

Texture Analysis of Non-Small Cell Lung Cancer on Unenhanced CT and Blood Flow Maps: a Potential Prognostic Tool

Serena Baiocco*✉, Domenico Barone†, Giampaolo Gavelli†, Alessandro Bevilacqua*

*University of Bologna, Bologna, Italy

{s.baiocco, alessandro.bevilacqua}@unibo.it

†IRCCS - Istituto Scientifico Romagnolo per lo Studio e la Cura dei Tumori (IRST), Meldola (FC), Italy

{d.barone, g.gavelli}@irst.emr.it

Abstract—The presence of tumour heterogeneity makes the clinical oncological practice very challenging, since introduces a great variability in tumours’ response to available therapies. For this reason, in the last decade, quantifying the salient features of the intra-tumoural heterogeneity has gained a great attention, also leading to a re-emerging of the texture analysis. Tumour heterogeneity represents the complex biology of tumour micro-environment, characterised by both spatial and temporal variability, increased by the presence of chaotic blood vessels within tumour tissue. Computed Tomography (CT) has always been considered one of the reference technologies for morphological analysis of organs and tissues, permitting to capture the “in vivo” spatial tumour heterogeneity. The need to also detect hemodynamic tumour features has stimulated the use of CT perfusion (CTp), a promising functional imaging technique in the oncological field. CTp allows detecting the presence of tumour abnormal hemodynamic patterns, by analysing the tissue temporal variations occurring after an intravenous administration of contrast medium. This work presents the extraction of meaningful statistical and texture features from both baseline CT images and perfusion maps of lung tumours, which could work as prognostic image-based biomarkers.

I. INTRODUCTION

Non-small-cell lung cancer (NSCLC) is the leading cause of cancer mortality worldwide [1]. Despite the introduction of novel advanced therapies, NSCLC remains associated with poor prognosis and overall survival (OS) [2]. Tumour staging is probably the most important prognostic parameter for survival, though different outcomes within the same stage grouping suggest that also other factors influence survival [3]. One of the challenges in this age of personalized medicine remains the identification of more effective markers of prognosis for a better stratification of NSCLC patients [4], with strong implications in patient management, particularly as regards treatment choices [5].

NSCLC tissues are highly heterogeneous at different scale, reflecting the presence of angiogenesis, hypoxia, high cell density, necrosis and hemorrhage [6], [7]. Angiogenesis is a process that consists in the chaotic sprouting of blood vessels from pre-existing ones [8], [9]. With the enlarging of the vascularization, the tumour tissue is able to receive nutrients, also creating paths for cells to leave or enter the blood circulation. This process leads to the tumour growth associated with an increased interstitial pressure, which may lead to local reductions in blood flow (BF) and, consequently,

to an inadequate oxygen delivery for cells viability. This condition is known as hypoxia and has a decisive role in the tumour response to treatment [10].

By its very nature, quantitative tumour evaluation through molecular characterisation, involving biopsies or invasive surgeries to get small tissue specimens, cannot adequately represent the features of the tumour as a whole [11]. On the contrary, imaging is able to provide voxel-based information for the whole tumour [12]. It has the great potential to capture the *in vivo* intra-tumoural heterogeneity non-invasively [13], paving the way for a routine tumour monitoring and therapy response evaluation.

Generally, the way images are clinically interpreted is mostly visually [14]. The aim of the clinician is to detect potential abnormalities, recognising patterns and linking the perceived patterns with possible diagnoses. Therefore, the success of a correct image interpretation strongly depends on the skills of the clinician [15]. Consequently, a qualitative heterogeneity evaluation will be prone to subjective variability. As a result, despite tissue heterogeneity has been proved to be strongly related to tumour aggressiveness [16], prognosis [17] and survival [18], its quantitative evaluation is still not included in the clinical practice.

In the last decade, several studies have focused on the analysis of tumour heterogeneity to identify more accurate prognostic biomarkers through texture analysis [14], [19], [20]. Texture analysis includes several different techniques, which are not new in the medical field but date back to the early 70s, when texture measures were suggested for the detection of lung disease patterns [21]. The major advantage in clinical practice is that these mathematical approaches can offer information not visible by the human eye in an objective way, exploiting data routinely acquired [22].

Recently, CTp has been accepted as a clinical technique in the oncologic field, primarily for the early evaluation of tumour response to anti-angiogenic therapies (i.e. therapies aiming at reducing tumour vasculature) [23]. Indeed, CTp allows capturing tumours vascular patterns through the analysis of the time-concentration curves (TCCs), representing the tissue density temporal variations [24], directly proportional to the quantity of contrast medium within the tissue. One of the most significant perfusion parameters is the BF, which permits to detect the earliest functional changes on tumour vasculature

even within the first week of anti-angiogenic therapy [25]. Moreover, this parameter has proved a high correlation with the tissue biomarker micro-vessel density (MVD) [26], [27].

