



OPEN ACCESS

EDITED BY

Andreas Kanavos,
Ionian University, Greece

REVIEWED BY

Emre Sefer,
Özyegin University, Türkiye
Hamidreza Bolhasani,
Islamic Azad University, Iran

*CORRESPONDENCE

Lloy Pinedo
✉ lloy.pinedo@uwiener.edu.pe

RECEIVED 03 August 2025

REVISED 25 September 2025

ACCEPTED 10 November 2025

PUBLISHED 02 January 2026

CITATION

Valles-Coral MA, Pinedo L, Rodríguez C,
Rodríguez D, Sánchez-Dávila K,
Arévalo-Fasanando L and Reátegui-Lozano N
(2026) Application of artificial intelligence in
cervical cytology: a systematic review of deep
learning models, datasets, and reported
metrics. *Front. Big Data* 8:1678863.
doi: 10.3389/fdata.2025.1678863

COPYRIGHT

© 2026 Valles-Coral, Pinedo, Rodríguez,
Rodríguez, Sánchez-Dávila,
Arévalo-Fasanando and Reátegui-Lozano.
This is an open-access article distributed
under the terms of the [Creative Commons
Attribution License \(CC BY\)](#). The use,
distribution or reproduction in other forums is
permitted, provided the original author(s) and
the copyright owner(s) are credited and that
the original publication in this journal is cited,
in accordance with accepted academic
practice. No use, distribution or reproduction
is permitted which does not comply with
these terms.

Application of artificial intelligence in cervical cytology: a systematic review of deep learning models, datasets, and reported metrics

Miguel Angel Valles-Coral^{1,2}, Lloy Pinedo^{3*}, Ciro Rodríguez¹,
Diego Rodríguez⁴, Keller Sánchez-Dávila⁵,
Lolita Arévalo-Fasanando⁶ and Nelly Reátegui-Lozano⁶

¹Facultad de Ingeniería de Sistemas e Informática, Universidad Nacional Mayor de San Marcos, Lima, Peru, ²Facultad de Ingeniería de Sistemas e Informática, Universidad Nacional de San Martín, Tarapoto, Peru, ³Facultad de Ingeniería y Negocios, Universidad Privada Norbert Wiener, Lima, Peru, ⁴Facultad de Ciencias de la Salud, Medicine, Universidad Peruana de Ciencias Aplicadas (UPC), Lima, Peru, ⁵Facultad de Medicina Humana, Universidad Nacional de San Martín, Tarapoto, Peru, ⁶Facultad de Ciencias de la Salud, Universidad Nacional de San Martín, Tarapoto, Peru

Introduction: The use of artificial intelligence (AI) in cervical cytology has increased substantially due to the need for automated tools that support the early detection of precancerous lesions.

Methods: This systematic review examined deep learning models applied to cervical cytology images, focusing on the architectures used, the datasets employed, and the performance metrics reported. Articles published between 2022 and 2025 were retrieved from Scopus using PRISMA methodology. After applying inclusion criteria and full-text screening, 77 studies were included for RQ1 (models), 75 for RQ2 (datasets), and 71 for RQ3 (metrics).

Results: Hybrid models were the most prevalent (56%), followed by convolutional neural networks (CNNs) and a growing number of Vision Transformer (ViT)-based approaches. SIPaKMed and Herlev were the most frequently used datasets, although the use of private datasets is increasing. Accuracy was the most commonly reported metric (mean 87.76%), followed by precision, recall, and F1-score. Several hybrid and ViT-based models exceeded 92% accuracy. Identified limitations included limited cross-validation, reduced clinical representativeness of datasets, and inconsistent diagnostic criteria.

Discussion: This review synthesizes current trends in AI-based cervical cytology, highlights common methodological limitations, and proposes directions for future research to enhance clinical applicability and standardization.

KEYWORDS

cervical cytology, cancer, deep learning, models, datasets, metrics

1 Introduction

Cervical cancer remains one of the leading causes of death among women worldwide, particularly in countries with limited access to healthcare services (Torres-Roman et al., 2021). The Papanicolaou test has, for decades, enabled early detection of cellular abnormalities in the cervix, helping to prevent their progression to invasive cancer

([Papanicolaou and Traut, 1941](#)). This review aims to systematically analyze studies that apply artificial intelligence (AI) in cervical cytology, focusing on the models and datasets used, as well as the main performance outcomes.

Computational solutions in medicine have evolved from simple heuristic systems based on rule sets to more complex deep learning models, particularly in medical imaging ([Cabral et al., 2025](#); [Bolia and Joshi, 2025](#)). Currently, with reduced computational costs, there is increasing interest in hybrid architectures that combine convolutional neural networks (CNNs) with vision transformer (ViT)-based models, which exhibit superior ability to identify complex patterns in cellular images and aid in cytopathological diagnosis ([Maurya et al., 2023](#); [Hong et al., 2024](#)). Recent advances in AI have also demonstrated applications beyond cytology, such as transcriptomic event inference in cancer cells and drug response prediction using graph-based models ([Eralp and Sefer, 2024](#); [Sefer, 2025](#)). These developments highlight the broad potential of AI in oncology and reinforce the relevance of its application to cervical cytology. Despite these advances, important limitations remain, such as the flawed assumption that cell classification alone is sufficient for cancer diagnosis, the poor quality and representativeness of the datasets used, and the excessive complexity of some models.

The use of AI in cervical cytology seeks to assist in the automatic identification of suspicious cytological lesions ([Tang et al., 2023](#); [Shinde et al., 2022](#)), as part of an automated pipeline for early detection of cellular abnormalities. This involves computer vision models analyzing microscopic images, classifying cells based on morphological features, and diagnosing cellular lesion levels, primarily using supervised learning algorithms ([Mohammed et al., 2022](#); [Wang et al., 2024](#); [Cheng et al., 2021](#)). While this automation can facilitate clinical workflows, it also poses risks. A poorly constructed, parametrized, or trained model could produce false positives or negatives, compromising patient care. Moreover, if healthcare providers distrust the model's outputs, they may avoid using it or use it improperly. Trust depends on how well the model's decisions are understood ([Martinez, 2005](#)). Thus, good technical performance alone is insufficient; clinical context and other considerations must also be addressed.

Recent years have seen a surge in research applying AI to cytological image analysis ([Dhawan et al., 2018](#); [Gorantla et al., 2019](#); [Kavitha et al., 2023](#)), with growing interest in CNNs, ViTs, regression-based models, or their combinations. These studies fall into two main categories: those that rely on public datasets such as Herlev, SIPaKMeD, or Mendeley LBC ([Ouh et al., 2024](#); [Chowdary et al., 2023](#); [Chauhan et al., 2023](#)), and those that propose new architectures tailored to specific tasks like detection, segmentation, or classification using proprietary datasets ([Cheng et al., 2021](#); [Kanavati et al., 2022](#)). While these contributions demonstrate that certain tasks traditionally performed by cytopathologists can be automated with reasonable accuracy, many of them rely on datasets that do not reflect real clinical contexts. Additionally, a frequent conceptual error is equating the detection of cytological anomalies with cancer diagnosis.

The earliest attempts to automate cervical cytology with deep learning relied heavily on CNN-based architectures applied to small, well-curated datasets such as Herlev or SIPaKMeD. These

studies demonstrated that automatic recognition of precancerous lesions was technically feasible, though often limited by overfitting and narrow class diversity ([Devi et al., 2023](#)). Refinements soon followed with optimized convolutional pipelines or hybridized CNN-GRU variants that improved sensitivity to complex cytological patterns ([Rohini and Kavitha, 2024](#)). Others explored tailored CNN designs for Pap smear images, reporting encouraging accuracies but mostly within closed datasets that lacked external validation ([Khozaimi and Mahmudy, 2024](#)).

From 2022 onwards, a new wave of studies emphasized hybrid pipelines that combined deep features with classical classifiers or optimization heuristics. For example, hybridization with fuzzy neural networks or ensemble learning improved robustness against inter-sample variability ([Kalbhor et al., 2023b](#)). In parallel, researchers began to incorporate non-traditional datasets, including liquid-based cytology and field-of-view tiles from whole-slide images ([Gao et al., 2022](#)). These approaches sought to move beyond isolated single-cell images, capturing contextual information closer to real practice, although issues of transparency and reproducibility persisted.

More recently, the field has witnessed the entrance of transformer-inspired backbones and knowledge distillation frameworks, aiming to capture long-range morphological dependencies and optimize computational costs ([Kang and Li, 2024](#)). Studies have also experimented with graph-based models and metaheuristic optimizations to enhance precancerous lesion detection, reporting near-perfect accuracies in benchmark datasets but with uncertain clinical transferability ([Song et al., 2024](#)). Taken together, these contributions reflect an energetic but fragmented landscape: while technical metrics frequently surpass 90% accuracy, the lack of dataset diversity, external validation, and standardized reporting highlights the persistent gap between benchmark-driven innovation and real-world clinical needs.

This review examines studies applying AI models to classification and diagnostic tasks in cervical cytology, emphasizing how models were constructed, what data they used, and what metrics were reported. In this context, AI refers to algorithms capable of learning from cytological images to identify cellular patterns associated with potential abnormalities. This field integrates computer vision, deep learning, public health, and cytopathology, aiming to develop practical solutions for the early identification of atypical cellular patterns. Common challenges include models that lack generalizability, data that fail to reflect the complexity of real cytology slides (often composed of isolated, well-selected cell images), and limited use of clinical variables that may impact diagnostic decisions ([Ssedyabane et al., 2024](#); [Gafeer et al., 2025](#)).

The number of publications in this area has increased notably, reflecting strong scientific interest. Within this context, the present review offers a comprehensive perspective that not only systematizes deep learning models applied to cervical cytology, but also compares the most widely used datasets and provides a cross-sectional analysis of reported performance metrics ([Alias et al., 2022](#); [Allahqoli et al., 2022](#)). Although some studies have recently begun addressing issues such as model explainability ([Hemalatha et al., 2023](#)), this remains in early stages. The wide variety of approaches and techniques makes it difficult to compare

TABLE 1 Application of the PICO method.

| Criterion | Description |
|------------------|--|
| Population (P) | Cervical cytology images. |
| Intervention (I) | AI models for image classification (CNN, ViT, hybrid models). |
| Comparison (C) | Comparison across different model architectures, classical vs. emerging datasets, and metrics. |
| Outcome (O) | Types of models used, datasets applied, reported performance metrics, and current trends. |

results and establish standards. There is still a need for a review that not only aggregates studies but also critically analyzes their technical limitations in light of the clinical settings where they might be applied.

A systematic review was conducted using the PRISMA methodology to examine AI applications in cervical cytology. Unlike prior reviews, this study offers a critical perspective, distinguishing between cellular lesion detection and cancer diagnosis, and interrogating the conceptual and ethical foundations of the evaluated models. Articles were retrieved from the Scopus database using well-defined inclusion and exclusion criteria, and the data were structured for both qualitative and quantitative analysis. This review aims to guide future research, enhance existing models, and promote responsible use of AI in clinical contexts. Its main contribution lies in a detailed characterization of the most frequently used architectures, datasets, and performance metrics, complemented by a cross-analysis linking model types, data sources, and diagnostic accuracy. Additionally, the review synthesizes recurring patterns and emerging trends to help guide future studies toward more efficient and clinically applicable AI solutions.

2 Methodology

This systematic review aims to identify and analyze studies that apply artificial intelligence (AI) to the classification of cervical cytology images, with special attention to the most commonly used models, the datasets employed, and the performance metrics reported. The review followed a three-phase process: planning, execution, and reporting, aligned with PRISMA guidelines and the PICO strategy. The review process was supported by the use of Mendeley (version 1.19.8) for reference management, Excel for tracking study selection, and draw.io (online version, accessed June 13, 2025) for creating the PRISMA flow diagram.

2.1 Review planning

To structure the search strategy and clarify the key terms, the PICO framework was adapted to the context of this review (Table 1).

In this phase, a strategy was designed to identify empirical studies that effectively address the research questions. Table 2 describes the questions guiding this systematic review.

