

PHÁT HIỆN BẤT THƯỜNG HÀNH VI LỘN TỪ CAMERA SỬ DỤNG HỌC SÂU

Dương Thành Nam*, Nguyễn Thị Kim Phương*, Trần Thị Hoa*, Trần Sơn Tùng*,
Dương Trần Đức*

* Trung tâm Nghiên cứu và Chuyển giao Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam

+ Học viện Công nghệ Bưu chính Viễn thông

Tóm tắt: Phát hiện bất thường hành vi lộn trong các trang trại là một hoạt động quan trọng trong theo dõi và giám sát sức khỏe lợn. Theo các nghiên cứu trước đây, lợn có vấn đề về sức khỏe thường có các biểu hiện khác thường trong hành vi. Việc sớm phát hiện được các bất thường trong hành vi lợn có thể giúp xác định sớm tình hình sức khỏe lợn và có các hoạt động khám bệnh và điều trị sớm, nhằm ngăn chặn tình trạng lây lan của dịch bệnh. Trong bài báo này, chúng tôi đề xuất phương pháp phát hiện một số tình huống bất thường trong hành vi lợn thông qua theo dõi và nhận diện tự động từ camera giám sát. Phương pháp tiếp cận bao gồm các bước phát hiện, theo dõi, nhận diện hành vi, và đánh giá tình huống bất thường trong hành vi. Các tình huống bất thường trong hành vi được phát hiện bao gồm các hành vi như nằm lâu, bỏ ăn, hoặc tăng vận động. Các tình huống được phát hiện theo các khung thời gian hoạt động, giúp giảm thời gian theo dõi tình huống mà vẫn đảm bảo phát hiện bất thường với độ chính xác cao. Các thực nghiệm cho thấy việc sử dụng thời gian theo dõi 30 phút vừa đảm bảo giảm thiểu lỗi sai định danh khi theo dõi, trong khi vẫn đảm bảo độ chính xác phát hiện bất thường lên tới 93.8%.

Từ khóa: phát hiện bất thường hành vi lợn, học sâu, theo dõi lợn.

I. MỞ ĐẦU

Theo dõi lợn là một hoạt động có vai trò quan trọng trong việc phát hiện các vấn đề về sức khỏe lợn. Việc theo dõi lợn thủ công có nhiều vấn đề như không thể theo dõi liên tục và khả năng kết nối thông tin giám sát kém. Đặc biệt, trong bối cảnh các trang trại lợn ngày càng được công nghiệp hóa và gia tăng về số lượng, việc theo dõi thủ công trên số lượng lớn là điều bất khả thi. Việc sử dụng camera giám sát theo cách thông thường cũng chỉ giúp tăng khả năng theo dõi từ xa và xem lại hình ảnh quá khứ khi có nhu cầu. Một phương pháp cho phép tự động theo dõi và nhận diện hành vi lợn từ camera giám sát sẽ giúp giải quyết các vấn đề trên. Đầu tiên, việc theo dõi tự động sẽ giúp quá trình theo dõi được thực hiện thường xuyên và không bị ngắt quãng. Hơn nữa, theo dõi tự động cũng giúp ghi nhận các biểu hiện hành vi của các cá thể lợn trong một khoảng thời gian, giúp khả năng phát hiện bất thường tốt hơn khi kết hợp chuỗi hành vi trong một khoảng thời gian thay vì chỉ xét trong một thời điểm.

Tác giả liên hệ: Dương Trần Đức,

Email: duongtranduc@gmail.com

Đến tòa soạn: 8/2023, chỉnh sửa: 9/2023, chấp nhận đăng: 10/2023.

Bài báo này trình bày phương pháp tự động theo dõi, nhận diện, và phát hiện bất thường hành vi lợn từ camera giám sát. Các hành vi nhận diện được có thể được lưu trữ và kết hợp để xem xét khả năng có sự bất thường hành vi trong một khoảng thời gian, từ đó có các dự báo về tình trạng sức khỏe lợn. Các thực nghiệm được thực hiện trên tập dữ liệu video lợn được thu thập từ trang trại thực tế với hơn 15.000 ảnh được trích xuất từ video và gắn nhãn (tổng cộng hơn 240.000 nhãn lợn). Các video thu thập trong điều kiện tự nhiên của trang trại nhưng vẫn cho kết quả tiềm năng với độ chính xác nhận diện, theo dõi, và phát hiện bất thường trên 93%.

Bài báo có cấu trúc như sau. Phần II trình bày về các nghiên cứu liên quan trong lĩnh vực tự động phát hiện, theo dõi, và phát hiện bất thường hành vi lợn. Phần III mô tả phương pháp. Phần IV trình bày về các kết quả và thảo luận. Cuối cùng, các kết luận sẽ được trình bày trong phần V của bài báo.

II. TỔNG QUAN

Tự động theo dõi hành vi lợn là một hướng nghiên cứu đã nhận được nhiều quan tâm trong các năm gần đây. Thời gian đầu, các nghiên cứu được thực hiện để theo dõi và nhận diện hành vi theo nhóm lợn (group-level). Mặc dù nhận diện hành vi theo nhóm lợn đã chứng minh tác dụng trong cảnh báo sức khỏe đàn, nhưng việc theo dõi và nhận diện hành vi theo cá thể (individual-level) đem lại nhiều ưu điểm khi có thể đánh giá được tình trạng của từng cá thể và có phương án điều trị riêng khi cần thiết [1]. Theo dõi tới từng cá thể là một tác vụ khó khăn hơn, tuy nhiên với những cải tiến về kỹ thuật phát hiện và theo dõi vật thể trong thời gian gần đây, các nghiên cứu về theo dõi và phát hiện lợn hiện tập trung chủ yếu vào hướng thực hiện theo cá thể.

