The 3D EdgeRunner Pipeline: a novel shape-based analysis for neoplasms characterization

Fernando Yepes-C, Rebecca Johnson, Yi Lao, Darryl Hwang, Julie Cologner, Felix Yap, Desai Bushan, Phillip Cheng, Inderbir Gill, Vinay Duddalwar, et al.

The 3D EdgeRunner Pipeline: A Novel Shape-Based Analysis for Neoplasms Characterization

Fernando Yepes-C[b,*], Rebecca Johnson[a,*], Yii-Lao[b], Darryl Hwang[c], Julie Coloigner[b], Felix Yap[c], Desai Bushan[c], Phillip Cheng[c], Inderbir Gill[d], Vinay Duddalwar[c], Natasha Lepore[a,b]

[a]University of Southern California, 900 W 34th St, Los Angeles-CA, USA;
[b]Children's Hospital Los Angeles, 4650 Sunset Blvd, Los Angeles-CA, USA;
[c]Dept of Radiology, Keck School of Medicine - USC, 1975 Zonal Av, Los Angeles-CA, USA;
[d]Institute of Urology, USC, 1441 Eastlake Ave, Los Angeles-CA, USA;

ABSTRACT

The characterization of tumors after being imaged is currently a qualitative process performed by skilled professionals. If we can aid their diagnosis by identifying quantifiable features associated with tumor classification, we may avoid invasive procedures such as biopsies and enhance efficiency. The aim of this paper is to describe the 3D EdgeRunner Pipeline which characterizes the shape of a tumor. Shape analysis is relevant as malignant tumors tend to be more lobular and benign ones tare generally more symmetrical. The method described considers the distance from each point on the edge of the tumor to the centre of a synthetically created field of view. The method then determines coordinates where the measured distances are rapidly changing (peaks) using a second derivative found by five point differentiation. The list of coordinates considered to be peaks can then be used as statistical data to compare tumors quantitatively. We have found this process effectively captures the peaks on a selection of kidney tumors.

Keywords: Image Processing, Computer Aided Diagnostics, Tumor Shape Analysis, EdgeRunner Pipeline

1. INTRODUCTION

As high-resolution medical imaging has become ubiquitous in the clinic, there has been growing interest in the medical community in finding reliable imaging-based tumor classification methods.[1,2] Of particular interest is the characterization of tumors as either malignant or benign, without the need for an invasive procedure such as a biopsy, which has its own procedure related expense, causes time delay and may increase morbidity.[3,4]

Conventional diagnostic imaging provides qualitative visual information which can be interpreted by expert radiologists. However, there is a lot of information held in the image data set that is not conventionally analyzed and could be harnessed as quantitative data to assist their decision making. Tumor shape is an example of this opportunity. It is a known fact that malignant tumors have a disorganized growth pattern.[5] In particular, qualitative differentiation of benign and malignant lung and breast masses currently exploits the fact that benign tumors usually have compact shapes and well-defined margins, whereas malignant cancers are more irregular and lobulated.[6]

We introduce the 3D EdgeRunner Pipeline (3DERP), a strategy that analyses tumor volumes as series of 2D slices where the edges are compared with ideal circular shapes in order to capture fluctuations in the surface of the tumor. The 3DERP currently produces highlighted zones of highly irregular shapes on each 2D slice and collects frequencies of these irregularities over the entire tumor. This method not only provides an accurate mechanism to localize surface irregularities in any tumor but also yields a numerical output where histological results are well followed.

2. MATERIALS AND METHODS

Initially, two kidney tumors were chosen for analysis as proofs of concept. Both belong to a larger database 3DERP will be extended to in which segmentation was performed by an expert radiologist. **Fig. 1** illustrates the pipeline proposed and is accompanied by **Table 1** with corresponding descriptions.

