

Clasificación del grado tumoral en imágenes histológicas de subtipos de cáncer de endometrio mediante análisis de color y textura con aprendizaje automático

Tumor grade classification in histological images of endometrial cancer subtypes using color and texture analysis with machine learning

HAYDEÉ ITZEL LIRA CASAS^a , NELLY GORDILLO CASTILLO^{a*} 

^aDepartamento de Ingeniería Eléctrica y Computación, Instituto de Ingeniería y Tecnología, Universidad Autónoma de Ciudad Juárez, México

*Autor de correspondencia. Correo electrónico: nelly.gordillo@uacj.mx

N.º de resumen	Formato
9CP25-11	Ponencia
Tema	Presentador
Procesamiento de señales digitales	Haydeé Itzel Lira Casas
Fecha de la presentación	Estatus
Mayo 20, 2025	Estudio terminado

Resumen

El cáncer de endometrio representa un reto diagnóstico importante dentro de la patología ginecológica, debido a la subjetividad en la evaluación del grado tumoral y la heterogeneidad morfológica entre subtipos moleculares. Esta investigación propone una metodología computacional robusta y automatizada para la clasificación del grado tumoral en los subtipos MSI, CN-LOW y POLE, a partir del análisis de imágenes histológicas del conjunto TCGA-UCEC. Se empleó una estrategia de preprocesamiento que incluyó eliminación de artefactos no relevantes mediante HistoQC, normalización de color con método de Macenko y mejora de contraste con CLAHE a 2.5x. Posteriormente, se segmentó el tejido tumoral mediante MiniBatchKMeans en el espacio de color LAB, y se generaron parches a 40x sobre regiones con al menos 30 % de contenido tumoral. A cada parche se le extrajeron características texturales (GLCM), estadísticas de color (momentos) y un descriptor híbrido RSHD, además de estimar la densidad nuclear con watershed y umbralización de Otsu. Para reducir la dimensionalidad, se aplicó NMF, seleccionando componentes óptimos por subtipo. La clasificación se realizó con XGBoost, utilizando validación estratificada a nivel de imagen (WSI). El modelo alcanzó una exactitud del 95.04 % en POLE, 85.67 % en MSI y 75.22 % en CN-LOW. Si bien la precisión varía según el subtipo y la distribución de grados, los resultados reflejan una diferenciación efectiva entre G1, G2 y G3. Esta propuesta reduce la subjetividad diagnóstica y aporta valor clínico al facilitar un análisis cuantitativo reproducible. Se considera una base sólida para integrar modelos automatizados en flujos de trabajo diagnósticos en histopatología.

Palabras clave: cáncer de endometrio; clasificación de grado tumoral; aprendizaje automático; procesamiento de imágenes.

Abstract

Endometrial cancer represents a significant diagnostic challenge in gynecological pathology due to the subjectivity in tumor grade assessment and the morphological heterogeneity among molecular subtypes. This study proposes a robust and automated computational methodology for tumor grade classification in the MSI, CN-LOW, and POLE subtypes using histological images from the TCGA-UCEC dataset. A preprocessing strategy was implemented, including artifact removal with HistoQC, color normalization using the Macenko method, and contrast enhancement via CLAHE at 2.5x magnification. Tumor tissue was subsequently segmented using MiniBatchKMeans in the LAB color space, and 40x magnification patches were extracted from regions with at least 30% tumor content. Each patch



was characterized by extracting texture features (GLCM), color statistics (moments), and a hybrid descriptor (RSHD), along with nuclear density estimation using Otsu thresholding and watershed segmentation. Dimensionality reduction was performed using NMF, selecting optimal components for each subtype. Classification was carried out with the XGBoost algorithm, using stratified validation at the whole slide image (WSI) level. The model achieved an accuracy of 95.04% for POLE, 85.67% for MSI, and 75.22% for CN-LOW. Although precision varies across subtypes and grade distributions, the results demonstrate effective differentiation among G1, G2, and G3. This approach reduces diagnostic subjectivity and offers clinical value by enabling a reproducible quantitative analysis. It is considered a solid foundation for integrating automated models into diagnostic workflows in histopathology.

Keywords: endometrial cancer; tumor grade classification; machine learning; image processing.

Entidad legal responsable del estudio

Universidad Autónoma de Ciudad Juárez

Financiamiento

Beca SECIHTI, CVU 1315380.

Conflictos de interés

Los autores declaran que no hay conflicto de intereses.