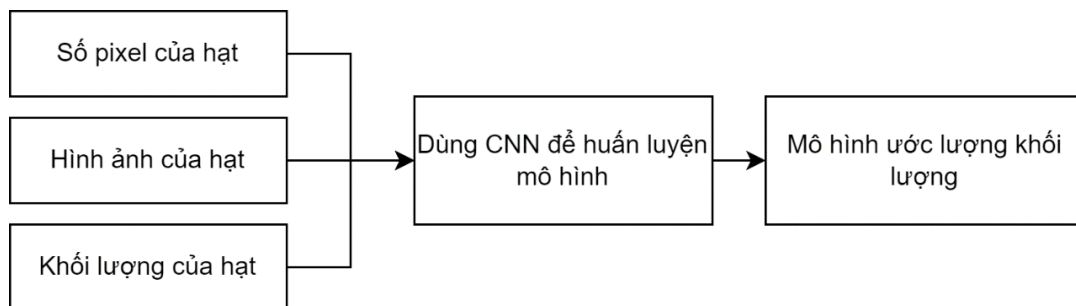
$$\text{KLTB của các hạt} = \frac{\text{Tổng khối lượng hạt}}{\text{Số lượng hạt}}$$

$$\text{PTB} = \frac{\text{Tổng số pixel}}{\text{Số lượng hạt}}$$

Khi đã có KLTB, PTB của các hạt trên một ảnh, sẽ thực hiện tính tiếp khối lượng của từng hạt trên ảnh đó như sau:

$$\text{Khối lượng từng hạt} = \frac{\text{Số pixel từng hạt} \times \text{KLTB của các hạt}}{\text{PTB}}$$

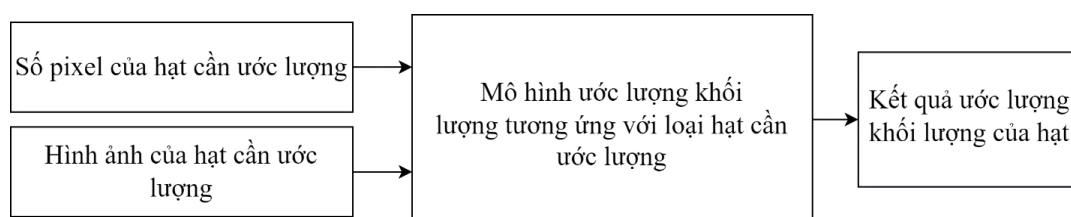
TDL để huấn luyện cho bài toán ULKL gồm có ảnh, số lượng Pixel và khối lượng của từng hạt theo từng loại hạt (Bảng 1), nhóm tác giả xây dựng được 12 TDL của 12 loại hạt và tiến hành huấn luyện 12 mô hình ULKL ứng với 12 loại hạt đó. Nếu chỉ dựa vào khối lượng và số Pixel của từng hạt thuộc 12 loại hạt trên 12 TDL để huấn luyện mô hình ULKL là chưa đủ, vì mỗi ảnh hạt cà phê có những hình dạng và cấu trúc khác nhau, đó là chỉ tính trong cùng một loại hạt, nhưng ở đây nhóm tác giả xét 12 loại hạt nên sẽ có những đặc trưng không chỉ các hạt cùng loại mà còn các hạt khác loại, do đó, phải cần thêm ảnh của từng hạt thuộc từng loại hạt để mô hình có thể học được những đặc trưng và hình dạng của các hạt thuộc từng loại hạt đó, để tạo ra được 12 mô hình ứng với 12 loại hạt. Nhóm sử dụng CNN để huấn luyện mô hình ULKL của từng loại hạt như Hình 12, các giá trị đầu vào để huấn luyện mô hình là số Pixel, khối lượng và ảnh từng hạt thuộc từng loại hạt. Để có được giá trị hồi quy, nhóm tác giả tiến hành cài đặt mạng nơ-ron tích chập ở tầng kết nối đầy đủ (Fully Connected – FC), lớp cuối cùng nhóm sử dụng Activation = “Linear”.



