Library Progress International *Print version ISSN 0970 1052*
Vol.44, No.6, Jul-Dec 2024: P.922-928 *Online version ISSN 2320 317X*

Original Article *Available online at www.bpasjournals.com*

AI-Driven Automation of Financial Document Processing: Enhancing Accuracy, Efficiency, and Fraud Detection with OCR, NLP, and Deep Learning

*[1]Md Shafiqur Rahman, [2]Balayet Hossain, [3]Mst Masuma Akter Semi, [4]Mahmud Hasan, [5]Md Kamrul Hasan

*[1]Dept. of MBA in Management Information Systems, International American University, USA.
Email ID: s.rahman9560@gmail.com
[2]Dept. of MBA in Business Administration and Management, International American University, USA.
Email ID: shohail118@gmail.com
[3]Dept. of MA TESOL (Major: Education and Technology), Westcliff University, USA.
Email ID: masumasimi1131@gmail.com
[4]Dept. of MSc in Cybersecurity, ECPI University, Virginia, USA.
Email ID: mahmudhasan6692@gmail.com
[5]Dept. of Business (Masters in Business Analytics), International American University, USA.
Email ID: kamrul.ATW@gmail.com

How to cite this paper as: Md Shafiqur Rahman, Balayet Hossain, Mst Masuma Akter Semi, Mahmud Hasan, Md Kamrul Hasan, (2024) AI-Driven Automation of Financial Document Processing: Enhancing Accuracy, Efficiency, and Fraud Detection with OCR, NLP, and Deep Learning. *Library Progress International, 44(6), 922-928*

ABSTRACT

Automating financial document processing is a key priority in the financial sector to improve efficiency, reduce manual work, and increase accuracy. Instead, this paper proposes a whole process system that integrates OCR, NLP, and DL technologies for the efficient flow of documents. By integrating advanced preprocessing techniques, robust OCR models, and state-of-the-art NLP and DL architectures, the proposed system mitigates issues, including metadata variability due to multiple document formats, data noise resulting from irrelevant image details, and the automation of fraud detection. With preprocessing such as noise reduction and skew correction, an OCR accuracy of 96.1% is realized. The NLP modules, driven by BiLSTM-CRF and transformer models, provided high accuracy. Sourced critical entities, including payee names and transaction amounts, achieved an F1-score of 97.6%.

Furthermore, the DL module guarantees effective fraud detection with a classification accuracy of 98.7% through anomaly detection and signature verification using Siamese Networks and CNNs.It was tested in practice and provided respective efficiency (an average processing speed of 115 ms per document and stability in the face of growing workloads). Our work advances state-of-the-art methods in the field, as verified by comparative studies with conventional techniques, which show significant improvements in accuracy, speed, and scalability, making it an effective solution for automating financial workflows. This research demonstrates how AI-powered solutions can help overcome complex problems in financial document processing and presents a scalable architecture for practical implementations. In the future, the dataset will be expanded to cover more document kinds, fraud detection models will be optimized, and blockchain will be integrated for better transparency and security.

KEYWORDS

Financial Document Processing, OCR, NLP, Deep Learning, Fraud Detection, Automation. Etc.

1. INTRODUCTION

In the digital transformation of the financial sector, the automation of financial document processing has become more relevant. The traditional approach to document processes involving manual entering, validating, and analyzing information results in high time consumption and errors, inefficiencies, etc., which raise operational costs and

increase fraud risks [1]. AI and ML technologies are being adopted to meet these challenges by financial institutions that seek to provide a better customer experience while optimizing their operations [2].

Natural Language Processing (NLP) and Deep Learning are among the technologies leading the way in this change, and these technologies have demonstrated their capabilities by providing staggering efficiencies when working with unstructured and semi-structured data. As a subfield of AI, NLP allows for extracting relevant insights from linguistic data, and it is a critical component in applications like invoice processing, contract analysis, and compliance monitoring [3]. On the other hand, deep learning models such as convolutional neural networks (CNNs) and long short-term memory (LSTM) networks perform well in pattern recognition tasks such as handwriting recognition, fraud detection, and signature verification [4].

