SSRG International Journal of Electronics and Communication Engineering
ISSN: 2348-8549/ https://doi.org/10.14445/23488549/IJECE-V13I8P109

Volume 13 Issue 8, 131-149, August 2026

Review Article

Evaluation of Pressure Ulcers using Convolutional Neural Networks: A Systematic Review of the Recent Scientific Literature

Brigitte Torres-Pinedo[1], Sebastián Ramos-Cosi[1], Juan Morales[2], Ana Huamani-Huaracca[1*]

[1]*Image Processing Research Laboratory (INTI-Lab), Universidad de Ciencias y Humanidades, Lima, Perú.*
[2]*E-Health Research Center, Universidad de Ciencias y Humanidades, Lima Perú.*

[1*]*Corresponding Author : ahuamani@uch.edu.pe*

Received: 01 May 2026 Revised: 10 June 2026 Accepted: 11 August 2026 Published: 31 August 2026

Abstract - *The clinical setting faces a significant problem with Pressure Ulcers (PUs), and artificial intelligence has emerged as a promising tool for assessing them. Therefore, the aim of this systematic review is to comprehensively examine the scientific information published in the last twelve years (2013 - 2025) on the use of Convolutional Neural Networks (CNNs) to evaluate this type of PUs. A set of 42 articles were selected using the PRISMA methodology. These articles were chosen and examined based on the previously established criteria, using the VOSviewer program and Bibliometrix to create bibliometric maps and thus facilitate an exhaustive analysis. China was one of the countries that published the most articles during that period of time analyzed, and that has had the most collaborations with other national countries. In addition, it was evident that from 2020 onwards it was more noticeable, as more publication was shown, which shows a greater interest in the application of deep learning techniques in the clinical environment for the evaluation of PUs. In conclusion, the integration of technologies that employ artificial intelligence in the assessment of PUs can improve diagnostic accuracy, but to achieve effective clinical implementation, it is crucial to move towards reliable models and validation in real-world environments.*

Keywords - *Pressure ulcers, Convolutional Neural Networks, Deep Learning, Application.*

1. Introduction

Globally, Pressure Ulcers (PUs) are a common clinical problem of great relevance and persistent despite the fact that they are considered, to a large extent, injuries that can be prevented. Despite the fact that most cases are considered preventable with timely follow-up and appropriate care, their prevalence is still a major challenge for health systems. According to the European Advisory Panel on Pressure Ulcers (EPUAP), the National Advisory Panel on Pressure Ulcers (NPIAP) and the International Alliance for the Prevention of Skin Lesions (PPPIA), the frequency of pressure ulcer cases in healthcare settings ranges from 6% to more than 18%, depending on the category in which patients and the care setting are [1]. In addition, millions of patients in different countries suffer from these types of complications every year, leading to an increase in hospital stays, treatment prices and the risk of adverse outcomes.

On the other hand, patients with obesity, neurological disorders, spinal cord injuries, or prolonged immobilization are the people who are at a higher risk of developing pressure ulcers [2]. According to the World Health Organization, the number of people requiring rehabilitation has exceeded 2.4 billion worldwide, indicating the magnitude of the problem of long-term care and its associated complications [3]. In recent years, teleworking has also increased; Therefore, the long time spent in a sedentary position could contribute to the development of tissue damage from the absence of regular changes in posture.

It is also known that PUs is mainly caused by prolonged pressure on the soft tissues, along with aspects such as shear, friction, and alterations in blood flow [4]. This causes tissue ischemia and necrosis; in difficult situations, it can compromise even deeper layers like muscles and bones. These wounds not only greatly affect the patient's daily life, but also increase the possibility that infections, sepsis, and even the probability of death may arise [5].

Faced with this dilemma, in the last ten years, there has been an increasing interest in including technologies based on Artificial Intelligence (AI), especially through neural networks, to optimize the evaluation, categorization, and prediction of this type of injury [6]. These tools make it possible to study extensive amounts of clinical information and medical visualizations, recognize complex patterns that are not visible to the naked eye, thus facilitating decision-making in the clinical field in a clear and timely manner.

Convolutional Neural Networks (CNNs), in this situation, have emerged as an instrument with great potential to detect PUs in time and identify PUs, providing the possibility of reducing the frequency of severe injuries, maximizing the use of health resources, and increasing clinical results worldwide. To a large extent, the studies are based on reduced data sets or images obtained in idealized situations, which complicates the possibility of extending the findings and applying them appropriately in real-world contexts.

Despite the advances made, the contemporary literature does not have a common scheme to evaluate models that use CNNs in the treatment of PUs [7]. Similarly, numerous investigations focus on a single method without comparing it to other available options, while tests are usually carried out under controlled conditions, which creates a lack of information about its effectiveness in authentic and varied clinical situations [8]. Based on this problem, the research questions that will guide this study are presented: What methods based on neural networks have been implemented in recent years for the assessment of pressure ulcers? How effective have these models been in terms of diagnostic accuracy and reliability? and What future challenges, limitations, and opportunities do the implementation of neural networks in clinical settings represent?

Therefore, the objective of this systematic review was to fully examine the scientific information published in the last twelve years on the evaluation of PUs using CNNs. In addition, it was essential to identify the current applications, the methods used and the reported performance results, as well as the restrictions that limit their use and acceptance in real clinical scenarios. Unlike other articles that deal with the use of AI in wound management broadly or that concentrate on specific applications, this systematic review differs by focusing on the assessment of PUs through CNNs, examining in an orderly manner the recently designed methods. The purpose of this study is to contribute to the development of more effective tools to support healthcare professionals. In this sense, the review offers a complete perspective of the current state of the problem and thus directs new studies and facilitates informed decision-making that offers more solid, comparable and usable solutions in real care contexts.

2. Literature Review

Deep learning methods in medical imaging have recently significantly expanded their application in various contexts due to their potential to automatically extract complex features from images without the need for manual feature engineering. According to the study [8], the authors conducted in-depth research and found findings that deep learning models have already been successfully used to diagnose various diseases, including oncological, dermatological, and ophthalmological pathologies, demonstrating high accuracy rates. They analyzed approximately 25 studies and reported that, in some cases, the values of the area under the ROC curve (AUC) exceeded 0.90, which would confirm the high potential and capacity of these technologies to support clinical diagnosis.

In the field of pressure ulcer research, computer vision methods are also playing a key role. For example, in one study [9], they proposed to develop a system to automatically classify pressure ulcer stages based on CNN, demonstrating that its use can significantly improve diagnostic accuracy compared to other traditional methods of image analysis. The researchers also indicated that early detection of tissue damage can facilitate early medical intervention and reduce the risk of complications.

On the other hand, there are other studies that have focused only on the use of pre-trained deep learning models. As evidenced by the study by [10], which demonstrated the effectiveness of transfer learning to classify medical images with very limited data. Despite this, the results confirmed that architectures such as ResNet and DenseNet have the ability to achieve a high accuracy of 90% even with a small number of labeled medical images.

Beyond classifying images, current research also explores the use of AI to comprehensively support clinical decision-making. In the study of [11], the authors demonstrated that integrating deep learning algorithms into processes to assess wounds and thereby improve diagnostic objectivity and reduce variability between experts. In addition, they highlight the potential of automated systems for monitoring chronic wounds and predicting their future evolution.

Despite the significant progress of research addressing this area, some of them still disagree on the datasets used, CNN architectures, and performance evaluation criteria. In addition, other studies only focus on individual models or small patient samples, making it difficult to get a holistic view of the current state of research. Therefore, there is a need to systematize the latest scientific achievements in relation to the use of CNN to assess and classify pressure ulcers.

3. Materials and Methods

In order to rigorously synthesize the already available scientific evidence on the use of CNNs to assess PUs, this study used a systematic review method. This method was chosen for its ability to minimize bias and ensure an organized and replicable procedure, in contrast to other types of reviews that are more descriptive [12]. To achieve this, a method was developed that consisted of formulating research questions, systematically searching for and selecting relevant studies, as well as synthesizing and summarizing the results, offering a structured perspective on the current situation of knowledge in this area.

In addition, to ensure clarity, accuracy, and reproducibility of the evaluation procedure, the current study was conducted according to the guidelines of the protocol Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). This method facilitated the orderly organization of the phases of identification, selection, suitability and incorporation of the research, in addition to accurately recording the procedure for reviewing the literature and the identification of the most and least significant topics for the study [13].

3.1. Studio Design

This study is based on the collection and structured research of scientific literature related to the use of CNNs to verify PUs. By examining the selected research, an attempt was made to recognize fields of study that have not yet been adequately addressed, in addition to common obstacles in the suggested methodologies. The study also included research conducted in various healthcare settings, such as hospitals, patients in long-term care areas that suggest or examine technology options aimed at optimizing the early identification and clinical analysis of these lesions, in order to facilitate decision-making and minimize the progression of PUs.

On the other hand, it was carried out from a qualitative approach on the technologies that have been developed and proposed to identify and evaluate PUs, paying special attention to those that are based on CNNs. The articles analyzed mainly use deep learning models such as CNNs in medical imaging, as well as methods that combine additional clinical information. Likewise, the assessment of the accuracy of these tools was carried out through a descriptive compilation of the measures presented in each research, including precision, accuracy, f1-score and Recall, without carrying out joint statistical analyses.

3.2. Data Sources

To determine which studies were relevant, a thorough search was carried out in the scientific literature through four databases that stood out for their academic strength and scope in various disciplines, which were Scopus, IEEE Xplore, PubMed, and Web of Science. In particular, the selection of these databases was made because it was necessary to include research in the fields of biomedical engineering and informatics and health sciences, since they are directly linked to the use of CNNs to evaluate PUs [14]. This approach ensured extensive and representative coverage of literary publications, diminishing the possibility of neglecting important research.