Hình 12. Mô hình ước lượng khối lượng sử dụng CNN và các giá trị đầu vào

Do có 12 loại hạt như trong Bảng 1 nên sẽ có 12 mô hình ULKL (MHULKL) cho từng loại hạt đó, như: MHULKL cho hạt GOOD, MHULKL cho hạt LIGHTFM, MHULKL cho hạt CHERRY...

Sau khi đã có được một mô hình ước lượng của 12 loại hạt thì cách để thực hiện ULKL của một ảnh hạt cà phê mới, ta cần phải tính số Pixel của hạt cà phê mới và ảnh của hạt đó, đồng thời phải biết hạt đó thuộc loại hạt nào trong 12 loại hạt mô tả trong Bảng 1 để chọn mô hình ước lượng tương ứng, như trong sơ đồ Hình 13 dưới đây.



Hình 13. Mô hình ước lượng khối lượng hạt cà phê

Kết quả đánh giá mô hình huấn luyện 12 hạt cà phê bằng CNN được trình bày trong Bảng 11 và kết quả thực nghiệm trên 5 mẫu để so sánh kết quả thực tế và dự báo của từng loại hạt được trình bày trong Bảng 12.

Bảng 11.

Kết quả của 12 loại hạt

Tên hạt	MAE	MSE	MAPE	RMSE	R2
BigBroken	0,0303	0,0025	0,0897	0,0500	0,9028
Black	0,0265	0,0013	0,0764	0,0372	0,9369
Brown	0,0281	0,0013	0,0674	0,0370	0,9343
LightFM	0,0449	0,1160	0,1303	0,3406	0,7721
Cherry	0,0243	0,0014	0,0667	0,0375	0,9281
Good	0,0145	0,0003	0,0354	0,0187	0,9789
PartlyBlack	0,0210	0,0007	0,0578	0,0280	0,9588
Mold	0,0311	0,0016	0,1132	0,0411	0,8946
HeavyFM	0,0220	0,0019	0,1127	0,0439	0,9298
Broken	0,0233	0	0,0645	0,0058	0,9487
Immature	0,0280	0,0013	0,0877	0,0361	0,9321
Insect	0,0211	0,0008	0,0615	0,0293	0,9383

Bảng 12.

Kết quả thực nghiệm trên 5 mẫu

Tên hạt		Mẫu 1	Mẫu 2	Mẫu 3	Mẫu 4	Mẫu 5
BigBroken	Thực tế	0,1857	0,1777	0,3362	0,5648	0,6065
	Dự báo	0,2040	0,1838	0,3512	0,5251	0,5277
Black	Thực tế	0,1635	0,3820	0,2000	0,3939	0,2348
	Dự báo	0,1773	0,3834	0,1987	0,3693	0,3032
Brown	Thực tế	0,5550	0,4081	0,5743	0,2465	0,5403
	Dự báo	0,4774	0,4142	0,5724	0,3534	0,4911
LightFM	Thực tế	0,0551	0,0491	0,0589	0,0808	0,5811
	Dự báo	0,0546	0,0564	0,0641	0,0792	0,4582
Cherry	Thực tế	0,2231	0,3601	0,2757	0,2483	0,2276
	Dự báo	0,2111	0,3545	0,2701	0,2451	0,2837
Good	Thực tế	0,5302	0,5144	0,4678	0,2546	0,3393
	Dự báo	0,5044	0,4896	0,4256	0,2665	0,3273
PartlyBlack	Thực tế	0,5274	0,2912	0,3137	0,3808	0,2282
	Dự báo	0,6073	0,2676	0,2896	0,3415	0,2261
Mold	Thực tế	0,3255	0,1432	0,2212	0,2671	0,3873
	Dự báo	0,3577	0,1487	0,2970	0,2904	0,3673
HeavyFM	Thực tế	0,5030	0,5073	0,0670	0,0514	0,0712
	Dự báo	0,4654	0,4578	0,0685	0,0522	0,0747
Broken	Thực tế	0,3887	0,2484	0,6333	0,4062	0,2557
	Dự báo	0,3780	0,2382	0,6120	0,4354	0,3085
Immature	Thực tế	0,3680	0,3051	0,3096	0,5317	0,6714
	Dự báo	0,3860	0,2908	0,3204	0,4758	0,6924
Insect	Thực tế	0,3010	0,2234	0,5538	0,3454	0,2556
	Dự báo	0,3129	0,2333	0,5412	0,3006	0,2527