TABLE 2 Research questions and objectives.

| Research question (RQ) | Objective |
|---|--|
| RQ1: What types of AI models (monolithic, hybrid, or others) have been used in cervical cytology image classification, and what trends are observed in recent approaches? | To identify the most applied AI models in cervical cytology and characterize current trends. |
| RQ2: What datasets have been most frequently used in studies applying AI to cervical cytology, and what new datasets are emerging? | To determine the most used datasets and describe the characteristics of emerging datasets. |
| RQ3: What performance metrics are most frequently reported in studies applying AI to cervical cytology image classification, and what are the typical values? | To identify the most common performance metrics and analyze reported values in recent studies. |

2.2 Search strategy and criteria

The search strategy was defined based on the three research questions posed in the planning phase, with specific search strings developed for each (see Table 3). These queries were executed independently in the Scopus database using the Advanced Search option. For each question, key terms and relevant synonyms were defined to maximize retrieval and ensure reproducibility. Additionally, inclusion and exclusion criteria were established as follows:

Each query was designed to independently address one of the review's research questions (RQ1, RQ2, RQ3), ensuring traceability, reproducibility, and alignment with the review's objectives.

2.3 Keywords and relation to research questions

Table 4 summarizes the keywords used in the search strategy and their relation to the research questions.

2.4 Relevant fields for data extraction

During the data extraction phase, key metadata fields were defined to systematize the analysis of selected studies (see Table 5).

As part of the data extraction process, AI models were categorized according to their structural nature. Monolithic models were defined as those relying on a single deep learning architecture, while hybrid models were defined as architectures that integrate two or more complementary computational strategies within the same pipeline. Examples of hybrid approaches include CNNs combined with traditional classifiers (e.g., SVM, Random Forest, XGBoost), CNNs enhanced with fuzzy logic or evolutionary algorithms, and CNNs integrated with transformer or attention modules. This categorization allowed us to systematically compare single-architecture strategies with more complex, multi-stage approaches.

2.5 Information sources

The literature search was conducted in the Scopus database, including only peer-reviewed original research articles. Search

TABLE 3 Search strings per research question.

| Key criteria | Description |
|-----------------------|---|
| Search string | <p>RQ1: TITLE-ABS-KEY("artificial intelligence" OR "machine learning" OR "deep learning" OR "convolutional neural network" OR CNN OR "vision transformer" OR ViT OR "transformer-based model") AND TITLE-ABS-KEY("cervical cytology" OR "pap smear" OR "cervical cells" OR "cervical cytological images") AND TITLE-ABS-KEY("classification" OR "cell classification" OR "lesion classification" OR "lesion detection" OR "automatic diagnosis")</p> <p>RQ2: TITLE-ABS-KEY("artificial intelligence" OR "machine learning" OR "deep learning" OR "convolutional neural network" OR CNN OR "vision transformer" OR ViT OR "transformer-based model") AND TITLE-ABS-KEY("cervical cytology" OR "pap smear" OR "cervical cells" OR "cervical cytological images") AND TITLE-ABS-KEY("dataset" OR "database" OR "image collection" OR "image repository" OR "public dataset" OR "private dataset" OR "labeled data")</p> <p>RQ3: TITLE-ABS-KEY("artificial intelligence" OR "machine learning" OR "deep learning" OR "convolutional neural network" OR CNN OR "vision transformer" OR ViT OR "transformer-based model") AND TITLE-ABS-KEY("cervical cytology" OR "pap smear" OR "cervical cells" OR "cervical cytological images") AND TITLE-ABS-KEY("classification" OR "cell classification" OR "lesion classification" OR "lesion detection" OR "automatic diagnosis") AND TITLE-ABS-KEY("performance metric" OR "accuracy" OR "precision" OR "recall" OR "sensitivity" OR "specificity" OR "F1-score" OR "AUC" OR "ROC curve")</p> |
| Intervention criteria | Peer-reviewed journal articles. Full-text available in English or Spanish. Keywords related to AI in cervical cytology, datasets, and performance metrics. Publication date between 2019 and 2024 |
| Exclusion criteria | Articles published before 2022. Studies not using AI models. Clinical studies without computational models. Reviews without new data. Works without full-text access |
| Search mode | Applied to title, abstract, and keywords. |

TABLE 4 Keywords, synonyms, and related research questions.

| Keyword/term | Synonyms/variants | Related to |
|------------------------------|--|---------------|
| Artificial intelligence | Machine learning, deep learning, IA, DL | RQ1, RQ2, RQ3 |
| Convolutional neural network | CNN, ConvNet | RQ1 |
| Vision transformer | ViT, transformer | RQ1 |
| Cervical cytology | Pap smear, cervical cells, cervical cytological images | RQ1, RQ2, RQ3 |
| Dataset | Database, image collection, image repository | RQ2 |
| Dataset quality | Public dataset, private dataset, balanced dataset | RQ2 |
| Performance metrics | Accuracy, precision, recall, F1-score, AUC, sensitivity, specificity | RQ3 |
| Model comparison | Architecture comparison, model evaluation | RQ1 |
| Classification | Cell classification, lesion classification, lesion detection | RQ1, RQ3 |

strategies were independently defined for each research question (RQ1, RQ2, and RQ3) and executed in June 2025. Filters were applied based on publication year, document type (scientific articles only), and access type (Gold Open Access and Hybrid Gold). The selected records were exported in RIS format for further analysis. To avoid duplicates, cross-checking was performed between result sets from each research question. Full-text evaluation was then conducted to ensure each article met the established inclusion criteria.

The full-text evaluation results for each research question were systematically recorded in dedicated spreadsheets. These final datasets, corresponding to RQ1, RQ2, and RQ3, are available as [Supplementary material](#) in the files RQ11.xlsx, RQ21.xlsx, and RQ31.xlsx, respectively.

TABLE 5 Relevant data extraction fields.

| Field | Description |
|---------------------|---|
| Reference | Article title, authors, and citation. |
| Publication | Type of publication (journal article, conference, preprint). |
| Year | Year of publication. |
| Dataset | Name, type (public/private), size, and characteristics. |
| Dataset type | Public or private. |
| Classical dataset | Indicates whether the dataset is classical (Y/N). |
| Main idea | Summary of the core concept of the study. |
| Main contributions | Novel contributions of the article to AI in cervical cytology. |
| Gaps or limitations | Limitations, problems, or research gaps identified in the study. |
| Methodology | Methods and models used (e.g., neural networks, training techniques, hybrid architectures). |
| Results | Model performance outcomes (metrics). |
| AI models | List of AI models used (CNN, ViT, hybrid, etc.). |
| Model type | Indicates if the model is monolithic or hybrid. |
| Datasets | List of datasets used in the study. |
| Performance metrics | Reported metrics and their values (Accuracy, F1-score, AUC, etc.). |

2.6 Quality appraisal

To evaluate the methodological robustness and potential risk of bias of the included studies, we implemented a structured quality appraisal adapted from the QUADAS-2 framework, which is widely used for diagnostic accuracy research.

In addition, the appraisal was structured as a checklist aligned with key criteria frequently recommended for AI studies in

medical imaging: dataset transparency, validation protocol rigor, and completeness of statistical reporting.

In the context of artificial intelligence applied to cervical cytology, the appraisal focused on four key aspects: the representativeness of the datasets, the validation strategy employed the extent to which performance metrics were reported beyond accuracy, and the transparency of model description and training procedures.

Each article was assessed according to these domains, and an overall risk of bias was subsequently assigned as low, high, or unclear, based on predefined decision rules. Studies were considered low risk when they met most of the criteria satisfactorily, high risk when multiple domains were judged inadequate, and unclear when reporting was insufficient to permit confident assessment.

The per-study evaluations are available in [Supplementary Table RQ31.xls](#), where additional columns have been included to document the quality appraisal domains and the overall risk of bias.

3 Results

3.1 Selected articles and general characteristics

Following the search strategies defined for each research question, 534 records were initially identified for RQ1, 456 for RQ2, and 381 for RQ3. Filters were subsequently applied based on publication year (2022–2025), document type (journal articles only), and access type (Gold and Hybrid Gold), which reduced the datasets to 91 articles for RQ1, 94 for RQ2, and 75 for RQ3.

A crosschecking process was then carried out to remove duplicate records across the three sets. 68 duplicates were found between RQ1 and RQ2, 57 between RQ1 and RQ3, and 54 between RQ2 and RQ3. Moreover, 47 articles were common to all three search results. After removing duplicates, a consolidated set of 117 unique articles was obtained.

Each article was reviewed in full text to confirm its relevance. As a result, 14 articles were excluded from RQ1, 18 from RQ2, and 4 from RQ3. The main reasons for exclusion included: focus on colposcopic images, the use of non-cytological diagnostic modalities, or lack of relevant information aligned with this review's objectives. [Table 6](#) summarizes the distribution of articles by research question and filtering stage.

A PRISMA 2020 flow diagram was generated to illustrate the study identification, screening, and inclusion process ([Figure 1](#)).

[Table 7](#) shows the distribution of the selected articles by year of publication. The year 2024 accounts for the highest number of articles across all three research questions, reflecting growing scientific interest in the application of AI to cervical cytology in recent years.

This annual distribution also reveals a rising trend in scientific output from 2022 to 2024 in the areas related to lesion classification, specialized dataset usage, and the evaluation of AI model performance metrics in cervical cytology imaging.

3.2 RQ1: artificial intelligence models applied to cervical cytology

The analysis of the 77 articles selected for RQ1 revealed a wide range of approaches for applying artificial intelligence (AI) models to the classification of cervical cytology images. Most studies implemented models based on convolutional neural networks (CNNs), followed by more recent architectures such as Vision Transformers (ViTs) and, to a lesser extent, transfer learning techniques with pretrained models. There is also growing interest in hybrid models that combine multiple techniques or stages, such as feature fusion via CNNs with traditional classifiers (e.g., SVM, XGBoost), or the integration of sequential models with attention modules.

[Table 8](#) summarizes the frequency with which different types of models were reported in the analyzed studies. While CNNs remain prevalent, there has been a significant increase in the use of hybrid architectures over the past 3 years, suggesting a trend toward more complex and adaptive solutions.

To complement the quantitative distribution shown in [Table 8](#), [Figure 2](#) illustrates a taxonomy of the AI models reported in the reviewed studies. This schematic representation highlights the hierarchical organization of the main categories—CNN-based models, Hybrid models, Ensembles, and Decision Trees—together with their most frequently used sub-architectures (e.g., ResNet, DenseNet, EfficientNet for CNNs, and CNN+SVM, CNN+XGBoost, CNN+Fuzzy for Hybrids). The figure provides a conceptual overview that facilitates understanding of how different computational strategies have been applied to cervical cytology, and emphasizes the predominance of hybrid approaches, which accounted for 61% of all included studies.

Additionally, the models were categorized based on their structural nature into two groups: monolithic, referring to those using a single deep learning architecture; and hybrid, defined in this review as architectures that integrate two or more complementary computational strategies within the same pipeline. Hybrid approaches included, for example, CNNs combined with traditional classifiers (e.g., SVM, Random Forest, XGBoost), CNNs enhanced with fuzzy logic or evolutionary algorithms, and CNNs integrated with transformer or attention modules. Hybrid models accounted for 61% of the reviewed articles, surpassing monolithic approaches (44%). This finding reflects a growing preference for composite strategies, which better address variability in cytological images, combine multiple feature sources, and improve classification accuracy.

