



Daniel Alberto Cruz Moreno
Maestría en Inteligencia Analítica,
CADIT
dcruz@arttech.com.mx



María del Carmen Villar Patiño
Facultad de Ingeniería, CADIT
maria.villar@anahuac.mx

Introducción

El cáncer de cérvix es un problema que puede tratarse si se detecta en su evolución temprana y, comparado con otros tipos de cáncer, no representa un gasto significativo. La clave para el tratamiento adecuado es clasificar el tipo de cérvix dentro de tres posibilidades. Sin embargo, éste es uno de los principales problemas en las sociedades de mediano a bajos recursos debido a la falta de especialistas [1]. Para automatizar dicha clasificación empleando la visión por computadora, es importante que en las imágenes se seleccione la región de interés (ROI, por sus siglas en inglés), en este caso un área rectangular en la imagen médica que contenga al cérvix, la cuál será la entrada de un modelo predictivo [2].

El ser humano utiliza los ojos y el cerebro para captar su entorno sin esfuerzo. Por ejemplo, en la Figura 1 sabemos qué parte es el cérvix y cuál no (ruido). La visión por computadora no ha alcanzado totalmente este objetivo pero, gracias a los avances recientes, hay buenos resultados para automatizar la extracción, análisis y entendimiento de información en imágenes.

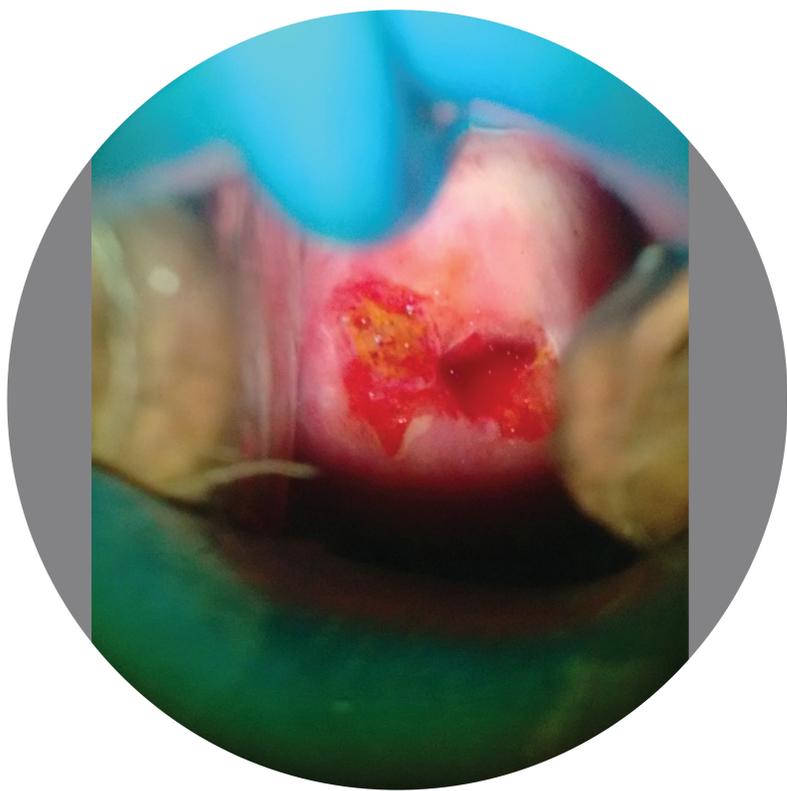


Figura 1.

Material y Método

Para demarcar el ROI, se empleó como característica discriminante los modelos de color: RGB, YCrCb, CIELAB y HSV, en la biblioteca OpenCV versión 6.3 del lenguaje Python versión 3.5 [3].

Una vez calibrados los componentes por modelo de color, se realizaron diferentes experimentos para generar el ROI del cérvix a una muestra de 705 imágenes, entre 1475 (1) con:

1. Cada modelo de color.
2. Cada modelo aplicado a un clúster de color creado con *K-means*.
3. Combinación de tres modelos.