5. Phân tích và đánh giá

- Bài toán nhận diện: Nhóm sử dụng mAP để đánh giá, từ kết quả trong Bảng 8 cho thấy tham số Best Validation mAP của Yolov5 (0,8793) cao hơn Faster R-CNN (0,256) nên trong bài toán này nhóm chọn Yolov5 để triển khai.

- Bài toán phân loại: Nhóm dựa vào Precision, Recall và F1 để đánh giá cho bài toán này. Dựa vào kết quả huấn luyện mô hình bằng YOLOv5, Faster R-CNN và EfficientNet trong Bảng 10 thì kết quả EfficientNet cho kết quả tốt nhất ở 3 tiêu chí lần lượt là 0,9762, 0,9754, 0,9758. Vì thế, trong bài toán phân loại hạt cà phê ở trong nghiên cứu này, EfficientNet được chọn.

- Bài toán ULKL: Nhóm sử dụng mô hình CNN có hồi quy với 5 tiêu chí để đánh giá mô hình là MAE, MSE, MAPE, RMSE, R2. Trong đó, R2 là tiêu chí được chọn để đánh giá chính cho mô hình này và Bảng 11 là kết quả R2 của 12 loại hạt đã được huấn luyện. Ở bài toán này, sai số ước lượng ở khoảng 10–15%, cho thấy vẫn còn hạn chế nhưng có thể ứng dụng được.

6. Kết luận

Sau tất cả những thực nghiệm, có thể kết luận rằng các mục tiêu của nghiên cứu đề ra đã hoàn thành. Mục tiêu thứ nhất là thu thập, xử lý hai bộ dữ liệu là Coffee84k và Coffee2k. Mục tiêu thứ hai là bài toán nhận diện đối tượng hạt cà phê trên tập Coffee84k, kết quả so sánh giữa YOLOv5 và Faster R-CNN được thể hiện trong Bảng 8. Trong bảng này, YOLOv5 cho kết quả tốt hơn Faster R-CNN. Mục tiêu thứ ba là bài toán phân loại các hạt cà phê, kết quả của mục tiêu này được thể hiện trong Bảng 10, theo bảng này, mô hình EfficientNet cho kết quả tốt nhất so với hai mô hình còn lại với TDL Coffee2k. Mục tiêu cuối cùng là bài toán ULKL của từng hạt được thể hiện trong Bảng 11, kết quả chung các hạt với $R^2 > 0,9$, riêng hạt LightFM cho kết quả thấp hơn với $R^2 = 0,77$.

Với những kết quả trên thì các mục tiêu của từng bài toán đều có thể đáp ứng tốt trong điều kiện thực tế: Bài toán phân loại có thể giúp phân loại ra từng loại hạt cà phê được thể hiện trong một ảnh, đồng thời từ đó có thể đánh giá nhanh chất lượng hạt cà phê thu được, xem số lượng hạt tốt và hạt lỗi đạt bao nhiêu phần trăm; bài toán ước lượng có thể tính toán, ước lượng được khối lượng các loại hạt cà phê nhân, điều này sẽ làm cho việc định giá trong thu mua được nhanh hơn rất nhiều. Nghiên cứu này giúp rút ngắn thời gian phân tích, đánh giá chất lượng hạt cà phê xuống chỉ còn vài phút và ngay tại điểm thu mua ngoài hiện trường, có thể định giá nhanh chất lượng hạt cà phê thông qua ULKL các hạt bằng AI thay vì phải thu mua cà phê trước và đưa về trụ sở, vài ngày sau cà phê đó mới được nhân viên quản lý chất lượng trực tiếp kiểm tra bằng mắt người cho tất cả các lô hàng nhập về.

