



ỨNG DỤNG AI CHATBOT TRONG HỖ TRỢ KIỂM TRA VÀ XÁC THỰC BỘ CHỨNG TỪ NGOẠI THƯƠNG TRONG DOANH NGHIỆP VỪA VÀ NHỎ

APPLYING AI CHATBOTS TO SUPPORT THE VERIFICATION AND VALIDATION OF FOREIGN TRADE DOCUMENTS IN SMALL AND MEDIUM-SIZED ENTERPRISES

Nguyễn Đình Vũ Hoàng

Trường Đại học Giao thông vận tải TP. Hồ Chí Minh

DOI: <https://doi.org/10.65934/mkusj.2026.01-e.827>

Email: nguyendinhvuhoang00@gmail.com

Ngày nhận bài: 13/12/2025

Ngày phản biện: 31/12/2025

Ngày duyệt bài: 01/06/2026

TÓM TẮT

Trong bối cảnh thế giới biến động nhanh bởi Cách mạng công nghiệp 4.0, biến đổi khí hậu và bất ổn chính trị, Việt Nam đứng trước yêu cầu đổi mới mô hình tăng trưởng để hướng đến phát triển bền vững. Trong đó, ngoại thương giữ vai trò then chốt khi vừa thúc đẩy kinh tế vừa mở rộng năng lực cạnh tranh quốc gia. Tuy nhiên, quy trình lập và kiểm tra bộ chứng từ ngoại thương vẫn còn phức tạp, phụ thuộc lớn vào thủ công, đặc biệt là ở doanh nghiệp vừa và nhỏ. Bộ chứng từ như Hợp đồng, Hóa đơn, Packing List, Vận đơn hay Giấy chứng nhận xuất xứ không chỉ mang tính pháp lý mà còn là cơ sở để thông quan và thanh toán quốc tế, nhất là dưới phương thức tín dụng chứng từ (L/C). Mọi sai sót, dù nhỏ, đều có thể gây chậm trễ thông quan, phát sinh chi phí hoặc bị ngân hàng từ chối thanh toán, ảnh hưởng trực tiếp đến tiến độ giao hàng và dòng vốn doanh nghiệp. Trước thực trạng đó, trí tuệ nhân tạo (AI) nổi lên như công cụ hiệu quả nhằm giảm lỗi con người và nâng cao tốc độ – độ chính xác trong kiểm tra chứng từ. Vì vậy, nhóm nghiên cứu lựa chọn đề tài “Ứng dụng AI Chatbot trong hỗ trợ kiểm tra – xác thực bộ chứng từ ngoại thương cho doanh nghiệp vừa và nhỏ” nhằm đề xuất giải pháp hỗ trợ nghiệp vụ ngoại thương trong điều kiện chuyển đổi số hiện nay.

Từ khóa: Ngoại thương, Chứng từ xuất nhập khẩu, Trí tuệ nhân tạo (AI), Chatbot kiểm tra chứng từ

ABSTRACT

In the context of rapid global changes driven by the Fourth Industrial Revolution, climate challenges, and political instability, Vietnam faces increasing pressure to reform its growth model toward sustainable development. Foreign trade plays a pivotal role in this process by fostering economic expansion and enhancing national competitiveness. However, the preparation and verification of trade documentation remain complex and heavily manual, especially for small and medium-sized enterprises. Documents such as Contracts, Invoices, Packing Lists, Bills of Lading, and Certificates of Origin are not only legal evidence of transactions but also essential for customs clearance and international payment, particularly under the Letter of Credit (L/C) method. Even minor errors can lead to clearance delays, additional costs, or bank refusal of payment, directly affecting delivery schedules and business cash flow. In response to these challenges, artificial intelligence (AI) has emerged as an effective tool to minimize human error and improve the speed and accuracy of document checking. Therefore, the research team selected the topic “Applying AI Chatbots to Support Verification and Validation of Foreign Trade Documents for Small and Medium-Sized Enterprises” to propose a practical solution for enhancing trade operations in the digital transformation era.