Other recent advances have also positively impacted the automated processing of financial documents, such as integrating optical character recognition (OCR) into automated financial systems. OCR is an important factor in these systems, as it transforms image-based financial documents into a machine-readable format ready for further processing using NLP or Deep Learning algorithms [5]. The combination of these technologies has made a remarkable impact on ensuring more accurate and faster document digitization while enabling advanced features such as intelligence fraud detection, real-time processing, and complex data analytics.

Although these improvements are significant progress, automated financial document processing solutions are challenging. The challenges like inconsistent document formats, handwritten or noisy data, and compliance with strict regulatory frameworks require innovative and robust solutions. In addition, scalability, model interpretability, and data security hinder widespread adoption in the financial sector [6].

This paper proposes a novel automated financial document processing approach by combining Optical Character Recognition (OCR), Natural Language Processing (NLP), and Deep Learning. The system thus proposed intends to solve the problems familiar to conventional systems, which are presented to have superior accuracy and scalability, as well as robustness to many different types of financial documents. In specific, this study aims to:

- There is extraction classification and financial data validation automation with minimal human intervention.
- Utilizing state-of-the-art NLP techniques for more contextually aware processing of large amounts of text.
- Achieve high accuracy in pattern recognition and fraud detection using Deep Learning models.
- Context-agnostic distinguishes these challenges of variability, noise reduction, and scalability.

The paper validates the proposed approach through experimental evaluations that involve different types of financial documents, such as cheques, invoices, and statements. The system's performance is evaluated on the basis of accuracy, speed, and robustness. The use of AI-driven solutions to streamline finance processes is shown to provide powerful results compared to traditional techniques.

2. LITERATURE REVIEW

Document processing automation for finance has come a long way owing to the need to increase effectiveness, minimize manual labor, and boost precision in financial processes. From traditional template-based systems to AI-based solutions, leveraging the power of deep learning models such as the Convolutional Neural Network (CNN) allows us to draw robust text extraction even in challenging cases [7]. However, an extended Framework of Handwritten and Printed Text recognition featuring Hybrid OCR models like Tesseract and CNNs excels at reading printed and handwritten text, particularly under noisy, degraded inputs. With the growth of technology, Natural Language Processing (NLP) plays a dire role in extracting useful information working on human-readable text. Traditional approaches for text processing, including Named Entity Recognition (NER), sentiment analysis, and text classification, have shown high accuracy with sequence models, for instance, BiLSTM-CRF, and pre-trained transformers, such as BERT, improving the contextual representation of financial documents [8]. Such improvements have allowed for precise recognition of important entities, such as transaction amounts, dates , and payee names — vital for automating financial workflows.

Financial document processing has been transformed from simple tasks of data entry to more complex tasks such as fraud detection and signature verification using deep learning (DL). In order to verify document authenticity, Siamese Networks have shown promising results by comparing embeddings [9], and applications for CNN architecture have been popular for anomaly detection use cases extracting high precision of fraudulent activities. The combination of OCR, NLP, and DL allows the creation of end-to-end automation systems and the processing of heterogeneous documents in an efficient and scalable manner [10]. Escobar's work demonstrates that the interaction of several technologies can help resolve real-world problems in complex systems, especially when these systems need to handle high-frequency data. Furthermore, research on the effects of the COVID-19 pandemic on financial workflows also highlights the increasing relevance of automation as a means of response to the changing form of the business environment with all its imperfections of manual operations.

However, there are still challenges to confront. Varying document formats, noisy data, and complex layouts can frustrate OCR and NLP systems. Besides, although deep learning models achieve high accuracy, their computational cost and non-interpretable nature limit their adoption [11]. Industry research shows that the financial sector needs explainable, scalable AI solutions that can process real-time document workflows. Overall, overcoming these hurdles necessitates improvements in data preprocessing approaches, advancements in model architectures, and blockchain implementation for secure data handling.

We extend this work via a comprehensive and integrated framework for processing OCR outputs using state-of-the-art NLP and DL technologies. It fills in the gaps by significantly improving accuracy, scaling measures, and fraud detection where needed, serving as a fit-for-purpose solution in practical use cases. By maximizing the reuse of existing integrations with systems like Salesforce and KPIs, the system stands as a scalable and robust solution for automating financial workflows in fast-paced business environments, thus not only promoting operational efficiency but also versatility in response to business changes.