Kết quả nghiên cứu có thể được triển khai tại bộ phận thu mua của các công ty thu mua hạt cà phê thông qua một ứng dụng di động. Những người thu mua sẽ đến các hộ nông dân, dùng ứng dụng để kiểm tra chất lượng và phản hồi chất lượng cà phê một cách khách quan, các số liệu hình ảnh xử lý sẽ được ghi nhận lại trên Server. Ngoài ra, việc theo dõi đặc điểm chất lượng hạt cà phê của từng vùng địa lý, từng hộ dân cũng giúp việc đề xuất các kiến nghị thay đổi cách thức canh tác, chế biến sau thu hoạch nhằm nâng cao chất lượng và lựa chọn nơi thu mua nguyên liệu đáp ứng yêu cầu xuất khẩu của từng thị trường nước ngoài. Ở cấp độ quản lý, dữ liệu hình ảnh và dữ liệu đặc điểm của vùng trồng, hộ dân, thời điểm thu hoạch, số lô hàng... được lưu trữ ở Server sẽ được sử dụng để tiếp tục tái huấn luyện nhằm cải thiện mô hình.

Bên cạnh đó, ứng dụng này có thể được mở rộng ra để vận dụng cho các loại nông sản khác đang là thế mạnh xuất khẩu của Việt Nam hiện nay như: Các loại đậu, hạt điều, hạt tiêu...

Đối với bài toán ULKL hạt cà phê trong nghiên cứu này vẫn còn sai số cao (10–15%), vì vậy, có thể tối ưu việc tách đường bao quanh hạt cà phê và ứng dụng các thuật toán khác để nâng cao độ chính xác. Ngoài ra, nghiên cứu này được thực hiện với định hướng thực nghiệm từ nhận dạng vấn đề thực

tế trong kiểm tra chất lượng và thu mua hạt cà phê, từ đó xây dựng bài toán và thiết kế quy trình nghiên cứu từ thu thập dữ liệu, tiền xử lý, xây dựng thuật toán nhận dạng đối tượng, phân loại đối tượng và ULKL hạt cà phê cho từng loại hạt trên mỗi tập mẫu kiểm tra chất lượng hạt cà phê nên có hạn chế trong việc kiểm định đầy đủ các giả định về dữ liệu của mô hình cũng như kiểm định đánh giá lựa chọn mô hình tối ưu nhất cho bài toán.

Hướng phát triển trong tương lai là nghiên cứu nhiều yếu tố khác, chẳng hạn như xây dựng, mở rộng, tăng cường thêm ảnh chất lượng cao cho TDL; từ đó cấu trúc của mô hình thuật toán sẽ được tinh chỉnh. Đồng thời, thực hiện thêm nhiều mô hình thuật toán khác nhau để phù hợp với TDL mở rộng, từ đó mang lại hiệu suất dự đoán tốt hơn. Tuy nhiên, để ứng dụng được hoàn thiện, nhóm nghiên cứu cần thu thập dữ liệu đa dạng hơn, tiến hành kiểm định các giả thuyết của mô hình và đánh giá lựa chọn mô hình tối ưu.