3.3. Search Strategy

Techniques were applied to search for articles that are significant for this research, for which Boolean operators were used that were adapted to the structure and characteristics of each database. Linguistic variations were integrated by combining keywords related to PUs and CNNs to increase the scope of the search. Similarly, words related to identification, categorization, analysis and prognosis processes were added, in order to obtain important research from clinical and technological angles. All searches were carried out through sections such as title, abstract and keywords, according to the alternatives available on each platform, giving preference to the detection of research that is directly linked to the purpose of the study. related to the subject. The following are the sets of keywords used in each of the databases:

- Scopus: (TITLE ("pressure ulcer*" OR "pressure injury*" OR "bedsore*")) AND (TITLE-ABS-KEY ("neural network*" OR "deep learning" OR "artificial intelligence")) AND (TITLE-ABS-KEY (detect* OR assess* OR classif* OR predict*)) AND (TITLE-ABS-KEY (system* OR application* OR tool* OR framework*))
- IEEE Xplore: ("Document Title":p ressure ulcer*) OR ("Document Title":p ressure injury*) AND ("Abstract": convolutional neural network) AND ("Abstract": image) OR ("Abstract":d etection) AND ("Abstract": accuracy) AND ("Abstract": segmentation) AND ("Abstract": clinical)
- Pubmed: ((pressure ulcer[Title] OR pressure injury[Title] OR pressure ulcer[Title/Abstract] OR pressure injury[Title/Abstract]) AND (deep learning[Title/Abstract] OR convolutional neural network[Title/Abstract] OR CNN[Title/Abstract]) AND (accuracy[Title/Abstract] OR sensitivity[Title/Abstract] OR validation[Title/Abstract]))
- Web of Science (WoS): Title=("pressure ulcer*" OR "pressure injury*" OR bedsore) AND Abstract=("convolutional neural network" OR CNN) AND Abstract=image* AND Abstract=(detect* OR segment*) AND Abstract=(accuracy OR validation) AND Abstract=application

3.4. Study Selection Process

Figure 1 shows the outline of the research selection method used in this systematic review, which sequentially exposes the identification, filtering, screening and inclusion phases. In a first phase (Identification), a total of 11,411 entries were obtained from four databases: 111 from Scopus, 853 from IEEE Xplore, 2,901 from Pubmed and 7,546 from WoS. This made it easier to form a wide variety of publications that could be important for the research topic.

Then, in the filtered stage, criteria were implemented to include and exclude studies in order to clean up the initial records. In this phase, the format of the document was taken into account, as well as whether they have a direct link to the subject of the study of the PUs and the implementation of CNNs, in addition to the possibility of accessing the full text of the article. As a result of this procedure, in the next screening phase, it was possible to obtain: 37 from Scopus, 64 from IEEE Xplore, 28 from Pubmed and 83 from WoS. In addition, an exhaustive analysis of the titles and abstracts of the chosen articles was carried out, from which a total of 47 investigations were collected for subsequent analysis, and 165 articles were excluded because they did not comply with the requirements of the established criteria. Subsequently, in the inclusion phase, 5 duplicate articles from the four databases were identified, so they were eliminated. Finally, a set of 42 articles relevant to the study of this systematic review was obtained. With this, this systematic review becomes a useful source of information for future new research on how technology is supporting health professionals to be able to detect and classify PUs in time.

Likewise, of the articles collected, they were systematically organized to make their analysis and review easier. For each study, important bibliographic data were obtained, including authors, date of publication of the scientific journal, Journal and other numerical indicators required to analyze scientific information. This facilitated an orderly analysis of the methods that use CNNs in research on PUs, in addition to contrasting the various approaches used for their identification, categorization, evaluation or anticipation. Similarly, the geographical dispersion of the studies was examined in order to recognize which countries show greater involvement and scientific contributions in the creation of solutions based on deep learning for this clinical question.

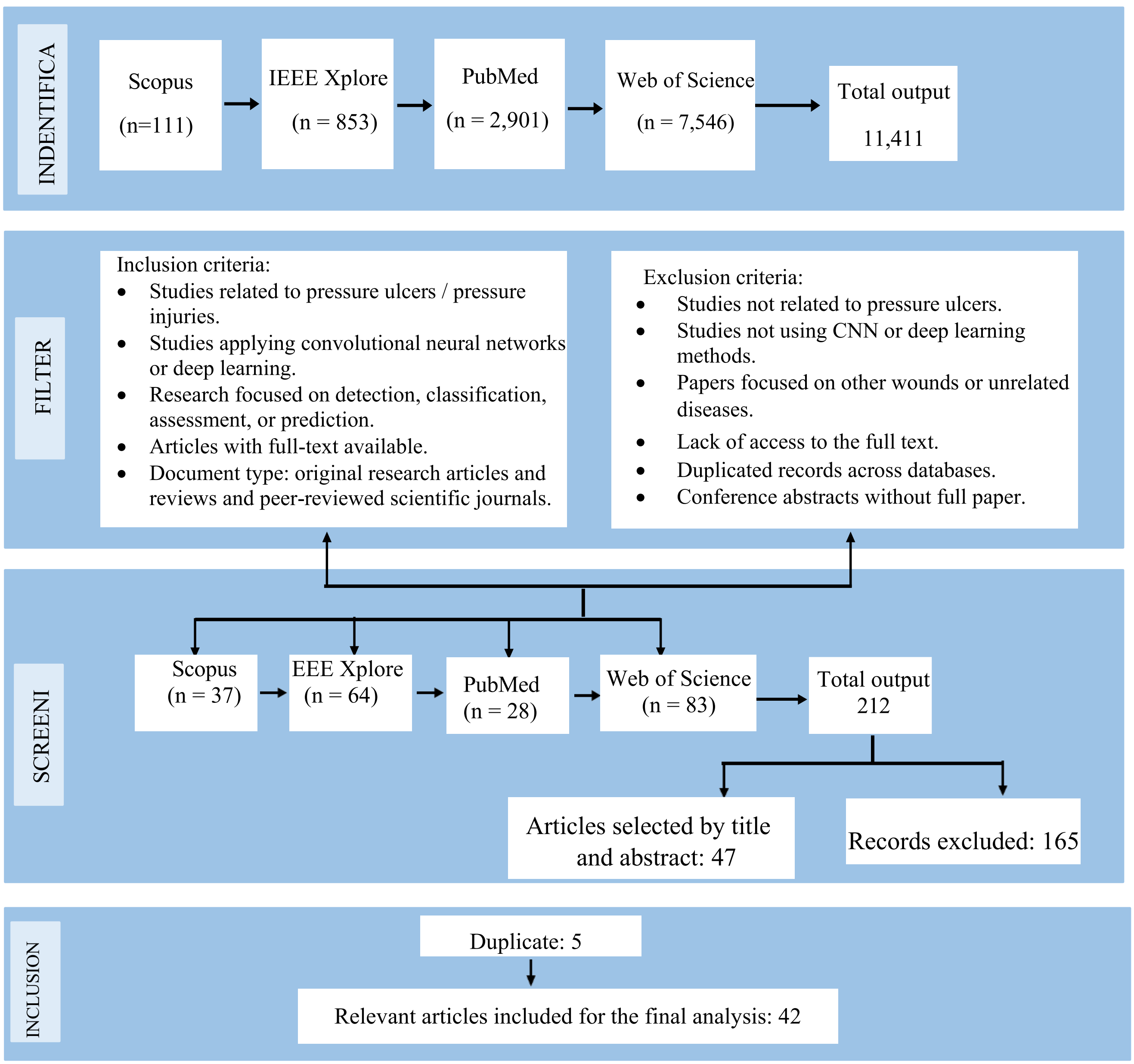

Fig. 1 PRISMA-based flowchart

On the other hand, in order to manage and keep an order of all the articles collected, the Mendeley program was used, which made it possible to save the records obtained, analyze and eliminate duplicates, in addition to simplifying the systematic review of the chosen articles. Similarly, the VOSviewer program was used to create bibliometric maps and collaborative networks, which facilitated the visual analysis of scientific productions [15]. In addition, Bibliometrix, which is located in the R environment, was used to be able to carry out additional bibliometric analyses [16]. The joint use of these tools enhanced the methodological soundness of the research and facilitated an exhaustive and organized analysis of the chosen literature.

In summary, the selection method carried out in this study facilitated the gradual reduction of a large amount of scientific literature, thus establishing a final set of research that is very relevant and aligned with the purposes of this systematic review. The use of inclusion and exclusion standards at each stage ensured that the selected studies focused directly on the topic. In addition, it made it possible to identify applications that have already been carried out in the clinical field, evaluate their level of precision and know their limitations in detecting ulcers and classifying at what level the patient is, so that future research can improve what exists or develop better proposals with those observations collected, given that technological advances may facilitate the development of more robust approaches in future studies. Therefore, the final set of research forms a firm basis and interprets the current situation regarding the use of deep learning techniques to identify PUs.

4. Results

The findings of this systematic review showed a continued increase in scientific interest towards the use of deep learning methods, especially CNNs, to detect PUs. In addition, the reviewed study indicated that these informatics methodologies have acquired a fundamental role in supporting activities such as early detection, automatic categorization, risk analysis and projection of the progress of injuries, thus facing a clinical challenge that is quite common and complicated [17].

In general terms, the research presented showed an inclination towards the use of models based on images, biomechanical signals collected through clinical sensors, which facilitates the identification of significant patterns that cannot always be detected with conventional evaluation methods [18]. In this framework, CNNs are distinguished by their ability to assimilate discriminative features automatically, reaching encouraging degrees of precision in various clinical situations [19].

On the other hand, the findings make it possible to recognize variations in the methods used, the types of information used, and the level of clinical validation achieved by the initiatives, in addition to an inequality in the geographical location of scientific research [20]. These results offer a comprehensive perspective of the present state of research and lay the foundation for analysis detailed bibliometric and technical study developed in the following subsections.

