



ING. MECATRÓNICA

**Trabajo de Integración Curricular previo a la obtención del título de
Mechatronics Engineer.**

Autor: Julian Andre Guapaz Pozo

Tutor: Lorena de los Ángeles Guachi Guachi, Ph.D.

Desarrollo de un sistema de visión artificial para la obtención de imágenes ecográficas hepáticas de relevancia clínica
Development of an Artificial Vision System for the Acquisition of Clinically Relevant Hepatic Ultrasound Images

Quito-2026

Resumen

Las imágenes ecográficas se utilizan ampliamente por su bajo costo, portabilidad y carácter no invasivo; sin embargo, su adquisición depende en gran medida de la experiencia del operador. Este trabajo presenta un sistema de visión artificial para asistir a usuarios con experiencia limitada en la adquisición de cuatro vistas ecográficas hepáticas: *longitudinal*, *transversal*, *oblicua* y *hepatorrenal*. El sistema emplea dos estrategias para identificar *frames* informativos. Para la *vista longitudinal*, DeepLabV3+ segmenta la región ecográfica de interés (ROI) y el hígado, mientras que U-Net segmenta el lumen anecoico (LA); a partir de estas regiones se extraen características morfológicas, estadísticas y de textura para determinar la informatividad del *frame*. Para las vistas *transversal*, *oblicua* y *hepatorrenal*, DINOv2-Small extrae directamente de cada *frame* representaciones visuales profundas, posteriormente clasificadas mediante Random Forest, regresión logística y k-NN, respectivamente.

El sistema se desarrolló con 4,040 *frames* de tres voluntarios y se evaluó independientemente con 12 secuencias de aproximadamente 3 s, adquiridas por un voluntario sin experiencia tras aproximadamente 15 min de entrenamiento. El análisis comprendió 303 decisiones individuales de la *vista longitudinal* y 36 decisiones sobre ventanas no solapadas de DINOv2-Small. La evaluación binaria utilizó casos *clear* y *blurry* revisados por el radiólogo experto; los casos *medium* se reservaron para analizar incertidumbre. La *vista longitudinal* aceptó 0/101 *frames* informativos y rechazó 101/101 no informativos, con sensibilidad de 0,0%, especificidad de 100,0%, macro-F1 de 33,3% y *balanced accuracy* de 50,0%; por tanto, no generalizó al participante independiente. La *vista transversal* aceptó 4/4 ventanas informativas y rechazó 1/4 no informativas, con macro-F1 de 56,4% y *balanced accuracy* de 62,5%. Las vistas *oblicua* y *hepatorrenal* clasificaron correctamente 8/8 ventanas cada una, con macro-F1 y *balanced accuracy* de 100,0%, pero cada resultado procedió de solo un video informativo y uno no informativo y no demuestra generalización. No se combinaron *frames* y ventanas en una métrica global. La evaluación revela un colapso longitudinal asociado con LA y falsos positivos transversales, mientras que la evidencia de las otras dos vistas sigue siendo insuficiente.

Palabras clave: ecografía hepática, visión artificial, identificación de imágenes informativas.

Abstract

Ultrasound images are widely used because of their low cost, portability, and non-invasive nature; however, their acquisition depends heavily on the operator's experience. This work presents a computer vision system to assist users with limited experience in acquiring four hepatic ultrasound views: *longitudinal*, *transverse*, *oblique*, and *hepatorenal*. The system uses two strategies to identify informative *frames*. For the *longitudinal view*, DeepLabV3+ segments the ultrasound region of interest (ROI) and the liver, while U-Net segments the anechoic lumen (LA); morphological, statistical, and texture features are then extracted from these regions to determine *frame* informativity. For the *transverse*, *oblique*, and *hepatorenal views*, DINOv2-Small directly extracts deep visual representations from each *frame*, which are subsequently classified using Random Forest, logistic regression, and k-NN, respectively.

The system was developed using 4,040 *frames* from three volunteers and independently evaluated using 12 sequences of approximately 3 s, acquired by a novice volunteer after approximately 15 min of training. The analysis comprised 303 individual *longitudinal* decisions and 36 non-overlapping DINOv2-Small window decisions. Binary evaluation used expert-radiologist-reviewed *clear* and *blurry* cases; *medium* cases were reserved for uncertainty analysis. The *longitudinal view* accepted 0/101 informative *frames* and rejected 101/101 non-informative *frames*, yielding 0.0% sensitivity, 100.0% specificity, 33.3% macro-F1, and 50.0% balanced accuracy; it therefore did not generalize to the independent participant. The *transverse view* accepted 4/4 informative windows and rejected 1/4 non-informative windows, with 56.4% macro-F1 and 62.5% balanced accuracy. The *oblique* and *hepatorenal views* each classified 8/8 windows correctly, with 100.0% macro-F1 and balanced accuracy, but each result came from only one informative and one non-informative video and does not demonstrate generalization. Frames and windows were not combined into a global metric. The evaluation reveals longitudinal collapse associated with LA and transverse false positives, while evidence for the other two views remains insufficient.

Keywords: hepatic ultrasound, computer vision, informative-image identification.