The analysis of BF maps could allow detecting abnormal vascular patterns within tumour tissue. However, several factors may affect the reliability of BF maps. Among these, cardiac and respiratory motion artefacts, physics-based artefacts as well as partial volume artefacts may alter the calculation of perfusion parameters. Therefore, a proper denoising is needed to properly characterise the hemodynamic heterogeneity and, finally, to derive quantitative features [28].

In this work, we analyse the intra-tumoral heterogeneity of NSCLC, at diagnosis stage, emerging from unenhanced CT images and BF maps. To this purpose, BF maps were computed according to the Maximum-Slope method. A proper reliability analysis of the BF maps was performed to automatically detect and remove those pixels and regions undergoing high computing errors due to artefacts. The correlations between the OS parameter and statistical feature-pairs computed on CT images and BF maps were explored to assess whether these features could work as prognostic image-based biomarkers for NSCLC. The discrimination ability of the feature-pair was analysed in comparison with the staging, since is commonly associated with OS. The features derived from functional BF maps proved a strong correlation with OS, resulting more effective than those directly derived from CT images, this suggesting a promising prognostic clinical application of CTP. Actually, also the features extracted from non-contrast enhanced CT images proved to be quite promising, being still more effective than the stage parameter but also easily derivable from data routinely acquired in clinical practice.

II. MATERIALS AND METHODS

A. Study population

As reported in a previous work [29], this study was approved by the Institutional Review Board that waived informed consent for the retrospective data analysis of the patients. 36 consecutive patients (25 men, 11 women; age range 36-81 years) with primary NSCLC, subdivided in 28 adenocarcinoma (AC), 6 squamous-cell carcinoma (SCC) and 2 large-cell carcinoma (LCC), were enrolled for data analysis between September 2010 and December 2012. According to the TNM classification (seventh edition) of malignant tumours 2 patients were diagnosed stage IB, 3 patients IIIA, 4 patients IIIB and 27 patients IV tumour stage. Patients inclusion criteria were: (a) over eighteen years old; (b) with histologically verified NSCLC; (c) no prior history of chemotherapy, surgery or thoracic radiation therapy.

Exclusion criteria were: (a) clinically significant cardiovascular disease; (b) pregnancy or lactation; (c) a known history of deep vein thrombus or pulmonary embolus; (d) patients lost to follow-up for unknown reasons ($n = 8$); (e) patients alive at the time of the study ($n = 4$); (f) not having the longest axial diameter of the lesion greater than 10 mm in at least one slice ($n = 3$); and (g) examination hardly affected by physics-based artefact ($n = 2$).

Finally, 19 patients were considered for the analysis. TNM staging was considered as the reference to predict survival.

This variable was dichotomised so as to divide in early stage patients (stage I – IIIA) likely receiving curative surgery or curative chemoradiotherapy and advanced stage patients (stage IIIB – IV) likely undergoing non-curative chemotherapy, radiotherapy, or molecular therapies [29].

Table I reports the previous information, also including patients histological subtype (ACC, SCC, LCC) and OS data, which is defined as the time interval in months between the date of the baseline CT examination and the date of death.

TABLE I. SUMMARY OF THE HISTOLOGICAL DIAGNOSIS, TUMOUR STAGE AND OS DATA (MONTHS) RELATIVE TO EACH PATIENT.

Patient ID	OS	Diagnosis	Stage
ID1	6	LCC	IIIB
ID2	4	AC	IV
ID3	14	SCC	IB
ID4	13	AC	IV
ID5	5	AC	IV
ID6	5	AC	IB
ID7	6	AC	IV
ID8	10	AC	IV
ID9	17	SCC	IV
ID10	52	SCC	IV
ID11	6	AC	IV
ID12	6	AC	IV
ID13	4	AC	IV
ID14	8	AC	IV
ID15	12	AC	IV
ID16	11	SCC	IIIB
ID17	12	AC	IV
ID18	0	AC	IV
ID19	7	AC	IV

B. CTP protocol

At diagnosis, patients underwent axial CTP performed on a 256-slice CT system (Brilliance iCT, Philips Medical System, Best, The Netherlands) and laid in the supine (feet first) position. An initial unenhanced low-dose full-body axial CT scan was performed to identify the target lesion at base-line condition. Then, an intravenous 50 mL bolus of contrast agent (Iomeron, Bracco, Milan, Italy) was injected at 5 mL/s for axial cine contrast enhanced CT. CTP scan of 25-second duration was performed for each patient under breath-hold condition. This protocol yielded 20 scans, each corresponding to different sampling instants, with 55 mm of z-coverage (11 slices \times 5-mm slice thickness, 0.4-sec rotation time, at 80 kV, 250 mA). Image data are reconstructed to 220 cine images (512 \times 512 pixel, 11 slices, 350 mm \times 350-mm, 5-mm slice spacing, 1.25-sec temporal resolution).