Resultados

En la Figura 2 se observa un ejemplo de ROI mal y bien generado por el primer experimento. En la Figura 3, se observa una imagen transformada con *K-means* y su ROI. El tercer experimento empleó los tres mejores modelos arrojados por el primero.

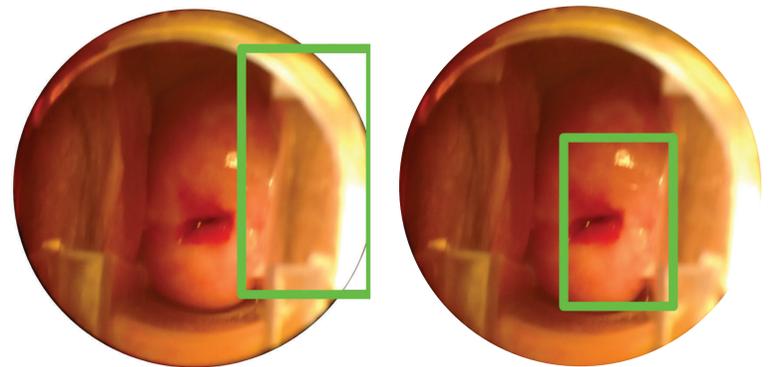


Figura 2.

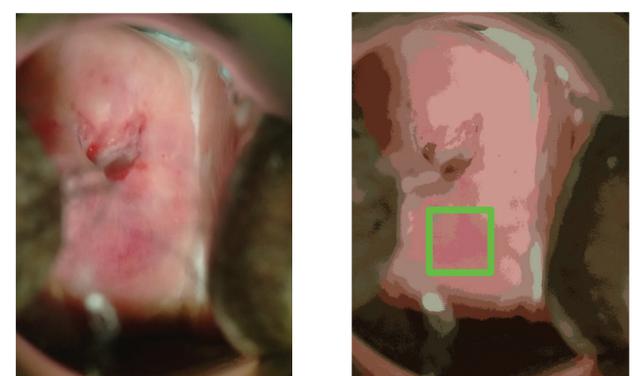


Figura 3.

EXPERIMENTO	MODELO	PRECISIÓN	TIEMPO (Seg)
1	RGB	72.06%	102
1	HSV	87.08%	102
1	CIELAB	92.91%	102
1	YCrCb	89.08%	102
2	RGB	82.84%	102
2	HSV	66.38%	102
2	CIELAB	83.26%	102
2	YCrCb	73.76%	102
3	HSV - CIELAB - YCrCb	94.81%	153

Cuadro 1. Comparación de precisión y tiempo en los tres experimentos.

Discusión

El tercer experimento es el mejor para seleccionar automáticamente el ROI que contiene menos ruido. Resaltar colores en gama de rojo, creando un clúster con *K-means*, no superó los resultados de los otros experimentos.

El tiempo de procesamiento entre uno y tres modelos no fue significativo, requirió de 51 segundos más procesar las 705 imágenes y obtener 1.9% de precisión extra.

Con estos resultados, el siguiente paso es utilizar un modelo o un ensamble de modelos predictivos para lograr la clasificación del tipo de cérvix y que éste pueda aprender de forma autónoma con nuevos datos.

Referencias

1. Intel & MobileODT Cervical Cancer Screening [Internet]. USA [consultado el 22 de marzo de 2017]. Disponible en: <https://www.kaggle.com/c/intel-mobileodt-cervical-cancer-screening>
2. Greenspan H, Gordon S, Zimmerman, G, Lotenberg S, Jeronimo J, Antani S, et al. Automatic Detection of Anatomical Landmarks in Uterine Cervix Images. IEEE Transactions on Medical Imaging, 2009;28(3):454-468.
3. Color spaces in OpenCV (C++/Python) Learn OpenCV. [Internet]. USA [consultado el 5 de junio de 2017]. Disponible en: <http://www.learnopencv.com/color-spaces-in-opencv-cpp-python/>