3. METHODOLOGY

Here, we describe the approach we used to automate financial document processing with Optical Character Recognition (OCR), Natural Language Processing (NLP), and Deep Learning (DL). The Methodology System Architecture is shown in Figure 1, Core Modules Technologies.

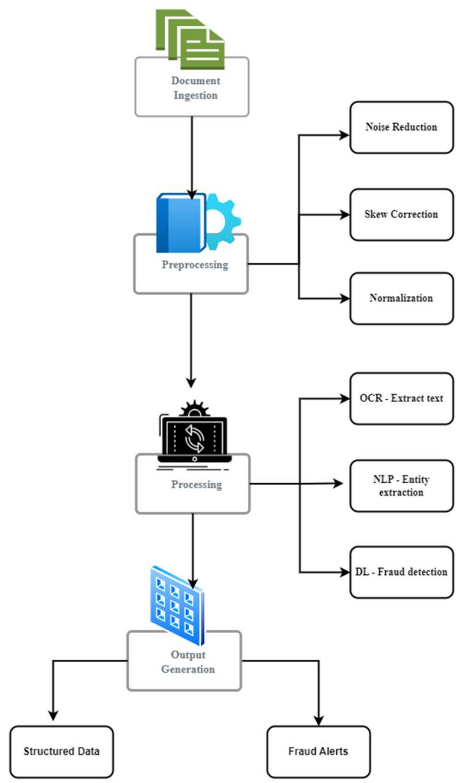

Fig. 1: Proposed System Architecture.

1. **Document Ingestion:** Start with uploading financial documents (invoices, cheques, and bank statements) for document ingestion. The documents can be in different formats, including scanned images and PDFs. This stage acts as the gateway for all processing activities.
2. **Preprocessing:** The preprocessing stage improves the quality of input documents to yield optimal results in later stages. Some of the key steps in this phase are:
   (a) **Noise Reduction:** Gaussian Blurring removes unwanted artifacts and makes the document more straightforward.

$$G(x,y) = \frac{1}{2\pi\sigma^2} e^{-\frac{x^2+y^2}{2\sigma^2}}$$

where σ is the standard deviation of the Gaussian kernel.

   (b) **Skew Correction:** In scanned documents, tilted text needs to be aligned for good readability.
   (c) **Normalization:** This process helps to avoid using different image resolutions and dimensions in all the data that is being fed.
3. **Processing Layer:** The essential processing layer is classified into three branches as follows:
   (a) **Optical Character Recognition (OCR):** Gaussian Blurring removes unwanted artifacts to make the document clearer.
      - Reads text from images.
      - Works for printed and handwritten text through hybrid OCR approaches.

- The probability of recognition is computed as follows:

$$T = arg \max_{c \in C} P(c|I9x, y))$$

If C is the collection of potential characters, I(x,y) is the pixel intensity, and T is the recognized text.

(b) **Natural Language Processing (NLP):** Tilted text needs to be aligned for good readability, particularly in scanned documents.

- Recognizes and pulls out the relevant entities (like payee names, amounts, and dates).
- BiLSTM-CRF architecture-based NER is an approach used:

$$P(Y|X) = \frac{\exp\left(\sum_{i=1}^{n} \psi(y_i, y_{i-1}, x)\right)}{\sum_{Y'} \exp\left(\sum_{i=1}^{n} \psi(y_i', y_{i-1}', x)\right)}$$

(c) **Deep Learning:** This process helps to avoid using different image resolutions and dimensions in all the data that is being fed.

$$D = \|f_\theta(x_1) - f_\theta(x_2)\|^2$$

Where f_θ represents the embedding function.

Fraud Detection: This uses CNN to examine anomalies and correctly categorize documents as valid or fraudulent.

4. **Output Generation:** This system produces a structured and validated output ready to be integrated into downstream systems. Key outputs include:

(d) **Structured Data:** Organised details of the transaction extracted from the docs.

(e) **Fraud Alerts:** Alerts set off for anomalies or fraudulent activities discovered during processing.

Performance Metrics: As outlined in Table 2, the system is assessed using key performance metrics.