* correspondant author: fyepes@chla.usc.edu

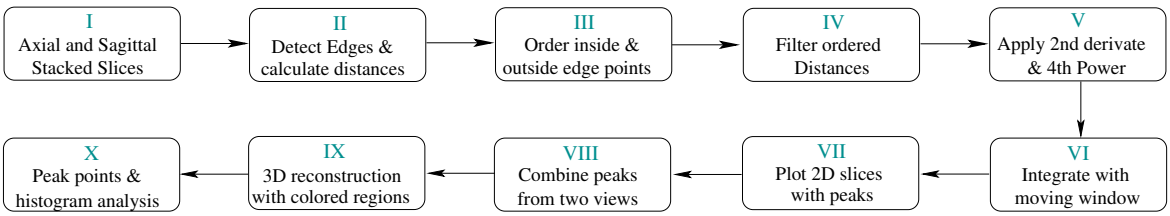


Figure 1: Sequence of steps in the 3D EdgeRunner Pipeline (3DERP)

No	Description
I	Segmented slices are stacked axially with coordinates in the x,y, plane. Data is then rotated to a sagittal view of the x,z, plane for two sets of slices.
II	Each mask's edges are detected automatically using a Canny edge detector.[7] Distance from each edge point to the center is also associated with the coordinate.
III	Points belonging to the edge (and associated distances) are ordered from the top, rotating clockwise, to standarize analysis of how distance is changing. Both an inside edge and outside edge are traced so as to not miss divergences away from the center in either direction.
IV	The highest frequencies are filtered out until the spectral power reaches 80 percent of it's original value. This frequency-cutting value is presented by,[8] as the passing-power spectrum window where the important content of the signal is held.
V	Five point differentiation is performed twice to find the second derivative. Let the list of points to be differentiated be of the form $p_n = (s_n, d_n)$ where s is a step size between each distance (d), and is always exactly $s_{(n+1)} + 1$. Differentiated array points p', will be of the form in **Eq. 1**. Each point is then raised to the 4th power.
VI	Moving window integration selects peaks in the exponential second derivative. Given points $p(s) = (s, d)$, the area (A) under the curve section with widge (N), is calculated for step (s), as is shown in **Eq. 2**. If $A(s) > T$ for a chosen value of T, s will be added to this list of locations of peaks.
VII	For each slice, plot the inside edge and outside edge tracings with peak points highlighted.
VIII	List of peak coordinates are stored for each slice and collected into one representative list per tumor.
IX	List of peaks associated with the tumor can be used to generate a coloring system to highlight perturbations from a sphere.
X	The peaks list for the tumor can be used as quantitative data for tumor characterization.

Table 1: Descriptions of the steps in the 3DERP as they are shown in **Fig. 1**

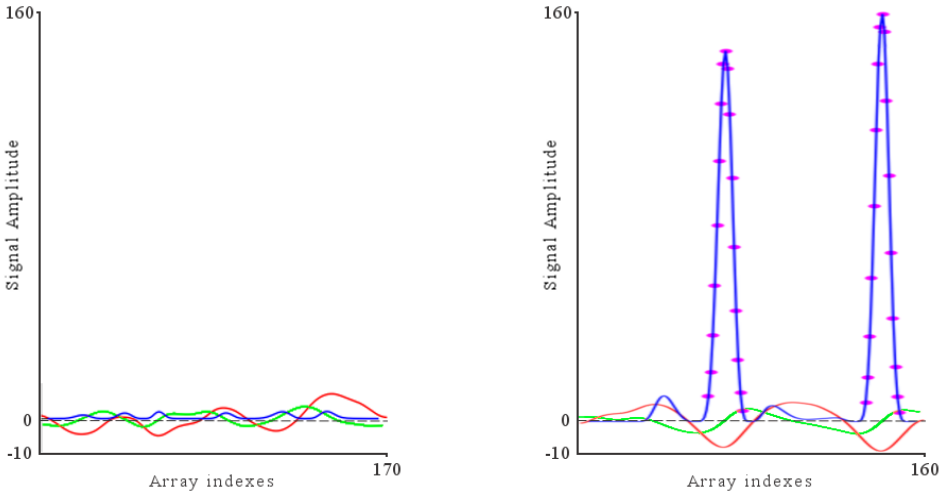
$$p'_n = [(d_n - d_{n-2}) + (d_n - d_{n-1}) + (d_{n+1} - d_n) + (d_{n+2} - d_n)]/4 \tag{1}$$

$$A(s) = \sum_{i=-\frac{N}{2}}^{\frac{N}{2}} p[(\frac{sN}{2}) - (\frac{N}{2} + i)] \tag{2}$$

Once the 3DERP was succesfully tested, a cohort of 10 tumors randomly selected from clinical database that includes the same number of samples of Clear Cell Renal Carcinoma (CCRC) and Renal Oncocytoma (RO) specimenes were studied under 3DERP methods. These tumors are histologically classified. Some results of this exercise are presented in the following section.