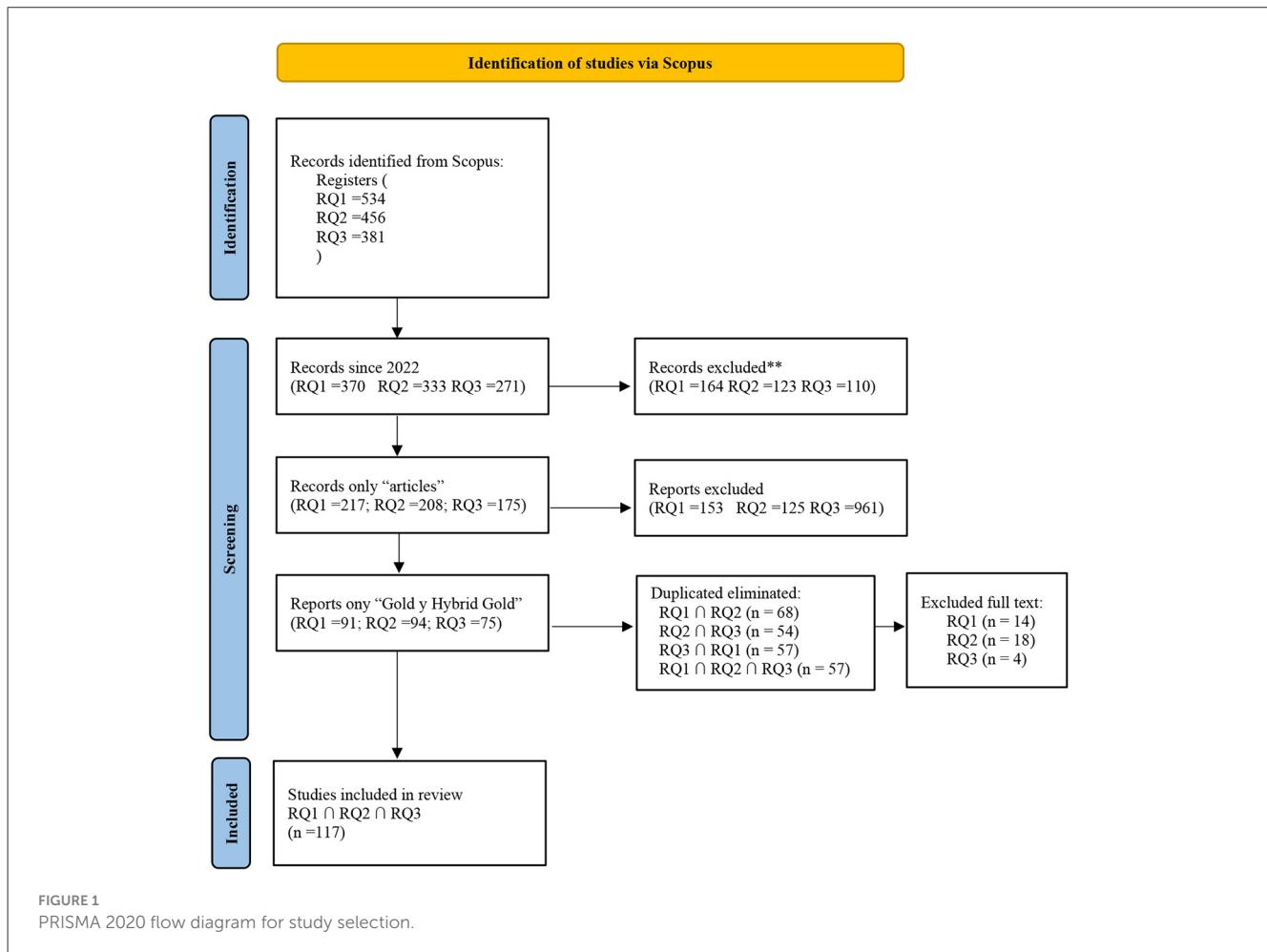
In terms of temporal trends, a clear shift was observed from traditional CNN-based models to more sophisticated hybrid architectures. Between 2022 and 2024, there was an increase in the incorporation of attention modules, transformer layers, and ensemble strategies, highlighting the influence of recent advances in computer vision.

Several studies also emphasized the specific advantages of hybrid models, such as greater robustness to intercellular variability and improvements in performance metrics when combining classifiers. However, they also acknowledged limitations, including increased computational complexity, reduced reproducibility, and

TABLE 6 Distribution of articles by research question.

| Research question | Identified (unfiltered) | From 2022 | Journal articles | Gold + hybrid gold |
|-------------------|-------------------------|-----------|------------------|--------------------|
| RQ1 | 534 | 370 | 217 | 91 |
| RQ2 | 456 | 333 | 208 | 94 |
| RQ3 | 381 | 271 | 175 | 75 |

The column “Journal Articles” refers to documents that meet the required publication type (peer-reviewed journals), and the “Gold + Hybrid Gold” column shows the final set selected for analysis.



the need for larger annotated datasets to effectively train the additional modules.

3.3 RQ2: datasets used in the studies

A detailed review of the selected articles revealed 74 records in which the datasets used were explicitly stated. The findings indicate a strong reliance on classic datasets, particularly Herlev and SIPaKMeD, which were used in 16 and 15 studies, respectively. Additionally, 10 studies combined both datasets, likely to increase class diversity or improve training performance. This dominance can be attributed to their public availability, well-structured annotations, and broad dissemination within the scientific community.

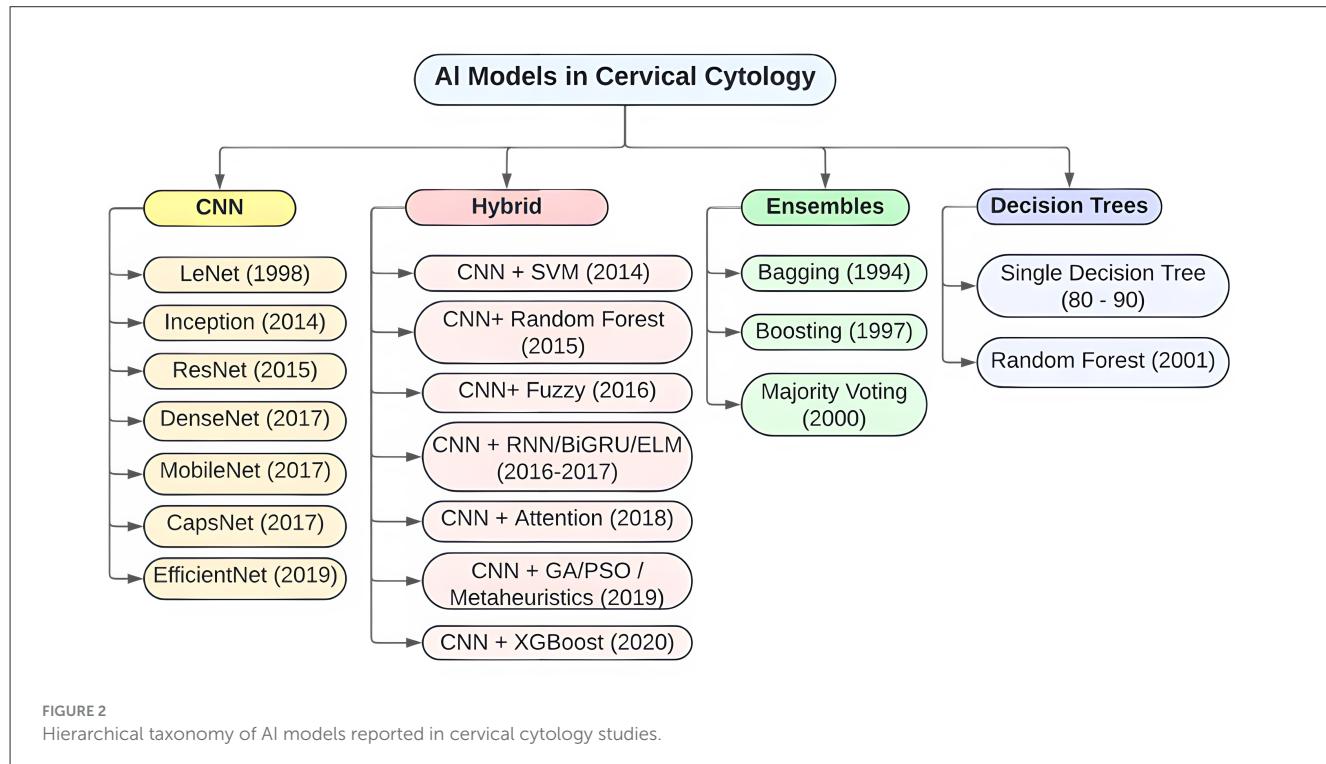
TABLE 7 Distribution of selected articles by year and research question (2022–2025).

| Year | RQ1 | RQ2 | RQ3 |
|-------|-----|-----|-----|
| 2022 | 22 | 16 | 22 |
| 2023 | 23 | 24 | 19 |
| 2024 | 29 | 29 | 27 |
| 2025 | 4 | 6 | 3 |
| Total | 77 | 75 | 71 |

In contrast, there is a growing trend toward the use of emerging or proprietary datasets. Eight studies reported the use of private datasets generated by the authors themselves, highlighting

TABLE 8 Frequency of AI models used in cervical cytology studies.

| Model type | Frequency | Percentage | References |
|----------------|-----------|------------|---|
| Decision trees | 2 | 3% | Devi et al., 2023; Kalbhor et al., 2022 |
| CNN | 24 | 31% | Khozaimi and Mahmudy, 2024; Mazroa et al., 2023; Rasheed et al., 2023; Wong et al., 2023; Resmi et al., 2024; Attallah, 2023; Tan et al., 2024; Skerrett et al., 2022; Ahmed et al., 2024; Xu et al., 2022; Utomo et al., 2025; Shandilya et al., 2024; Priya and Bai, 2024; Zammataro, 2024; Shiny and Parasuraman, 2023; Akash et al., 2024; Wubineh et al., 2024a; Alohal et al., 2024; Zhang et al., 2025; Rodríguez et al., 2024; Tian et al., 2024; Anandavally and Bai, 2024; Kurita et al., 2023; Fang et al., 2022 |
| Hybrid | 47 | 61% | Tang et al., 2023; Shinde et al., 2022; Wang et al., 2024; Rohini and Kavitha, 2024; Gao et al., 2022; Song et al., 2024; Kalbhor et al., 2023c; Alquran et al., 2022a; Pang et al., 2025; Ahishakiye and Kanobe, 2024; AlMohimeed et al., 2024; Jain et al., 2024; Maurya et al., 2024; Alsalat et al., 2022; Lilhore et al., 2022; Gangrade et al., 2025; Chauhan et al., 2023; Wang et al., 2025; Anupama et al., 2022; Mansouri and Ragab, 2023; Bora et al., 2022; Battula and Sai Chandana, 2023; Kalbhor et al., 2023a; Fekri-Ershad and Alsaaffar, 2023; Battula and Chandana, 2022; AbuKhilil et al., 2022; Joy nab et al., 2024; Du et al., 2023; Alquran et al., 2022b; Diniz et al., 2022; Yi et al., 2024; Chowdary et al., 2023; Chen et al., 2022; Stegmüller et al., 2024; Alsalat et al., 2023; Benhari and Hossseini, 2022; Mahmoud et al., 2022; Ji et al., 2022; Bera et al., 2024; Ando et al., 2024; Sahoo et al., 2023; Waly et al., 2022; Akbar et al., 2024; Fahad et al., 2024; Mathivanan et al., 2024; Nour et al., 2024; Fang et al., 2024 |
| Ensembles | 3 | 4% | Karamti et al., 2023; Khanarsa and Kitsiranuwat, 2024; Kurniawati and Prabowo, 2022 |
| Total | 77 | 100% | |

FIGURE 2
Hierarchical taxonomy of AI models reported in cervical cytology studies.

efforts to develop data contextualized to specific clinical cases or newer acquisition technologies (e.g., whole slide imaging or liquid-based cytology). Other datasets such as Mendeley LBC, ComparisonDetector, ISBI-2014/2015, and CRIC are also gaining attention for their variety of cell types and complex annotations.

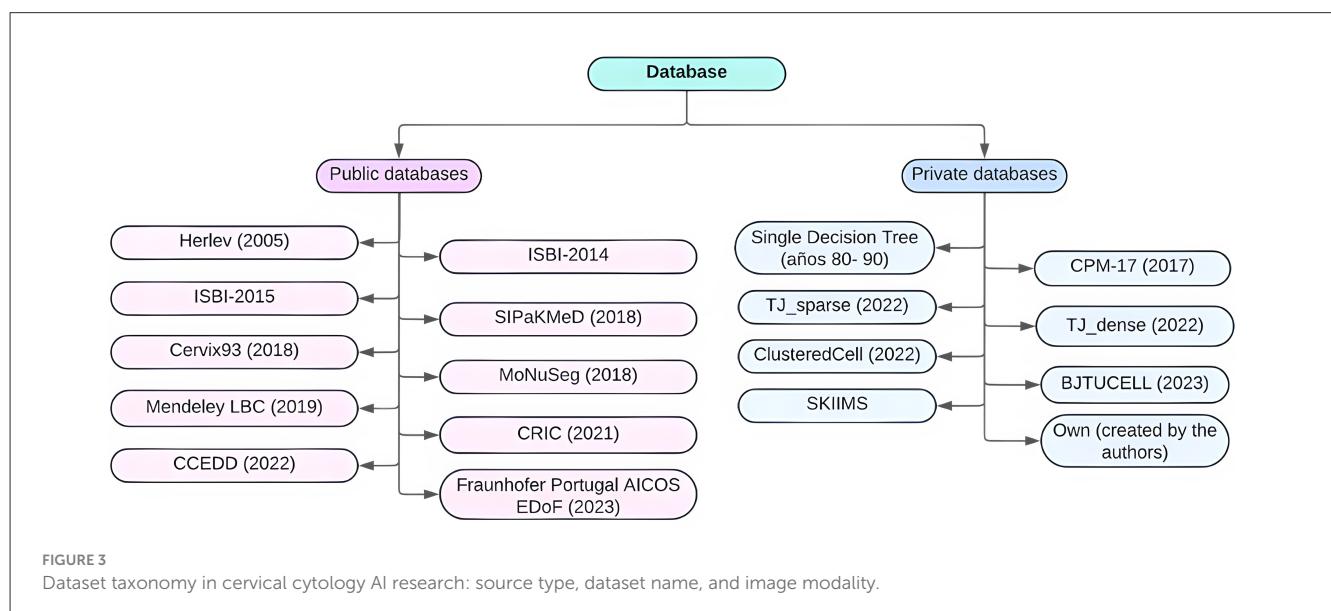
Table 9 summarizes the frequency of dataset usage, with less frequently used datasets grouped under “Others.”