Theo dõi lợn tự động từ camera là hướng nghiên cứu được quan tâm rộng rãi gần đây, do giá thành thực hiện thấp và dễ dàng lắp đặt, triển khai so với các phương pháp giám sát bằng thiết bị đeo (thể tai, thiết bị đeo chân lợn). Một số nghiên cứu sử dụng các cảm biến có chiều sâu như camera 3D để đo đặc chiều sâu đối tượng trong ảnh và phát hiện ra hành vi lợn là đang đứng hay nằm [2-5]. Tuy nhiên, việc sử dụng các thiết bị có khả năng đo chiều sâu này làm giá thành tăng lên, đồng thời không áp dụng để nhận diện các hành vi đa dạng khác như ăn, uống, đi, chạy, v.v. Các nghiên cứu gần đây tập trung vào sử dụng kỹ thuật học sâu để phát hiện, theo dõi và nhận diện đa dạng các hành vi lợn. Dựa trên kiến trúc mạng nơ ron tích chập (CNN – Convolutional Neural Network), các thuật toán phát hiện đa đối tượng đã được phát triển và sử dụng rộng rãi như

Faster R-CNN [6], SSD [7], YOLO [8]. Các thuật toán này có thể xác định vị trí và phân loại các đối tượng phát hiện được trong ảnh, nhờ đó có thể áp dụng để xác định vị trí và phân loại tư thế lộn trong các khung hình video của camera giám sát. Phương pháp này có thể được áp dụng trên các camera 2D thông thường và có thể phát hiện các hành vi đa dạng như đứng, nằm, ăn, uống, đi lại, v.v. [4, 5, 9-14] hoặc phân biệt giữa hành vi di chuyển hay đứng yên [1]. Các hành vi tĩnh (không di chuyển) có thể được nhận diện trong giai đoạn phát hiện lộn bằng cách phân chia lớp đối tượng trong mô hình phát hiện lộn thành các lớp con như lộn đứng, lộn nằm, lộn ăn v.v. [1, 10, 11] hoặc sử dụng một mô hình phân loại độc lập bổ sung để phân loại sau khi đã phát hiện ra đối tượng lộn [13]. Các hành vi dịch chuyển như đi hoặc chạy có thể được nhận diện trong quá trình theo dõi bằng cách đo khoảng cách dịch chuyển của cá thể lộn trong các khung hình liên tiếp từ video [1, 13]. Khi các hành vi được nhận diện trong quá trình theo dõi, sự thay đổi hoặc bất thường trong hành vi có thể được phát hiện thông qua phân tích thời lượng của các hành vi [10, 13] hoặc thông qua phân tích tổng thời gian dành cho các hành vi di chuyển hoặc đứng yên và khoảng cách dịch chuyển [11].

Một trong các vấn đề khó khăn của theo dõi lộn tự động là các lỗi định danh gây ra khi cá thể lộn bị đảo hoặc thay đổi định danh trong quá trình theo dõi đa đối tượng. Một số nghiên cứu trước đây đã cố gắng cải tiến vấn đề này thông qua các phương pháp như bổ sung thêm các bước tái khớp đối tượng mới xuất hiện và bị mất dấu [15, 16]. Mặc dù vậy, với điều kiện của các trang trại hiện nay như mật độ lộn cao, thiếu ánh sáng, góc nhìn camera không tối ưu, đồng thời do đặc điểm giống nhau về hình dáng lộn, việc theo dõi một cá thể lộn trong một thời gian lâu dài với tỷ lệ lỗi nhằm định danh thấp vẫn là một tác vụ khó khăn.

Trong nghiên cứu này, chúng tôi đề xuất một phương pháp phát hiện bất thường hành vi lộn theo từng giai đoạn hoạt động trong ngày. Phương pháp tổng thể bao gồm các giai đoạn phát hiện, theo dõi, nhận diện hành vi và phát hiện bất thường hành vi lộn. Để giảm thiểu tác động của lỗi nhằm định danh trong quá trình theo dõi lên kết quả tổng thể, nghiên cứu đề xuất phương pháp phát hiện bất thường trong khoảng thời gian 30 phút. Ngoài ra, để đảm bảo khả năng phát hiện bất thường hành vi trong thời gian ngắn, nghiên cứu cũng đề xuất việc áp dụng phát hiện bất thường theo giai đoạn hoạt động. Theo nghiên cứu của chúng tôi, lộn thường có các hành vi điển hình trong các giai đoạn hoạt động khác nhau trong ngày như giai đoạn ăn, nghỉ, chơi và việc phát hiện bất thường hành vi theo từng giai đoạn có thể được thực hiện với thời gian ngắn mà không cần phải theo dõi quá lâu. Để tìm hiểu của chúng tôi, chưa có nghiên cứu nào trước đây áp dụng cách tiếp cận tương tự. Phần tiếp theo sẽ trình bày chi tiết hơn về phương pháp được thực hiện trong nghiên cứu.

III. PHƯƠNG PHÁP

Để có thể phát hiện bất thường trong hành vi của từng cá thể lộn, đầu tiên các cá thể phải được nhận diện trong các khung hình video. Các cá thể lộn được phát hiện sau đó được sử dụng để theo dõi và nhận diện hành vi. Trong phương pháp đề xuất, các cá thể lộn khi được nhận diện sẽ được nhận kèm các hành vi tĩnh (không chuyển động) như đứng, nằm, ăn. Để xác định các hành vi liên quan đến chuyển động, cần phải nhận diện dựa trên vị trí lộn trong các khung hình liên tục trong video. Các hành vi (bao gồm cả tĩnh và động) sau khi được nhận diện trong các khung hình sẽ được phân tích để xác định các bất thường trong

hành vi. Chi tiết về các thuật toán được áp dụng trong các giai đoạn này được trình bày ở các phần tiếp theo.