C. Perfusion maps generation

The lesions and the arterial input were selected in consensus by two radiologists (D.B. and G.G., > 25-year experience each). For each lesion, the radiologists manually drew on a reference slice the related region of interest (ROI). Then, a 3D rigid alignment was performed according to the following procedure. The lesion ROI was translated on the slices of the reference sequence, so as the contour of the lesion visually

match. When opportune, adjacent slice levels were chosen by the radiologists to achieve the best match with the reference slice [30] and, consequently, the “best” sequence.

BF values were computed by considering the first-pass phase [31] and modelling each voxel as a single-compartment including both the intravascular and the extra-vascular spaces. This permitted to perform short-time examinations, thus reducing motion artefact effects, since the patients had the possibility to stay in apnoea. The model adopted is characterised by a single inlet, the arterial input, with an arterial blood plasma temporal concentration $C_A(t)$, and a single outlet, the venous outlet, with a venous blood plasma temporal concentration $C_V(t)$. $C_T(t)$ represents the TCC of the tissue. The relation between $C_A(t)$, $C_V(t)$ and $C_T(t)$ can be expressed through the Fick’s principle representing the conservation of mass:

$$\frac{dC_T(t)}{dt} = \frac{BF}{V_T} [C_A(t) - C_V(t)] \quad (1)$$

The assumption of no venous outflow during the measurement period $C_V(t) \simeq 0$ (i.e., no contrast agent has reached the venous side of the circulation yet) leads to the following approximation:

$$\frac{dC_T(t)}{dt} = \frac{BF}{V_T} [C_A(t) - C_V(t)] \simeq \frac{BF}{V_T} C_A(t) \quad (2)$$

This implies that the slope of the tissue curve, $dC_T(t)/dt$, reaches a maximum (steepest slope) when the TCC of the arterial input, $C_A(t)$, reached its peak density [32]. Thus, the BF, expressed in mL/min/100 g, is given by

$$\frac{BF}{V_T} \simeq \frac{[\frac{dC_T(t)}{dt}]_{max}}{[C_A(t)]_{max}} \quad (3)$$

This is the single-compartment formulation also known as Mullani-Gould model. The simplification adopted might not hold for organs characterised by a complex microcirculation, such as spleen and kidneys [33] but is adequate for most clinical applications [34].

The BF values of the reference slice are represented using functional colorimetric maps, as shown in Fig. 1. Blue regions point out low perfusion values, while regions characterised by hottest colours highlight the presence of highest perfusion areas. These maps provide a useful view of the perfusion and the functional heterogeneity of the tumour, highlighting for instance the presence of hypodense regions, hemorrhages, blood vessels and bronchi.

In order to obtain the TCC signal for each voxel we adopted a sigmoid-shape model, as reported in Fig. 2. Sigmoid is modelled through the *Hill Equation* in Eq. (4), able to robustly fit the main trend of the concentration samples [35]:

$$\bar{Y}_{TCC}(t) = E_0 + (E_{max} - E_0) \frac{t^\alpha}{(EC_{50} + t)^\alpha} \quad (4)$$

where E_0 is the baseline concentration of the tissue, that is before the arrival of the tracer, E_{max} is the saturation value of the

concentration reached after its arrival, EC_{50} is the time instant at the half-maximum value of the concentration, and α is the parameter which mainly affects the slope of the sigmoid curve. The curve fitting is achieved using an in-house fitting algorithm based on the nonlinear, least squares, Levenberg-Marquardt minimization algorithm (lsqcurvefit, Matlab[®]; MathWorks, Natick, MA, USA).

D. Perfusion maps denoising

Denoising is a necessary step to be performed before the features extraction. Essentially, it consists in detecting and excluding unreliable BF values. BF values strictly lower than 1 mL/min/100 g were automatically excluded, being considered improbable as physiological values and more likely attributable to numerical errors, as the perfusion computation algorithm forces BF values to be positive.