To complement the descriptive distribution presented in Table 9, Figure 3 illustrates a hierarchical taxonomy of datasets applied in cervical cytology studies. The classification begins by distinguishing between public and private datasets, then

specifies the individual datasets most frequently reported (e.g., Herlev, SIPaKMeD, Mendeley LBC, ISBI-2014/2015, and others), and finally maps the type of image analyzed in each case (individual cells, partial microscope fields, or combined approaches). This visualization highlights the dominance of public datasets, particularly Herlev and SIPaKMeD, but also reveals an emerging contribution of private institutional collections that integrate partial fields or whole-slide derivatives. Such a taxonomy not only clarifies the methodological landscape but also underscores the heterogeneity in data sources and image modalities, which

TABLE 9 Distribution of datasets used in AI studies for cervical cytology.

| Dataset | Individual Cells (A) | Partial Microscope Fields (B) | A + B | Frequency | References |
|----------------------|----------------------|-------------------------------|-------|-----------|---|
| Herlev | 16 | 0 | 0 | 16 | Song et al., 2024; Resmi et al., 2024; Attallah, 2023; Tan et al., 2024; Priya and Bai, 2024; Akash et al., 2024; Anandavally and Bai, 2024; Jain et al., 2024; Anupama et al., 2022; Ando et al., 2024; Waly et al., 2022; Nour et al., 2024; Kurniawati and Prabowo, 2022; Janani and Christopher, 2023; Lotfi and Momenzadeh, 2022; Qin et al., 2022 |
| SIPaKMeD | 6 | 1 | 8 | 15 | Tang et al., 2023; Wang et al., 2024; Chauhan et al., 2023; Khozaimi and Mahmudy, 2024; Gao et al., 2022; Utomo et al., 2025; Shandilya et al., 2024; Wubineh et al., 2024a; Alohal et al., 2024; Kurita et al., 2023; Maurya et al., 2024; Gangrade et al., 2025; Fekri-Ershad and Alsaffar, 2023; Ontor et al., 2023; Mathivanan et al., 2024 |
| SIPaKMeD + Herlev | 0 | 2 | 8 | 10 | Shinde et al., 2022; Kalbhor et al., 2023b; Zhang et al., 2025; Fang et al., 2022; AlMohimeed et al., 2024; Chowdary et al., 2023; Benhari and Hosseini, 2022; Chen et al., 2022; Zhao et al., 2023; Fahad et al., 2024 |
| Private | 2 | 1 | 5 | 8 | Bora et al., 2022; Joynab et al., 2024; Yi et al., 2024; Stegmüller et al., 2024; Liu et al., 2022; Riana et al., 2023; Yin et al., 2024; An et al., 2023 |
| Mendeley LBC | 3 | 0 | 2 | 5 | Wong et al., 2023; Shiny and Parasuraman, 2023; Rodríguez et al., 2024; Battula and Chandana, 2022; Sahoo et al., 2023 |
| ISBI-2014/2015 | 2 | 1 | 2 | 5 | Zhang et al., 2024a; Mahyari and Dansereau, 2022; Zhang et al., 2024b; Kalbhor et al., 2023a; Rasheed et al., 2023 |
| Comparison Detector | 0 | 1 | 2 | 3 | Pang et al., 2025; Cheng et al., 2025; Li et al., 2025 |
| SIPaKMeD + CRIC | 0 | 0 | 2 | 2 | Battula and Sai Chandana, 2023; Fang et al., 2024 |
| Others (11 datasets) | 3 | 2 | 5 | 10 | Kang and Li, 2024; Xu et al., 2022; Kalbhor et al., 2023c; Yang et al., 2022; Galande et al., 2024; Albuquerque et al., 2023; Ji et al., 2023; Nazir et al., 2023; Luo et al., 2022; Wu et al., 2023a |
| Total | 13 | 6 | 45 | 74 | |



directly affects the comparability and generalizability of AI models in cervical cytology.

Regarding dataset type, the analysis reveals that 69% of the studies relied on **public datasets**, while 31% used **private or self-generated datasets**. This distribution underscores the importance of open data in scientific reproducibility, while also

emphasizing the need to expand the diversity—both demographic and technological—of training sets.

In terms of image types, three major approaches were identified: **Individual Cell Images (A)** – most common. **Partial Microscopy Fields (B)** – derived from WSI (whole slide images), capturing spatial and contextual features. **Combined A + B** – integrating both

approaches. Although the use of tiles from WSIs remains emerging, it is seen as a growing trend, especially with the advent of models that are more sophisticated and the need for clinical scalability.

While public datasets enhance comparability across studies, they also present risks of overfitting, limited class diversity and poor representativeness of morphological variants from different populations. On the other hand, private datasets face challenges in access and validation but offer opportunities for personalized diagnostic solutions tailored to real-world clinical settings.

3.4 RQ3: performance metrics and results obtained

Although most studies reported high accuracy values, the quality assessment revealed frequent methodological limitations. The main issues included overreliance on classical public datasets, lack of external validation, incomplete reporting of class-level metrics, and insufficient description of training procedures. These patterns suggest that the reported performance must be interpreted with caution. Detailed assessments for each study are provided in the [Supplementary material](#) (see [Supplementary Table RQ31.xls](#), with extended fields for quality appraisal).

The analysis of the studies included in this systematic review reveals a predominant use of traditional classification metrics to evaluate the performance of artificial intelligence (AI) models applied to cervical cytology images. The most frequently reported metrics were accuracy, precision, recall, F1-score, specificity, and area under the ROC curve (AUC). This selection reflects a focus not only on overall classification accuracy but also on the models' ability to detect minority classes, which is critical in clinical contexts.

Among the 71 reviewed articles, 93.9% reported accuracy as their primary metric, both in binary and multiclass classification schemes. Accuracy values ranged from 63.08 to 100%, with the highest performances associated with Vision Transformer (ViT)-based models and hybrid architectures. Quantitative analysis yielded a mean accuracy of 87.76%, making it the recurrent metric across the 121 recorded performance entries.

Precision and recall were reported in approximately 65% of the studies, highlighting growing attention to class-level performance and the trade-off between true positives and false negatives. The mean precision was 87.01%, while recall averaged 78.06%, with wide variation across models, suggesting differences in how class imbalance was handled. The F1-score, used in 54% of the studies, had an average of 64.65%, but reached values close to 99% in well-optimized multiclass models, especially those evaluated on datasets such as SIPaKMeD.

[Table 10](#) presents a statistical summary of the most frequently reported performance metrics, including frequency, mean, and observed minimum and maximum values, along with references to the studies that used them as primary metrics.

Specificity was reported less frequently (15%), typically in models that incorporated probabilistic outputs or clinical attention modules. It was more common in studies designed to simulate real-world medical validation. Metrics such as balanced accuracy, Matthews's correlation coefficient, and AUC were less common, but their appearance has increased in studies published since 2023—an

indication of evolving practices toward more clinically meaningful and balanced evaluations.

From a comparative perspective, hybrid models (e.g., CNN-ViT combinations or architectures with attention mechanisms) achieved the highest average accuracy (96.63%), followed by CNN-based models (e.g., ResNet, DenseNet) with an average of 94.91%. In contrast, ensemble and classical models as if Random Forest exhibited lower performance, with average accuracy around 63–83%, depending on the dataset used ([Table 11](#)).

In summary, the metrics reported reveal a favorable outlook for AI-based models in cervical cytology, with performance levels that match or even exceed human-level diagnosis in specific tasks. Nonetheless, common limitations persist, including inconsistent result reporting, lack of external cross-validation, and limited discussion of the statistical significance of differences between models. These aspects must be addressed in future research to ensure reliable, clinically robust, and ethically sound AI implementations.

3.5 Cross-analysis: relationships between models, datasets, and metrics

The cross-analysis of model types, datasets used, and reported performance metrics reveals emerging patterns and significant associations that characterize the current development of AI-based models in cervical cytology.

A predominance of convolutional neural networks (CNNs) and hybrid architectures (e.g., CNN + Transformer, CNN + RNN, or attention-enhanced models) was observed. These models were most frequently applied to classic datasets such as SIPaKMeD and Herlev. CNNs trained on SIPaKMeD achieved an average accuracy of 99.12%, while on Herlev the average dropped to 87.44%. This suggests a higher affinity between CNN architectures and the visual characteristics of SIPaKMeD, possibly due to its well-defined class structure and standardized preprocessing.

Hybrid models also achieved high average performance—97.31% on SIPaKMeD and 95.30% on Mendeley LBC—surpassing pure CNNs and demonstrating their ability to capture complex morphological relationships. Notably, in more heterogeneous datasets like Cervix93, which include greater variability and less uniformity in the samples, hybrid models still maintained high accuracy levels (up to 99.01%), highlighting their robustness.

In contrast, decision tree models and those based on traditional machine learning techniques showed lower average performance (83.00% in mixed datasets) and appeared less frequently in recent studies, likely due to their limitations when dealing with complex, multiclass cytological images.

There is a clear tendency to report better metrics when using classic datasets like SIPaKMeD and Herlev. These datasets are not only widely used, but also offer more consistency in terms of resolution, annotation, and class balance—factors that favor the training and evaluation of deep learning models. However, this dependency on classic datasets poses a significant limitation for clinical generalization, as they do not capture the full variability of real-world cytological environments.

TABLE 10 Frequency and statistical values of performance metrics reported in the reviewed studies.