A. Phát hiện lộn trong ảnh

Trong giai đoạn này, thuật toán YOLO v7 [17] được sử dụng để phát hiện các vùng ảnh có cá thể lộn trong một ảnh toàn cảnh thu được từ camera. YOLO là một lớp thuật toán phát hiện đa vật thể trong ảnh được sử dụng rất rộng rãi hiện nay, cho kết quả tốt cả về khía cạnh độ chính xác và tốc độ thực hiện. Để ứng dụng YOLO v7 trong nghiên cứu này, đầu tiên chúng tôi tiến hành hoạt động tiền huấn luyện (pre-train) mạng này với tập dữ liệu chung là COCO [20], sau đó tiến hành hoạt động học chuyển giao (transfer-learning) với tập dữ liệu riêng. Hoạt động tiền huấn luyện là một hoạt động cần thiết với hầu hết các mạng nơ ron tích chập (CNN) khi sử dụng trong bài toán phân loại hoặc phát hiện vật thể, nhằm tạo ra các trọng số khởi tạo tốt cho các mạng nơ ron dựa trên các tập dữ liệu cực kỳ lớn với hàng triệu ảnh.

Trong giai đoạn tiếp theo, mô hình sẽ được huấn luyện tiếp trên tập dữ liệu riêng. Mục tiêu của mô hình phát hiện lộn là không chỉ nhận diện cá thể lộn cùng vị trí của chúng mà còn nhận diện kèm các hành vi tĩnh như đứng, nằm, ăn.

B. Theo dõi lộn trong video

Từ các cá thể lộn phát hiện được trong các khung hình và vị trí của chúng, chúng tôi sử dụng DeepSORT [18], một thuật toán theo dõi đa đối tượng sử dụng phổ biến hiện nay, để theo dõi lộn trong video. DeepSORT là một bản cải tiến của thuật toán SORT, vốn sử dụng bộ lọc Kalman [21] cùng với thuật toán Hungarian [22] để khớp tập các đối tượng phát hiện được giữa các video frame, từ đó hình thành quá trình theo dõi từng cá thể lộn. Cải tiến chính của DeepSORT nằm ở chỗ nó sử dụng độ tương đồng hình ảnh học được từ một mạng học sâu để xác định hai ảnh cá thể lộn có phải là từ một cá thể lộn hay không. Cải tiến này làm tăng khả năng xác định đúng một đối tượng phát hiện ở video frame kế tiếp là đối tượng nào đã được phát hiện ở video frame trước hay là một đối tượng mới được phát hiện. Điều này sẽ giúp làm giảm các tình huống nhận diện sai đối tượng khi xảy ra trường hợp bị mất dấu (do khuất tầm camera hoặc bị lỗi trong quá trình phát hiện). Để xây dựng độ đo tương đồng như nói ở trên, chúng tôi sử dụng mạng OSNet (Omni-Scale Network) [23], một mạng được tiền huấn luyện trên tập dữ liệu tái nhận diện người và sau đó được huấn luyện bổ sung trên tập tái nhận diện lộn do chúng tôi tự xây dựng.

Thuật toán DeepSORT nguyên bản được phát triển cho bối cảnh nhận diện đa đối tượng tổng quát, trong đó số lượng các đối tượng là không biết trước và có thể biến động (ví dụ số lượng người đi bộ trong video có thể thay đổi khi có người đi ra khỏi khung hình hoặc người mới đi vào khung hình). Khi ứng dụng trên tác vụ theo dõi lộn trong trang trại với bối cảnh số lượng đối tượng theo dõi là các cá thể lộn trong một chuồng có độ ổn định hơn, thuật toán DeepSORT có thể gán các mã ID khác nhau cho cùng một đối tượng là một cá thể lộn được theo dõi khi các khung hình trong video tăng lên. Điều này sẽ gây ra lỗi nhằm định danh và làm tăng mã ID vượt quá số lượng lộn thực tế, dẫn đến việc trích xuất hành vi của các cá thể lộn bị giảm độ chính xác. Để giải quyết vấn đề này, chúng tôi sử dụng thuật toán DeepSORT cải tiến được đề xuất bởi S. Tu et al. [16], trong đó bước tái khớp đối tượng (re-matching) được bổ sung cho các đối tượng bị mất dấu và các đối tượng mới tạo ra sử dụng cả dữ liệu quỹ đạo và độ đo tương đồng hình

ảnh. Bước bổ sung này đã giúp làm tăng đáng kể độ chính xác theo dõi và giảm các lỗi nhầm định danh như mô tả trong phần kết quả của hoạt động theo dõi lợn.

C. Phát hiện bất thường trong hành vi lợn

Trong giai đoạn này, các hành vi của lợn sẽ được nhận diện và phân tích, nhằm phát hiện các bất thường trong hành vi. Từ kết quả của các giai đoạn phát hiện và theo dõi lợn, các hành vi của các thể lợn này có thể được nhận diện trong từng video frame hoặc trên một chuỗi các video frame liên tiếp. Trong trường hợp tư thế của cá thể lợn được nhận diện là nằm hoặc ăn, hành vi được xác định trong video frame tương ứng với tư thế đó. Tuy nhiên, nếu tư thế được nhận diện là đứng, thì cần tiếp tục xác định trong các video frame tiếp theo xem hành vi là đứng hay dịch chuyển. Điều này được xác định dựa trên vị trí tương đối của cá thể lợn trong các video frame liên tiếp. Nếu vị trí của cá thể lợn không thay đổi thì hành vi được xác định là đứng, ngược lại hành vi được xác định là di chuyển.