Keywords: Foreign Trade, Import–Export Documentation, Artificial Intelligence (AI), Document-Checking Chatbot

GIỚI THIỆU NGHIÊN CỨU

Việt Nam đang đứng trước yêu cầu cấp thiết phải đổi mới mô hình tăng trưởng nhằm đáp ứng mục tiêu phát triển bền vững và nâng cao năng lực cạnh tranh của nền kinh tế. Trong tiến trình đó, ngoại thương giữ vai trò then chốt khi vừa thúc đẩy tăng trưởng, vừa mở rộng cơ hội

nhập kinh tế quốc tế. Tuy nhiên, hoạt động lập và kiểm tra bộ chứng từ ngoại thương hiện nay vẫn còn nhiều bất cập. Các loại chứng từ như Hợp đồng Thương mại, Hóa đơn, Phiếu đóng gói, Vận đơn hay Giấy chứng nhận xuất xứ đều đòi hỏi độ chính xác tuyệt đối, bởi chúng là căn cứ pháp lý để thông quan hàng hóa và thực hiện

thanh toán quốc tế, đặc biệt trong phương thức tín dụng chứng từ (L/C) (Nguyễn & Nguyễn, 2016).

Việc kiểm tra thủ công, đối chiếu chéo nhiều loại tài liệu khác nhau khiến quy trình này dễ xảy ra sai sót, từ chênh lệch số liệu, mô tả hàng hóa cho đến lỗi điều khoản vận chuyển. Những sai sót nhỏ cũng có thể dẫn đến chậm trễ thông quan, phát sinh chi phí hoặc bị ngân hàng từ chối thanh toán, gây ảnh hưởng trực tiếp đến tiến độ giao hàng và dòng tiền của doanh nghiệp. Tình trạng này đặc biệt nghiêm trọng đối với doanh nghiệp vừa, nhỏ và doanh nghiệp khởi nghiệp - nơi nguồn lực nhân sự hạn chế và quy trình xử lý chứng từ vẫn phụ thuộc hoàn toàn vào con người (Bùi & Phạm, 2005).

Bài báo tập trung phân tích các điểm yếu trong quy trình kiểm tra chứng từ, đề xuất mô hình ứng dụng AI Chatbot và tiến hành thử nghiệm trên bộ chứng từ thực tế để đánh giá hiệu quả. Kết quả nghiên cứu kỳ vọng đóng góp cơ sở khoa học và giải pháp thực tiễn giúp doanh nghiệp Việt Nam giảm thiểu rủi ro, tăng tốc độ xử lý và nâng cao chất lượng quản trị trong hoạt động xuất nhập khẩu.

Nghiên cứu sử dụng kết hợp nhiều phương pháp nhằm bảo đảm tính khoa học và thực tiễn. Nhóm tiến hành tổng hợp và phân tích tài liệu để xây dựng nền tảng lý thuyết về nghiệp vụ ngoại thương, nhận diện các thách thức trong quản lý chứng từ và tham khảo các nghiên cứu chuyên ngành. Đồng thời, phân tích số liệu và đánh giá

tình huống tại một số doanh nghiệp xuất nhập khẩu được thực hiện để phản ánh sát thực nhu cầu của cộng đồng doanh nghiệp Việt Nam.

Bên cạnh đó, nhóm xây dựng mô hình AI Chatbot nhằm minh chứng khả năng hỗ trợ kiểm tra chứng từ trong thực tiễn. Mô hình cho phép đánh giá cả yếu tố định lượng (mức độ trùng khớp dữ liệu, tần suất lỗi) và định tính (tính hợp lý nội dung, khả năng phát hiện bất thường). Mục tiêu cuối cùng là tạo ra công cụ thử nghiệm, qua đó kiểm chứng tính khả thi, độ chính xác và tiềm năng ứng dụng của AI trong quy trình ngoại thương, làm cơ sở đề xuất giải pháp phù hợp cho doanh nghiệp xuất nhập khẩu Việt Nam.

1. Cơ sở lý thuyết

Việc kiểm tra bộ chứng từ là bước quen thuộc trong giao dịch ngoại thương, nhưng phần lớn doanh nghiệp vừa và nhỏ tại Việt Nam và nhiều nước ASEAN vẫn thực hiện thủ công. Nhân viên xuất nhập khẩu phải đối chiếu từng dòng thông tin giữa Sales Contract, Invoice và Packing List, sau đó sửa lỗi và trao đổi lại với đối tác khi phát sinh sai lệch. Một số doanh nghiệp lớn đã áp dụng Excel macro hoặc các hệ thống ERP như SAP, Oracle hay Odoo để hỗ trợ đối chiếu (Nguyễn, Vũ, & Trần, 2025). Tuy nhiên, việc xác nhận cuối cùng vẫn phải do con người thực hiện, do dữ liệu chứng từ được tạo từ nhiều chủ thể khác nhau và thiếu tiêu chuẩn hóa, khiến hệ thống khó tự động kiểm tra hoàn toàn chính xác.