For each TCC signal, the quality of the fitting has been assessed by considering the residuals (ϵ), computed as the difference between the observed samples Y_{TCC} and the relative values obtained through the fitting \bar{Y}_{TCC} for each time instant t . ϵ is defined as in Eq. (5):

$$\epsilon(t) = |Y_{TCC}(t) - \bar{Y}_{TCC}(t)| \quad (5)$$

As a goodness-of-fit index for the single TCC, the temporal mean value μ_ϵ of the residual ϵ relative to the considered pixel has been computed. The distribution of this index has been used to detect the BF values characterised by high fitting errors in order to exclude them from the analysis, since they are computed on TCCs not correctly fitted.

High errors are usually associated to the presence of noise, artefacts and anatomical structures, such as bony regions, blood vessels and bronchi. Through the use of this index, these structures can be automatically excluded. As shown in Fig. 1, at the end of this automatic denoising process, the removed BF values appear highlighted in the colorimetric maps with the pink colour. As one can see, the denoising method removes also unreliable regions as the borders, which are affected by partial volume effect, still keeping a wide range of BF values [36].

E. Texture analysis

Feature generation and selection were performed using an in-house software based on Matlab[®] (MathWorks, Natick, MA, USA). First- (histogram) and second-order (grey-level co-occurrence matrix) statistical features were computed for the reference slices on unenhanced CT images and BF maps.

First-order statistics are based on the histogram of pixels intensity values in the image. The features derived from values distribution depend only on single pixel values and not on the interaction or co-occurrence of neighbouring pixel values. They are global features, which describe how intensity values within a ROI are distributed. Let $x = \{x_1, x_2, \dots, x_N\}$ be the set of L values of the N voxels in the image. Let $h = \{h_1, h_2, \dots, h_L\}$ be the histogram with count n_i of each intensity level i in x . The occurrence probability p_i for each intensity level is $p_i = \frac{h_i}{N}$. Accordingly, we can have the following definitions:

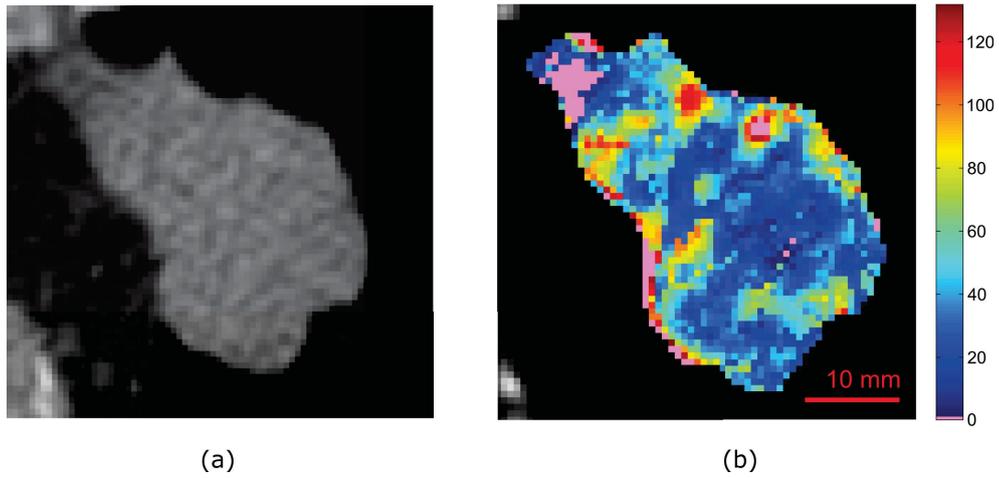


Fig. 1. Unenhanced CT image (a) and corresponding BF map (b) of a lung tumour (squamous-cell carcinoma)