Dựa trên cách thức này, các hành vi của từng cá thể lợn sẽ được xác định trong quá trình phát hiện và theo dõi. Thời gian của mỗi hành vi sẽ được tính toán dựa trên hành vi nhận diện được trên mỗi khung hình và được phân tích để xác định có sự bất thường trong hành vi của lợn trong một khoảng thời gian nhất định hay không. Trong nghiên cứu này, chúng tôi thực nghiệm phân tích phát hiện ba loại bất thường hành vi, bao gồm:

- Lợn nằm lâu không hoạt động
- Lợn bỏ ăn
- Lợn tăng vận động

Để giảm thời gian theo dõi nhưng vẫn đảm bảo độ chính xác phát hiện bất thường, chúng tôi đưa ra khái niệm về “giai đoạn hoạt động” của lợn trong ngày. Một giai đoạn hoạt động của lợn chỉ khoảng thời gian lợn chủ yếu thực hiện một số hành vi điển hình. Trong nghiên cứu này, chúng tôi đưa ra ba giai đoạn hoạt động của lợn là giai đoạn nghỉ, giai đoạn ăn, giai đoạn chơi như trong bảng số 1.

Các giai đoạn và thời gian của mỗi giai đoạn được đưa ra dựa trên tham khảo từ các trang trại trong thực tế (theo thời gian biểu và quan sát của các trang trại). Lưu ý rằng, thời gian biểu của mỗi trang có thể thay đổi và trong các hệ thống triển khai thương mại, các chủ trang trại có thể thiết lập khung giờ hoạt động cho trang trại hoặc cho từng chuồng riêng. Hệ thống sẽ tiến hành thu thập dữ liệu và thực hiện phân tích dựa trên các thiết lập này.

Bảng 1. Các giai đoạn hoạt động của lợn trong ngày

Giai đoạn	Thời gian	Mô tả
Giai đoạn nghỉ	0h-6h, 8h-10h, 14h-16h, 18h-24h	Lợn chủ yếu dành thời gian cho các hoạt động nghỉ, ngủ
Giai đoạn ăn	7h-8h, 13h-14h, 17h-18h	Lợn chủ yếu dành thời gian để ăn
Giai đoạn chơi	6h-7h, 16-17h	Lợn chủ yếu dành thời gian để đi lại, tìm thức ăn

Với việc chia thời gian trong ngày của lợn ra thành các giai đoạn hoạt động, việc phát hiện các bất thường hành vi sẽ được thực hiện theo các quy tắc trên các khung giờ này

để đảm bảo phát hiện ra sự bất thường trên các hành vi điển hình trong mỗi giai đoạn. Nhờ việc phát hiện bất thường hành vi theo hành vi điển hình trong mỗi giai đoạn, thời gian theo dõi phát hiện bất thường có thể được rút ngắn thay vì phải theo dõi trong một thời gian dài như nửa ngày hoặc một ngày. Trong nghiên cứu này, chúng tôi lựa chọn thời gian theo dõi phát hiện bất thường là 30 phút. Khoảng thời gian theo dõi phát hiện bất thường 30 phút vừa đảm bảo giảm thiểu các lỗi định danh khi theo dõi lợn (càng tăng khi thời gian theo dõi càng dài), nhưng vẫn đảm bảo kết quả theo dõi chính xác do việc phát hiện được thực hiện dựa trên phân tích hành vi điển hình trong mỗi giai đoạn hoạt động. Cụ thể, việc phát hiện các bất thường dựa trên phân tích hành vi điển hình trong các giai đoạn hoạt động được thực hiện theo quy tắc sau:

- Nếu thời gian lợn nằm vượt ngưỡng trong giai đoạn ăn hoặc giai đoạn chơi, lợn có thể được coi là có bất thường hành vi nằm quá lâu.
- Nếu thời gian lợn ăn thấp hơn ngưỡng trong giai đoạn ăn, lợn có thể được coi là bất thường hành vi bỏ ăn.
- Nếu thời gian lợn di chuyển vượt ngưỡng trong giai đoạn nghỉ hoặc ăn, lợn có thể được coi là có bất thường hành vi tăng vận động.

IV. THỰC NGHIỆM VÀ KẾT QUẢ

A. Dữ liệu và môi trường thực nghiệm

Các tập dữ liệu thực nghiệm được sử dụng trong nghiên cứu này được thu thập từ các chuồng lợn tại sáu trang trại lợn trong thực tế. Trong tập dữ liệu này, một số chuồng chứa các cá thể lợn khỏe và một số chuồng chứa các cá thể lợn ốm (được gom từ các chuồng khác nhằm mục đích cách ly). Các videos từ các chuồng đều được sử dụng để tạo các tập dữ liệu huấn luyện cho mô hình phát hiện và theo dõi lợn, cũng như để thực nghiệm mô hình phát hiện bất thường hành vi lợn.

Các video được thu thập sử dụng các camera HIKVISION (resolution: 2560 x 1440, focal length: 4.0mm, max frame rate: 25 FPS) và EZVIZ (resolution: 1920 x 1080, focal length: 4.0mm, max frame rate: 15 FPS), trên các chuồng chứa 15-18 cá thể lợn với tuổi từ 3 đến 4 tháng và có trọng lượng từ 40 đến 50 kg. Các video được chọn để làm tập dữ liệu có độ phân giải 1280 x 720, được chọn trong các khung giờ khác nhau để đảm bảo độ khách quan của tập dữ liệu. Từ các video này, các video frame được trích ra để xây dựng các tập dữ liệu phục vụ phát hiện lợn, theo dõi lợn, tái nhận diện lợn, và phát hiện bất thường hành vi lợn.