| Metric | Frequency | Avg (%) | Min (%) | Max (%) | References |
|-------------|-----------|---------|---------|---------|---|
| Accuracy | 64 | 94.95 | 58.00 | 100.0 | Tang et al., 2023; Rohini and Kavitha, 2024; Khozaimi and Mahmudy, 2024; Gao et al., 2022; Song et al., 2024; Kalbhor et al., 2022; Wong et al., 2023; Attallah, 2023; Tan et al., 2024; Ahmed et al., 2024; Shandilya et al., 2024; Priya and Bai, 2024; Zammataro, 2024; Shiny and Parasuraman, 2023; Akash et al., 2024; Wubineh et al., 2024a; Zhang et al., 2025; Rodríguez et al., 2024; Tian et al., 2024; Anandavally and Bai, 2024; Kurita et al., 2023; Kalbhor et al., 2023c; Alquran et al., 2022a; Pang et al., 2025; Ahishakiye and Kanobe, 2024; AlMohimeed et al., 2024; Jain et al., 2024; Maurya et al., 2024; Alsalatiet al., 2022; Gangrade et al., 2025; Chauhan et al., 2023; Wang et al., 2025; Bora et al., 2022; Battula and Sai Chandana, 2023; Kalbhor et al., 2023a; Fekri-Ershad and Alsaffar, 2023; Battula and Chandana, 2022; AbuKhalil et al., 2022; Joynab et al., 2024; Alquran et al., 2022b; Diniz et al., 2022; Yi et al., 2024; Alsalatiet al., 2023; Benhari and Hosseini, 2022; Sahoo et al., 2023; Waly et al., 2022; Akbar et al., 2024; Karamti et al., 2023; Khanarsa and Kitsirawan, 2024; Kurniawati and Prabowo, 2022; Lotfi and Momenzadeh, 2022; Mathivanan et al., 2024; Chen et al., 2022; Fahad et al., 2024; Alohaliet al., 2024; Sudhakar et al., 2023; Petrov and Sokolov, 2023; Lilhore et al., 2022; Yang et al., 2024; Harsono et al., 2022; Haridas and Jayamalar, 2023; Wubineh et al., 2024b; Shinde et al., 2022 |
| Precision | 25 | 92.61 | 60.78 | 100.0 | Tang et al., 2023; Chowdary et al., 2023; Wong et al., 2023; Priya and Bai, 2024; Shiny and Parasuraman, 2023; Tian et al., 2024; Anandavally and Bai, 2024; Fang et al., 2022; Ahishakiye and Kanobe, 2024; AlMohimeed et al., 2024; Battula and Sai Chandana, 2023; AbuKhalil et al., 2022; Diniz et al., 2022; Mahmoud et al., 2022; Waly et al., 2022; Karamti et al., 2023; Kurniawati and Prabowo, 2022; Chen et al., 2022; Kalbhor et al., 2023a; Luo et al., 2022; Alohaliet al., 2024; Sudhakar et al., 2023; Harsono et al., 2022; Haridas and Jayamalar, 2023; Shinde et al., 2022 |
| Recall | 35 | 93.46 | 66.10 | 100.0 | Tang et al., 2023; Chowdary et al., 2023; Gao et al., 2022; Song et al., 2024; Attallah, 2023; Xu et al., 2022; Priya and Bai, 2024; Shiny and Parasuraman, 2023; Rodríguez et al., 2024; Tian et al., 2024; Anandavally and Bai, 2024; Fang et al., 2022; Ahishakiye and Kanobe, 2024; AlMohimeed et al., 2024; Maurya et al., 2024; Battula and Sai Chandana, 2023; Battula and Chandana, 2022; AbuKhalil et al., 2022; Diniz et al., 2022; Benhari and Hosseini, 2022; Mahmoud et al., 2022; Sahoo et al., 2023; Waly et al., 2022; Karamti et al., 2023; Kurniawati and Prabowo, 2022; Chen et al., 2022; Kalbhor et al., 2023a; Luo et al., 2022; Alohaliet al., 2024; Sudhakar et al., 2023; Petrov and Sokolov, 2023; Lilhore et al., 2022; Harsono et al., 2022; Haridas and Jayamalar, 2023; Shinde et al., 2022 |
| Specificity | 14 | 89.40 | 48.80 | 99.09 | Chowdary et al., 2023; Gao et al., 2022; Song et al., 2024; Rodríguez et al., 2024; Maurya et al., 2024; Battula and Chandana, 2022; Benhari and Hosseini, 2022; Ando et al., 2024; Chen et al., 2022; Sudhakar et al., 2023; Petrov and Sokolov, 2023; Harsono et al., 2022; Haridas and Jayamalar, 2023; Wubineh et al., 2024b |
| F1-score | 23 | 94.14 | 62.82 | 100.0 | Tang et al., 2023; Song et al., 2024; Attallah, 2023; Priya and Bai, 2024; Shiny and Parasuraman, 2023; Tian et al., 2024; Anandavally and Bai, 2024; Fang et al., 2022; Ahishakiye and Kanobe, 2024; AlMohimeed et al., 2024; Maurya et al., 2024; Battula and Sai Chandana, 2023; AbuKhalil et al., 2022; Diniz et al., 2022; Mahmoud et al., 2022; Sahoo et al., 2023; Waly et al., 2022; Karamti et al., 2023; Kurniawati and Prabowo, 2022; Kalbhor et al., 2023a; Lilhore et al., 2022; Haridas and Jayamalar, 2023; Shinde et al., 2022 |

On the other hand, models trained on private or non-traditional datasets have shown competitive metrics, but the lack of public availability and inconsistent annotation standards hinder direct comparison and limit the reproducibility of results.

This analysis indicates that although hybrid architectures offer superior performance in controlled scenarios, there remains an overreliance on a small set of classic datasets. Future research must prioritize evaluating these models in real clinical settings, incorporating whole-slide images (WSI) and multisource data.

Additionally, there is a need to develop standardized multiclass and multimodal benchmarks, and to encourage the open publication of expert-annotated datasets.

Furthermore, the research community should move toward the systematic use of complementary metrics (e.g., balanced accuracy, negative predictive value, Kappa coefficient) and ensure external cross-validation and the reporting of confidence intervals. These practices are essential to promote transparency, reproducibility, and clinical applicability of the proposed AI models.

TABLE 11 Comparison of metrics by architecture type.

| Metric | Accuracy | | | Precision | | | Recall | | | F1 Score | | |
|----------|----------|-------|--------|-----------|-------|--------|--------|-------|--------|----------|-------|--------|
| | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max | Min | Avg | Max |
| CNN | 72.80 | 94.91 | 100.00 | 80.00 | 92.06 | 99.00 | 77.40 | 93.36 | 100.00 | 82.50 | 92.98 | 100.00 |
| Ensemble | 63.08 | 63.08 | 63.08 | 60.78 | 60.78 | 60.78 | 66.10 | 66.10 | 66.10 | 62.82 | 62.82 | 62.82 |
| Hybrid | 81.11 | 96.63 | 100.00 | 68.00 | 95.15 | 100.00 | 70.00 | 95.46 | 100.00 | 89.00 | 96.61 | 100.00 |
| VIT | 99.11 | 99.11 | 99.11 | 99.12 | 99.12 | 99.12 | 99.11 | 99.11 | 99.11 | 99.11 | 99.11 | 99.11 |

The encouraging results of AI models in cervical cytology should be considered in light of their methodological limitations. Our quality appraisal revealed frequent risks of bias, including reliance on small or homogeneous datasets, absence of external validation, and incomplete reporting of clinically relevant metrics. These issues may inflate reported performance values and limit generalizability. Future research should prioritize representative datasets, standardized reporting frameworks, and external validation to ensure robust and clinically reliable evidence.

4 Discussion

4.1 AI Models: from CNN predominance to hybrid strategies

The most striking finding of this review is the shift from CNN dominance toward hybrid architectures and, more recently, Vision Transformers (AlMohimeed et al., 2024; Yamagishi and Hanaoka, 2025; Muksimova et al., 2024). This is not a trivial transition: CNNs have demonstrated robustness but also clear limitations in capturing global dependencies within cytology images, as also noted by Muksimova et al. (2025) and Mustafa et al. (2025). The fact that more than 60% of recent studies rely on hybrid combinations shows how the field is trying to address morphological variability (Table 8). However, this increasing sophistication comes with trade-offs: while hybrid models can boost metrics, they often do so at the cost of reproducibility, transparency, and computational feasibility in low-resource settings. The tension between technical precision and clinical applicability remains unresolved in the literature.

4.2 Datasets: the paradox of public versus private

Regarding datasets, the field still depends heavily on SIPaKMeD and Herlev. These collections are valuable as benchmarks, but their overuse introduces an evident bias: they fail to represent the population diversity and preparation variability encountered in real-world practice (Ybatesa-Medina et al., 2025). It is telling that even the most sophisticated models can reach near-perfect accuracy on these “clean” datasets, while performance drops when evaluated on more heterogeneous data (Pang et al., 2025; Joynab et al., 2024; Wu et al., 2023a; Khan et al., 2023; Wu et al., 2023b). Attempts to develop private or institutional datasets are commendable because they move closer to clinical contexts, but their lack of

public availability prevents replication and fair comparison. This gap seriously undermines the community’s ability to establish robust standards.

4.3 Metrics and evaluation practices: beyond accuracy

Although accuracy remains the most frequently reported metric (Table 10), this emphasis is problematic. Global accuracy can inflate perceptions of success while masking poor performance in minority but clinically critical classes, such as HSIL or SCC (Ando et al., 2024; Cheng et al., 2025; Suksmono et al., 2021). The fact that fewer than 20% of studies reported metrics such as specificity, negative predictive value, or balanced accuracy reflects insufficient maturity in evaluation design. This shortfall is not merely technical: it has direct consequences for patient safety, as a model that maximizes accuracy at the expense of sensitivity in HSIL cannot be trusted in clinical decision-making. Future research must therefore standardize validation protocols, incorporate external validation, and adopt metrics that capture the real clinical cost of misclassification.

4.4 Cross-analysis: patterns and warnings

The cross-analysis of models, datasets, and metrics uncovers a paradox that cannot be overlooked: the best results cluster around classical datasets with relatively simple structures, while more realistic scenarios — whole-slide images and heterogeneous institutional collections — remain underexplored, as also noted in the review by Jiang et al. (2023). This reveals a persistent gap between academic research and clinical application. Moving forward, the field must prioritize multicenter benchmarks with diverse data and more rigorous evaluation criteria. Only then, can AI in cervical cytology move beyond being a promising academic exercise and evolve into a clinically reliable and ethically sound tool.

5 Conclusions

This systematic review provides an integrative perspective on the application of artificial intelligence in cervical cytology, focusing on deep learning models, datasets, and performance outcomes. Through the analysis of 77 peer-reviewed articles published between 2022 and 2025, we identified a clear predominance of convolutional neural networks and hybrid

architectures—particularly those combining CNNs with attention mechanisms or transformer-based models—as the core computational strategies for lesion classification.

In terms of data usage, the review revealed a significant dependency on a small number of publicly available datasets, particularly SIPaKMeD and Herlev. While these datasets offer consistency and facilitate benchmarking across studies, their limited clinical variability poses a challenge for real-world generalizability. The emergence of private or custom datasets represents an important effort to diversify data sources, although lack of accessibility and annotation standards hinders replication and external validation.

Regarding model performance, most studies reported high levels of accuracy, precision, and recall, especially those employing hybrid models trained on curated datasets. However, inconsistent reporting practices, limited use of external cross-validation, and underutilization of clinically meaningful metrics such as specificity, balanced accuracy, and AUC indicate the need for more robust evaluation protocols.

To our knowledge, this is the first systematic review to conduct a cross-sectional analysis that jointly examines the relationships between deep learning architectures, dataset types, and diagnostic metrics in the context of cervical cytology. This integrative approach offers a broader understanding of current practices and challenges in the field, contributing valuable insights that may inform the development of more reliable, interpretable, and clinically aligned AI systems for early detection of cervical lesions.

In summary, the most relevant outcomes of this review are threefold: (i) the predominance of hybrid architectures, particularly CNNs combined with transformer or attention modules, as the emerging computational trend; (ii) the continued dependence on a small set of classical datasets (SIPaKMeD and Herlev), despite increasing interest in private and heterogeneous collections; and (iii) the overall performance patterns, with accuracies typically ranging between 87 and 95% and F1-scores between 64 and 96%, which underscore both the potential and the methodological limitations of current models. These findings provide a concrete reference point for future research and development of clinically applicable AI systems in cervical cytology.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

Author contributions

MV-C: Conceptualization, Data curation, Formal analysis, Funding acquisition, Investigation, Methodology, Project administration, Resources, Software, Supervision, Validation, Visualization, Writing – original draft, Writing – review & editing. LP: Data curation, Formal analysis, Investigation, Methodology, Writing – review & editing. CR: Formal analysis, Investigation,

Methodology, Writing – review & editing. DR: Data curation, Software, Writing – original draft. KS-D: Data curation, Formal analysis, Writing – review & editing. LA-F: Data curation, Writing – original draft. NR-L: Formal analysis, Writing – original draft.

Funding

The author(s) declare that financial support was received for the research and/or publication of this article. The Article Processing Charge (APC) was funded by the Universidad Privada Norbert Wiener.

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Generative AI statement

The author(s) declare that Gen AI was used in the creation of this manuscript.

The author(s) verify and take full responsibility for the use of generative AI in the preparation of this manuscript. Generative AI tools (specifically, ChatGPT by OpenAI) were used to support language refinement, translation, and formatting of selected sections of the manuscript. All content generated was thoroughly reviewed, verified, and edited by the author(s) to ensure accuracy, integrity, and scientific validity. No generative AI was used to produce or analyze the original data or to write the scientific content itself.