3. RESULTS

This section contains a sample of the results from running 3DERP over two tumors, Tum_1 and Tum_2. For each tumor, a representative axial slice was selected for demonstration of the 3DERP's middle stages. **Fig. 2** shows a plot of the distances arrays for the chosen slice, its first derivative, second derivative to the 4th power, and the points meeting the moving window threshold. In **Fig. 3**, see the 3D images and detected edge points (3DERP stage II) in addition to qualitative and quantitative results. Qualitative results include the slice's edge tracings with peak points highlighted. Quantitatively, there is a histogram showing the frequency with which peaks were found within all of the slices of the tumor.



(a) Tumor 1 (Tum_1) Signals. Slice 10

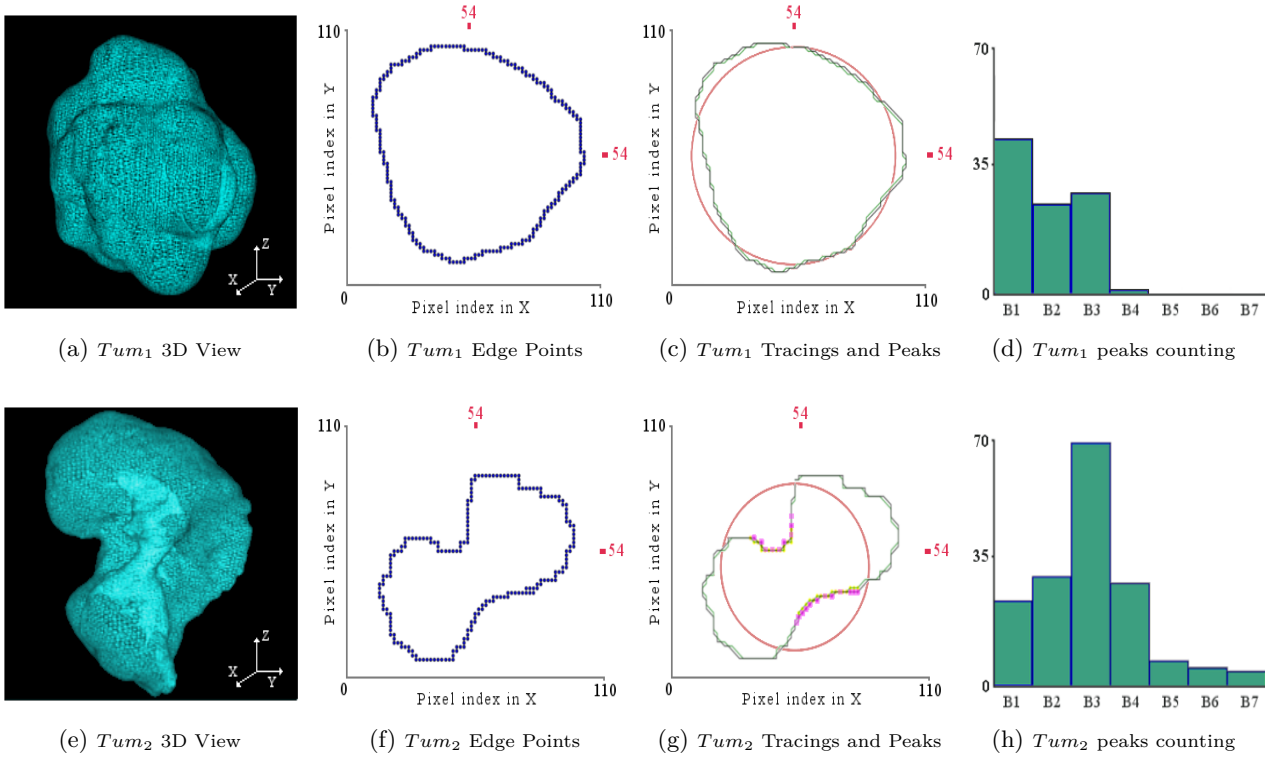
(b) Tumor 2 (Tum_2) Signals. Slice 34

Figure 2: Data generated from selected slices in Tum_1 and Tum_2. The ordered distances array is shown in red, the first derivative is shown in green, the second derivative raised to the 4th power (peak detector) is shown in blue. Additionally, any points for which the moving window threshold was met are shown in magenta. Note how the slice of Tum_1 did not have any points meeting the threshold.

Fig. 3 presents different stages of the 3DERP algorithm applied to the slices used as a proof of concept in **Fig. 2**. It also includes the final quantitative results after the peaks collection is completed in both tested tumors in the form of histograms.