Bảng 1. Quy trình kiểm tra chứng từ thông thường hiện nay

Chứng từ	Nội dung cần đối chiếu	Đối chiếu với
Sales Contract	Điều khoản mua bán, tên hàng, Incoterms, số lượng, giá, thanh toán	So với Invoice, Packing List
Commercial Invoice	Giá, mã hàng, điều kiện thanh toán, người bán - người mua	So với Sales Contract
Packing List	Số kiện, trọng lượng, quy cách đóng gói	So với Invoice & B/L
Bill of Lading (B/L)	Người gửi, người nhận, số container, số lượng hàng	So với Packing List & Invoice
C/O (Certificate of Origin)	Nước xuất xứ, tên hàng, mã HS	So với Invoice & Packing List
Tờ khai hải quan	Mã HS, trị giá tính thuế	So với Invoice & C/O

Trong tự động hóa kiểm tra chứng từ ngoại thương, một số công cụ như Shipamax, KlearStack và ABBYY FlexiCapture đã được triển khai nhằm giảm tải xử lý thủ công và nâng

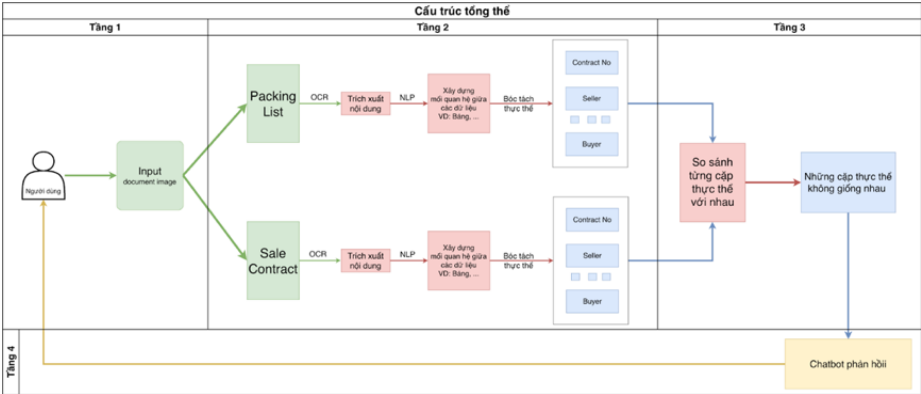
cao độ chính xác. Shipamax tự động trích xuất dữ liệu từ Bill of Lading, Invoice và Packing List và đối chiếu để phát hiện sai lệch. KlearStack ứng dụng OCR và AI để đọc, phân loại và thực

hiện 3-way matching giữa Sales Contract – Invoice – Packing List. ABBYY FlexiCapture sử dụng OCR và nhận diện bố cục để trích xuất dữ liệu có cấu trúc và kiểm tra tính hợp lệ bằng rules engine (Nguyễn, Vũ, & Trần, 2025).

Tuy nhiên, các giải pháp này chưa phổ biến do chi phí triển khai cao, sự đa dạng định dạng chứng từ, hạn chế dữ liệu huấn luyện AI và yêu cầu tuân thủ pháp lý liên quan đến chữ ký số, chuẩn XML hải quan và điều kiện L/C. Ngoài ra, việc tích hợp với ERP, WMS, VNACCS/VCIS, eB/L và các hệ thống khác còn phức tạp. Những hạn chế công nghệ hiện tại bao gồm độ chính xác OCR/NLP chưa tối ưu, rules engine khó xử lý ngoại lệ, phụ thuộc nhân sự kiểm duyệt và yêu cầu cao về bảo mật dữ liệu (Nguyễn, Vũ, & Trần, 2025).