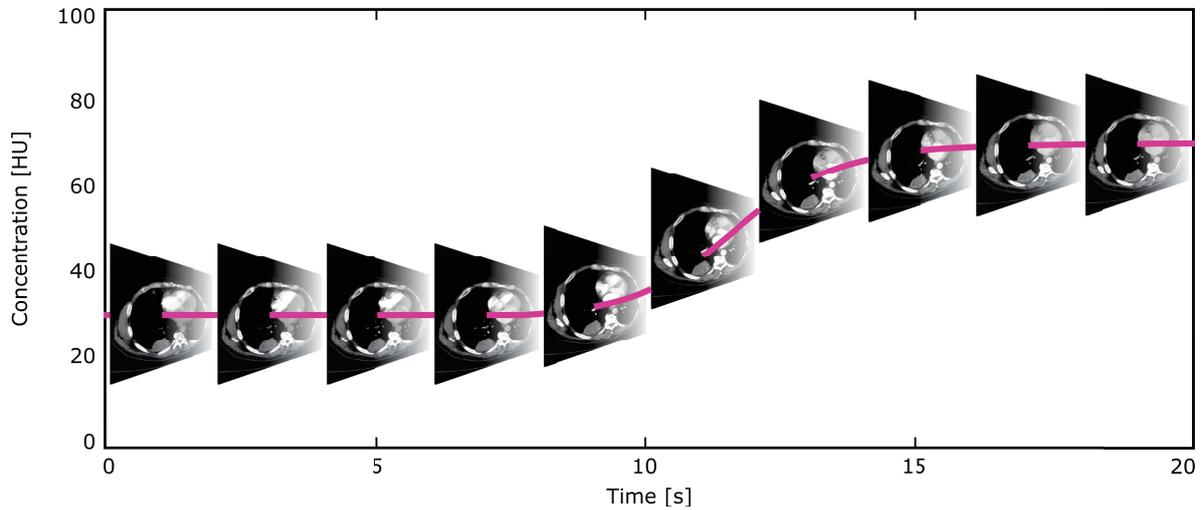


Fig. 2. Representation of the sigmoidal model used for the fitting procedure of Tumour voxel values, which belong to the temporal sequence of the reference slice

Mean μ

$$\mu = \frac{1}{N} \sum_{k=1}^N x_k \quad (6)$$

Standard deviation σ

$$\sigma = \left(\frac{1}{N} \sum_{k=1}^N (x_k - \mu)^2 \right)^{\frac{1}{2}} \quad (7)$$

Skewness S

$$S = \frac{\frac{1}{N} \sum_{k=1}^N (x_k - \mu)^3}{\left[\frac{1}{N} \sum_{k=1}^N (x_k - \mu)^2 \right]^{\frac{3}{2}}} \quad (8)$$

Kurtosis K

$$K = \frac{\frac{1}{N} \sum_{k=1}^N (x_k - \mu)^4}{\left[\frac{1}{N} \sum_{k=1}^N (x_k - \mu)^2 \right]^2} - 3 \quad (9)$$

Coefficient of variation CV

$$CV = \frac{\sigma}{\mu} \quad (10)$$

Entropy E also known as Shannon entropy

$$E = - \sum_{i=1}^L p_i \log_2 p_i \quad (11)$$

Uniformity U

$$U = \sum_{i=1}^L p_i^2 \quad (12)$$

Other parameters derived from the histogram analysis include: the *median*, the intensity level that divides a distribution in two halves, the *minimum* and the *maximum*, that is the lowest and the highest pixel values.

Second-order statistics are based on the likelihood of observing pixels pairs of specific intensity values. These features

belong to the textural features, since they can represent the spatial distribution of image values. More specifically, these features are based on the joint probabilities and provide co-occurrence measurements performed on matrices, known as grey level co-occurrence matrices (GLCMs), that reflect the spatial grey-level dependence. Actually, these matrices may contain any type of entity, besides grey levels. Each cell of these matrices represents how often a couple of pixel values (i, j) occurs, this being the basis for measuring the relationship between pixels. Therefore, these matrices provide valuable information about images since, as reported by *Haralick et al.*, tone and texture, which are always present in the images, bear an inextricable relationship [37].

The features that will be extracted from GLCM provide image statistical information regarding the distribution of pixels pairs along one of the image directions. GLCM provides a new image representation [38] and is an estimate of the second-order joint probability, p_{ij} , of the intensity values of two pixels i and j , separated by a distance δ ($\delta = 1$ for neighbouring pixels) along a given direction θ , (where usually θ can be horizontal, vertical, diagonal and anti-diagonal, i.e. $\theta = \{0^\circ, 45^\circ, 90^\circ, \text{ and } 135^\circ\}$). This joint probability can be represented with a square matrix p_{ij} , with row and column dimensions equal to the number of discrete intensity levels L in the image. Each p_{ij} element contains the frequency of a combination of intensity levels pairs, i and j , at a certain distance δ , along a specific direction θ . The resulting GLCM is a symmetric matrix, as shown in Fig. 3. For images tending to be flat, i.e.