4.1. Bibliometric Analysis

It was carried out in order to summarize and understand the scientific research associated with the use of CNNs for the study of PUs. This method facilitated the recognition of publication models, thematic trends and cooperation networks in this area of study.

To achieve this objective, the VOSviewer program was used, which facilitated the graphical visualization of co-occurrence of keywords, authors and countries and collaboration between countries, using bibliographic data extracted from databases such as Scopus, WoW, Pubmed and IEEE Xplore. This study made it possible to determine which are the most common topics and the most relevant lines of research, providing a broad perspective on the current situation of the area.

4.1.1. Keyword Co-Occurrence Map

A study on the co-occurrence of key terms was carried out based on the documents chosen from the four databases mentioned above, using the VOSviewer tool. This study also facilitated the detection of the most common connections between the keywords used in scientific research on the use of CNNs to investigate PUs. Figure 2 shows that the map generated a total of 41 keywords, which reflects that they are distributed in three fundamental groups, with 407 links and a total strength of 788 links. In the visual illustration, the size of the nodes indicates how often each word appears, the thickness of the links expresses how strong the connection between each term is, as well as the colors represent how the words are divided by thematic groups.

It can be seen that there are three main thematic groups, which represent the most important lines of research in the area. The first set, shown in green, is mainly related to computational strategies and methodologies based on deep learning. Likewise, in this set, it can be observed that terms such as "deep learning", "convolutional neural network", "classification", "feature extraction", "image segmentation", and "accuracy" stand out. This indicates that a considerable part of the literature focuses on the creation and assessment of automatic systems to study images of skin lesions linked to PUs.

The second set, indicated in blue, is deeply related to the clinical context and demographic of the disease. In this group, terms such as "pressure ulcer", "decubitus", "adult", "aged", "diagnosis", "human", "female" and "male" can be identified. This shows the focus of research on defining the characteristics of PUs in various ages and population groups, as well as their use in authentic clinical situations.

Finally, the set shown in red is related to general and epidemiological issues of the situation, highlighting concepts such as "pressure injury", "prevalence", "humans", "ulcers" and "journal article". In addition, this group indicates that a section of the literature deals with how PUs affect, how common they are, and how important they are, offering the framework required to create and validate responses supported by CNNs [21]. In summary, this graphical representation reveals a remarkable overlap between clinical methods and computational algorithms, indicating that contemporary research focuses on employing advanced deep learning methods to optimize the identification, categorization, and analysis of PUs [22].

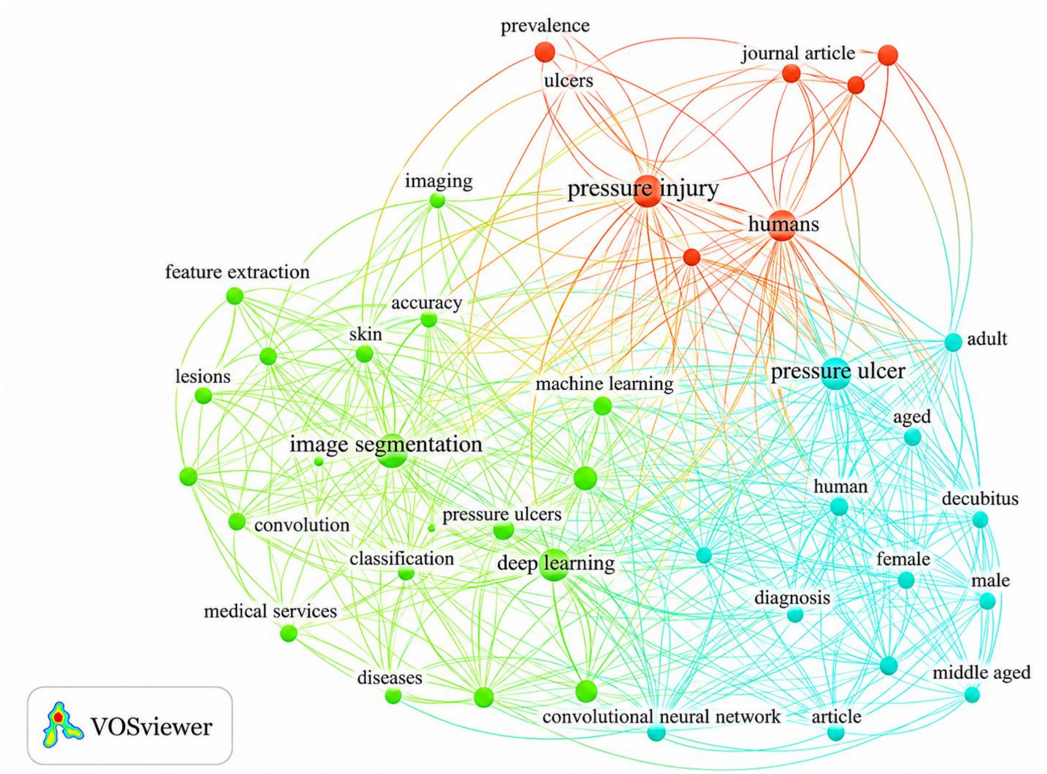

Fig. 2 Keyword concurrency map

4.1.2. Co-Authorship Map

The co-authorship graph makes it easier to analyze the joint work connections between the researchers participating in the study on the use of CNNs in PU research. Figure 3 shows that it generated a group of 5 authors with 10 links and a total strength of 20 links. On the other hand, it can be observed that the co-authorship system derived from the chosen articles, in which each dot symbolizes an author and the relationships between the dots indicate the presence of collaborative works.

In addition, the volume of the nodes indicates the degree of involvement of each author in the group studied, and the thickness of the connections shows the degree of collaboration. As well, the network reveals a limited number of authors who are connected to each other, which indicates that scientific activity in this area is centralized in certain study groups. On the other hand, it shows that among the authors who stand out are "Warraich, A.", "Roberts, D.", "Chalmers, C.", "Henderson, W." and "Fergus, P.", who establish various connections of cooperation between them. This network shows the presence of a dynamic research group that provides to continue researching the application of deep learning in the analysis of PUs.

Similarly, the organization of the map shows that the associations are mostly carried out within this central core, with fewer external connections, which implies a rather restricted collaboration network. On the other hand, this model could be related to the specificity of the topic, in which established research teams lead the advancement of methodologies and the confirmation of models based on CNNs.

In summary, the study on co-authorship demonstrates that research work in this sector is based on close associations between certain authors, which indicates both the specialization of knowledge and the value of cooperation to progress in the analysis of PUs using deep learning methods.

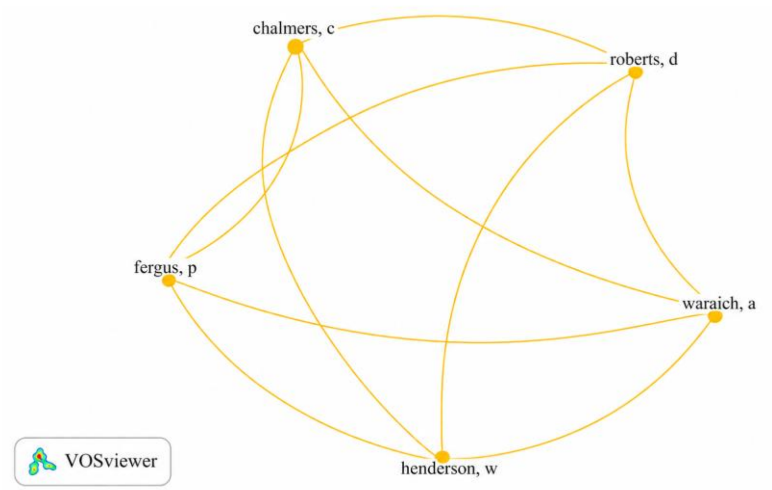

Fig. 3 Co-authorship map

4.1.3. Co-Authorship Map by Country

This graphical representation facilitates the study of global cooperation in research on the use of CNNs to evaluate PUs. Figure 3 shows that two groups were generated distributed in 4 countries with 4 links. In this graph, each node symbolizes a nation and the links indicate that works made jointly by authors from different countries have been published.

Likewise, the volume of the nodes indicates the amount of scientific productivity, while the relationships show the degree of collaboration at the international level.

The network shows that there are a few countries with international collaborations in co-authorship, indicating that most scientific research in this area is carried out in local teams. However, China is seen to play the role of central node in the network, forming direct partnerships with Hong Kong, the United Kingdom, and Pakistan, which indicates that most of the scientific work in this area is carried out in national teams.

In addition, the design of the map reveals the existence of two fundamental groups of cooperation, where China unifies various geographical areas, promoting the exchange of knowledge and collaboration in research projects. Similarly, the existence of restricted connections with other nations indicates that global joint work in the field of PUs supported by AI is still at an early stage.

The analysis of co-authorship worldwide, as a whole, shows a global collaboration network that is not dense, although a clearly predominant node is identified. This highlights the possibility of being able to improve cooperation between countries and thus promote multicenter research that advances in the development and validation of solutions supported by CNNs to investigate PUs.

4.1.4. Annual Evolution of Scientific Production

Figure 4 illustrates the annual progress in scientific research related to the use of CNNs in the analysis of PUs, which was carried out in the Bibliometrix program. In addition, the horizontal axis represents the years in which scientific research was published, and the vertical axis indicates the number of articles that published each year.

In the initial stage, which covers approximately 2013 to 2019, the generation of scientific research is quite scarce, with a reduced or non-existent number of published articles. The curve shows a behavior that indicates that, during those years, the use of deep learning methods for the study of PUs had not become a widely developed research path, remaining in an initial or exploratory phase. Since 2020, there has been a gradual increase in the number of publications, which is consistent with the fact that deep learning techniques are beginning to be adopted more widely in the medical field. Likewise, during the years 2021 to 2024, this increase becomes more noticeable, scientific production shows a constant upward trend, indicating a growing interest in using CNNs to identify, classify and assess PUs.