In **Fig. 3**, column 1 shows results for Tum_1 (top) and Tum_2 (bottom). The column 2 is a plot of the slice's edge in each tumor (3DERP stage II), the slices are the ones used in **Fig. 2**. Column 3 shows inside edge tracings in green and the outside edge tracings in black (3DERP stage III). The red circle has a radius equal to the average of the distances found. Also in column 3 are the points selected as peaks (3DERP stage VII). Finally, column 4 contains histograms created by running the 3DERP over the entire tumor and keeping track of the number of points selected as peaks for each slice. The heights of the bars are frequencies of number of peaks per slice. The bins in this proof of concept are equally separated along the span given by the maximum amount of peaks found in the tumor.

Willing to verify the findings in the proof of concept regarding the classification capabilities of the 3DERP, the pipeline was tested in samples of common malignant (CCRC) and benign (RO) masses. Results are presented in **Fig. 4**. Here as in the proof of concept, right sided bins tend to be filled while the tumor has a more spiculated surface.



(a) Tum_1 3D View (b) Tum_1 Edge Points (c) Tum_1 Tracings and Peaks (d) Tum_1 peaks counting

(e) Tum_2 3D View (f) Tum_2 Edge Points (g) Tum_2 Tracings and Peaks (h) Tum_2 peaks counting

Figure 3: Step by step results of the 3DERP pipeline in two randomly selected real tumors