1) Tập dữ liệu phát hiện lợn

Để tạo tập dữ liệu phát hiện lợn, các ảnh sẽ được trích xuất từ các video thu thập được và được gán nhãn thủ công bằng công cụ LabelImg []. Khi sử dụng công cụ này để gán nhãn, mỗi cá thể lợn sẽ được tạo một đường viền xung quanh và gán cho nó một trong ba nhãn là đứng, nằm, hoặc ăn (để gán nhãn ăn, đường viền phải bao không những cá thể lợn đang ăn mà bao cả máng ăn của lợn để mô hình có thể học được tư thế lợn đang ăn thức ăn trong máng hay không). Khi được gán nhãn, công cụ này sẽ tạo một tập dữ liệu chú giải lưu trữ vị trí của hình viền bao gồm tọa độ tâm và hai góc trên trái dưới phải của hình chữ nhật, cùng với nhãn của đối tượng được gán. Tổng cộng có 15.469 ảnh được trích xuất từ các videos và gán nhãn cho mục đích

huấn luyện và kiểm thử mô hình phát hiện lợn (248.532 cá thể lợn được gán nhãn trong các ảnh).



Hình 1. Một số ảnh mẫu trong tập dữ liệu

2) Tập dữ liệu theo dõi lợn

Để tạo tập dữ liệu phát hiện lợn, chỉ cần tạo viền và gán nhãn phân loại cho các cá thể lợn trong ảnh như nói ở trên. Tuy nhiên, để tạo tập dữ liệu theo dõi lợn, cần phải gán cho mỗi cá thể lợn một mã định danh giống nhau trên các khung hình. Các khung hình in tập dữ liệu theo dõi lợn được lựa chọn theo thứ tự chúng xuất hiện trong video nhưng không cần phải liên tục do tần suất khung hình (frame rate) của video rất nhanh, nên việc lấy và gán nhãn các khung hình liên tục là không cần thiết (vì trong hai khung hình liên tiếp, các cá thể lợn ở vị trí rất sát nhau, khó phân biệt bằng mắt thường). Trung bình, các khung hình được lựa chọn và gán nhãn bằng một phần tư tổng số khung hình trong mỗi video được chọn làm tập dữ liệu theo dõi lợn. Mô hình theo dõi lợn được xây dựng như mô tả ở phần III.B không cần phải huấn luyện cho bước theo dõi (chỉ huấn luyện cho bước tái nhận diện bằng tập dữ liệu tái nhận diện sẽ được mô tả ở bên dưới). Các tập dữ liệu theo dõi lợn được tạo ra như mô tả bên trên được sử dụng để đánh giá mô hình. Để đánh giá mô hình trên cả chuồng lợn khỏe và lợn ốm, chúng tôi đã tạo hai tập dữ liệu trên các video từ một chuồng lợn khỏe và một chuồng lợn ốm. Như đã nói ở trên, các khung hình trong một tập dữ liệu theo dõi cần phải được đặt theo thứ tự xuất hiện, do vậy chúng có thể được coi như một chuỗi khung hình (frame sequences). Bảng 2 mô tả về các tập dữ liệu theo dõi lợn được sử dụng để đánh giá mô hình theo dõi lợn trong nghiên cứu.

Bảng 2. Mô tả về tập dữ liệu theo dõi lợn

Tập dữ liệu	Độ dài (giây)	Số khung hình
Chuồng lợn khỏe	300	247
Chuồng lợn ốm	300	250

3) Tập dữ liệu tái nhận diện lợn

Tập dữ liệu tái nhận diện lợn chứa các ảnh cá thể lợn với mục đích nhận diện lợn không phân biệt hành vi. Do vậy, tập dữ liệu này được xây dựng từ các ảnh lợn được trích ra từ các khung hình trong video. Tổng cộng 50.460 ảnh được thu thập cho 90 cá thể lợn được sử dụng cho tập dữ liệu tái nhận diện lợn. Các ảnh có kích thước là 256 x 128 và được dùng để huấn luyện mạng học sâu tái nhận diện lợn như mô tả ở phần III.

4) Tập dữ liệu phát hiện bất thường hành vi lợn

Tập dữ liệu phát hiện bất thường hành vi lợn bao gồm sáu videos, mỗi video dài 30 phút được thu thập từ các chuồng lợn khỏe và lợn ốm, và trong ba giai đoạn hoạt động như đã nói ở trên. Các tập dữ liệu này nhằm mục đích để đánh giá mô hình phát hiện bất thường theo giai đoạn hoạt động. Mỗi video dùng để đo đạc, phân tích và phát hiện bất thường hành vi của lợn (khỏe hoặc ốm) trong một giai đoạn hoạt động (ăn, nghỉ, hoặc chơi).

B. Độ đo đánh giá

Để đánh giá mô hình nhận diện, đầu tiên độ chính xác vị trí được tính toán qua chỉ số IoU (Intersection of Union) của 2 hình chữ nhật viền của đối tượng được phát hiện và đối tượng thực tế theo công thức 1. Chỉ số IoU càng cao chứng tỏ vị trí được phát hiện càng chính xác. Một ngưỡng sẽ được đặt ra để cho biết mức độ khớp của 2 hình là có đạt không.

$$IoU = \frac{\text{Diện tích giao của 2 hình}}{\text{Diện tích hợp của 2 hình}} \quad (1)$$

Từ chỉ số IoU của mỗi cặp cá thể dự đoán được và thực tế, độ đo AP (Average Precision) sẽ được tính cho mỗi lớp đối tượng và cuối cùng độ đo mAP là trung bình của các AP trên mỗi lớp được sử dụng là độ đo đánh giá cuối cùng theo các công thức 2 và 3, trong đó $TP(c)$ và $FP(c)$ lần lượt là chỉ số dự đoán đúng dương và sai dương (True Positive và False Positive) còn $AP(c)$ là chỉ số AP của lớp c.