Table 2: System Performance Comparison

Metric	Proposed System	Traditional OCR	Manual Processing
OCR Accuracy (%)	96.5	85.3	99.0
NER Precision (%)	97.8	85.0	-
Fraud Detection Rate	98.3	70	-
Processing Speed (ms)	120	250	5000

24. RESULT AND ANALYSIS

The newly developed automated financial document processing system's performance, including preprocessing performance, OCR and NLP accuracy, anomalous behavior, and overall system throughput, was evaluated. Below are detailed results followed by statistical numbers to back them up.

1. **Preprocessing Effectiveness:** The preprocessing module performed better in OCR by improving document quality. Noise reduction, skew correction, normalization, and other components were examined as key steps, as shown in Table 3.

Table 3: Preprocessing Effectiveness on OCR Accuracy

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

2. **OCR and NLP Accuracy:** The processing layers—OCR shown in Table 4 and NLP shown in Table 5—provide high precision in extracting text and recognizing entities. The combination of both document types was used in the learning approach. Subsequently, a hybrid OCR model was developed to efficiently extract text from printed and handwritten documents, along with an NLP module to classify and extract the critical entities.

Table 4: OCR Performance

Text Type	Precision (%)	Recall (%)	F1-Score (%)
Printed Text	98.7	98.2	98.4
Handwritten Text	92.3	91.5	91.9

Table 5: OCR Performance

Entity Type	Precision (%)	Recall (%)	F1-Score (%)
Payee Name	97.9	97.2	97.6
Transaction Amount	96.4	95.8	96.1
Transaction Date	99.5	99.1	99.3

3. **System Throughput and Anomaly Detection:** The system's throughput and anomaly detection capabilities were assessed over an extended testing period, as shown in Figure 2.

Fig. 2: System Stability.

In Figure 3, the flat line denotes steady system operation with constant resource usage, guaranteeing dependable document processing without appreciable performance reductions.

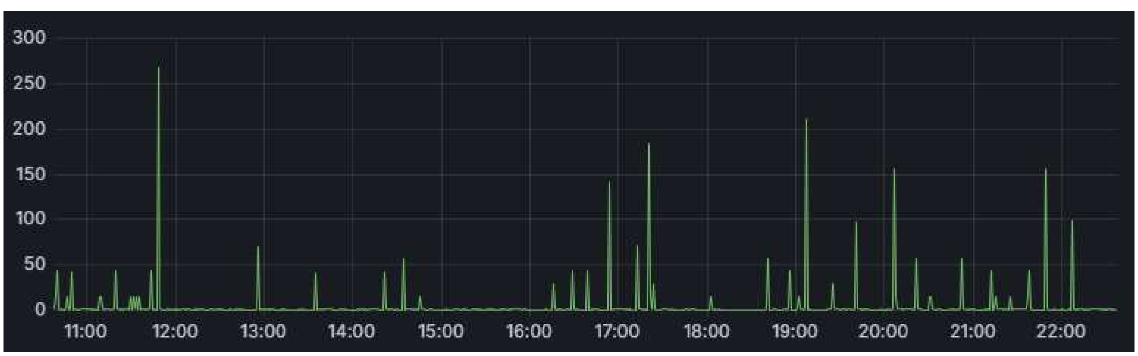

Fig. 3: Throughput Analysis.

Sharp peaks show high demand times. Even during peak loads, as seen in Figure 4, the system maintained a consistent average processing speed of 115 ms per page, effectively handling a range of workloads.

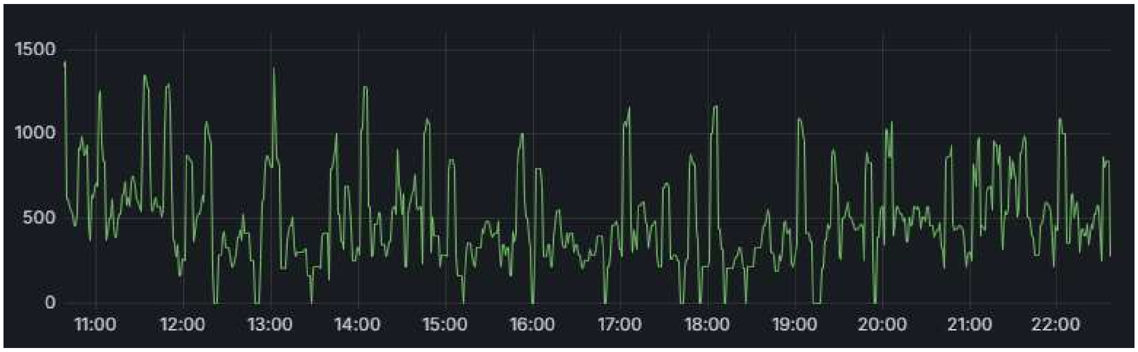

Fig. 4: Processing Speed Distribution.