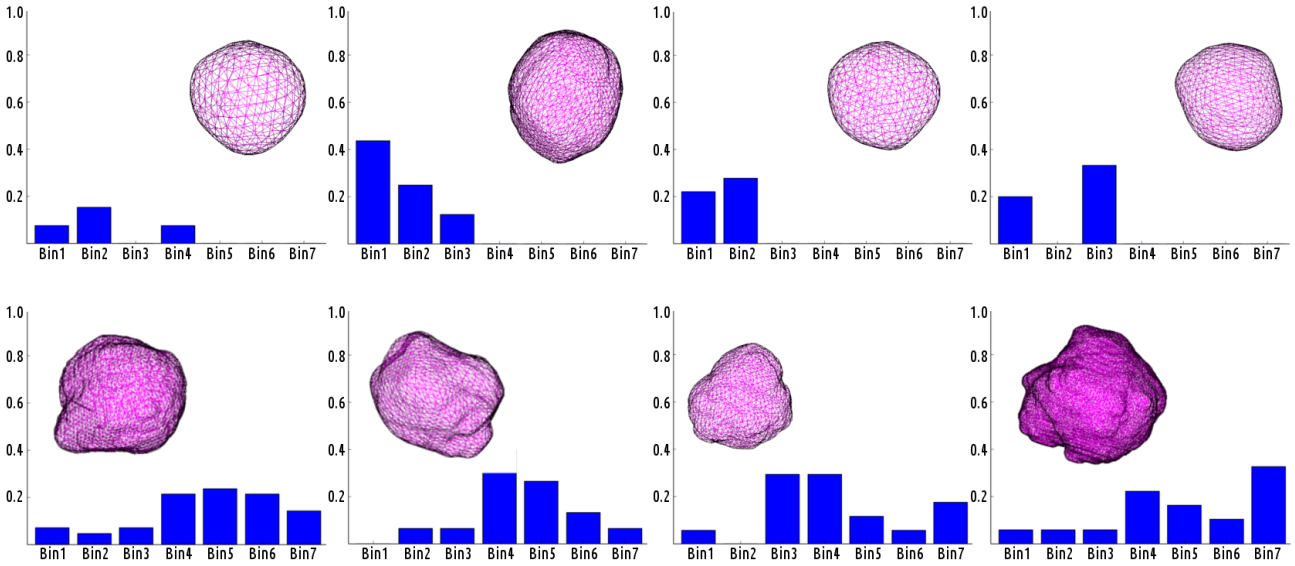


Figure 4: Cohort of tumors analyzed with the 3DERP. Top row shows results on benign masses (RO) while bottom row shows samples of CCRC, the most common type of renal malignant neomass. The y axes of the histograms have been normalized to enable fair comparisons.

4. DISCUSSION

In **Fig. 2** of the results section, it is visually discernible that the more spherical tumor (Tum_1) has a more circular slice chosen and no peaks were found, whereas the lumpier tumor (Tum_2) has two main inward dents from a circle form which were selected as regions with peaks. When this procedure is reproduced through out all the slices in a tumor, more peaks are recruited while the edges keep differing from a perfect circle. By applying this pipeline in different tumors one can create a qualitatively differentiation between highly spiculated masses, the ones mostly spherical and any other in between. This hypothesis was validated with the histologically classified samples shown in **Fig. 4**, where the 3DERP accurately populated the right side bins of CCRC's histograms while those of the RO mases are always empty. The two pass strategy used when running the algorithm in axial and coronal views allows the refinement of peaks selection by discarding those points that either do not overlap among the surface of the analyzed mass or appear as isolated spots after merging the results of the two views.

5. CONCLUSIONS

The 3DERP was created with the aim of helping to differentiate a malignant tumor from a benign tumor without the need for an invasive biopsy and all its associated risks. This method shows potential for implementation in the clinics and is being validated with a bigger testing sample. Applying the 3DERP methods has also been proposed for analyzing tumors that grown in other parts of the body.

REFERENCES

[1] Catalano, O. A., Samir, A. E., Sahani, V., and Hahn, P. F., "Pixel Distribution Analysis: Can It be Used to Distinguish Clear Cell Carcinomas from Angiomyolipomas with Minimal Fat?," *Genitourinary Imaging* **247**(3), 11 – 38 (2007).

[2] Arakeri, M. P. and Reddy, R. M., "A Novel CBIR Approach to Differential Diagnosis of Liver Tumor on Computed Tomography Images," *Procedia Engineering* **38**, 528 – 536 (2012).

[3] Bedossa, P., "Liver biopsy," *Grastroenterol Clin Bio* **32**, 4 – 7 (2008).

[4] Stigliano, R., Marelli, L., Yu, D., Davies, N., Patch, D., and Burroughs, A. K., "Seeding following percutaneous diagnostic and therapeutic approaches for hepatocellular carcinoma. what is the risk and the outcome? seeding risk for percutaneous approach of hcc.," *Cancer Treat Rev* **33**(5), 437–47 (2007).

[5] Ambrosi, D. and Mollica, F., "On the mechanics of a growing tumor," *International Journal of Engineering Science* **40**, 1297 – 1316 (2002).

[6] Erasmus, J. J., Connolly, J. E., McAdams, H. P., and Roggli, V. L., "Solitary pulmonary nodules: Part i. morphologic evaluation for differentiation of benign and malignant lesions," *Radiographics* **20**(1), 43–58 (2000).

[7] Canny, J., "A computational approach to edge detection," *Pattern Analysis and Machine Intelligence, IEEE Transactions on* **PAMI-8** (1986).

[8] SHANNON, C. E., "A mathematical theory of communication," *The Bell System Technical Journal* **27**, 379–423,623–656 (1948).