Any alternative text (alt text) provided alongside figures in this article has been generated by Frontiers with the support of artificial intelligence and reasonable efforts have been made to ensure accuracy, including review by the authors wherever possible. If you identify any issues, please contact us.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fdata.2025.1678863/full#supplementary-material>

References

AbuKhalil, T., Alqaralleh, B. A. Y., and Al-Omari, A. H. (2022). Optimal deep learning based inception model for cervical cancer diagnosis. *Comput. Mater. Contin.* 72, 57–71. doi: 10.32604/cmc.2022.024367

Ahishakiye, E., and Kanobe, F. (2024). Optimizing cervical cancer classification using transfer learning with deep gaussian processes and support vector machines. *Discov Artif Intell.* 4:73. doi: 10.1007/s44163-024-00185-6

Ahmed, R., Dahmani, N., Dahy, G., Hassanien, A. E., and Darwish, A. (2024). Early detection and categorization of cervical cancer cells using smoothing cross entropy-based multi-deep transfer learning. *IEEE Access* 12, 157838–157853. doi: 10.1109/ACCESS.2024.3485888

Akash, R. S., Islam, R., Badhon, S. M. S. I., and Hossain, K. S. M. T. (2024). CerviXpert: a multi-structural convolutional neural network for predicting cervix type and cervical cell abnormalities. *Digit. Heal.* 10. doi: 10.1177/20552076241295440

Akbar, A. H., Sitanggang, I. S., Agmalaro, M. A., Haryanto, T., Rulaningtyas, R., Husin, N. A., et al. (2024). Layer selection on residual network for feature extraction of pap smear images. *J. Adv. Res. Appl. Sci. Eng. Technol.* 36, 56–66. doi: 10.37934/araset.36.2.566

Albuquerque, T., Rosado, L., Cruz, R., Vasconcelos, M. J. M., Oliveira, T., Cardoso, J. S., et al. (2023). Rethinking low-cost microscopy workflow: image enhancement using deep based extended depth of field methods. *Intell. Syst. Appl.* 17:200170. doi: 10.1016/j.iswa.2022.200170

Alias, N. A., Mustafa, W. A., Jamlos, M. A., Alquran, H., Hanafi, H. F., Ismail, S., et al. (2022). Pap smear images classification using machine learning: a literature matrix. *Diagnostics* 12:2900. doi: 10.3390/diagnostics12122900

Allahqoli, L., Laganà, A. S., Mazidmoradi, A., Salehiniya, H., Günther, V., Chiantera, V., et al. (2022). Diagnosis of cervical cancer and pre-cancerous lesions by artificial intelligence: a systematic review. *Diagnostics* 12:2771. doi: 10.3390/diagnostics12112771

AlMohimeed, A., Shehata, M., El-Rashidy, N., Mostafa, S., Saleh, H., and Samy Talaat, A. (2024). ViT-PSO-SVM: cervical cancer predication based on integrating vision transformer with particle swarm optimization and support vector machine. *Bioengineering* 11:729. doi: 10.3390/bioengineering11070729

Alohal, M. A. M. A., El-Rashidy, N., Alaklabi, S., Elmannai, H., Alharbi, S., Saleh, H., et al. (2024). Swin-GA-RF: genetic algorithm-based swin transformer and random forest for enhancing cervical cancer classification. *Front. Oncol.* 14:1392301. doi: 10.3389/fonc.2024.1392301

Alquran, H., Abdi, R. A., Alsalat, M., Mustafa, W. A., and Ismail, A. R. (2022a). Cervical net: a novel cervical cancer classification using feature fusion. *Bioengineering* 9:578. doi: 10.3390/bioengineering9100578

Alquran, H., Qasmieh, I. A., Alqudah, A. M., Mustafa, W. A., Yacob, Y. M., Alsalat, M., et al. (2022b). Cervical cancer classification using combined machine learning and deep learning approach. *Comput. Mater. Contin.* 72, 5117–5134. doi: 10.32604/cmc.2022.025692

Alsalat, M., Alquran, H., Mustafa, W. A., Yacob, Y. M., and Alayed, A. A. (2022). Analysis of cytology pap smear images based on ensemble deep learning approach. *Diagnostics* 12:2756. doi: 10.3390/diagnostics12112756

Alsalat, M., Alquran, H., Zyout, A., Alqudah, A. M., Mustafa, W. A., Kaifi, R., et al. (2023). A new weighted deep learning feature using particle swarm and ant lion optimization for cervical cancer diagnosis on pap smear images. *Diagnostics* 13:2762. doi: 10.3390/diagnostics1312762

An, H., Ding, L., Ma, M., Huang, A., Gan, Y., Sheng, D., et al. (2023). Deep learning-based recognition of cervical squamous interepithelial lesions. *Diagnostics* 13:1720. doi: 10.3390/diagnostics13101720

Anandavally, P. S. N., and Bai, V. M. A. (2024). Deep neural network for the detection and classification of spontaneous abortion associated with cervical cancer. *J. Adv. Res. Appl. Sci. Eng. Technol.* 39, 19–36. doi: 10.37934/araset.39.2.1936

Ando, Y., Ko, S., Han, H., Cho, J., and Park, N. J. Y. (2024). Toward interpretable cell image representation and abnormality scoring for cervical cancer screening using pap smears. *Bioengineering* 11:567. doi: 10.3390/bioengineering11060567

Anupama, C. S. S., Benedict Jose, T. J., Eid, H. F., Aljehane, N. O., Al-Wesabi, F. N., Obayya, M., et al. (2022). Intelligent classification model for biomedical pap smear images on iot environment. *Comput. Mater. Contin.* 71, 3969–3983. doi: 10.32604/cmc.2022.022701

Attallah, O. (2023). Cervical cancer diagnosis based on multi-domain features using deep learning enhanced by handcrafted descriptors. *Appl. Sci. Switz.* 13:1916. doi: 10.3390/app13031916

Battula, K. P., and Chandana, B. S. (2022). Deep learning based cervical cancer classification and segmentation from pap smears images using an efficientnet. *Int. J. Adv. Comput. Sci. Appl.* 13, 899–908. doi: 10.14569/IJACSA.2022.01309104

Battula, K. P., and Sai Chandana, B. (2023). Multi-class cervical cancer classification using transfer learning-based optimized SE-ResNet152 model in pap smear whole slide images. *Int. J. Electr. Comput. Eng. Syst.* 14, 613–623. doi: 10.32985/ijeces.14.6.1

Benhari, M., and Hosseini, R. (2022). An improved fuzzy deep learning (IFDL) model for managing uncertainty in classification of pap-smear cell images. *Intell. Syst. Appl.* 16:200133. doi: 10.1016/j.iswa.2022.200133

Bera, A., Bhattacharjee, D., and Krejcar, O. (2024). PND-Net: plant nutrition deficiency and disease classification using graph convolutional network. *Sci. Rep.* 14:15537. doi: 10.1038/s41598-024-66543-7

Bolia, C., and Joshi, S. (2025). Optimized deep neural network for high-precision psoriasis classification from dermoscopic images. *Rev. Científica Sist e Informática* 5:e966. doi: 10.51252/rsci.v5i2.996

Bora, K., Borah, K., Chyrmang, G., Barua, B., Das, H. S., Mahanta, L. B., et al. (2022). Machine learning based approach for automated cervical dysplasia detection using multi-resolution transform domain features. *Mathematics* 10:4126. doi: 10.3390/math10214126

Cabral, B. P., Braga, L. A. M., Conte Filho, C. G., Penteado, B., and Freire de Castro Silva, S. L., et al. (2025). Future use of AI in diagnostic medicine: 2-wave cross-sectional survey study. *J. Med. Internet Res.* 27:e53892. doi: 10.2196/53892

Chauhan, N. K., Kumar, A., Singh, K., and Kolambakar, S. B. (2023). BHDFCN: a robust hybrid deep network based on feature concatenation for cervical cancer diagnosis on WSI Pap smear slides. *Biomed. Res. Int.* 2023. doi: 10.1155/2023/4214817

Chen, W., Shen, W., Gao, L., and Li, X. (2022). Hybrid Loss-Constrained Lightweight convolutional neural networks for cervical cell classification. *Sensors* 22:3272. doi: 10.3390/s22093272

Cheng, S., Liu, S., Yu, J., Rao, G., Xiao, Y., Han, W., et al. (2021). Robust whole slide image analysis for cervical cancer screening using deep learning. *Nat. Commun.* 12:5639. doi: 10.1038/s41467-021-25296-x

Cheng, Y., Wu, H., Wu, F., Wang, Y., Jiang, W., Xiong, M., et al. (2025). Fine-grained pathomorphology recognition of cervical lesions with a dropped multibranch swin transformer. *Quant. Imaging Med. Surg.* 15, 3551–3564. doi: 10.21037/qims-24-1590

Chowdary, G. J. J., Suganya, G., Premalatha, M., and Yogarajah P. (2023). Nucleus segmentation and classification using residual SE-UNet and feature concatenation approach incervical cytopathology cell images. *Technol. Cancer Res. Treat.* 22. doi: 10.1177/15330338221134833

Devi, S., Gaikwad, S. R., and Harikrishnan, R. (2023). Prediction and detection of cervical malignancy using machine learning models. *Asian Pacific J. Cancer Prev.* 24, 1419–1433. doi: 10.31557/APJCP.2023.24.4.1419

Dhawan, S., Singh, K., and Arora, M. (2018). Cervix image classification for prognosis of cervical cancer using deep neural network with transfer learning. *EAI Endorsed Trans. Pervasive Heal. Technol.* 169183:e5. doi: 10.4108/eai.12-4-2021.169183

Diniz, D. N., Keller, B. N. S., Bianchi, A. G. C., Luz, E. J. S., Souza, M. J. F., Rezende, M. T., et al. (2022). A cytopathologist eye assistant for cell screening. *Appliedmath* 2, 659–674. doi: 10.3390/appliedmath2040038

Du, H., Dai, W., Wang, C., Tang, J., Wu, R., Zhou, Q., et al. (2023). AI-assisted system improves the work efficiency of cytologists via excluding cytology-negative slides and accelerating the slide interpretation. *Front. Oncol.* 13:1290112. doi: 10.3389/fonc.2023.1290112

Erpal, B., and Sefer, E. (2024). Reference-free inferring of transcriptomic events in cancer cells on single-cell data. *BMC Cancer* 24:607. doi: 10.1186/s12885-024-12331-5

Fahad, N. M. N. M., Azam, S., Montaha, S., Mukta, M. S. H. M. S. H., Azam, S., Montaha, S., et al. (2024). Enhancing cervical cancer diagnosis with graph convolution network: AI-powered segmentation, feature analysis, and classification for early detection. *Multimed. Tools Appl.* 83, 75343–75367. doi: 10.1007/s11042-024-18608-y

Fang, M., Lei, X., Liao, B., and Wu, F. X. A. (2022). Deep neural network for cervical cell classification based on cytology images. *IEEE Access* 10, 130968–130980. doi: 10.1109/ACCESS.2022.3230280

Fang, M., Wu, F. X., Fu, M., Liao, B., and Lei, X. (2024). Deep integrated fusion of local and global features for cervical cell classification. *Comput. Biol. Med.* 171:108153. doi: 10.1016/j.combiomed.2024.108153

Fekri-Ershad, S., and Alsaffar, M. F. (2023). Developing a tuned three-layer perceptron fed with trained deep convolutional neural networks for cervical cancer diagnosis. *Diagnostics* 13:686. doi: 10.3390/diagnostics13040686

Gafeer, M. M., Alperstein, S., Appleby, R., Carniello, J., Heymann, J. J., Goyal, A., et al. (2025). Unsatisfactory pap test results: a critical patient management problem pre-analytically addressed by the cytopathology laboratory. *Diagn. Cytopathol.* 53, 10–17. doi: 10.1002/dc.25398

Galande, A. S., Thapa, V., Vijay, A., and John, R. (2024). High-resolution lensless holographic microscopy using a physics-aware deep network. *J. Biomed. Opt.* 29:106502. doi: 10.1117/1.JBO.29.10.106502

Gangrade, J., Kuthiala, R., Singh, Y. P., Solanki, S., Gangrade, S., and Manoj, R. (2025). A deep ensemble learning approach for squamous cell classification in cervical cancer. *Sci. Rep.* 15:7266. doi: 10.1038/s41598-025-91786-3

Gao, W., Xu, C., Li, G., Bai, N., Li, M., Zhang, Y., et al. (2022). Cervical cell image classification-based knowledge distillation. *Biomimetics* 7:195. doi: 10.3390/biomimetics7040195

Gorantla, R., Singh, R. K., Pandey, R., and Jain, M. (2019). "Cervical cancer diagnosis using CervixNet - a deep learning approach," In: 2019 IEEE 19th International Conference on Bioinformatics and Bioengineering (BIBE) (Piscataway, NJ: IEEE), 397–404. doi: 10.1109/BIBE.2019.00078

Haridas, S., and Jayamalar, T. A. (2023). Versatile detection of cervical cancer with i-WFCM and deep learning based RBM classification. *J. Mach. Comput.* 3, 238–250. doi: 10.53759/7669/jmc202303022

Harsono, A. B. A. B., Susiarno, H., Suardi, D., Owen, L., Fauzi, H., Kireina, J., et al. (2022). Cervical pre-cancerous lesion detection: development of smartphone-based VIA application using artificial intelligence. *BMC Res. Notes* 15:356. doi: 10.1186/s13104-022-06250-6