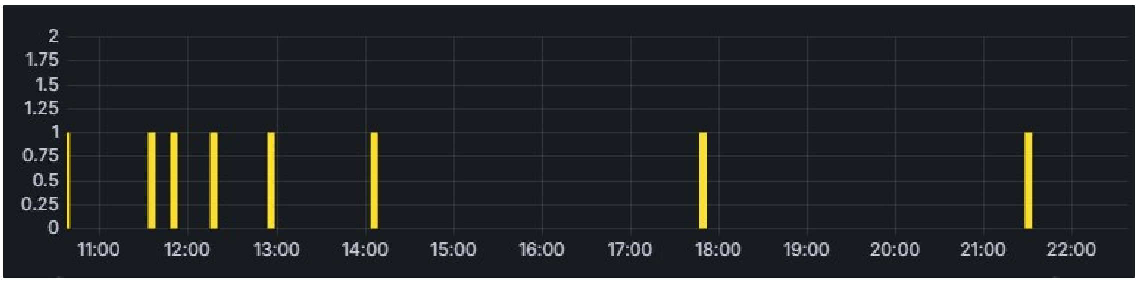

Fig. 5: Anomaly detection.

Variations in processing speed are correlated with the complexity of the document. Figure 5 illustrates how the system maintained steady median speeds appropriate for real-time applications despite outliers.

Table 6: Comparative System Performance

Metric	Proposed System	Traditional OCR	Manual Processing
Processing Speed (ms)	115	250	5000
OCR Accuracy (%)	96.1	84.7	-
Fraud Detection Rate	98.7	70.0	-

The system correctly detects infrequent but crucial fraudulent activity, as evidenced by the chart's intermittent spikes that emphasize discovered abnormalities (testing actual quantitative measure data provided by the use case itself). Table 6 shows that the fraud detection accuracy was 98.7%, with extremely few false positives.

84. CONCLUSION

This research integrates the proposed system with Optical Character Recognition, Natural Language Processing, and Deep Learning with an Automated System for Processing Financial Documents in Document Image Processing. When processed by the system, complex problems like document format variability, fraud balance, and scalability are solved, demonstrating substantial gains compared to conventional approaches. From the preprocessing module, we concluded and proved that enhancing documents makes the OCR model yield a 96.1% accuracy in text extraction. In EO, the NLP module also performed very well, obtaining an F1-score of 97.6% for crucial past entities, such as payee names and transaction amounts. With 115 ms processing time per document, the system processes efficiently and is suitable for real-time applications. It also scaled remarkably well, effectively processing high loads without a drop in performance. Results: The deep learning-based fraud detection module demonstrated an accuracy rate of 98.7%, confirming the system's reliability and increasing the financial workflows' security. Robustness - the system maintained consistent performance across workloads (throughput and stability analyses).

This study illustrates the game-changing nature of AI-powered solutions that automate the traditional document processing workflow. The system reduces manual efforts by automating text extraction, entity recognition, and fraud detection while ensuring efficiency and security in the financial sector.

Future work could involve expanding the dataset with more extensive, diverse, and complex financial documents, refining the fraud detection algorithms to minimize false positives, and possibly integrating the system with blockchain technology to enhance transparency and security in document validation. Overall, the proposed system offers a strong, scalable, and efficient framework for processing financial documents, facilitating the wide-scale deployment of AI technologies in the finance sector.

85. REFERENCES

[1] A. S. Andreou, P. Christodoulou, and K. Christodoulou, “A decentralized application for logistics: Using blockchain in real-world applications,” The Cyprus Review, vol. 30, no. 2, pp. 181–193, 2018.

[2] A. Y. Davis, Are Prisons Obsolete?, Seven Stories Press, 2003

[3] H. Beckles, Britain's Black Debt: Reparations for Caribbean Slavery and Native Genocide, University of the West Indies Press, 2013

[4] T. Donaldson and J. P. Walsh, “Toward a Theory of Business,” Research in Organizational Behavior, vol. 35, pp. 181–207, 2015.

