Deep Learning in Gynecologic Cancer Diagnosis: Current Advances, Challenges, and Future Directions

Anonymous Author(s)
Affiliation
Address
email

Abstract

Gynecologic cancers pose significant global health challenges, especially in low-resource settings with limited diagnostic expertise. Deep learning (DL) presents a promising avenue for automating image analysis, offering more consistent and accurate diagnoses. This review assesses current DL applications, challenges, and future prospects in gynecologic cancer diagnosis across diverse imaging modalities. Following PRISMA-2 guidelines, a literature review evaluated studies utilizing DL for diagnosing gynecologic cancers via MRI, CT scans, Pap smears, and colposcopy. Data extraction and quality assessment were conducted using the QUADAS-2 tool, with diagnostic performance evaluated using R software. Of 48 reviewed studies, 24 met inclusion criteria for meta-analysis. DL models, primarily ResNet, VGGNet, and UNet, demonstrated higher sensitivity (89.40%) but slightly lower specificity (87.6%) compared to traditional machine learning (ML) methods (sensitivity: 68.1%, specificity: 94.1%). DL models exhibited an AUC of 0.88, indicating high diagnostic accuracy. Challenges including study heterogeneity and methodological biases highlight the need for standardized protocols. Despite obstacles, DL shows promise in gynecologic cancer diagnosis, particularly in resource-limited settings. Addressing these challenges can enhance DL's clinical utility and improve patient outcomes.

1 Introduction

Gynecologic cancers, including cervical, ovarian, and endometrial cancers, pose significant health risks to women globally, especially in low-resource settings[11, 2, 3]. Traditional diagnostic methods relying on human interpretation are inconsistent and error-prone [2, 3, 7]. Recent advances in DL offer automated, consistent, and precise diagnostic capabilities using medical images [5, 10, 1, 12, 8, 4, 6]. This paper reviews recent DL applications in gynecologic cancer diagnosis, discusses their performance compared to conventional methods, and highlights future research directions aimed at overcoming current challenges.

2 Methods

A systematic review and meta-analysis following PRISMA-2[9] guidelines were conducted. The protocol was registered in PROSPERO under registration number 356104. We searched databases including PubMed, Embase, and Scopus for articles published between January 2018 and December 2022, focusing on DL applications for diagnosing several gynecologic cancers. Articles were screened and assessed using QUADAS-2 [13] for eligibility and quality. Data on diagnostic performance metrics, specifically sensitivity and specificity, were extracted and analyzed using the "meta" and "mada" packages. Pooled sensitivity, specificity, summary receiver operating characteristic (SROC) curves, and area under the curve (AUC) metrics were calculated using R software.

Submitted to 38th Conference on Neural Information Processing Systems (NeurIPS 2024). Do not distribute.

3 Experimental Results

The review included 48 studies, with 24 studies eligible for meta-analysis. The imaging modalities used across these studies included cytology (20 studies), colposcopy (15 studies), MRI (8 studies), CT scans (4 studies), and hysteroscopy (1 study). Popular DL models were ResNet, VGGNet, and UNet.

Figure 1: Overview of the diversity and popularity of models used in different studies for the diagnosis of gynecologic cancers

DL algorithms demonstrated high sensitivity (89.40%) but lower specificity (87.6%) compared to ML methods, which had lower sensitivity (68.1%) and higher specificity (94.1%). The AUC for DL algorithms was 0.88, indicating good diagnostic accuracy. 3D-UNet and EfficientNet-B3 were top-performing models, with EfficientNet-B3 achieving a classification accuracy of 99.01% for cervical cancer detection and 3D-UNet demonstrating high segmentation capabilities with an average Dice Similarity Coefficient (DSC) score of 0.93. Despite the promising results, significant heterogeneity

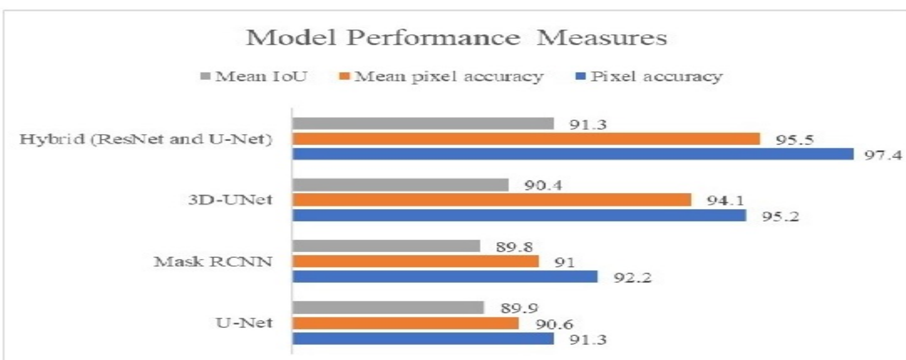

Figure 2: A Deep learning models used for abnormality detection

and risk of bias were noted, primarily due to variability in patient selection, imaging techniques, and DL model implementation.