$$AP(c) = \frac{TP(c)}{TP(c)+FP(c)} \quad (2)$$

$$mAP = \frac{1}{\text{classes}} \sum_{\text{classes}} AP(c) \quad (3)$$

Độ đo đánh giá cho mô hình theo dõi lợn được sử dụng trong nghiên cứu là độ đo MOTA (Multi-Object Tracking Accuracy). Đây là độ đo được sử dụng phổ biến trong các nghiên cứu về theo dõi đa đối tượng trước đây, điển hình là trong các cuộc thi MOT Challenge. Độ đo này sử dụng công thức dựa trên 3 loại lỗi có thể xảy ra trong theo dõi đa đối tượng là số sai âm (FN - False Negative), số sai dương (FP - False Positive) và số cá thể bị nhầm (IDSW - Identity Switch) như công thức số 4.

$$MOTA = 1 - \frac{\sum_i FN_i + FP_i + IDSW_i}{\sum_i GT_i} \quad (4)$$

Trong công thức trên, FN_i là cá thể lợn có thực nhưng không được theo dõi, FP_i là cá thể lợn không có thực nhưng lại được theo dõi, $IDSW_i$ là số cặp cá thể bị nhầm định danh, và GT_i là cá thể lợn thực tế.

Bên cạnh MOTA, độ đo F1 cũng được sử dụng để đánh giá mô hình theo dõi lợn. Độ đo này tập trung vào đánh giá vấn đề nhầm định danh hơn so với độ đo MOTA. Độ đo

này được tính như trong công thức (5), trong đó $IDTP$, $IDFP$, $IDFN$ là các chỉ số True Positive, False Positive, False Negative trên các lỗi nhầm định danh.

$$IDF_1 = \frac{2IDTP}{2IDTP+IDFP+IDFN} \quad (5)$$

Để đánh giá mô hình phát hiện bất thường dựa trên phân tích hành vi, độ đo chính xác thông thường (Accuracy) được sử dụng như trong công thức (6).

$$Accuracy = \frac{Số\ trường\ hợp\ phát\ hiện\ đúng}{Tổng\ số\ trường\ hợp} \quad (6)$$

C. Kết quả thực nghiệm

Bảng 3 trình bày kết quả của mô hình phát hiện lộn. Lưu ý rằng bên cạnh độ đo mAP như đã trình bày ở phần trên, chúng tôi báo cáo thêm kết quả nhận diện theo các độ đo Precision (P) và Recall (R).

Bảng 3. Kết quả của mô hình phát hiện lộn

Lớp	mAP	Precision (%)	Recall (%)
Đúng	99.3	98.3	98.5
Nằm	99.6	98.6	99.1
Ăn	98.9	97.0	97.0
Trung bình	99.3	97.9	98.2

Từ bảng 3, có thể thấy mô hình phát hiện lộn có kết quả rất tốt ở tất cả các độ đo, xét trên điều kiện của các trang trại thương mại tại Việt Nam (được thu thập làm dữ liệu thực nghiệm của nghiên cứu). Trong các trang trại thương mại hiện nay, trần của các chuồng lợn được thiết kế thấp nhằm mục đích tiết kiệm do vậy các camera không có được góc nhìn tốt và các cá thể lợn dễ bị che khuất khi chúng đứng sát nhau hoặc đi ra xa vị trí camera. Độ chính xác phát hiện lộn có thể tăng lên nếu các trang trại được thiết kế phù hợp hơn cho hệ thống theo dõi tự động.

Kết quả của giai đoạn theo dõi lộn được biểu thị trong bảng 4. Ngoài các độ đo MOTA và IDF1 như đã trình bày ở trên, chúng tôi cũng báo cáo kết quả theo độ đo identity switches (IDS). Độ đo này tính toán số lần các cá thể được theo dõi bị nhầm lẫn với cá thể khác. Đây cũng là độ đo quan trọng, thể hiện số lỗi nhầm định danh trong quá trình theo dõi lộn.

Bảng 4. Kết quả của mô hình theo dõi lộn

	DeepSORT			DeepSORT cải tiến		
	MOTA (%)	IDF1 (%)	IDS	MOTA (%)	IDF1 (%)	IDS
Chuồng khỏe	91.5	93.4	13	92.8	95.6	10
Chuồng ốm	92.5	94.0	10	94.3	96.1	8
Trung bình	92.0	93.7	11.5	93.6	95.9	9

Các kết quả trung bình ở độ đo MOTA (93.6%) và IDF1 (95.9%) là các kết quả tốt. Tập dữ liệu thu thập ở chuồng lộn ốm có kết quả cao hơn do trong chuồng lộn ốm, các cá thể lợn ít di chuyển gây xáo trộn hơn so với tập dữ liệu ở chuồng lộn khỏe.

Chỉ số IDS trung bình là 9, có nghĩa là mỗi cá thể lợn bị thay đổi định danh 0.4 lần trong suốt quá trình theo dõi. Các kết quả của thuật toán DeepSORT cải tiến cũng cao hơn so với thuật toán DeepSORT gốc đặc biệt là chỉ số IDF1 và IDS cho thấy hiệu quả của thuật toán này trong việc giải quyết vấn đề giảm lỗi nhầm định danh.