2. Kết quả nghiên cứu và đề xuất giải pháp

Hiện nay, việc rà soát và đối chiếu chứng từ xuất nhập khẩu vẫn chủ yếu làm thủ công, tốn nhiều thời gian và phụ thuộc vào nhân sự có kinh nghiệm. Để khắc phục hạn chế này, nghiên cứu đề xuất ứng dụng Chatbot AI tích hợp OCR, NLP và Data Matching nhằm tự động trích xuất, hiểu và đối chiếu dữ liệu trong bộ chứng từ ngoại thương (Bibhu Dash, 2021). OCR chuyển văn bản dạng hình ảnh thành dữ liệu số; NLP phân tích ngữ nghĩa và phân loại thông tin; còn Data Matching đối chiếu và xác thực dữ liệu giữa các chứng từ, đồng thời phát hiện sai lệch hoặc bất thường. Khi tích hợp trong một nền tảng Chatbot AI, giải pháp giúp giảm tải kiểm tra thủ công, tăng độ chính xác và thúc đẩy tự động hóa trong logistics và ngoại thương.



Hình 1. Kiến trúc tổng thể của hệ thống Chatbot AI kiểm tra chứng từ

Tầng 1 Nhập liệu và xác thực pháp lý ban đầu: Người dùng tải ảnh chụp, bản scan hoặc chứng từ điện tử lên hệ thống. Trước khi đưa vào xử lý OCR, hệ thống thực hiện bước xác thực sơ bộ nhằm đánh giá độ tin cậy của chứng từ đầu vào. Cụ thể, hệ thống kiểm tra nguồn phát hành chứng từ (doanh nghiệp phát hành, hãng tàu, cơ quan cấp C/O, ngân hàng hoặc cơ quan hải quan), kiểm tra sự hiện diện của các dấu hiệu pháp lý như chữ ký, con dấu, mã tham chiếu, số chứng từ, ngày phát hành và định dạng biểu mẫu. Đối với chứng từ điện tử, hệ thống có thể đối

chiếu thêm mã QR, metadata tệp, chữ ký số hoặc thông tin xác thực từ nguồn phát hành nếu có.

Trên cơ sở đó, mỗi chứng từ được gán một mức độ tin cậy đầu vào (input trust score). Những chứng từ thiếu dấu hiệu xác thực, sai định dạng, mờ/nhòe hoặc có dấu hiệu chỉnh sửa sẽ được gán cờ cảnh báo để chuyển sang bước kiểm tra thủ công trước khi đưa vào đối chiếu tự động. Cơ chế này giúp bảo đảm rằng dữ liệu đầu vào đủ độ tin cậy về mặt pháp lý và nguồn gốc trước khi hệ thống thực hiện trích xuất và đối chiếu nội dung.

Bảng 2. Tiêu chí xác thực pháp lý và nguồn gốc chứng từ đầu vào

Nhóm tiêu chí	Nội dung kiểm tra
Nguồn phát hành	Tên đơn vị phát hành, mẫu biểu, mã tham chiếu
Tính toàn vẹn	Có dấu hiệu chỉnh sửa, cắt ghép, thiếu trang

Dấu hiệu pháp lý	Chữ ký, con dấu, chữ ký số, mã QR
Chất lượng ảnh	Độ phân giải, lệch góc, mờ, nhiễu
Tính hợp lệ hình thức	Đúng loại chứng từ, đúng bố cục cơ bản

Tầng 2 Trích xuất và xử lý dữ liệu (OCR + NLP):

Hình ảnh chứng từ được xử lý bằng OCR để trích xuất chữ và chuyển thành dữ liệu có cấu trúc. Sau đó, NLP – đặc biệt là NER – phân tích văn bản, nhận diện các thực thể như số hợp đồng, người bán, người mua, ngày ký, điều khoản giao hàng... giúp chuẩn hóa dữ liệu phục vụ bước đối chiếu tiếp theo (Bibhu Dash, 2021).