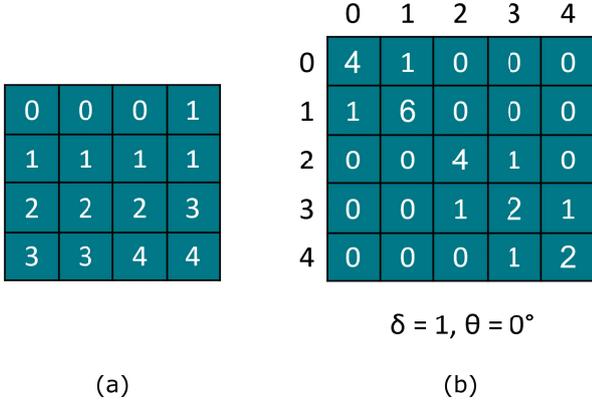


Fig. 3. Small image (a) and relative co-occurrence matrix (b) for $\delta = 1$ and $\theta = 0^\circ$. The joint probability is obtained dividing the matrix in (b) for the total number of possible pixel pairs.

uniform with no texture, the resulting GLCM would be almost diagonal. Contrarily, for images with increased local intensity variations, i.e. with increased texture, the resulting GLCM will result characterised by more off-diagonal values.

The features that will be derived from GLCM are local features that should encode in an efficient way, useful for classification tasks, the relevant information present in GLCM structures. Some of GLCM features reported below have a clear correspondence with human texture perception (e.g., coarseness, smoothness), while some others do not represent some specific visual properties, but they still encode texture information with high discriminatory power [39].

Joint maximum M_J is the probability corresponding to the

most common GLCM co-occurrence and is defined as

$$M_J = \max(p_{ij}) \quad (13)$$

Joint average μ_J is the weighted sum of the joint probabilities and is defined as

$$\mu_J = \sum_{i=1}^L \sum_{j=1}^L i p_{ij} \quad (14)$$

Joint variance σ_J^2 is a measure of the variability of the GLCM co-occurrences and is defined as

$$\sigma_J^2 = \sum_{i=1}^L \sum_{j=1}^L (i - \mu)^2 p_{ij} \quad (15)$$

where $\mu = \text{joint average}$.

Second order entropy E_J is defined as

$$E_J = - \sum_{i=1}^L \sum_{j=1}^L p_{ij} \log_2 p_{ij} \quad (16)$$

Angular second moment *ASM* [37], also called energy and uniformity, is defined as

$$ASM = \sum_{i=1}^L \sum_{j=1}^L p_{ij}^2 \quad (17)$$

Contrast C , a measure of the intensity levels variations, is defined as

$$C = \sum_{i=1}^L \sum_{j=1}^L (i - j)^2 p_{ij} \quad (18)$$

Inverse difference *IDF* is defined as

$$IDF = \sum_{i=1}^L \sum_{j=1}^L \frac{p_{ij}}{1 + |i - j|} \quad (19)$$

Inverse difference normalised *IDFN* is defined as

$$IDFN = \sum_{i=1}^L \sum_{j=1}^L \frac{p_{ij}}{1 + |i - j|/L} \quad (20)$$

Inverse difference moment *IDM*, a measure of the local homogeneity, is defined as

$$IDM = \sum_{i=1}^L \sum_{j=1}^L \frac{p_{ij}}{1 + (i - j)^2} \quad (21)$$

Homogeneity normalised *HN* is defined as

$$HN = \sum_{i=1}^L \sum_{j=1}^L \frac{p_{ij}}{1 + (i-j)^2/L^2} \quad (22)$$

Correlation $CORR$ [37] is defined as

$$CORR = \frac{\sum_{i=1}^L \sum_{j=1}^L i j p_{ij} - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (23)$$