4 Conclusion

Deep learning techniques offer substantial improvements in the accuracy and efficiency of diagnosing gynecological cancers using image-based data, surpassing traditional ML methods. However, challenges related to study heterogeneity and model biases need addressing to enhance the clinical applicability of DL models. Future research should focus on refining these algorithms, ensuring their robustness and generalizability, and addressing ethical considerations to fully leverage the potential of DL in gynecological cancer care. Collaborations between researchers, clinicians, and technologists are essential for advancing DL applications in gynecologic cancer diagnostics, ultimately improving early detection and patient outcomes.

References

[1] Hugo JWL Aerts, Emmanuel Rios Velazquez, Ralph TH Leijenaar, Chintan Parmar, Patrick Grossmann, Sara Carvalho, Johan Bussink, René Monshouwer, Benjamin Haibe-Kains, Derek Rietveld, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nature communications*, 5(1):4006, 2014.

[2] Rifat Atun, David A Jaffray, Michael B Barton, Freddie Bray, Michael Baumann, Bhadrasain Vikram, Timothy P Hanna, Felicia M Knaul, Yolande Lievens, Tracey YM Lui, et al. Expanding global access to radiotherapy. *The lancet oncology*, 16(10):1153–1186, 2015.

[3] Marc Brisson, Jane J Kim, Karen Canfell, Mélanie Drolet, Guillaume Gingras, Emily A Burger, Dave Martin, Kate T Simms, Élodie Bénard, Marie-Claude Boily, et al. Impact of hpv vaccination and cervical screening on cervical cancer elimination: a comparative modelling analysis in 78 low-income and lower-middle-income countries. *The Lancet*, 395(10224):575–590, 2020.

[4] Thibaud P Coroller, Patrick Grossmann, Ying Hou, Emmanuel Rios Velazquez, Ralph TH Leijenaar, Gretchen Hermann, Philippe Lambin, Benjamin Haibe-Kains, Raymond H Mak, and Hugo JWL Aerts. Ct-based radiomic signature predicts distant metastasis in lung adenocarcinoma. *Radiotherapy and Oncology*, 114(3):345–350, 2015.

[5] Oraianthi Fiste, Michalis Liontos, Flora Zagouri, Georgios Stamatakos, and Meletios Athanasios Dimopoulos. Machine learning applications in gynecological cancer: a critical review. *Critical Reviews in Oncology/Hematology*, 179:103808, 2022.

[6] Wentian Guo, Hui Li, Yitan Zhu, Li Lan, Shengjie Yang, Karen Drukker, Elizabeth Morris, Elizabeth Burnside, Gary Whitman, Maryellen L Giger, et al. Prediction of clinical phenotypes in invasive breast carcinomas from the integration of radiomics and genomics data. *Journal of medical imaging*, 2(4):041007–041007, 2015.

[7] Robert M Haralick and Linda G Shapiro. Image segmentation techniques. *Computer vision, graphics, and image processing*, 29(1):100–132, 1985.

[8] Philippe Lambin, Emmanuel Rios-Velazquez, Ralph Leijenaar, Sara Carvalho, Ruud GPM Van Stiphout, Patrick Granton, Catharina ML Zegers, Robert Gillies, Ronald Boellard, André Dekker, et al. Radiomics: extracting more information from medical images using advanced feature analysis. *European journal of cancer*, 48(4):441–446, 2012.

[9] Matthew J Page, Joanne E McKenzie, Patrick M Bossuyt, Isabelle Boutron, Tammy C Hoffmann, Cynthia D Mulrow, Larissa Shamseer, Jennifer M Tetzlaff, Elie A Akl, Sue E Brennan, et al. The prisma 2020 statement: an updated guideline for reporting systematic reviews. *Bmj*, 372, 2021.

[10] Neeraj Sharma and Lalit M Aggarwal. Automated medical image segmentation techniques. *Journal of medical physics*, 35(1):3–14, 2010.

[11] Hyuna Sung, Jacques Ferlay, Rebecca L Siegel, Mathieu Laversanne, Isabelle Soerjomataram, Ahmedin Jemal, and Freddie Bray. Global cancer statistics 2020: Globocan estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: a cancer journal for clinicians*, 71(3):209–249, 2021.

[12] Martin Vallières, Emily Kay-Rivest, Léo Jean Perrin, Xavier Liem, Christophe Furstoss, Hugo JWL Aerts, Nader Khaouam, Phuc Felix Nguyen-Tan, Chang-Shu Wang, Khalil Sultanem, et al. Radiomics strategies for risk assessment of tumour failure in head-and-neck cancer. *Scientific reports*, 7(1):10117, 2017.

[13] Penny F Whiting, Anne WS Rutjes, Marie E Westwood, Susan Mallett, Jonathan J Deeks, Johannes B Reitsma, Mariska MG Leeflang, Jonathan AC Sterne, Patrick MM Bossuyt, and QUADAS-2 Group*. Quadas-2: a revised tool for the quality assessment of diagnostic accuracy studies. *Annals of internal medicine*, 155(8):529–536, 2011.