Tầng 3 Đối chiếu chéo và kiểm tra nghiệp vụ chuyên sâu:

Ngoài việc so khớp dữ liệu văn bản giữa Invoice, Packing List, Contract, B/L và C/O, hệ thống cần được mở rộng bằng các quy tắc đối chiếu chuyên biệt theo từng nhóm hàng hóa. Với nhóm nông sản, hệ thống ưu tiên kiểm tra các thông tin liên quan đến mùa vụ, chủng loại, quy cách đóng gói, trọng lượng tịnh/trọng lượng cả

bì, xuất xứ vùng trồng và sự thống nhất giữa mô tả hàng hóa với chứng nhận kiểm dịch hoặc chứng nhận xuất xứ. Với nhóm thủy sản, hệ thống cần kiểm tra sâu hơn về điều kiện bảo quản lạnh, nhiệt độ vận chuyển, dạng sản phẩm, mã lô hàng, ngày sản xuất/hạn sử dụng và sự phù hợp với các yêu cầu kiểm dịch, an toàn thực phẩm. Với nhóm hàng công nghiệp, hệ thống tập trung đối chiếu mã hàng, model, thông số kỹ thuật, mã HS, số serial, quy cách kiện hàng và mức độ phù hợp giữa hợp đồng, hóa đơn và packing list.

Việc bổ sung tập quy tắc theo ngành hàng giúp hệ thống không chỉ phát hiện sai lệch bề mặt về văn bản mà còn nhận diện các bất nhất nghiệp vụ mang tính thực tế, từ đó nâng cao độ chính xác và khả năng ứng dụng trong môi trường ngoại thương đa dạng (Bibhu Dash, 2021).

Bảng 3. Quy tắc đối chiếu chéo mở rộng theo nhóm hàng hóa

Nhóm hàng	Trường đối chiếu chính	Quy tắc kiểm tra chuyên sâu
Nông sản	Tên hàng, quy cách, trọng lượng, xuất xứ	Đối chiếu mùa vụ, vùng trồng, quy cách bao bì, số lượng kiện
Thủy sản	Tên hàng, khối lượng, lô hàng, ngày SX/HSD	Kiểm tra nhiệt độ bảo quản, điều kiện vận chuyển, thông tin kiểm dịch
Công nghiệp	Model, mã hàng, số lượng, mã HS	Đối chiếu thông số kỹ thuật, serial, packing spec, consistency giữa Contract–Invoice–PL

Tầng 4: Chatbot phản hồi: Chatbot cung cấp thông báo lỗi, giải thích nguyên nhân sai lệch như sai số hợp đồng, nhầm tên người mua, khác điều khoản Incoterms, đồng thời đề xuất chỉnh sửa để người dùng xử lý nhanh và chính xác.

Việc bổ sung tầng xác thực pháp lý ban đầu là cần thiết vì một hệ thống AI dù có khả năng trích xuất và đối chiếu tốt vẫn có thể cho ra kết quả sai nếu dữ liệu đầu vào không bảo đảm tính xác thực, tính toàn vẹn và nguồn gốc phát hành.

Đánh giá hiệu quả công nghệ

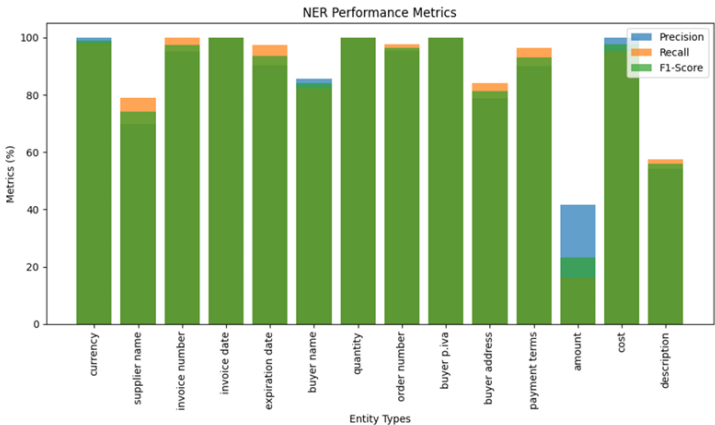
Bảng 4. Hiệu quả của các bước tiền xử lý đối với độ chính xác OCR

Preprocessing Step	OCR Accuracy (%)	Improvement (%)
Without Preprocessing	82.3	—
Noise Reduction	90.1	+7.8
Skew Correction	93.4	+3.3