On the other hand, in 2025, it reaches the maximum peak of scientific production, with the highest amount of research published. This peak shows that the topic has become an active field of research, thanks to advances in computer vision, the availability of clinical data, and the motivation to optimize automated diagnostic aid techniques.

Finally, the most recent year has seen a reduction in the number of publications, which may be due to the fact that the year is not yet over or that some of the recent work is being reviewed or published. All this shows that the set of research that was carried out during this period of time has shown that PUs based on CNNs is an area that has growth potential, where it reflects that its development is more accentuated in recent years.

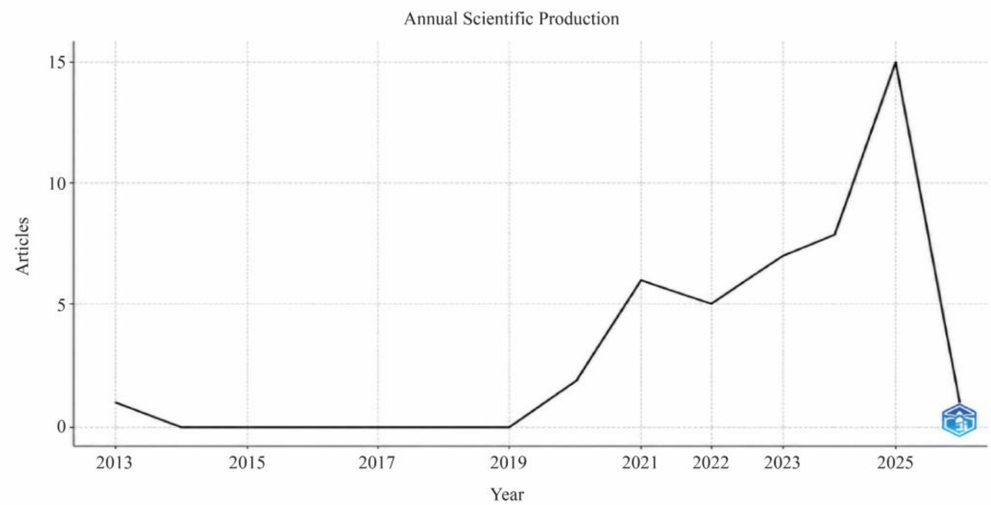

Fig. 4 Evolution of publications by year of scientific journals

4.1.5. Distribution of Publications by Journals

Figure 5 represents the journals published in relation to the PUs linked to the use of CNNs, indicating information regarding the most relevant journals and the number of articles published by each journal. The findings indicate that scientific research is spread across multiple publications, showing the interdisciplinary nature of the topic, including fields such as health informatics, health technology, nursing care, and practical engineering.

The publications of the journal "Studies in Health Technology and Informatics" stand out for being the most articles present a total of five, indicating its importance as a commonplace for the dissemination of studies that combine technology and health. Similarly, the publications of the journals "Applied Sciences (Switzerland)" and "JMIR Medical Informatics" have three articles each, establishing themselves as significant resources in the area. These publications commonly deal with studies focused on the use of new technologies, AI and computer systems in the clinical context, which makes them ideal for research related to CNNs and the examination of medical images [23].

Demonstrating a continuing interest from various disciplines in implementing deep learning methods to investigate PUs, other journals such as "Frontiers in Medical Technology", "JMIR AI", "Nursing Reports" and "Sensors" publish two issues each. Finally, it is identified that journals such as "Advances in Transdisciplinary Engineering" and "BMC Medical Informatics and Decision Making", have only released one study, which indicates that there are concrete but significant contributions in particular situations.

The initial question seeks to determine which techniques based on the use of CNNs have been used in the last twelve years to evaluate PUs. In summary, this distribution indicates that the study of PUs based on CNNs is not limited to a single publication, but extends to several specialized media, which shows the interconnection of the topic and the growing interest of various scientific communities.

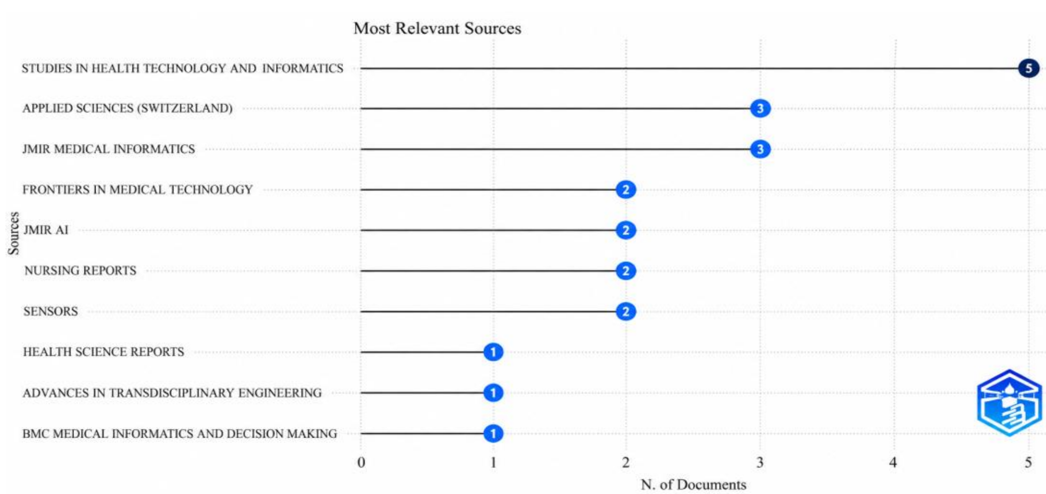

Fig. 5 Distribution by title of published articles

4.2. Content Analysis

The present study included a total of 42 articles selected through the PRISMA methodology, which were chosen and analyzed according to the criteria that had been previously determined. Likewise, this evaluation focused on the review of studies in order to know the level of precision of the proposals or development of technological applications to be able to detect and categorize PUs, that is, to analyze the reports that have been given in the last twelve years.

In order to methodically organize the evaluation of the contents and make it easier to understand the results, the study was organized based on three key questions that guided the collection, categorization and synthesis of the data. In addition, this method facilitated the comparison of the various techniques suggested in the writings, the identification of shared elements and the prevention of thematic dispersion, thus guaranteeing a clear and consistent evaluation.

The initial question seeks to determine which techniques based on the use of CNNs have been used in the last twelve years to evaluate PUs. In this way, this analysis allows the identification of how the procedure for the application of CNNs, image segmentation and the identification of patient lesion levels, in addition to the structures and methods used in different clinical scenarios. The second question investigates the effectiveness of these models in terms of their diagnostic accuracy and confidence level. To this end, performance indicators presented in the research were analyzed, including precision, sensitivity, specificity and accuracy, which facilitates the analysis of the degree of reliability of the suggested alternatives and their possible use in real health care contexts.

Finally, the third question focuses on the future opportunities, challenges, and constraints related to the implementation of CNNs in the clinical field. In this stage, factors such as access and quality of information, the ability to generalize models, the clarity of results, and obstacles to their implementation in health care are studied. Optimization possibilities are also considered to create more robust and usable systems.

Figure 6 shows the structure used for the detailed study of the elements of the chosen research. From the final set of articles, the procedure for reviewing and combining information, which was structured in three key areas corresponding to the purposes of the study. It also provides an overview of how ulcer detection is being carried out and how technologies are supporting health professionals. In addition, it will be known about the advantages or limitations of the applications that have been applied in the clinical environment in the last twelve years. With this scheme, it was possible to logically order the discoveries and facilitate the comparison between the various methods presented in the writings.

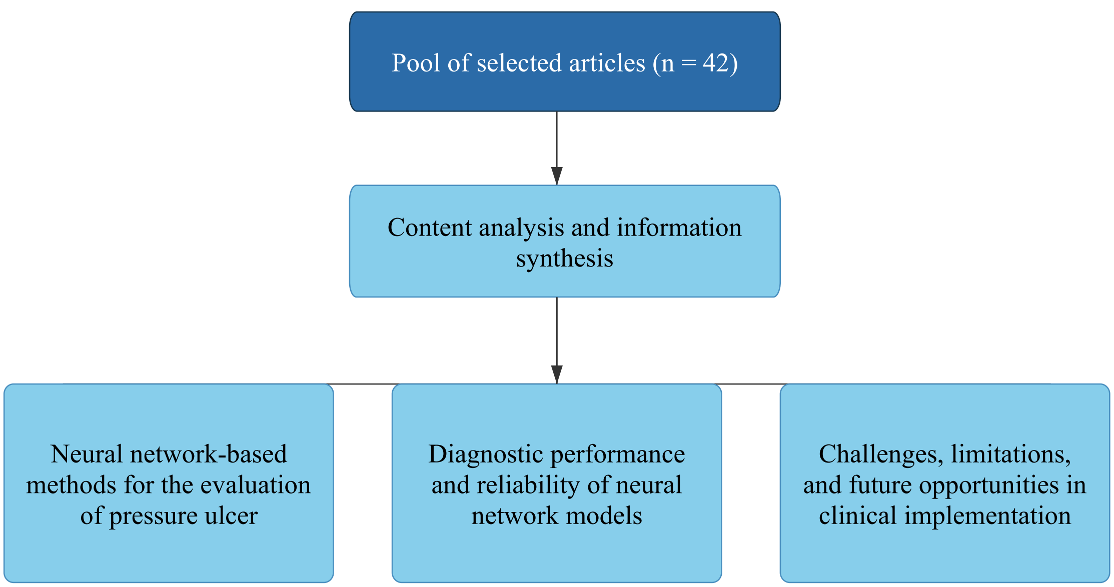

Fig. 6 Structure of systematic review of the contents

4.2.1. CNN-based Methods for Pressure Ulcer Assessment

Traditionally, the clinical assessment of PUs has historically been a process based on observation with the naked eye and the experience of the health professional, which generates a notable level of subjectivity and variation among those who perform the evaluation [24]. In the last twelve years, the development of deep learning methods has provided the opportunity to mechanize this procedure, making it possible to evaluate clinical images objectively, consistently, and with computational support. In this sense, neural networks, especially CNNs, have become one of the most researched technologies to detect, classify and analyze PUs [25].