where

$$\mu_x = \sum_{i=1}^L i \sum_{j=1}^L p_{ij} \quad (24)$$

$$\mu_y = \sum_{j=1}^L j \sum_{i=1}^L p_{ij} \quad (25)$$

$$\sigma_x = \sum_{i=1}^L (i - \mu_x)^2 \sum_{j=1}^L p_{ij} \quad (26)$$

$$\sigma_y = \sum_{j=1}^L (j - \mu_y)^2 \sum_{i=1}^L p_{ij} \quad (27)$$

Autocorrelation A is defined as

$$A = \sum_{i=1}^L \sum_{j=1}^L i j p_{ij} \quad (28)$$

F. Statistical Analysis

The principal purpose of this work was to determine whether statistical and textural features directly extracted from non-contrast enhanced CT images and denoised BF maps could work as surrogate prognostic biomarkers. Given the sample size, to avoid overfitting, no more than two features were analysed jointly. Feature selection was performed through a k -means ($k = 2$) clustering algorithm, using the Squared Euclidean distance as similarity measure. Then, to identify the most promising prognostic feature-pairs, the correlation with the OS parameters was analysed automatically. Mean survival for the entire cohort of patients ($\mu_{OS} = 10.5$) was used as cutoff to separate patients with highest ($OS \geq \mu_{OS}$) and lowest ($OS < \mu_{OS}$) survival expectation. In order to determine to what extent the feature means differ between the groups identified by the clustering algorithm, a multivariate analysis of variance (MANOVA) was performed (p -value < 0.001). Sensitivity (SE), specificity (SP), positive predictive value (PPV), negative predictive value (NPV) and accuracy (ACC) were calculated to quantify the discrimination ability of both features selected and staging, i.e. the variable used as reference for survival prediction. In particular, a low survival expectation and a high survival expectation were defined as condition true and condition false, respectively. Advanced stage condition is considered the positive prediction while the early stage is considered the negative one. Consequently, patients with advanced stage and low survival expectation belong to the true positive (TP), patients with early stage and high survival expectation are the true negative (TN). On the other hand,

patients with advanced stage and a high survival expectation are considered the false positive (FP), while patients with an early stage and a low survival expectation as the false negative (FN).

Statistical analysis was performed using Matlab[®] (Math-Works, Natick, MA, USA).

III. RESULTS

Feature-pairs derived from BF maps resulted characterised by the greatest discriminatory power in distinguish patients with different survival expectations. More specifically, first-order entropy (E) and inverse difference moment (IDM) derived from BF maps resulted highly effective in separating patients with different survival expectations (SE=86%, SP=75%, PPV=68%, NPV=90%, ACC=79%). This feature-pair resulted also the best pair correlating with the OS parameter among those derived from unenhanced CT images (SE=67%, SP=71%, PPV=80%, NPV=56%, ACC=68%). In addition, among the feature-pairs derived from BF maps, comparable performance can be obtained also with first-order entropy and second-order uniformity (ASM) (SE=86%, SP=75%, PPV=68%, NPV=90%, ACC=79%).

E and IDM as well as E and ASM computed on BF maps are able to group patients with the lowest ($OS < \mu_{OS}$) and highest ($OS \geq \mu_{OS}$) survival expectation. Fig. 4 graphically shows the distribution of the features values extracted from BF maps, for low and high survival expectation. As one can see, high E and low IDM or low ASM enclose the patients with the worst prognosis ($OS < \mu_{OS}$), while low E and high IDM or high ASM characterise the patients with the best prognosis in this study. Also for E and IDM derived from unenhanced CT images, relative box plots for low ($OS < \mu_{OS}$) and high survival expectation ($OS \geq \mu_{OS}$) are reported (Fig. 5). As one can see, analogously to the most promising BF-based features, high E and low IDM enclose the patients with the worst prognosis ($OS < \mu_{OS}$), while low E and high IDM those with the best prognosis.

The outcome of MANOVA proves that both BF-based E and IDM as well as E and ASM differ significantly for the two groups of patients, with p -values of $6.4 \cdot 10^{-5}$ and $1.5 \cdot 10^{-4}$, respectively. Also CT-based E and IDM differ significantly for the two groups of patients, with p -values of $9.8 \cdot 10^{-6}$.

TNM staging (SE=92%, SP=14%, PPV=65%, NPV=50%, ACC=63%) performed worse as survival predictor.

IV. CONCLUSION

Although only essential parameters are routinely quantified in the clinical practice, this study shows that statistical and textural features, considered to capture the intra-tumoral heterogeneity, could work as a prognostic image-based biomarker for NSCLC. In particular, two feature-pairs derived from denoised BF map resulted strongly associated with the OS parameter. As expected, these features being derived from BF maps carry a great functional information content, which mainly reflects the presence of hypoxic regions and vascular abnormalities. Therefore, this finding represents a promising approach for the utilization of CTp in the clinical practice.