Normalization	96.1	+2.7
---------------	------	------

Trong xử lý chứng từ ngoại thương, OCR quyết định hệ thống đọc chính xác tài liệu đến mức nào. Nếu không có tiền xử lý, độ chính xác chỉ đạt 82,3%, dễ gây sai dữ liệu và làm Chatbot phản hồi sai. Khi áp dụng các bước giảm nhiễu, chỉnh nghiêng và chuẩn hóa ảnh, độ chính xác tăng lên 96,1% – mức an toàn để chuyển đổi tài liệu scan sang văn bản tin cậy, đảm bảo NLP hoạt động ổn định (Biniam Abraha Masa, 2023). Sau OCR, dữ liệu được phân tích bằng NLP với

mô hình NER. Một mô hình NER tùy chỉnh cho hóa đơn thương mại đạt 86,38% độ chính xác tổng thể: các trường cố định như ngày phát hành, mã số thuế, số lượng, tiền tệ đạt 96,39%–100%, trong khi các trường phức tạp như Amount chỉ đạt 41,67% do phụ thuộc cấu trúc bảng và lỗi OCR. Điều này cho thấy chuỗi OCR → NER vẫn còn hạn chế và cần bổ sung Data Matching và mô hình Deep Learning phân tích bảng để cải thiện độ chính xác (Biniam Abraha Masa, 2023).

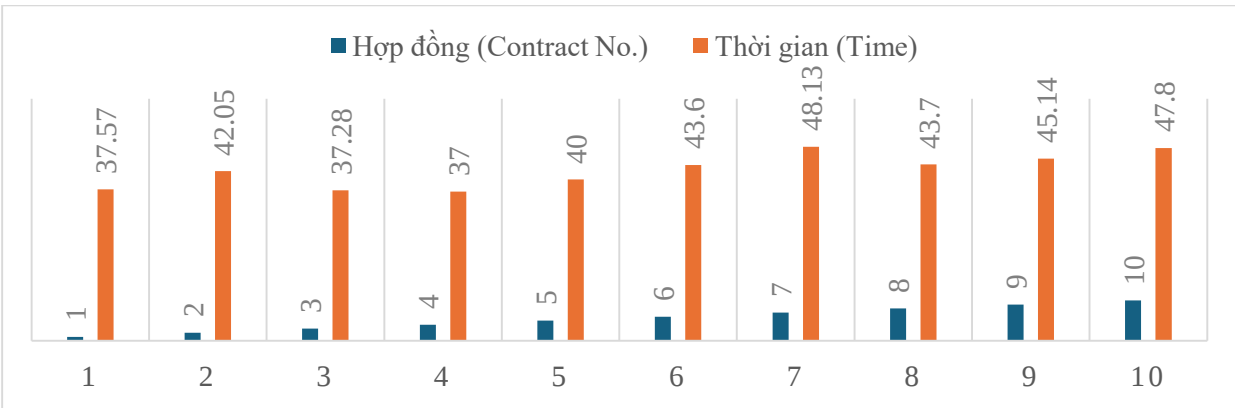


Hình 3. Hiệu suất mô hình NER theo từng loại thực thể

Hệ thống tích hợp OCR + NLP + Deep Learning cho thấy ưu thế vượt trội so với OCR truyền thống và thao tác thủ công. Kết quả nghiên cứu cho thấy khả năng phát hiện gian lận và bất thường trong chứng từ tài chính đạt tới 98,7%, chứng minh vai trò quan trọng của Data Matching kết hợp mô hình DL trong việc nhận

diện sai sót và rủi ro (Biniam Abraha Masa, 2023).Điều này khẳng định rằng OCR hoặc NLP đơn lẻ là chưa đủ; chỉ một kiến trúc đa tầng mới đáp ứng được yêu cầu về độ chính xác, tính an toàn và sự tin cậy trong xử lý chứng từ ngoại thương hiện đại.

Đánh giá hiệu quả của mô hình Chatbot AI



Hình 4. Thời gian chạy thử nghiệm 10 bộ chứng từ ngoại thương

Đề đánh giá hiệu quả thực tế, nhóm đã xây dựng mô hình AI Chatbot tích hợp OCR - NLP - Data Matching và thử nghiệm trên 10 bộ chứng từ ngoại thương. Kết quả cho thấy thời gian xử

lý trung bình chỉ 42,23 giây/bộ, nhanh hơn đáng kể so với quy trình thủ công mất 10-20 phút. Về độ chính xác, hệ thống đạt 75,76% khi đối chiếu tổng thể giữa Commercial Invoice (CI) và