Bảng 5 cho thấy kết quả của mô hình phát hiện bất thường hành vi lợn với ngưỡng là 15/30 phút. Do các vấn đề về quản lý không cho phép chúng tôi thu thập được các video tại các chuồng có cả lợn khỏe và lợn ốm, chúng tôi tiến hành thực nghiệm phát hiện bất thường trên các video tại các chuồng chỉ có lợn khỏe hoặc chỉ có lợn ốm, như đã trình bày ở phần trước. Quá trình phát hiện bất thường được xem là thành công nếu không phát hiện bất thường trong video lợn khỏe, nhưng phát hiện ra bất thường trong video lợn ốm. Theo kết quả trong bảng 5, ba video lợn khỏe trong ba giai đoạn đều chỉ cảnh báo bất thường từ 1 đến 2 cá thể trên tổng số 16 cá thể. Tỷ lệ chính xác là 87.5% đến 93.8%. Hai video lợn ốm trong giai đoạn ăn và chơi có tỷ lệ phát hiện bất thường là 14 và 13 trên tổng số 15, đạt độ chính xác 86.7% đến 93.3%. Video lợn ốm trong giai đoạn nghỉ không phát hiện bất thường tăng vận động do trong tập dữ liệu video lợn ốm sử dụng trong nghiên cứu, không có cá thể nào bị nhiễm loại bệnh có tính tăng vận động (hầu hết lợn ốm đều mệt và nằm nghỉ trong giai đoạn này). Do vậy, thực nghiệm phát hiện bất thường ở giai đoạn này có kết quả cao do không có mẫu lợn bệnh đúng. Mặc dù vậy, với độ chính xác của mô hình theo dõi và nhận diện hành vi, việc tính toán thời gian vận động của lợn cũng sẽ chính xác và cho kết quả tốt khi có mẫu lợn bệnh có tính chất tăng vận động.

Bảng 5. Kết quả của mô hình phát hiện bất thường hành vi lợn

Video	Số lợn	Loại chuồng	Giai đoạn hoạt động	Loại bất thường	Số lợn phát hiện bất thường	Tỷ lệ đúng
1	16	Khỏe	Giai đoạn ăn	Bỏ ăn, Nằm lâu	1	93.8%
2	16	Khỏe	Giai đoạn nghỉ	Tăng vận động	2	87.5%
3	16	Khỏe	Giai đoạn chơi	Nằm lâu	1	93.8%
4	15	Ốm	Giai đoạn ăn	Bỏ ăn, Nằm lâu	14	93.3%
5	15	Ốm	Giai đoạn nghỉ	Tăng vận động	0	100%
6	15	Ốm	Giai đoạn chơi	Nằm lâu	13	86.7%

V. KẾT LUẬN

Trong nghiên cứu này, chúng tôi đã trình bày phương pháp phát hiện, theo dõi, nhận diện và phát hiện bất thường hành vi

lợn trong điều kiện tự nhiên tại các trang trại sử dụng học sâu. Nghiên cứu đề xuất phương pháp phát hiện bất thường theo giai đoạn hoạt động, qua đó giảm thời gian theo dõi nhưng vẫn đảm bảo độ chính xác trong phát hiện bất thường. Các thực nghiệm đã được tiến hành trên các tập dữ liệu tự thu thập từ các trang trại lợn thương mại, với các kịch bản thực nghiệm khác nhau cho các kết quả khả quan. Điều này cho thấy tiềm năng của việc ứng dụng phương pháp trong thực tế, đặc biệt là chỉ cần sử dụng các thiết bị thông thường và rẻ tiền như camera giám sát.

Nghiên cứu có thể tiếp tục được cải tiến với các hướng tiềm năng như tự động xác định thời gian cho các giai đoạn hoạt động cho lợn hay xác định các tình huống bất thường phức tạp hơn nhằm có các đánh giá tốt hơn về tình hình sức khỏe lợn.

LỜI CẢM ƠN

Nghiên cứu này được hỗ trợ bởi Viện Hàn lâm Khoa học và Công nghệ Việt Nam trong đề tài mã số CN4000.01/21-23.

TÀI LIỆU THAM KHẢO

- [1] J. Cowton, I. Kyriazakis, and J. Bacardit, "Automated Individual Pig Localisation, Tracking and Behaviour Metric Extraction Using Deep Learning," IEEE Access, vol. 7, pp. 108049-108060, 2019, doi: 10.1109/ACCESS.2019.2933060.
- [2] M. Mittek, E. Psota, L. Pérez, T. Schmidt, and B. Mote, "Health monitoring of group-housed pigs using depth-enabled multi-object tracking," in Proc. Int Conf. Pattern Recognit., Workshop Vis. Observ. Anal. Vertebrate Insect Behav., 2016.
- [3] J. Sa, Y. Choi, H. Lee, Y. Chung, D. Park, and J. Cho, "Fast pig detection with a top-view camera under various illumination conditions," Symmetry, vol. 11, no. 2, p. 266, 2019.
- [4] J. Kim et al., "Depth-Based Detection of Standing-Pigs in Moving Noise Environments," Sensors 17, 2757, 2017, doi: 10.3390/s17122757
- [5] J. Kim et al., "Lying-Pig Detection using Depth Information," in Proc. ICACS '17, pp. 40-43, 2017, doi: 10.1145/3127942.3127949
- [6] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, "You only look once: Unified, real-time object detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., pp. 779-788, 2016.
- [7] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," in Proc. IEEE Trans. Pattern Anal. 2017, 39, 1137-1149
- [8] W. Liu et al., "Ssd: Single shot multibox detector," Lect. Notes Comput. Sci. 2016, 9905, 21-37.
- [9] Q. Yang, D. Xiao, and S. Lin, "Feeding behavior recognition for group-housed pigs with the faster r-cnn," Comput. Electron. Agric. 155, 453-460, 2018.
- [10] A. Alameer, I. Kyriazakis, and J. Bacardit, "Automated recognition of postures and drinking behaviour for the detection of compromised health in pigs," Sci Rep 10, 13665, 2020, doi: 10.1038/s41598-020-70688-6
- [11] A. Alameer, I. Kyriazakis, H. Dalton, A. Miller, and J. Bacardit, "Automatic recognition of feeding and foraging behaviour in pigs using deep learning," Biosyst. Eng. 197, 91-104, 2020.
- [12] M. Riekert, A. Klein, F. Adrion, C. Hoffmann, and E.

Gallmann, "Automatically detecting pig position and posture by 2d camera imaging and deep learning," Comput. Electron. Agric. 174, 105391, 2020.