Hemalatha, K., Vetriselvi, V., Meignanamoorthi, D., and Aruna Gladys, A. (2023). CervixFuzzyFusion for cervical cancer cell image classification. *Biomed. Signal Process Control* 85:104920. doi: 10.1016/j.bspc.2023.104920

Hong, Z., Xiong, J., Yang, H., and Mo, Y. K. (2024). Lightweight low-rank adaptation vision transformer framework for cervical cancer detection and cervix type classification. *Bioengineering* 11:468. doi: 10.3390/bioengineering11050468

Jain, S., Jain, A., Jangid, M., and Shetty, S. (2024). Metaheuristic driven framework for classifying cervical cancer on smear images using deep learning approach. *IEEE Access* 12, 160805–160821. doi: 10.1109/ACCESS.2024.3482975

Janani, S., and Christopher, D. F. X. (2023). Conditional super resolution generative adversarial network for cervical cell image enhancement. *SSRG Int. J. Electr. Electron. Eng.* 10, 70–76. doi: 10.14445/23488379/IJEEE-V10I4P107

Ji, J., Zhang, W., Dong, Y., Lin, R., Geng, Y., Hong, L., et al. (2023). Automated cervical cell segmentation using deep ensemble learning. *BMC Med. Imaging* 23:137. doi: 10.1186/s12880-023-01096-1

Ji, M., Xue, R., Su, W., Kong, Y., Fei, Y., Ma, J., et al. (2022). Early detection of cervical cancer by fluorescence lifetime imaging microscopy combined with unsupervised machine learning. *Int. J. Mol. Sci.* 23:11476. doi: 10.3390/ijms231911476

Jiang, P., Chen, Y., Wang, L., Chen, H., Feng, J., Liu, J., et al. (2023). A systematic review of deep learning-based cervical cytology screening: from cell identification to whole slide image analysis. *Artif. Intell. Rev.* 56, 2687–2758. doi: 10.1007/s10462-023-10588-z

Joynab, N. S., Islam, M. N., Aliya, R. R., Hasan, A. S. M. R., Khan, N. I., Sarker, I. H., et al. (2024). A federated learning aided system for classifying cervical cancer using PAP-SMEAR images. *Informatics Med. Unlocked* 47:101496. doi: 10.1016/j.imu.2024.101496

Kalbhor, M., Shinde, S., Lahade, S., and Choudhury, T. (2023a). DeepCerviCancer-deep learning-based cervical image classification using colposcopy and cytology images. *EAI Endorsed Trans. Pervasive Heal. Technol.* 9, 1–24. doi: 10.4108/eptpt.9.3473

Kalbhor, M., Shinde, S., Popescu, D. E., and Hemanth, D. J. (2023b). Hybridization of deep learning pre-trained models with machine learning classifiers and fuzzy min-max neural network for cervical cancer diagnosis. *Diagnostics* 13:1363. doi: 10.3390/diagnostics13071363

Kalbhor, M., Shinde, S., Wajire, P., and Jude, H. (2023c). CerviCell-detector: An object detection approach for identifying the cancerous cells in pap smear images of cervical cancer. *Heliyon* 9:e22324. doi: 10.1016/j.heliyon.2023.e22324

Kalbhor, M., Shinde, S. V., and Jude, H. (2022). Cervical cancer diagnosis based on cytology pap smear image classification using fractional coefficient and machine learning classifiers. *Telkomnika Telecommun Comput Electron Control* 20, 1091–1102. doi: 10.12928/telkomnika.v20i5.22440

Kanavati, F., Hirose, N., Ishii, T., Fukuda, A., Ichihara, S., Tsuneki, M. A., et al. (2022). Deep learning model for cervical cancer screening on liquid-based cytology specimens in whole slide images. *Cancers* 14:1159. doi: 10.3390/cancers14051159

Kang, J., and Li, N. (2024). CerviSegNet-DistillPlus: an efficient knowledge distillation model for enhancing early detection of cervical cancer pathology. *IEEE Access* 12, 85134–85149. doi: 10.1109/ACCESS.2024.3415395

Karamti, H., Alharthi, R., and Al Anizi, A., Alhebshi, R. M., Eshmawi, A. A., Alsulai, S., et al. (2023). Improving prediction of cervical cancer using KNN imputed SMOTE features and multi-model ensemble learning approach. *Cancers* 15:4412. doi: 10.3390/cancers15174412

Kavitha, R., Jothi, D. K. K., Saravanan, K., Swain, M. P. M. P., González, J. L. A., J. L. A., Bhardwaj, R. J. R. J., et al. (2023). Ant colony optimization-enabled CNN deep learning technique for accurate detection of cervical cancer. *Biomed. Res. Int.* 2023, 1–9. doi: 10.1155/2023/1742891

Khan, A., Han, S., Ilyas, N., Lee, B., and Lee, Y. M. (2023). CervixFormer: a multi-scale swin transformer-based cervical pap-smear WSI classification framework. *Comput. Methods Programs Biomed.* 240:107718. doi: 10.1016/jcmpb.2023.107718

Khanarsa, P., and Kitsirawan, S. (2024). Deep learning-based ensemble approach for conventional pap smear image classification. *Ecti Trans. Comput. Inf. Technol.* 18, 101–111. doi: 10.37936/ecti-cit.2024181.254621

Khozaimi, A., and Mahmudy, W. F. (2024). New insight in cervical cancer diagnosis using convolution neural network architecture. *IAES Int. J. Artif. Intell.* 13, 3092–3100. doi: 10.11591/ijai.v13.i3.pp3092-3100

Kurita, Y., Meguro, S., Kosugi, I., Enomoto, Y., Iwashita, T., Tsuyama, N., et al. (2023). Accurate deep learning model using semi-supervised learning and noisy student for cervical cancer screening in low magnification images. *PLoS One* 18:e0285996. doi: 10.1371/journal.pone.0285996

Kurniawati, Y. E., and Prabowo, Y. D. (2022). Model optimisation of class imbalanced learning using ensemble classifier on over-sampling data. *IAES Int. J. Artif. Intell.* 11, 276–283. doi: 10.11591/ijai.v11.i1.pp276-283

Li, G., Fan, X., Xu, C., Lv, P., Wang, R., Ruan, Z., et al. (2025). Detection of cervical cell based on multi-scale spatial information. *Sci. Rep.* 15:3117. doi: 10.1038/s41598-025-87165-7

Lilhore, U. K. U. K., Poongodi, M., Hamdi, M., Kaur, A., Simaiya, S., Algarni, A. D. A. D., et al. (2022). Hybrid model for detection of cervical cancer using causal analysis and machine learning techniques. *Comput. Math. Methods Med.* 2022, 1–17. doi: 10.1155/2022/4688327

Liu, J., Fan, H., Wang, Q., Li, W., Tang, Y., Wang, D., et al. (2022). Local label point correction for edge detection of overlapping cervical cells. *Front. Neuroinform.* 16:895290. doi: 10.3389/fninf.2022.895290

Lotfi, M., and Momenzadeh, M. (2022). Detection of cervical precancerous cells from Pap-smear images using ensemble classification. *Med. J. Tabriz Univ. Med. Sci.* 44, 281–289. doi: 10.34172/mj.tabriz.2022.034

Luo, D., Kang, H., Quan, T., Liu, X., Long, J., Zhang, J., et al. (2022). Dual supervised sampling networks for real-time segmentation of cervical cell nucleus. *Comput. Struct. Biotechnol. J.* 20, 4360–4368. doi: 10.1016/j.csbj.2022.08.023

Mahmoud, H. A. H., Alarfaj, A. A., and Hafez, A. M. A. (2022). Fast hybrid classification algorithm with feature reduction for medical images. *Appl. Bionics. Biomech.* 2022:1367366. doi: 10.1155/2022/1367366

Mahyari, T. L., and Dansereau, R. M. (2022). Multi-layer random walker image segmentation for overlapped cervical cells using probabilistic deep learning methods. *IET Image Process* 16, 2959–2972. doi: 10.1049/ipt2.12531

Mansouri, A. R., Ragab, M. (2023). Equilibrium optimization algorithm with ensemble learning based cervical precancerous lesion classification model. *Healthc. Switz.* 11:55. doi: 10.3390/healthcare11010055

Martínez, R. G. (2005). "What's wrong with me?": cervical cancer in Venezuela-living in the borderlands of health, disease, and illness. *Soc. Sci. Med.* 61, 797–808. doi: 10.1016/j.socscimed.2004.08.050

Mathivanan, S. K. S. K., Francis, D., Srinivasan, S., Khatavkar, V., and Shah, P. K. (2024). Enhancing cervical cancer detection and robust classification through a fusion of deep learning models. *Sci. Rep.* 14:10812. doi: 10.1038/s41598-024-61063-w

Maurya, R., Nath Pandey, N., and Kishore Dutta, M. (2023). VisionCervix: Papanicolaou cervical smears classification using novel CNN-Vision ensemble approach. *Biomed. Signal Process Control* 79:104156. doi: 10.1016/j.bspc.2022.104156

Maurya, R., Rajput, L., and Mahapatra, S. (2024). Deep and domain specific feature-based cervical cancer classification using support vector machine optimized with particle swarm optimization. *IEEE Access* 12, 193960–193971. doi: 10.1109/ACCESS.2024.3519806

Mazroa, A. A., Ishak, M. K., Aljarbouh, A., and Mostafa, S. M. (2023). Improved bald eagle search optimization with deep learning-based cervical cancer detection and classification. *IEEE Access* 11, 135175–135184. doi: 10.1109/ACCESS.2023.3337032

Mohammed, B. A., Senan, E. M., Al-Mekhlafi, Z. G., Alazmi, M., Alayba, A. M., Alanazi, A. A., et al. (2022). Hybrid techniques for diagnosis with wsis for early detection of cervical cancer based on fusion features. *Appl. Sci.* 12:8836. doi: 10.3390/app12178836

Muksimova, S., Umirzakova, S., Cho, Y. I., and Baltayev, J. (2025). RL-Cervix.Net: a hybrid lightweight model integrating reinforcement learning for cervical cell classification. *Diagnostics* 15:364. doi: 10.3390/diagnostics15030364

Muksimova, S., Umirzakova, S., Shoraimov, K., Baltayev, J., Cho, Y. I. Y. I., Shoraimov, K., et al. (2024). Novelty classification model use in reinforcement learning for cervical cancer. *Cancers* 16:3782. doi: 10.3390/cancers16223782

Mustafa, W. A., Khiruddin, K., Khusairi, F. Y., Ismail S., and Jamaludin, K. R. (2025). Comparative analysis of cervical cell classification using machine learning algorithms. *J. Electron. Electromed. Eng. Med.* 7, 646–662. doi: 10.35882/jeeemi.v7i3.829

Nazir, N., Sarwar, A., Saini, B. S., and Shams, R. A. (2023). Robust deep learning approach for accurate segmentation of cytoplasm and nucleus in noisy pap smear images. *Comput. Math. Methods Med.* 11:195. doi: 10.3390/computation11100195

Nour, M. K., Issaoui, I., Edris, A., Mahmud, A., Assiri, M., Ibrahim, S. S., et al. (2024). Computer aided cervical cancer diagnosis using gazelle optimization algorithm with deep learning model. *IEEE Access* 12, 13046–13054. doi: 10.1109/ACCESS.2024.3351883

Ontor, M. Z. H., Ali, M. M., Ahmed, K., Bui, F. M., Al-Zahrani, F. A., Hasan Mahmud, S. M., et al. (2023). Early-stage cervical cancerous cell detection

from cervix images using YOLOv5. *Comput. Mater. Contin.* 74, 3727–3741. doi: 10.32604/cmc.2023.032794

Ouh, Y. T., Kim, T. J., Ju, W., Kim, S. W., Jeon, S., Kim, S. N., et al. (2024). Development and validation of artificial intelligence-based analysis software to support screening system of cervical intraepithelial neoplasia. *Sci. Rep.* 14:1957. doi: 10.1038/s41598-024-51880-4

Pang, W., Jiang, H., Yu, Q., and Ma, Y. (2025). Cells grouping detection and confusing labels correction on cervical pathology images. *Bioengineering* 12:23. doi: 10.3390/bioengineering12010023

Papanicolau, G. N., and Traut, H. F. (1941). The diagnostic value of vaginal smears in carcinoma of the uterus. *Am. J. Obstet. Gynecol.* 42, 193–206. doi: 10.1016/S0002-9378(16)40621-6