On the other hand, it was evidenced that 18 of the 42 articles selected for this review directly deal with the use of CNN-based techniques to evaluate PUs. Figure 7 presents a summary of the classification of these studies based on the type of method used, showing a notable superiority of convolutional architectures. The most significant set refers to the CNNs models used in classification activities, which are in 14 publications. In these cases, the nets are mainly used to recognize the appearance of PUs, to differentiate between various clinical levels or to distinguish between skin lesions that are similar. In addition, these methods are based on clinical images acquired in real-world situations, highlighting their ability to identify complex visual patterns linked to transformations in the skin.

On the other hand, 8 studies use CNN for segmentation tasks, which have as their main goal to accurately define the area impacted by the injury. This vision modality is particularly important for clinical follow-up because it facilitates the measurement of ulcer size, examines its progress over time, and supports therapeutic decisions based on objective parameters [26]. Similarly, 5 articles investigate deep learning methods to extract features automatically, using CNNs as instruments to detect significant qualities of the image, such as texture, color or contours, which are then used in classification processes or clinical assessment. These analyses underscore the benefit of eliminating the need to design visual descriptors manually, which decreases human influence on evaluation [27].

Finally, a smaller set of three studies suggests hybrid or combined models, in which CNNs are merged with other machine learning methodologies or clinical guidelines. These suggestions attempt to utilize the advantages of different approaches, strengthening the robustness of the system and its ability to adjust to various clinical situations [28].

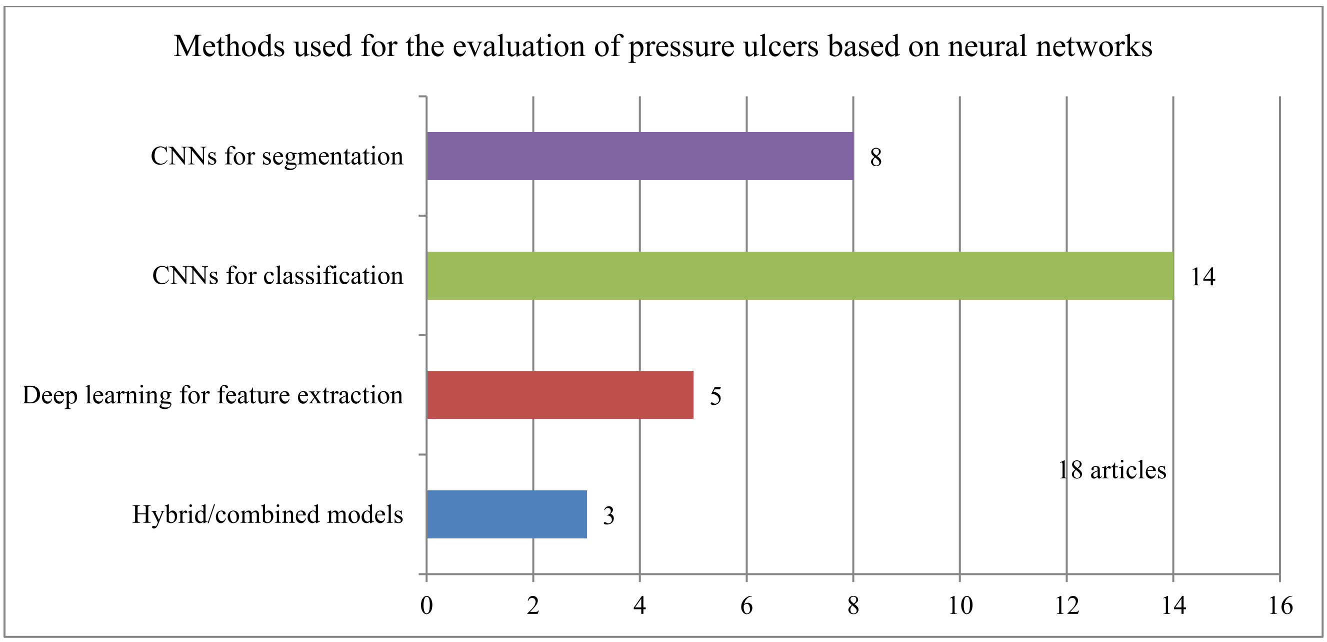

Fig. 7 Distribution of journals on UPP evaluation methods

4.2.2. Diagnostic Performance and Reliability of Neural Network Models

Regardless of the architectural style used, a key factor for the implementation of models based on CNNs in the clinical environment is their effectiveness and diagnostic accuracy. In the field of PUs, these systems need not only to detect lesions, but to do so with an accuracy that exceeds that of conventional clinical evaluation, ensuring that they are consistent, reproducible and useful in real healthcare situations.

Likewise, the analysis of the incorporated research shows that 15 of the 42 documents clearly analyze the effectiveness and reliability of the suggested models. Figure 8 shows how these studies are grouped according to the most prominent evaluation criteria mentioned in the papers.

The most common data is accuracy, which is found in 15 publications, which shows how important it is for the scientific community to measure the overall performance of models. On the other hand, these investigations indicate that CNNs can identify significant visual patterns in photographs of pressure ulcers, reaching constant classifications between the different categories or stages of the injury [29, 30]. In addition, it can be seen that, in nine studies, the evaluation is carried out through measures of sensitivity and specificity, which are seen as essential in the clinical context. Sensitivity helps to be able to measure how well the model can recognize true lesions, while specificity determines whether its ability can prevent incorrect results [31]. Therefore, the incorporation of these measures indicates an interest in reducing significant clinical errors, especially in the initial stages of the disease.

Similarly, it can be observed that 7 studies include clinical validation processes, in which the results obtained from the models are compared with diagnoses made by health experts or with defined clinical standards. This type of verification provides an extra degree of reliability, as it facilitates the evaluation of the system's performance in situations more similar to authentic medical practice. Finally, 5 studies carry out a direct comparative analysis between automated models and manual evaluation, investigating the similarities and differences in the results. These analogies underscore the ability of CNNs to minimize the subjectivity that often accompanies visual assessment, as well as function as auxiliary resources in the clinical decision process, particularly in situations where the evaluator's experience may differ [32].

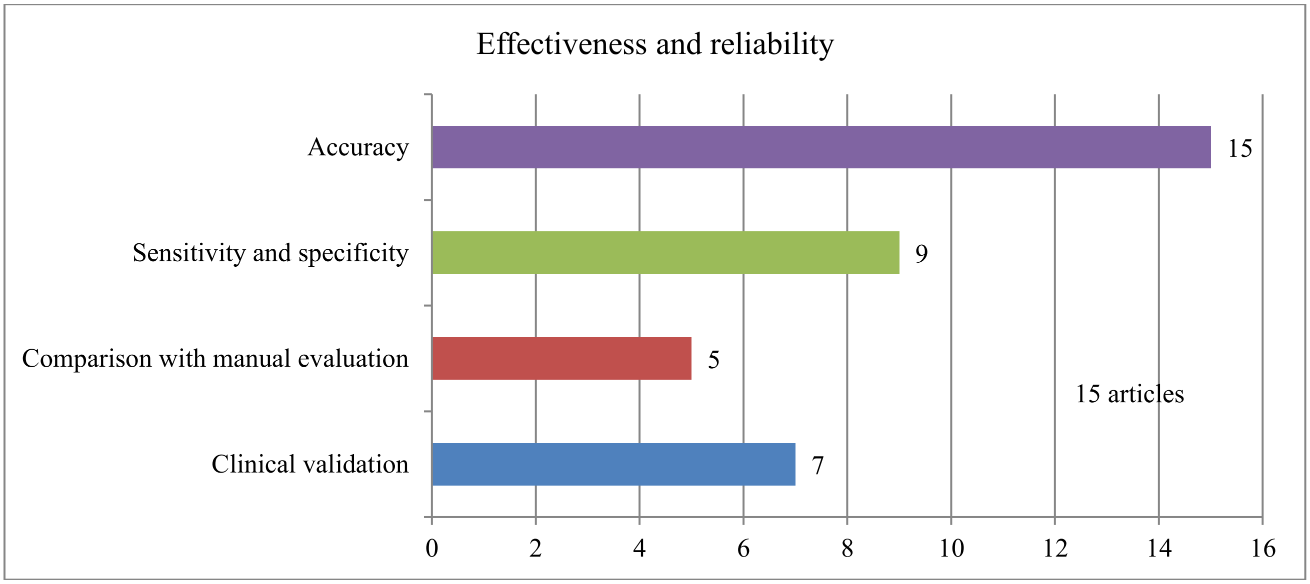

Fig. 8 Distribution of article by effectiveness and reliability of the model

4.2.3. Future Challenges, Limitations, and Opportunities in Clinical Implementation

Although there has been significant progress in the application of CNNs to be able to assess PUs, sources agree that their application still faces several obstacles that restrict their full use in real clinical contexts. Assessing these constraints is crucial not only to understand the current state of research, but also to detect specific opportunities for advancement and future areas of evolution in both technology and clinic.

Of all the studies reviewed, 9 articles directly address the challenges and opportunities linked to the use of models based on CNNs. Figure 9 illustrates how these studies are organized based on the key issues and critical issues identified. The most common limitation has to do with the accessibility and quality level of clinical data, noted in six publications [33]. On the other hand, research indicates that datasets tend to be small, unbalanced, or collected in specific clinical situations, which complicates the task of training robust and representative models. In addition, differences in light, skin type, lesion status, and imaging circumstances directly affect model performance [34].