- [13] L. Bergamini et al., "Extracting accurate long-term behavior changes from a large pig dataset," in Proc. 16th VISIGRAPP, Vol. 4, pp. 524-533, 2021, doi: 10.5220/0010288405240533.
- [14] S. Matthews, A. Miller, T. Plötz, and I. Kyriazakis, "Automated tracking to measure behavioural changes in pigs for health and welfare monitoring," Sci. Rep. 7, 17582, 2017.
- [15] L. Zhang, H. Gray, X. Ye, L. Collins, and N. Allinson, "Automatic individual pig detection and tracking in pig farms," Sensors 19, 1188, 2019.
- [16] S. Tu et al., "Automated Behavior Recognition and Tracking of Group-Housed Pigs with an Improved DeepSORT Method," Agriculture, vol. 12, no. 11, p. 1907, Nov. 2022, doi: 10.3390/agriculture12111907.
- [17] C. Wang, A. Bochkovskiy, H. Liao, "YOLOv7: Trainable bag-offreebies sets new state-of-the-art for real-time object detectors," arXiv 2022, arXiv:2207.02696.
- [18] N. Wojke, A. Bewley, and D. Paulus, "Simple Online and Realtime Tracking with a Deep Association Metric," arXiv, 2017, doi: 10.48550/ARXIV.1703.07402.
- [19] Tzutalin, "LabelImg," Git code, 2015. <https://github.com/tzutalin/labelImg>
- [20] G. Bernal, et al., "Coco", 2018.
- [21] R. Kalman, "A new approach to linear filtering and prediction problems," Trans. ASME, D, J. Basic Eng., vol. 82, pp. 35-45, 1960.
- [22] H. Kuhn, "The Hungarian method for the assignment problem," Naval Res. Logistics Quart., vol. 2, nos. 1_2, pp. 83_97, 1955.
- [23] K. Zhou, "Omni-Scale Feature Learning for Person Re-Identification," 2019, <https://arxiv.org/abs/1905.00953>
- [24] A. Milan, L. Leal-Taixé, L. Reid, S. Roth, and K. Schindler, "Mot16: A benchmark for multi-object tracking," arXiv preprint arXiv:1603.00831, 2016.

PIG BEHAVIOR ABNORMALITY DETECTION FROM CAMERA USING DEEP LEARNING

Abstract: Automatically detection and tracking pigs from surveillance cameras in pig farms is an important task in abnormal detection of pig health and welfare. Previous researches showed that pigs with problems in health often have abnormalities in behaviors. Monitoring the pig behaviors can help detect the abnormality timely and take early medical examination and treatment to prevent disease from spreading. In this paper, we presented a method for automatically detection of pig behavior abnormality based on behavior tracking and identification from camera using deep learning. Our approach includes a stages such as pig detection, tracking, behavior identification, and behavior abnormality detection. The behavior abnormalities detected include lying too long, eating less, being excited. The behavior abnormalities are detected according to different activity periods, which helps reduce the time to monitor, while still ensuring the high accuracy. Experiments show that the use of a 30-minute monitoring period ensures a reduction in identification errors when tracking, while still ensuring an abnormality detection accuracy of up to 93.8%.

Keywords: pig behavior abnormality detection, deep learning, pig tracking.



Dương Thành Nam Tốt nghiệp Đại học chuyên ngành Thổ nhưỡng năm 2001 tại Trường Đại học KHTN Hà Nội; bằng Thạc sĩ chuyên ngành Thổ nhưỡng năm 2005 tại trường Đại học Nông nghiệp I Hà Nội, Thạc sĩ ngành Khoa học Môi trường năm 2006 tại trường Đại học KHTN Hà Nội và tiến sĩ ngành Đất và Dinh dưỡng cây trồng năm 2011 tại Đại học Nông nghiệp Hà Nội. Hiện tại, ông là Phó Giám đốc Trung tâm Phát triển và Dịch vụ Đo lường, Trung tâm Nghiên cứu và Chuyển giao Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.



Nguyễn Thị Kim Phương nhận bằng cử nhân chuyên ngành Công nghệ Kỹ thuật Môi trường năm 2020 và đang theo học Thạc sĩ chuyên ngành Khoa học Môi trường tại Trường Đại học Tài nguyên và Môi trường Hà Nội. Hiện tại đang công tác tại Viện Kiểm định Công nghệ và Môi trường.



Trần Thị Hoa nhận bằng cử nhân chuyên ngành Hoá hữu cơ và Sinh hoá hữu cơ năm 2017; bằng thạc sĩ chuyên ngành Hoá phân tích năm 2020 tại Trường Đại học Tổng hợp quốc gia Voronezh, Liên Bang Nga. Hiện đang công tác tại Trung tâm Phát triển và Dịch vụ Đo lường, Trung tâm Nghiên cứu và Chuyển giao Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.



Trần Sơn Tùng nhận bằng cử nhân chuyên ngành Môi trường năm 2010 tại Trường Đại học Thái Nguyên; bằng Thạc sĩ chuyên ngành Công nghệ Môi trường năm 2013 tại trường Đại học Khoa học Tự nhiên Hà Nội. Hiện đang công tác tại Trung tâm Phát triển và Dịch vụ Đo lường, Trung tâm Nghiên cứu và Chuyển giao Công nghệ, Viện Hàn lâm Khoa học và Công nghệ Việt Nam.



Dương Trần Đức Tốt nghiệp Đại học KHTN, Đại học Quốc gia Hà Nội ngành Công nghệ thông tin năm 1999, Thạc sĩ chuyên ngành Hệ thống thông tin tại Đại học Tổng hợp Leeds, Vương Quốc Anh năm 2004, và Tiến sĩ chuyên ngành Kỹ thuật máy tính tại Học viện Công nghệ Bưu chính Viễn thông năm 2018. Hiện đang công tác tại Khoa Công nghệ Thông tin, Học viện Công nghệ Bưu chính Viễn thông.