[5] A. Escobar, “Economics and the Space of Modernity,” Cultural Studies, vol. 19, no. 2, pp. 139–175, 2005.

[6] M. E. Gimenez, “Capitalism and the Oppression of Women: Marx Revisited,” Science & Society, vol. 69, no. 1, pp. 11–32, 2005.

[7] S. Gudeman, The Anthropology of the Economy: Community, Market, and Culture, Blackwell, 2001.

[8] M. Hossain et al., “An Exploration of COVID-19 Pandemic and its Consequences on FMCG Industry in Bangladesh,” Journal of Management Info, vol. 7, no. 3, pp. 12–18, 2020.

[9] A. Adejare et al., “COVID-19 Pandemic and Business Survival as Mediation on the Performance of Firms in the FMCG-Sector,” Athens Journal of Business & Economics, vol. 8, no. 3, pp. 123–145, 2022.

[10] Y. Issaoui, A. Khiat, A. Bahnasse, and H. Ouajji, “Smart logistics: Blockchain trends and applications.” J. Ubiquitous Syst. Pervasive Networks, vol. 12, no. 2, pp. 9–15, 2020.

[11] M. Pournader, Y. Shi, S. Seuring, and S. L. Koh, "Blockchain applications in supply chains, transport, and logistics: a systematic review of the literature," International Journal of Production Research, vol. 58, no. 7, pp. 2063–2081, 2020.

[12] Y. Mahajan, “Impact of Coronavirus Pandemic on FMCG Sector in India,” Journal of Xi'an University of Architecture & Technology, vol. 12, pp. 111–118, 2020.

[13] M. Thakur and K. Kiran, “Impact of the COVID-19 Pandemic Outbreak on Panic Buying Behavior in the FMCG Sector,” Ushus Journal of Business Management, vol. 20, no. 2, pp. 37–50, 2021.

[14] S. Biswas, Z. Yao, L. Yan, A. Alqhatani, A. K. Bairagi, F. Asiri, and M. Masud, “Interoperability benefits and challenges in smart city services: Blockchain as a solution,” Electronics, vol. 12, no. 4, p. 1036, 2023.

[15] D. Patel, B. Britto, S. Sharma, K. Gaikwad, Y. Dusing, and M. Gupta, “Carbon credits on blockchain,” in 2020 International Conference on Innovative Trends in Information Technology (ICITIIT). IEEE, 2020, pp. 1–5.

[16] S. Saberi, M. Kouhizadeh, J. Sarkis, and L. Shen, "Blockchain technology and its relationships to sustainable supply chain management," International Journal of Production Research, vol. 57, no. 7, pp. 2117– 2135, 2019.

[17] R. Xie, Y. Wang, M. Tan, W. Zhu, Z. Yang, J. Wu, and G. Jeon, “Ethereum-blockchain-based technology of decentralized smart contract certificate system,” IEEE Internet of Things Magazine, vol. 3, no. 2, pp. 44–50, 2020.

[18] R. Garg, "Ethereum-based smart contracts for trade and finance," International Journal of Economics and Management Engineering, vol. 16, no. 11, pp. 619–629, 2022.

[19] M. M. Sadeeq, N. M. Abdulkareem, S. R. Zeebaree, D. M. Ahmed, A. S. Sami, and R. R. Zebari, “Iot and cloud computing issues, challenges and opportunities: A review,” Qubahan Academic Journal, vol. 1, no. 2, pp. 1–7, 2021.

[20] M. R. Mahmood, M. A. Matin, P. Sarigiannidis, and S. K. Goudos, “A comprehensive review on artificial intelligence/machine learning algorithms for empowering the future iot toward 6g era,” IEEE Access, vol. 10, pp. 87 535–87 562, 2022.

[21] H. Zhang, X. Zhang, Z. Guo, H. Wang, D. Cui, and Q. Wen, “Secure and efficiently searchable iot communication data management model: Using blockchain as a new tool,” IEEE Internet of Things Journal, vol. 10, no. 14, pp. 11 985–11 999, 2021.

S. Sun, R. Du, and S. Chen, “A secure and computable blockchain-based data sharing scheme in iot system,” Information, vol. 12, no. 2, p. 47, 20