Another major challenge is the ability of models to generalize, mentioned in 5 posts. Several studies indicate that, although models can perform well in specific datasets, their effectiveness is often reduced when used in different populations or clinical situations that do not resemble those of their training [35]. Even the performance of the models is directly affected by the variability of the lighting, the type of skin, the state of the lesion and the conditions in which the images are taken [36].

In addition, differences in light, skin type, lesion status, and the circumstances under which the images are taken directly affect the effectiveness of the models. The "black box" feature of numerous deep learning models makes it difficult for healthcare workers to understand the process behind system decisions, which can influence clinical confidence and restrict their implementation [37]. Along these lines, several writers stress the relevance of adding explanatory methods that make it possible to see which areas of the image have an impact on diagnosis [38].

Finally, three papers discuss the challenges associated with implementation in real-world situations, covering issues such as connection to health systems, computing constraints, processing time, and approval by health personnel. These analyses emphasize that the success of these innovations is not only based on their accuracy, but also on their functionality in daily work and their simplicity of use in everyday healthcare [39].

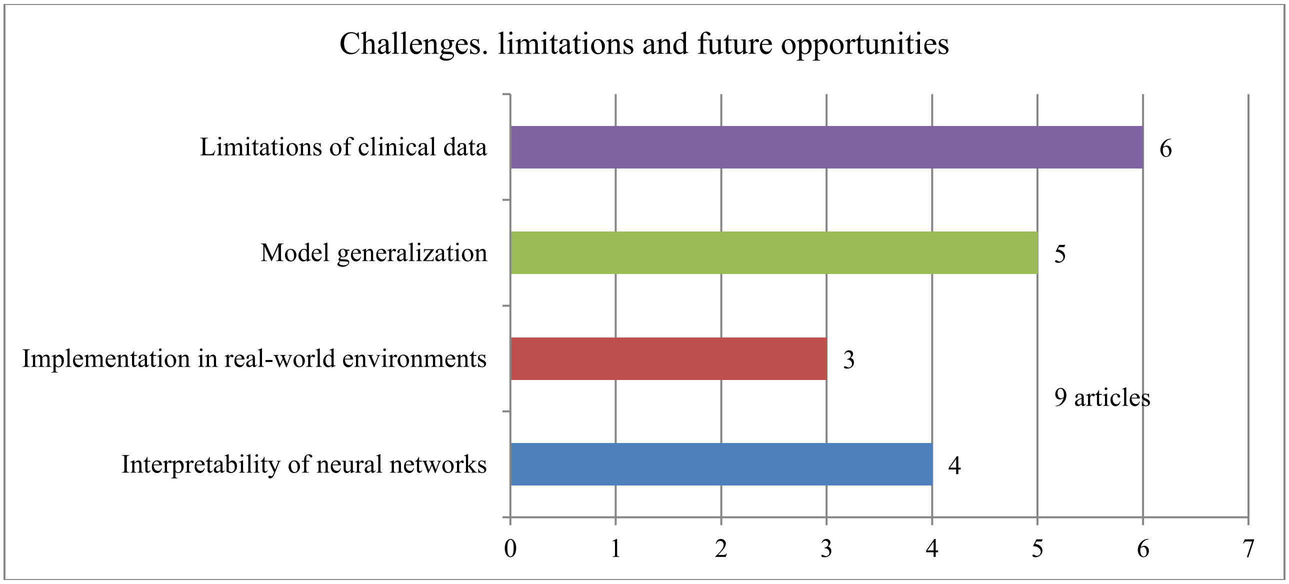

Fig. 9 Item distribution by future challenges, constraints, and opportunities

4.3. Features y Performance of the Proposed Models for UPPs

In order to coherently organize the most relevant technological contributions found in this study, a comparative examination of research presenting models based on CNN for the prevention, identification and classification of PUs was carried out. Table 1 shows the main qualities of each of the technological projects created, covering the approach used, the clinical activity treated, and the performance metrics mentioned.

This organization makes it possible to objectively measure how developed the suggested solutions are, and also makes it easy to compare their effectiveness and reliability based on metrics that are widely recognized in the field of machine learning and medical research.

In general, according to the findings found in Table 1, most of the proposed models have demonstrated high performance, since it can be visualized that there are accuracy rates greater than 95% in several studies. Likewise, the best results were obtained by Song et al. [29], where the Random Forest model reached an accuracy of 99.88%. On the other hand, image-based approaches such as Inception-ResNet-v2 [10], Vision Transformer [34] and DenseNet121 [32] have demonstrated a high performance of 93%.

The analysis also revealed a high interest in alternative approaches, such as thermography [6], 3D wound modelling [27] and mobile applications based on YOLO [22, 46]. However, despite the results found, the vast majority of studies share common limitations in relation to small datasets, the use of images from a single health centre, and insufficient external validation. This would indicate the need for more research to improve the generalizability of the models and their implementation in real clinical settings.

5. Discussions

The systematic review presented examined and contrasted recent work using AI and machine learning methods to identify, classify, assess, and predict PUs. It also showed remarkable progress in the application of models based on deep learning, especially CNNs, as well as machine learning methods that are easy to understand [41].

It should be noted that the results obtained are consistent with previous studies presented in the literature review section. The study of [9] reported that state-of-the-art deep learning models are capable of achieving AUC values above 0.90 in various medical imaging tasks, which is consistent with the high throughput rates found in this review. Similarly, the studies of [10, 11] demonstrated the potential of convolutional neural networks and transfer learning methods for automated analysis of medical images, but the results of the present analysis showed that the CNN, DenseNet, Vision Transformer, and Inception-ResNet-v2 architectures regularly achieve accuracy rates above 90% for assessing and classifying pressure ulcers. In addition, the findings they obtained [12] regarding the clinical value of intelligent decision support systems are also supported by the studies reviewed, as they demonstrate the great potential of AI for improving diagnostic objectivity and timely detection of pressure ulcer risk.

Similarly, some research has included sophisticated methods for detection and segmentation, such as Faster R-CNN, Mask R-CNN, and U-Net, which seek not only to identify the lesion, but also to locate the affected area and measure important clinical features [42, 43]. In these situations, indicators such as Dice score and the IoU turned out to be more suitable than traditional accuracy, justifying the variety seen in the presentation of metrics across the studies examined. On the other hand, this variability prevents the models from being compared directly. However, it shows the variety of clinical objectives that are treated (structural evaluation, segmentation, detection or classification).

On the other hand, we found studies that applied different imaging techniques, such as ultrasound and infrared thermography, mainly focused on the early identification of deep lesions and their prevention [44, 45]. These analyses also demonstrate that AI is able to identify changes under the skin that cannot be observed in a superficial review, which is a significant clinical benefit. However, these methodologies often need specific tools and reduced data sets, limiting their direct use in resource-poor environments.

Regarding prognostic models that use organized clinical data, such as Random Forest, XGBoost and various machine learning algorithms, very high returns were recorded, with figures that almost reach 100% in certain metrics [46]. However, it is important to analyze these findings with caution, as multiple studies point to the possibility of excessive adjustment, the use of data from a single center, and the lack of external validation. Despite this, these models are recognized for their ability to predict the appearance of PUs, which could be very valuable in advance notification mechanisms.

An important point made in this review is the limited uniformity in datasets, assessment metrics, and validation protocols, which makes it difficult to make consistent comparisons between different investigations. In addition, many of the studies have a small number of participants, retrospective analyses and internal verifications, which restricts the possibility of applying the findings to authentic clinical situations and to different population groups.

Finally, although the results are encouraging, the information indicates that the general adoption of these systems in the clinical setting still needs prospective validations in multiple centers, adaptation to hospital procedures and analysis of the real effect on medical decision-making. On the other hand, the inclusion of understandable methods, such as SHAP, and the creation of simple mobile applications are key areas of research to promote acceptance in clinical practice and build trust in health professionals.

However, it is important to analyze these findings with caution, as multiple studies point to the possibility of over-tuning, the use of data from a single center, and the lack of external validation. Despite this, these models are recognized for their ability to predict the appearance of PUs, which could be very valuable in advance notification mechanisms.

An important point made in this review is the limited uniformity in datasets, assessment metrics, and validation protocols, which makes it difficult to make consistent comparisons between different investigations. In addition, many of the studies have a small number of participants, retrospective analyses and internal verifications, which restricts the possibility of applying the findings to authentic clinical situations and to different population groups.

Finally, although the results are encouraging, the information indicates that the general adoption of these systems in the clinical setting still needs prospective validations in multiple centers, adaptation to hospital procedures and analysis of the real effect on medical decision-making.

6. Conclusion

This systematic review shows that AI techniques, especially models based on CNNs, constitute a tool with great potential to detect, classify, evaluate and predict PUs. Likewise, the findings that were analyzed indicate that strategies that use clinical images in two dimensions achieve excellent performance in the automatic classification of stages, while models to detect and segment provide a more exhaustive assessment of the lesion and its scope.

Similarly, the use of other types of imaging, such as ultrasound and infrared thermal imaging, showed a great ability to identify deep lesions in early stages, providing subcutaneous data that cannot be seen with simple visual observation. However, its use in medical practice is restricted due to the requirement for specific equipment and the small volume of data used in current research.

On the other hand, it was evidenced that predictive algorithms, such as Random Forest and XGBoost, which use organized clinical data, demonstrated an outstanding performance in the early detection of the risk of forming PUs. However, it is important to consider these results with caution due to the risk of overfitting, the lack of external validations, and the fact that the data come exclusively from a single center. Despite the progress made, significant challenges remain linked to the lack of standardization in datasets, assessment metrics, and validation protocols, which complicates direct comparison between investigations and restricts the generalizability of findings. Likewise, the need for multicenter and prospective research is emphasized because most of the studies analyzed are based on small samples and retrospective evaluations.