Petrov, M., and Sokolov, I. (2023). Machine learning allows for distinguishing precancerous and cancerous human epithelial cervical cells using high-resolution AFM imaging of adhesion maps. *Cells* 12:2536. doi: 10.3390/cells12212536

Priya, S. A., and Bai, V. M. A. (2024). Variable kernel feature fusion and transfer learning for pap smear image-based cervical cancer classification. *SSRG Int. J. Electron. Commun. Eng.* 11, 228–243. doi: 10.14445/23488549/IJECE-V11I11P119

Qin, M., Gao, Y., Sun, H., Gou, M., Zhou, N., Yao, Y., et al. (2022). Efficient cervical cell lesion recognition method based on dual path network. *Wirel. Commun. Mob. Comput.* 2022. doi: 10.1155/2022/8496751

Rasheed, A., Shirazi, S. H., Umar, A. I., Shahzad, M., Yousaf, W., Khan, Z., et al. (2023). Cervical cell's nucleus segmentation through an improved UNet architecture. *PLoS One* 18:e0283568. doi: 10.1371/journal.pone.0283568

Resmi, S., Singh, R. P., and Palaniappan, K. (2024). Automated cervical cytology image cell segmentation using enhanced multiresUNet with DCT and spectral domain attention mechanisms. *IEEE Access* 12, 189387–189408. doi: 10.1109/ACCESS.2024.3516935

Riana, D., Jamil, M., Hadianti, S., Na'am, J., Sutanto, H., Sukwadi, R. (2023). Model of watershed segmentation in deep learning method to improve identification of cervical cancer at overlay cells. *TEM J.* 12, 813–819. doi: 10.18421/TEM122-26

Rodríguez, M., Córdova, C., San Martín, S., and Benjumeda, I. (2024). Automated cervical cancer screening using single-cell segmentation and deep learning: enhanced performance with liquid-based cytology. *Computation* 12:232. doi: 10.3390/computation12120232

Rohini, D., and Kavitha, M. (2024). ABC-optimized CNN-GRU algorithm for improved cervical cancer detection and classification using multimodal data. *Int. J. Adv. Comput. Sci. Appl.* 15, 701–714. doi: 10.14569/IJACSA.2024.0150971

Sahoo, P., Saha, S., Mondal, S., Seera, M., Sharma, S. K., Kumar, M., et al. (2023). Enhancing computer-aided cervical cancer detection using a novel fuzzy rank-based fusion. *IEEE Access* 11, 145281–145294. doi: 10.1109/ACCESS.2023.3346764

Sefer, E. (2025). DRGAT: predicting drug responses via diffusion-based graph attention network. *J. Comput. Biol.* 32, 330–350. doi: 10.1089/cmb.2024.0807

Shandilya, G., Gupta, S., Bharany, S., Almogren, A., Altameem, A., Rehman, A. U. A., et al. (2024). Enhancing advanced cervical cell categorization with cluster-based intelligent systems by a novel integrated CNN approach with skip mechanisms and GAN-based augmentation. *Sci. Rep.* 14:29040. doi: 10.1038/s41598-024-80260-1

Shinde, S., Kalbhor, M., and Wajire, P. (2022). DeepCyto: a hybrid framework for cervical cancer classification by using deep feature fusion of cytology images. *Math. Biosci. Eng.* 19, 6415–6434. doi: 10.3934/mbe.2022301

Shiny, T. L., and Parasuraman, K. A. (2023). Graph-Cut guided ROI segmentation algorithm with lightweight deep learning framework for cervical cancer classification. *Int. J. Adv. Comput. Sci. Appl.* 14, 779–792. doi: 10.14569/IJACSA.2023.0141280

Skerritt, E., Crouch, B., Ramanujam, N., Miao, Z., Qiu, Q., Asiedu, M. N., et al. (2022). Multicolumn pocket colposcopy cervical cancer diagnostic algorithm for referral populations. *BME Front.* 2022:9823184. doi: 10.34133/2022/9823184

Song, J., Wang, L., Zhang, Y., Yan, J., and Feng, Y. (2024). Enhancing cervical precancerous lesion detection using African vulture optimization algorithm with deep learning model. *Biomed. Signal Process Control* 97:106665. doi: 10.1016/j.bspc.2024.106665

Ssedyabane, F., Niyonzima, N., Nambi Najjuma, J., Birungi, A., Atwine, R., Tusbira, D., et al. (2024). Prevalence of cervical intraepithelial lesions and associated factors among women attending a cervical cancer clinic in Western Uganda; results based on Pap smear cytology. *SAGE Open Med.* 12. doi: 10.1177/20503121241252265

Stegmüller, T., Abbet, C., Bozorgtabar, B., Thiran, J. P., Clarke, H., Petignat, P., et al. (2024). Self-supervised learning-based cervical cytology for the triage of HPV-positive women in resource-limited settings and low-data regime. *Comput. Biol. Med.* 169:107809. doi: 10.1016/j.combiomed.2023.107809

Sudhakar, K., Saravanan, D., Hariharan, G., Sanaj, M. S., Kumar, S., Shaik, M., et al. (2023). Optimised feature selection-driven convolutional neural network using gray level co-occurrence matrix for detection of cervical cancer. *Open Life Sci.* 18:20220770. doi: 10.1515/biol-2022-0770

Suksmono, A. B., Ismayanto, D. F., Rulaningtyas, R., Nabila, A. N. L., Maharani, R. N., and Katherine, et al. (2021). Classification of adeno carcinoma, high squamous intraepithelial lesion, and squamous cell carcinoma in Pap smear images based on extreme learning machine. *Comput. Methods Biomed. Eng. Imaging Vis.* 9, 115–120. doi: 10.1080/21681163.2020.1817793

Tan, S. L., Selvachandran, G., Ding, W., Paramesran, R., and Kotecha, K. (2024). Cervical Cancer classification from pap smear images using deep convolutional neural network models. *Interdiscip. Sci. Comput. Life Sci.* 16, 16–38. doi: 10.1007/s12539-023-00589-5

Tang, J., Zhang, T., Gong, Z., Huang, X., Zhang, T., Huang, X., et al. (2023). High precision cervical precancerous lesion classification method based on ConvNeXt. *Bioengineering* 10:1424. doi: 10.3390/bioengineering10121424

Tian, C., Liu, X., Bai, J., Zeng, S., Cheng, S., Chen, L., et al. (2024). Disentanglement of content and style features in multi-center cytology images via contrastive self-supervised learning. *Biomed. Signal Process Control* 95:106395. doi: 10.1016/j.bspc.2024.106395

Torres-Roman, J. S., Ronceros-Cardenas, L., Valcarcel, B., Arce-Huamani, M. A., Bazalar-Palacios, J., Ybasaeta-Medina, J., et al. (2021). Cervical cancer mortality in Peru: regional trend analysis from 2008–2017. *BMC Public Health* 21:219. doi: 10.1186/s12889-021-10274-1

Utomo, C. P., Suhaeni, N., Insani, N., Suherlan, E., Fathurachman, M., Rahmah, N. A., et al. (2025). Optimizing image preprocessing for AI-driven cervical cancer diagnosis. *Adv. Sustain. Sci. Eng. Technol.* 7:0250111-01–0250111-011. doi: 10.26877/asset.v7i1.1128

Waly, M. I., Sikkandar, M. Y., Aboamer, M. A., Kadry, S., and Thinnukool, O. (2022). Optimal deep convolution neural network for cervical cancer diagnosis model. *Comput. Mater. Contin.* 70, 3297–3309. doi: 10.32604/cmc.2022.020713

Wang, J., Yu, Y., Tan, Y., Wan, H., Zheng, N., He, Z., et al. (2024). Artificial intelligence enables precision diagnosis of cervical cytology grades and cervical cancer. *Nat. Commun.* 15:4369. doi: 10.1038/s41467-024-48705-3

Wang, K., Fei, X., Su, L., Fang, T., and Shen, H. (2025). Auxiliary meta-learning strategy for cancer recognition leveraging external data and optimized feature mapping. *BMC Cancer* 25:367. doi: 10.1186/s12885-025-13740-w

Wong, L., Ccoba, A., Diaz, E., Valcarcel, S., Mauricio, D., Villoslada, V., et al. (2023). Deep learning and transfer learning methods to effectively diagnose cervical cancer from liquid-based cytology pap smear images. *Int. J. Online Biomed. Eng.* 19, 77–93. doi: 10.3991/ijoe.v19i04.37437

Wu, N., Jia, D., Zhang, C., and Li, Z. (2023a). Cervical cell extraction network based on optimized yolo. *Math. Biosci. Eng.* 20, 2364–2381. doi: 10.3934/mbe.2023111

Wu, N., Jia, D., Zhang, C., and Li, Z. (2023b). Cervical cell classification based on strong feature CNN-LSTM network using adaboost optimization. *J. Intell. Fuzzy Syst.* 44, 4335–4355. doi: 10.3233/JIFS-221604

Wubineh, B. Z., Rusiecki, A., and Halawa, K. (2024a). Classification of cervical cells from the Pap smear image using the RES_DCGAN data augmentation and ResNet50V2 with self-attention architecture. *Neural Comput. Appl.* 36, 21801–21815. doi: 10.1007/s00521-024-10404-x

Wubineh, B. Z., Rusiecki, A., and Halawa, K. (2024b). Segmentation and classification techniques for pap smear images in detecting cervical cancer: a systematic review. *IEEE Access* 12, 118195–118213. doi: 10.1109/ACCESS.2024.3447887

Xu, C., Li, M., Li, G., Sun, C., Bai, N., Zhang, Y., et al. (2022). Cervical cell/clumps detection in cytology images using transfer learning. *Diagnostics* 12:2477. doi: 10.3390/diagnostics12102477

Yamagishi, Y., and Hanaoka, S. (2025). “Benchmarking image models including CNNs, transformers, and hybrid architectures for cervical cell classification” in *Proceedings International Symposium on Biomedical Imaging* (Piscataway, NJ: IEEE). doi: 10.1109/ISBI60581.2025.10980788

Yang, G., Huang, J., He, Y., Chen, Y., Wang, T., Jin, C., et al. (2022). GCP-Net: a gating context-aware pooling network for cervical cell nuclei segmentation. *Mob. Inf. Syst.* 2022. doi: 10.1155/2022/7511905

Yang, H., Aydi, W., Innab, N., Ghoneim, M. E., and Ferrara, M. (2024). Classification of cervical cancer using dense CapsNet with Seg-UNet and denoising autoencoders. *Sci. Rep.* 14:31764. doi: 10.1038/s41598-024-82489-2

Ybasaeta-Medina, J., Ybasaeta-Soto, L., Ossco-Torres, O., Aquije-Paredes, C., and Hernández-Huariapauar, E. (2025). Sociodemographic, behavioral, and clinical risk factors associated with cervical dysplasia: a case-control study. *Medwave* 25, e3015. doi: 10.5867/medwave.2025.01.3015

Yi, J., Liu, X., Zeng, S., Cheng, S., and Chen, L. (2024). Multi-scale window transformer for cervical cytopathology image recognition. *Comput. Struct. Biotechnol.* J. 24, 314–321. doi: 10.1016/j.csbj.2024.04.028

Yin, J., Zhang, Q., Xi X, Liu, M., Lu, W., and Tu, H. (2024). Enhancing cervical cell detection through weakly supervised learning with local distillation mechanism. *IEEE Access* 12, 77104–77113. doi: 10.1109/ACCESS.2024.3407066

Zammataro, L. (2024). CINNAMON-GUI. *Revolutionizing pap smear analysis with CNN-based digital pathology image classification.* F1000research 13:897. doi: 10.12688/f1000research.154455.1

Zhang, B., Jiang, X., and Zhao, W. (2024a). An enhanced mask transformer for overlapping cervical cell segmentation based on DETR. *IEEE Access* 12, 176586–176597. doi: 10.1109/ACCESS.2024.3505616

Zhang, B., Wang, W., Zhao, W., Jiang, X., and Patnaik, L. M. (2024b). An improved approach for automated cervical cell segmentation with PointRend. *Sci. Rep.* 14:14210. doi: 10.1038/s41598-024-64583-7

Zhang, Y., Ning, C., and Yang, W. (2025). An automatic cervical cell classification model based on improved DenseNet121. *Sci. Rep.* 15:3240. doi: 10.1038/s41598-025-87953-1

Zhao, Y., Fu, C., Zhang, W., Ye, C., Wang, Z., Ma, H. F., et al. (2023). Automatic segmentation of cervical cells based on star-convex polygons in pap smear images. *Bioengineering* 10:47. doi: 10.3390/bioengineering10010047