Packing List (PL), trong đó CI đạt 82,76% còn PL thấp hơn ở mức 65,85% do chứa nhiều bảng biểu và bố cục không đồng nhất. Một số trường dữ liệu phức tạp vẫn bị nhận diện sai hoặc bỏ sót, như In Words (lỗi 7/10 bộ), Số/Ngày Invoice trong PL (thiếu 7/10 bộ), và các trường như Contract No., Swift Code, Tên ngân hàng (thiếu 3/10 bộ). Những hạn chế này cho thấy hệ thống còn gặp khó khăn với dữ liệu bán cấu trúc, phi cấu trúc và OCR bảng biểu, nhưng đồng thời cũng mở ra hướng nâng cấp rõ ràng cho việc cải thiện độ chính xác và tự động hóa trong giai đoạn tiếp theo.

4. Đề xuất lộ trình triển khai và kiến nghị

Nhóm đề xuất ba hướng cải tiến chính để nâng cao hiệu quả hệ thống. (1) Chuẩn hóa dữ liệu đầu vào, vì nhiều tài liệu thử nghiệm chưa đạt chất lượng, gây lỗi OCR; cần đặt chuẩn về độ phân giải, bố cục và áp dụng các bước tiền xử lý như giảm nhiễu, chỉnh nghiêng. (2) Nâng cấp mô hình phát hiện bất thường trong hợp đồng,

bổ sung kiểm tra ngữ nghĩa và anomaly detection để nhận diện sai logic và giá trị bất thường ngay ở tài liệu gốc. (3) Mở rộng phạm vi xử lý sang toàn bộ bộ chứng từ ngoại thương như Tờ khai hải quan, B/L, C/O nhằm tăng khả năng đối chiếu liên hồ sơ và phát hiện lỗi toàn diện hơn.

Để ứng dụng AI Chatbot hiệu quả, doanh nghiệp cần đào tạo nhân sự về kỹ năng số hóa và sử dụng OCR–Chatbot, đồng thời chuẩn hóa quy trình quản lý chứng từ, thử nghiệm trước với các chứng từ phổ biến như Invoice và Packing List (Phú Quý, 2025).

Cơ quan quản lý cần xây dựng chuẩn dữ liệu điện tử thống nhất, hoàn thiện khung pháp lý cho ứng dụng AI trong logistics và hỗ trợ doanh nghiệp nhỏ và vừa chuyển đổi số. Sự phối hợp giữa nhà nước, doanh nghiệp và đơn vị công nghệ sẽ giúp AI Chatbot phát huy hiệu quả và nâng cao minh bạch trong ngoại thương (Phú Quý, 2025).

Bảng 5. Định hướng mở rộng tập dữ liệu thử nghiệm

Tiêu chí mở rộng	Nội dung
Loại chứng từ	Contract, Invoice, Packing List, B/L, C/O, Customs Declaration
Loại hàng	Nông sản, thủy sản, công nghiệp
Tình huống lỗi	Sai số lượng, sai mô tả, lệch Incoterms, thiếu trường, sai mã HS
Chất lượng tệp	Scan rõ, mờ, nghiêng, nhiễu, ảnh chụp điện thoại
Mục tiêu đánh giá	Độ chính xác, độ ổn định, khả năng tổng quát hóa

Để tăng độ tin cậy khoa học, tập dữ liệu thử nghiệm nên được phân nhóm theo loại hàng hóa, loại chứng từ và dạng lỗi; trong đó mỗi nhóm cần có cả mẫu hợp lệ và mẫu chứa sai lệch. Cách thiết kế này giúp đánh giá không chỉ độ chính xác trung bình mà còn đo được khả năng phát hiện lỗi theo từng tình huống nghiệp vụ cụ thể.

KẾT LUẬN

Nghiên cứu của nhóm tập trung tối ưu hóa quy trình kiểm tra chứng từ ngoại thương trong bối cảnh doanh nghiệp cần tốc độ, độ chính xác và chuyển đổi số. Qua phân tích quy trình thủ công, nhóm nhận thấy nhiều hạn chế về thời gian, độ tin cậy và sự phụ thuộc vào kinh nghiệm cá nhân, cho thấy nhu cầu cấp thiết của một giải pháp tự động hóa.