In summary, the inclusion of systems that use AI in the evaluation of PUs can increase the accuracy of diagnoses, facilitate the clinical decision-making process and reinforce prevention measures. However, to achieve successful clinical adoption, it is essential to progress towards models that are robust, understandable, and approved in real-world situations, as well as to encourage the creation of accessible solutions that can be safely and effectively incorporated into healthcare.

6.1. Gaps in the Literature

Although significant progress has been made in the application of AI and machine learning to assess PUs, the existing knowledge currently has several clinical, technical, and methodological gaps that hinder the adoption and generalization of these solutions in medical practice. First, there is a lack of standardization in the datasets used to validate and train the models. On the other hand, the direct comparison of results and the possibility of repeating the proposed models are hindered by the fact that the reviewed research uses diverse databases in terms of size, type of image (ultrasound, thermography, 2D), criteria for clinical inclusion and annotation.

In addition, there is a lack of agreement in the evaluation metrics that have been reported. While certain studies use traditional metrics such as precision, accuracy, recall, and F1 score, others opt for specific segmentation indicators such as the Dice score or the IoU, and even resort to clinical metrics that are not directly comparable. In addition, this variety makes it difficult to reach an unambiguous agreement on the true performance of the models and restricts the ability to make uniform quantitative comparisons. The lack of external validations and multicenter prospective research is another significant shortcoming. It was evidenced that the risk of overfitting increases and the ability to generalize the models to diverse populations and clinical contexts decreases, since most of the studies analyzed are based on retrospective evaluations, with small samples and from a single center. A restricted clinical integration of the suggested solutions is noted. While several studies show functional prototypes or mobile apps, not many systematically examine how these tools affect clinical decision-making, the reduction of PUs, or progress in patient outcomes. On the other hand, the explainability of the models continues to be an aspect that has not been dealt with much. Although some interpretation functions, such as SHAP, are integrated into certain jobs, most deep learning models function as "black boxes," which could create distrust among health experts and hinder their clinical acceptance. Ultimately, there is limited research of hybrid methods that incorporate structured clinical data, multimodal imaging, and contextual variables. On the other hand, most research focuses on a single type of data, which limits the use of the potential of multimodal models for a more complete and accurate assessment of PUs.

6.2. Recommendations for Future Studies

Conduct prospective and longitudinal research that examines the performance of AI models over time, in order to assess their ability to identify the evolution, improvement, or prevention of PUs in various care settings. Develop and use large, standardized datasets, with expert clinical annotations, to enable unbiased comparison between models and improve generalizability and reproducibility. External and multicenter validation of the models presented should be deepened, including varied populations, different imaging devices, as well as a diversity of clinical and home situations.

Investigate multimodal methods that incorporate clinical imaging, ultrasound, thermography, and organized clinical data to increase diagnostic accuracy and the predictive capacity of AI-based systems. Finally, integrate privacy, data protection and explainability procedures, ensuring that regulatory and ethical principles are complied with, and promoting clinical acceptance by healthcare personnel of these technologies.

Conflicts of Interest

The authors declare that there were no conflicts of interest during the preparation and publication of this study.

Funding Statement

This study was not funded by any entity.

Acknowledgments

Acknowledgements are given to the University of Sciences and Humanities (UCH) for its support for science, especially to the Research Directorate and the INTI-Lab Laboratory for their constant support.

References

[1] A Dimitri Beeckman et al., "EPUAP Classification System for Pressure Ulcers: European Reliability Study," *Journal of Advanced Nursing*, vol. 60, no. 6, pp. 682-691, 2007. [CrossRef] [Google Scholar] [Publisher Link]

[2] Susanne Coleman et al., "Patient Risk Factors for Pressure Ulcer Development: Systematic Review," *International Journal of Nursing Studies*, vol. 50, no. 7, pp. 974-1003, 2013. [CrossRef] [Google Scholar] [Publisher Link]

[3] Rehabilitation, World Health Organization (WHO), 2023. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/rehabilitation

[4] Ning Wang et al., "Biomarkers for the Early Detection of Pressure Injury: A Systematic Review and Meta-Analysis," *Journal of Tissue Viability*, vol. 31, no. 2, pp. 259-267, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[5] Pınar Küçükdemirci Kaya et al., "Sepsis Episodes Caused by Pressure Injuries in Critical Illness: A Retrospective Observational Cohort Study," *Wound Management & Prevention*, vol. 69, no. 4, pp. 4-9, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[6] Yu Wang et al., "Infrared Thermal Images Classification for Pressure Injury Prevention Incorporating the Convolutional Neural Networks," *IEEE Access*, vol. 9, pp. 15181-15190, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[7] Paul Fergus et al., “Pressure Ulcer Categorization and Reporting in Domiciliary Settings Using Deep Learning and Mobile Devices: A Clinical Trial to Evaluate End-to-End Performance,” *IEEE Access*, vol. 11, pp. 65138-65152, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[8] Xiaoxuan Liu et al., “A Comparison of Deep Learning Performance Against Health-Care Professionals in Detecting Diseases from Medical Imaging: A Systematic Review and Meta-Analysis,” *The Lancet Digital Health*, vol. 1, no. 6, pp. 271-297, 2019. [CrossRef] [Google Scholar] [Publisher Link]

[9] Jemin Kim et al., “Augmented Decision-Making in Wound Care: Evaluating the Clinical Utility of a Deep-Learning Model for Pressure Injury Staging,” *International Journal of Medical Informatics*, vol. 180, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[10] Betul Ay et al., “Deep Transfer Learning-Based Visual Classification of Pressure Injuries Stages,” *Neural Computing and Applications*, vol. 34, no. 18, pp. 16157-16168, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[11] Tom J. Liu et al., “A Pressure Ulcers Assessment System for Diagnosis and Decision Making using Convolutional Neural Networks,” *Journal of the Formosan Medical Association*, vol. 121, no. 11, pp. 2227-2236, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[12] Debajyoti Pati, and Lesa N. Lorusso, “How to Write a Systematic Review of the Literature,” *HERD: Health Environments Research & Design Journal*, vol. 11, no. 1, pp. 15-30, 2018. [CrossRef] [Google Scholar] [Publisher Link]

[13] Matthew J. Page et al., “The PRISMA 2020 Statement: An Updated Guideline for Reporting Systematic Reviews,” *BMJ*, vol. 372, pp. 1-9, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[14] Michael Gusenbauer, and Neal R. Haddaway, “Which Academic Search Systems are Suitable for Systematic Reviews or Meta-Analyses? Evaluating Retrieval Qualities of Google Scholar, PubMed, and 26 Other Resources,” *Research Synthesis Methods*, vol. 11, no. 2, pp. 181-217, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[15] Nees Jan van Eck, and Ludo Waltman, “Software Survey: VOSviewer, A Computer Program for Bibliometric Mapping,” *Scientometrics*, vol. 84, no. 2, pp. 523-538, 2010. [CrossRef] [Google Scholar] [Publisher Link]

[16] Massimo Aria, and Corrado Cuccurullo, “bibliometrix: An R-tool for Comprehensive Science Mapping Analysis,” *Journal of Informetrics*, vol. 11, no. 4, pp. 959-975, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[17] Borui Chen et al., “Trends and Hotspots in Research on Medical Images with Deep Learning: A Bibliometric Analysis from 2013 to 2023,” *Frontiers in Artificial Intelligence*, vol. 6, pp. 1-14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[18] Silvia Caggiari et al., “Biomechanical Monitoring and Machine Learning for the Detection Of Lying Postures,” *Clinical Biomechanics*, vol. 80, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[19] Dinggang Shen, Guorong Wu, and Heung-Il Suk, “Deep Learning in Medical Image Analysis,” *Annual Review of Biomedical Engineering*, vol. 19, pp. 221-248, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[20] Samia Rahman et al., “A Comprehensive Scoping Review on the Use of Point-Of-Care Infrared Thermography,” *International Wound Journal*, vol. 22, no. 8, pp. 1-27, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[21] Nikola Pupic et al., “Detecting Patient Position Using Bed-Reaction Forces for Pressure Injury Prevention and Management,” *Sensors*, vol. 24, no. 19, pp. 1-20, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[22] Yujee Chang, “Diagnosis of Pressure Ulcer Stage Using On-Device AI,” *Applied Sciences*, vol. 14, no. 16, pp. 1-14, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[23] Guo Li et al., “Advances and Trends in Pressure Ulcer Care Research Over the Last 20 Years: A Bibliometric and Visual Analysis,” *Heliyon*, vol. 10, no. 19, pp. 1-8, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[24] Eba’a Dasan Barghouthi et al., “Systematic Review for Risks of Pressure Injury and Prediction Models Using Machine Learning Algorithms,” *Diagnostics*, vol. 13, no. 17, pp. 1-16, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[25] Khalid Majjouti et al., “Nursing-Centered Development of an AI-based Decision Support System in Pressure Ulcer and Incontinence-Associated Dermatitis Management - A Mixed Methods Study,” *BMC Nursing*, vol. 24, no. 1, pp. 1-9, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[26] Che Wei Chang et al., “Deep Learning Approach Based on Superpixel Segmentation Assisted Labeling for Automatic Pressure Ulcer Diagnosis,” *Plos One*, vol. 17, no. 2, pp. 1-19, 2022. [CrossRef] [Google Scholar] [Publisher Link]

[27] Sofia Zahia, Begonya Garcia-Zapirain, and Adel Elmaghraby, “Integrating 3D Model Representation for an Accurate Non-Invasive Assessment of Pressure Injuries with Deep Learning,” *Sensors*, vol. 20, no. 10, pp. 1-15, 2020. [CrossRef] [Google Scholar] [Publisher Link]

