Supplementary material for Leijenaar R. T. H. et al. External validation of a prognostic CT-based radiomic signature in oropharyngeal squamous cell carcinoma, Acta Oncol 2015;54:1423–1429.

Supplementary Description of Radiomic Features

The radiomic signature described by Aerts et al. [1] consists of the following four features, derived from the GTV: (1) “First order statistics: Energy”, (2) “Shape: Compactness”, (3) “Gray level run length: Gray level non-uniformity”, and (4) Wavelet (HLH) “Gray level run length: Gray level non-uniformity”. In the below text we provide a detailed description of the aforementioned features.

1. First order statistics: Energy

First-order statistics describe the distribution of voxel intensities within the CT image. Let X denote the three dimensional image matrix with N voxels. Energy is then defined as:

$$energy = \sum_{i}^{N} X(i)^2$$

2. Shape: Compactness

Shape features quantify the three-dimensional size and shape of the delineated tumor region. Let V denote the volume and A the surface area of the volume of interest. Compactness is then defined as:

$$compactness = \frac{V}{\sqrt{\pi} A^{\frac{2}{3}}}$$

3. Gray level run length: Gray level non-uniformity

Run length metrics quantify gray level runs in an image [2]. A gray level run is defined as the length in number of pixels, of consecutive pixels that have the same gray level value. In a gray level run length (GLRL) matrix $p(i,j|\theta)$, the (i,j)th element describes the number of times j a gray level i appears consecutively in the direction specified by θ, and N_g is the number of discrete gray level intensities. As a two dimensional example, consider the following 5x5 image, with 5 discrete gray levels:

$$I = \begin{matrix} 5 & 2 & 5 & 4 & 4 \\ 3 & 3 & 3 & 1 & 3 \\ 2 & 1 & 1 & 1 & 3 \\ 4 & 2 & 2 & 2 & 3 \\ 3 & 5 & 3 & 3 & 2 \end{matrix}$$

The GLRL matrix for $\theta = 0$, where 0 degrees is the horizontal direction, then becomes:

$$p(0) = \begin{matrix} 1 & 0 & 1 & 0 & 0 \\ 3 & 0 & 1 & 0 & 0 \\ 4 & 1 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 3 & 0 & 0 & 0 & 0 \end{matrix}$$

In this study, a GLRL matrix was computed for each of the 13 directions in three dimensions, and feature values were averaged over all directions. Let $p(i,j|\theta)$ be the (i,j)th entry in the given run-length matrix p for a direction θ, N_g the number of discrete intensity values in the image, N_r the number of different run lengths, and N_p the number of voxels in the image. Gray level non-uniformity (GLN) is then defined as:

$$GLN = \frac{\sum_{i=1}^{N_g}\left[\sum_{j=1}^{N_r} p(i,j|\theta)\right]^2}{\sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j|\theta)}$$

4. Wavelet (HLH) Gray level run length: Gray level non-uniformity

Wavelet transform effectively decouples textural information by decomposing the original image, in a similar manner as Fourier analysis, in low –and high-frequencies. In this study a discrete, one-level and undecimated three dimensional wavelet transform was applied to each CT image, which decomposes the original image X into 8 decompositions. Consider L and H to be a low-pass (i.e. a scaling) and, respectively, a high-pass (i.e. a wavelet) function, and the wavelet decompositions of X to be labeled as X_{LLL}, $X_{LLH}, X_{LHL}, X_{LHH}, X_{HLL}, X_{HLH}$, X_{HHL} and X_{HHH}. For example, X_{LLH} is then interpreted as the high-pass sub band, resulting from directional filtering of X with a low-pass filter along x-direction, a low pas filter along y-direction and a high-pass filter along z-direction and is constructed as:

$$X_{LLH}(i,j,k) = \sum_{p=1}^{N_L}\sum_{q=1}^{N_L}\sum_{r=1}^{N_H} L(p)L(q)H(r)X(i+p,j+q,k+r)$$

Where N_L is the length of filter L and N_H is the length of filter H. The other decompositions are constructed in a similar manner, applying their respective ordering of low or high-pass filtering in x, y and z-direction. Wavelet decomposition of the image X is schematically depicted in **Figure S1**. Since the applied wavelet decomposition is undecimated, the size of each decomposition is equal to the original image and each decomposition is shift invariant. Because of these properties, the original region of interest delineation can be applied directly to the decompositions after wavelet transform. In this study, the “Coiflet 1” wavelet was applied on the original CT images. For the HLH wavelet decomposition we calculated the gray level run length feature “Gray level non-uniformity”, as described above.

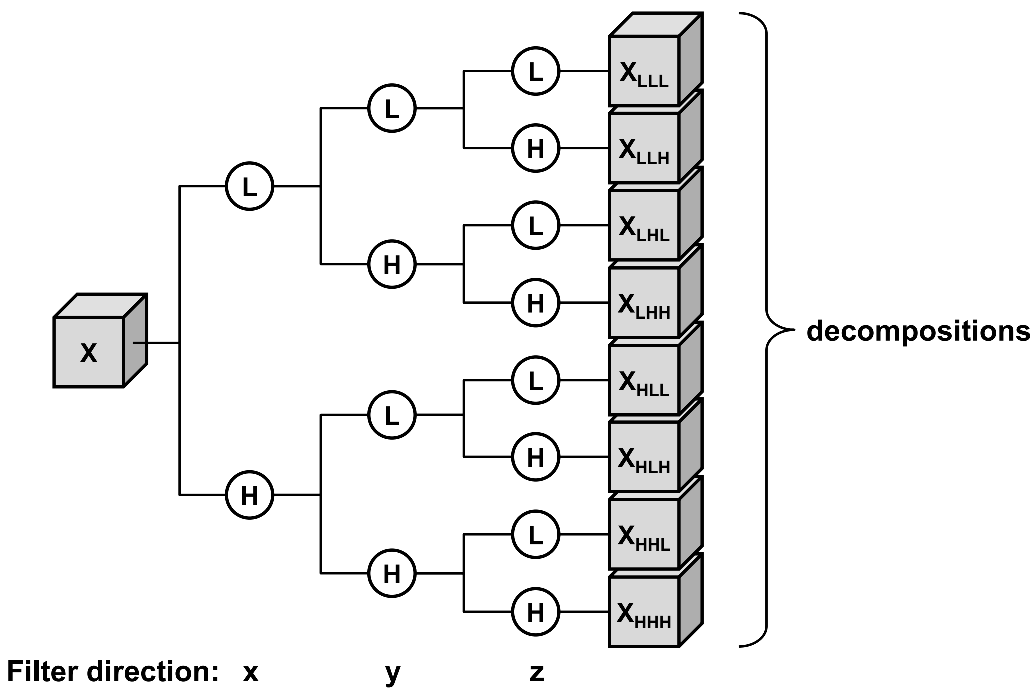

Figure S1: Schematic of the undecimated three dimensional wavelet transform applied to each CT image. The original image X is decomposed into 8 decompositions, by directional low-pass (i.e. a scaling) and high-pass (i.e. a wavelet) filtering: X_{LLL}, X_{LLH}, X_{LHL}, X_{LHH}, X_{HLL}, X_{HLH}, X_{HHL} and X_{HHH}.

Supplementary References

[1] Aerts HJWL, Velazquez ER, Leijenaar RTH, Parmar C, Grossmann P, Cavalho S, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nat Commun 2014;5.

[2] Galloway M. Texture analysis using gray level run lengths. Comput Vision Graph 1975;4: 172-179.