Tài liệu tham khảo

- Đức Anh. (2021). *10 nước sản xuất cà phê lớn nhất thế giới, Việt Nam đứng thứ 2*. Truy cập từ <https://vneconomy.vn/10-nuoc-san-xuat-ca-phe-lon-nhat-the-gioi-viet-nam-dung-thu-2.htm>
- Everingham, M., Eslami, S. M. A., van Gool, L., Williams, C. K. I., Winn, J., & Zisserman, A. (2015). The pascal visual object classes challenge: A retrospective. *International Journal of Computer Vision*, 111(1), 98–136. doi: 10.1007/s11263-014-0733-5
- Hà Mạnh Hùng, Trương Việt Hùng, Đinh Văn Thuật, & Vũ Quang Việt. (2020). Phương pháp xử lý hiện tượng mô hình quá khớp trong xây dựng mô hình học sâu để ước lượng khả năng chịu tải của giàn phi tuyến. *Tạp chí Khoa học Công nghệ Xây dựng*, 14(1V), 12–20. doi: 10.31814/stce.nuce2020-14(1V)-02
- Huang, Y., Cheng, Y., Bapna, A., Firat, O., Chen, M. X., Chen, D., . . . Chen, Z. (2018). *GPIPE: Efficient training of giant neural networks using pipeline parallelism*. Retrieved from <https://arxiv.org/abs/1811.06965>. doi: 10.48550/ARXIV.1811.06965
- Huy, T. V. (2021). *EfficientNet understanding*. Retrieved from <https://huytranvan2010.github.io/EfficientNet-Understanding/>
- Jumarlis, M., Mirfan, M., & Manga, A. (2022). Classification of Coffee Bean Defects Using Gray-Level Co-Occurrence Matrix and K-Nearest Neighbor. *ILKOM Jurnal Ilmiah*, 14, 1–9. doi: 10.33096/ilkom.v14i1.910.1-9
- Keles, M. C., Salmanoglu, B., Guzel, M. S., Gursoy, B., & Bostanci, G. E. (2022). *Evaluation of YOLO models with sliced inference for small object detection*. Retrieved from <https://arxiv.org/abs/2203.04799>. doi: 10.48550/ARXIV.2203.04799
- Lee, J. Y., & Jeong, Y. S. (2022). *Prediction of defect coffee beans using CNN*. Paper presented at the 2022 IEEE International Conference on Big Data and Smart Computing (BigComp).
- Lin, T.-Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., . . . Zitnick, C. L. (2014). *Microsoft COCO: Common objects in context*. Lecture Notes in Computer Science (pp. 740–755). doi: 10.1007/978-3-319-10602-1_48

- Luis, V. A. M., Quiñones, M. V. T., & Yumang, A. N. (2022). *Classification of defects in robusta green coffee beans using YOLO*. Paper presented at the 2022 IEEE International Conference on Artificial Intelligence in Engineering and Technology (IICAET).
- Piao, Y., Chen, A., & Shang, Z. (2019). *Housing price prediction based on CNN*. Paper presented at the 2019 9th International Conference on Information Science and Technology (ICIST).
- Pinto, C., Furukawa, J., Fukai, H., & Tamura, S. (2017). *Classification of Green coffee bean images basec on defect types using convolutional neural network (CNN)*. Paper presented at the 2017 International Conference on Advanced Informatics, Concepts, Theory, and Applications (ICAICTA).
- Ren, S., He, K., Girshick, R., & Sun, J. (2015). *Faster R-CNN: Towards real-time object detection with region proposal networks*. doi: 10.48550/ARXIV.1506.01497
- Tan, M., & Le, Q. V. (2019). *EfficientNet: Rethinking model scaling for convolutional neural networks*. doi:10.48550/ARXIV.1905.11946
- Trần Thanh Điện, Thái Nhựt Thanh, & Nguyễn Thái Nghe. (2019). Giải pháp phân loại bài báo khoa học bằng kỹ thuật máy học. *Tạp chí Khoa học Trường Đại học Cần Thơ*, 55(4A), 29–37. doi: 10.22144/ctu.jvn.2019.093
- Xu, R., Lin, H., Lu, K., Cao, L., & Liu, Y. (2021). A forest fire detection system based on ensemble learning. *Forests*, 12, 217. doi: 10.3390/f12020217
- Zhang, B., Guo, N., Huang, J., Gu, B., & Zhou, J. (2020). Computer vision estimation of the volume and weight of apples by using 3D reconstruction and noncontact measuring methods. *Journal of Sensors*, 2020, 1–12. doi: 10.1155/2020/5053407