[28] Xinghuo Ye, and Na Wang, “SLGMA-UNet: Comprehensive Feature Aggregation with Context-Sensitive Attention for Medical Image Segmentation,” *IEEE Access*, vol. 13, pp. 133457-133464, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[29] Jie Song et al., “The Random Forest Model Has the Best Accuracy Among the Four Pressure Ulcer Prediction Models Using Machine Learning Algorithms,” *Risk Management and Healthcare Policy*, vol. 14, pp. 1175-1187, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[30] Masaru Matsumoto et al., “Development of an Automatic Ultrasound Image Classification System for Pressure Injury based on Deep Learning,” *Applied Sciences*, vol. 11, no. 17, pp. 1-9, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[31] Cemil Zalluhoğlu et al., "Region-Based Semi-Two-Stream Convolutional Neural Networks for Pressure Ulcer Recognition," *Journal of Imaging Informatics in Medicine*, vol. 37, no. 2, pp. 801-813, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[32] Changbin Lei et al., "Convolutional Neural Network Models for Visual Classification of Pressure Ulcer Stages: Cross-Sectional Study," *JMIR Medical Informatics*, vol. 13, pp. 1-14, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[33] Shreenidhi Jogi et al., "Design and Development of an mHealth Application for Pressure Ulcer Care and Caregiver Support," *Frontiers in Digital Health*, vol. 7, pp. 1-19, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[34] Young-Bok Cho, and Hana Yoo, "Development of a Pressure Ulcer Stage Determination System for Community Healthcare Providers using a Vision Transformer Deep Learning Model," *Medicine*, vol. 104, no. 7, pp. 1-8, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[35] Yijie Qian et al., "Development of an Explainable Machine Learning Model for Predicting Device-Related Pressure Injuries in Clinical Settings," *BMC Medical Informatics and Decision Making*, vol. 25, no. 1, pp. 1-12, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[36] Marin Benčević et al., "Understanding Skin Color Bias in Deep Learning-Based Skin Lesion Segmentation," *Computer Methods and Programs in Biomedicine*, vol. 245, pp. 1-10, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[37] Dost Muhammad, and Malika Bendechache, "Unveiling the Black Box: A Systematic Review of Explainable Artificial Intelligence in Medical Image Analysis," *Computational and Structural Biotechnology Journal*, vol. 24, pp. 542-560, 2024. [CrossRef] [Google Scholar] [Publisher Link]

[38] Ramprasaath R. Selvaraju et al., "Grad-CAM: Visual Explanations from Deep Networks via Gradient-Based Localization," *2017 IEEE International Conference on Computer Vision (ICCV)*, Venice, Italy, pp. 618-626, 2017. [CrossRef] [Google Scholar] [Publisher Link]

[39] Molla Imaduddin Ahmed et al., "A Systematic Review of the Barriers to the Implementation of Artificial Intelligence in Healthcare," *Cureus*, vol. 15, no. 10, pp. 1-14, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[40] Fitri Anita, Nur Chayati, and Wantonoro Wantonoro, "Pressure-Relieving Devices to Prevent Pressure Injury in Bedridden Patients: A Literature Review," *Bali Medical Journal*, vol. 10, no. 3, pp. 1357-1363, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[41] Odai Y. Dweekat, Sarah S. Lam, and Lindsay McGrath, "Machine Learning Techniques, Applications, and Potential Future Opportunities in Pressure Injuries (Bedsores) Management: A Systematic Review," *International Journal of Environmental Research and Public Health*, vol. 20, no. 1, pp. 1-46, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[42] Tom J. Liu et al., "Automatic Segmentation and Measurement of Pressure Injuries using Deep Learning Models and a LiDAR Camera," *Scientific Reports*, vol. 13, pp. 1-13, 2023. [CrossRef] [Google Scholar] [Publisher Link]

[43] Abdulrahman Dira Khalaf et al., "Segmentation and Classification of Skin Cancer Diseases Based on Deep Learning: Challenges and Future Directions," *IEEE Access*, pp. 90163-90184, 2025. [CrossRef] [Google Scholar] [Publisher Link]

[44] Noriyuki Aoi et al., "Ultrasound Assessment of Deep Tissue Injury in Pressure Ulcers: Possible Prediction of Pressure Ulcer Progression," *Plastic and Reconstructive Surgery*, vol. 124, no. 2, pp. 540-550, 2009. [CrossRef] [Google Scholar] [Publisher Link]

[45] Fuman Cai MS et al., "Application of Infrared Thermography in the Early Warning of Pressure Injury: A Prospective Observational Study," *Journal of Clinical Nursing*, vol. 30, no. 3-4, pp. 559-571, 2021. [CrossRef] [Google Scholar] [Publisher Link]

[46] Chun Hon Lau et al., "An Artificial Intelligence-Enabled Smartphone App for Real-Time Pressure Injury Assessment," *Frontiers in Medical Technology*, vol. 4, pp. 1-10, 2022. [CrossRef] [Google Scholar] [Publisher Link]

Appendix

Table 1. Proposed model studies for UPPS

No.	Author	Year	Article Title	Technology Solution	Metrics				Limitations
					Accuracy	Accuracy	F1-score	Recall	
[34]	Young-Bok Cho and H. Yoo	2025	Development of a pressure ulcer stage determination system for community healthcare providers using a vision transformer deep learning model.	UPP Stadium Classification System with Vision Transformer (PUC-ViT) using real photos; proposes an approach to operate on low computing power (smart device/IoT).	97.76	96.02	95.46	90.83	Not fully validated on different camera models; uses 2D photos without depth (possible enhancement with 3D).
[32]	C. Lei et al.	2025	Convolutional Neural Network Models for Visual Classification of Pressure Ulcer Stages: Cross-Sectional Study.	Visual classification of pressure injury stages (I, II, III, IV, Unstageable, SDTI) comparing AlexNet, VGG16, ResNet18 and DenseNet121 with clinical images.	93.71	98.72	97.09	91.53	It includes normal skin due to segmentation, the wound needs to be made more prominent; does not integrate 3D/ thermography/fluids; limited prospective data and use of a single device.
[11]	T. J. Liu et al.	2022	A pressure ulcers assessment system for diagnosis and decision making using convolutional neural networks.	Clinical assessment system (PUAS) based on CNN Inception-ResNet-v2, integrated into a mobile app, which classifies erythema/non-erythema and degree of necrosis to support clinical decision-making in non-specialist caregivers.	98.5	97.7	90.96	95.5	Retrospective study with limited number of images; possible overfitting; it does not incorporate 3D information (depth, tunnels or sinuses); requires prospective studies and extended clinical validation.
[46]	C. H. Lau et al.	2022	An artificial intelligence-enabled smartphone app for real-time pressure injury assessment	Android mobile application for real-time detection and classification of UPP stadiums using YOLOv4 (object detection), trained with public images and	63.2	Not Registered	Not Registered	0.73	(

				validated in simulated scenarios of real use.					
[27]	S. Zahia, B. Garcia-Zapirain, and A. Elmaghraby	2020	Integrating 3D Model Representation for an Accurate Non-Invasive Assessment of Pressure Injuries with Deep Learning	End-to-end framework that combines Mask R-CNN (2D) + 3D mesh (Structure Sensor) to segment the lesion and calculate real measurements (depth, area, volume, axes).	Not Registered	0.87	0.83	0.85	Requires external 3D sensor (Structure Sensor + iPad); not all cases are suitable for volume (concave surfaces/locations such as heel/hip → invalid tights); ambiguous edges affect segmentation and can "widen" mask.
[6]	Y. Wang et al.	2021	Infrared Thermal Images Classification for Pressure Injury Prevention, Incorporating the Convolutional Neural Networks	Binary classification Normal vs PI using infrared thermal imaging (FLIR ONE PRO) + CNN for early detection (includes "1 day before" PI dataset).	95.2	Not Registered	Not Registered	96.67	Small dataset (82 PI + 82 normal), and the "1 day before" set was treated as PI by hypothesis; manual ROI segmentation (subjective and slow); They recommend collecting more images at different times.
[29]	J. Song et al.	2021	The Random Forest Model Has the Best Accuracy Among the Four Pressure Ulcer Prediction Models Using Machine Learning Algorithms	Models for predicting PU events in hospitalized patients (n=5814) using SVM, Decision Tree, Random Forest and ANN with 19 variables (48h previously). Better performance is reported with Random Forest.	99.88%	99.83%	99.88%	99.88%	Risk of overfitting (they themselves indicate it by "excessively high" values); single-center studio; EHR base with non-standardized descriptions (no images), and missing factors/risks not covered by the DB.
[22]	Yujee Chang et al.	2024	Diagnosis of Pressure Ulcer Stage Using On-Device AI	Mobile on-device platform (Android/Kotlin) with YOLOv8 model to detect the PU region and classify into 6 severities (stage 1–4, DTPI, unstageable); returns result in ~3 seconds.	0.846	0.897	0.894	0.891	They must improve robustness, reliability, and generalizability (improvements in lightweight models, processing, data augmentation, and collecting more diverse images in populations/camera conditions). In addition, they used public dataset.
[30]	M. Matsumoto et al.	2021	Development of an Automatic Ultrasound Image	Automatic system based on U-Net + ResNeXt-50 for	Not Registered	Not Registered	0.89	71.4	Small sample (73 images of 5 patients); confusion between cobblestone vs

			Classification System for Pressure Injury Based on Deep Learning	segmentation and classification of ultrasound findings of pressure injuries (unclear structure, cobblestone pattern, cloudy and anechoic pattern) aimed at detecting DTPIs by non-specialists.					cloud-like patterns; data only from Japanese older adults; assessment at a single time point (does not predict progression); unmatched by other SOTA models.
[35]	Y. Qian et al.	2025	Development of an explainable machine learning model for predicting device-related pressure injuries in clinical settings	Explainable model for predicting PRD using XGBoost (and compare RF, DT, LR, SVM, KNN). Interpretation with SHAP for the importance of variables (top: stay, type of instrument, emergency admission, material, duration).	0.919	Not Registered	Not Registered	0.957	Retrospective study, single center, and without large-scale external validation; limited generalization.