Từ đó, nhóm xây dựng mô hình AI Chatbot tích hợp OCR, NLP và Data Matching để trích xuất – hiểu – đối chiếu dữ liệu giữa các chứng từ. Kết quả thử nghiệm cho thấy hệ thống xử lý nhanh hơn đáng kể so với cách thủ công và có khả năng phát hiện sai lệch ở mức khá tốt, dù vẫn còn hạn chế với các trường dữ liệu phức tạp và tài liệu scan kém chất lượng.

Mặc dù kết quả thử nghiệm ban đầu trên 10 bộ chứng từ cho thấy tính khả thi của mô hình, quy mô dữ liệu hiện tại vẫn còn hạn chế và chưa phản ánh đầy đủ tính đa dạng của thực tiễn ngoại thương. Trong giai đoạn tiếp theo, nghiên cứu cần mở rộng tập dữ liệu theo cả chiều rộng và chiều sâu. Về chiều rộng, cần bổ sung thêm nhiều loại chứng từ như Bill of Lading, C/O, tờ khai hải quan, chứng thư kiểm dịch, chứng nhận

chất lượng và các chứng từ vận tải liên quan. Về chiều sâu, cần xây dựng các tình huống sai lệch đa dạng như sai số lượng, sai đơn giá, lệch điều kiện giao hàng, không thống nhất mã HS, thiếu trường thông tin, khác biệt mô tả hàng hóa hoặc lỗi do OCR. Đồng thời, tập dữ liệu cũng cần bao gồm các tài liệu có chất lượng hình ảnh khác nhau như scan rõ, scan mờ, ảnh chụp lệch góc và tài liệu có bố cục bảng phức tạp.

Việc mở rộng dữ liệu thử nghiệm sẽ giúp đánh giá tốt hơn độ ổn định, độ bền vững và khả năng tổng quát hóa của mô hình AI trong nhiều bối cảnh doanh nghiệp và nhiều loại bộ chứng từ khác nhau.

Tổng kết lại, nghiên cứu chứng minh rằng AI là hướng đi khả thi và cần thiết trong kiểm tra chứng từ ngoại thương. Với việc tiếp tục nâng cấp mô hình, cải thiện chất lượng dữ liệu đầu vào và mở rộng sang toàn bộ bộ chứng từ, AI Chatbot có thể trở thành công cụ hữu ích giúp doanh nghiệp giảm rủi ro, tăng năng suất và thúc đẩy chuyển đổi số trong logistics – ngoại thương Việt Nam.

TÀI LIỆU THAM KHẢO

- Bibhu Dash. (2021). A hybrid solution for extracting information from unstructured data using optical character recognition (OCR) with natural language processing (NLP). *ResearchGate*.
- Biniam Abraha Masa. (2023). Development of an artificial intelligence-based solution for document processing automation using machine learning and nlp techniques. *Alma mater studiorum – Aniversity of Bologna*, 74.
- Bùi, N. T., & Phạm, C. Q. (2005). Giáo trình kỹ thuật nghiệp vụ ngoại thương. Hà Nội: Nhà xuất bản Hà Nội.
- Nguyễn, P. M., Vũ, H. N., & Trần, H. H. (2025). Cải thiện quá trình xử lý tài liệu thanh toán quốc tế bằng chuyển tiền điện tín ứng dụng công nghệ OCR và các công nghệ tự động hóa: Nghiên cứu tại VPBank. *Tạp chí Kinh tế và Dự báo - Ấn phẩm của Tạp chí Kinh tế - Tài chính*.
- Nguyễn, T. V., & Nguyễn, H. T. (2016). Giáo trình thanh toán quốc tế & tài trợ ngoại thương. Trong *Giáo trình thanh toán quốc tế & tài trợ ngoại thương*. Hà Nội: Nhà xuất bản Lao Động.
- Phú Quý. (2025). *Chuyển đổi số: Chìa khóa giúp ngành logistics Việt Nam 'bứt phá'*. Báo Công Thương.
- Trần, Q. T. (2006). *Quản lý rủi ro trong hoạt động kinh doanh xuất nhập khẩu của các doanh nghiệp Việt Nam đáp ứng yêu cầu hội nhập kinh tế quốc tế*. Hà Nội: Trường Đại học Ngoại Thương.