BF-based features

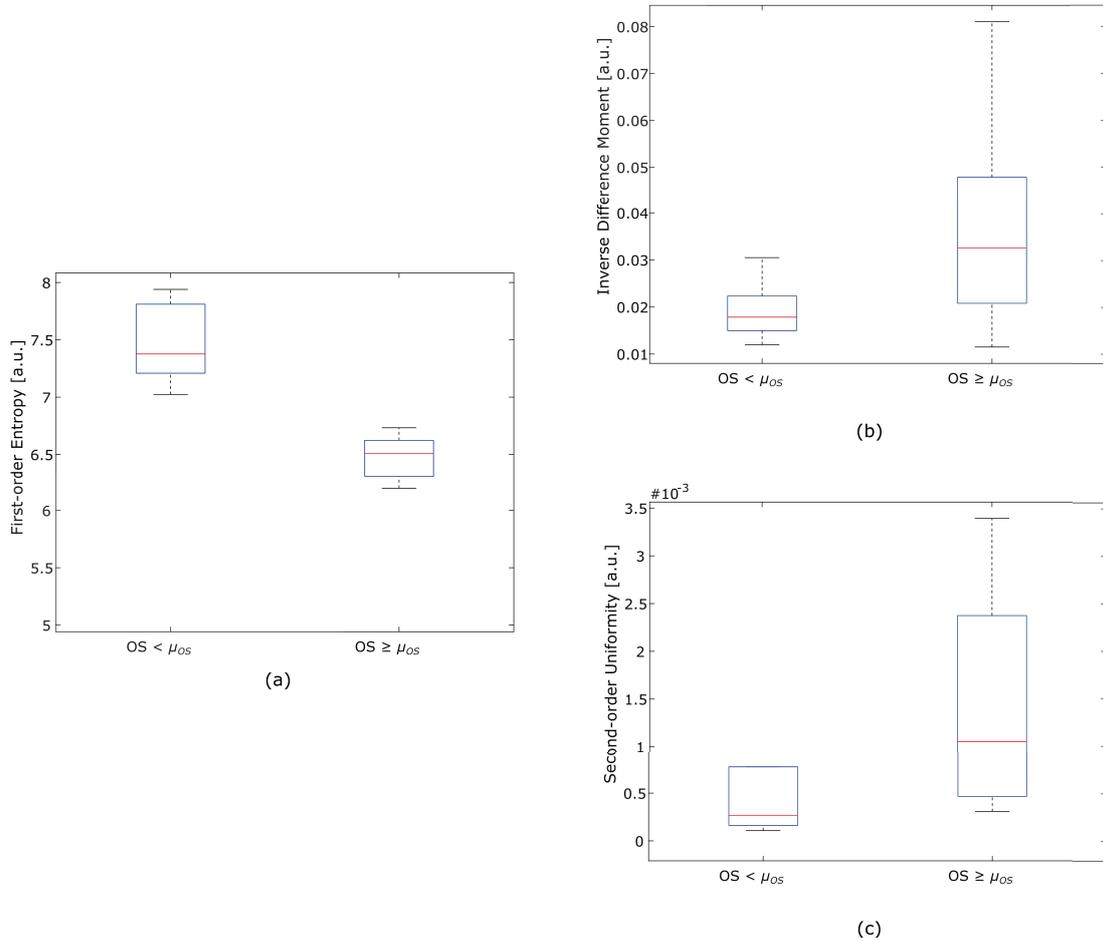


Fig. 4. Box plots showing first-order entropy (a), inverse difference moment (b) and second-order uniformity derived from BF maps for $OS < \mu_{OS}$ (left) and ($OS \geq \mu_{OS}$) (right). The median is indicated as a red line in the boxes, while the vertical box size gives the interquartile range

Despite being less promising than the BF-based features, it is worth noting that a feature-pair, derived from unenhanced CT images, resulted more specific than the TNM staging in discriminating patients with different survival expectation. It is also noticeable that this feature-pair is the same resulting promising for BF maps, this highlighting that the intra-tumoral heterogeneity for NSCLC, the phenomenon we investigated, is detectable independently from the modality adopted. This is an encouraging outcome since these features may be easily derived from conventional CT examinations that, contrarily to CTp examinations, are routinely acquired in the clinical practice.

The results of this study coherently showed that a greater intra-tumoral heterogeneity is associated to more aggressive tumours and strongly reflects on the OS parameter. It is worth noting that global entropy (E) resulted promising both if derived from BF maps and unenhanced CT images. This is not really surprising being E a meaningful feature representing values irregularity, better known as the measure of the information content. High entropy is associated to a worse prognosis. Analogously, great local homogeneity and uniformity correlate

with a better prognosis.

The encouraging outcomes achieved in this work should push for further studies, involving larger patients' cohorts as well as repeatability and reproducibility analyses. The results obtained also encourage the deepening of the unknown relationship between imaging findings and biologic features. This would permit to comprehensively understand the specific cancer biology of individual patient and infer phenotypic signatures containing predictive or prognostic values. Unlike biopsies, quantitative imaging could permit a longitudinal monitoring of the intra-tumour heterogeneity and, consequently, of the "specific" cell habitats and their changing over time, since clinical examinations can be easily repeated during the therapy [40].

In conclusion, the analysis of the tumour heterogeneity, able to detect valuable *in vivo* characteristics, should be further explored to achieve a personalisation of medicine. Indeed, having the potential to identify the patients who would benefit most from consideration of alternative therapies and treatment intensification, it could lead to a customisation of the therapy.

CT-based features

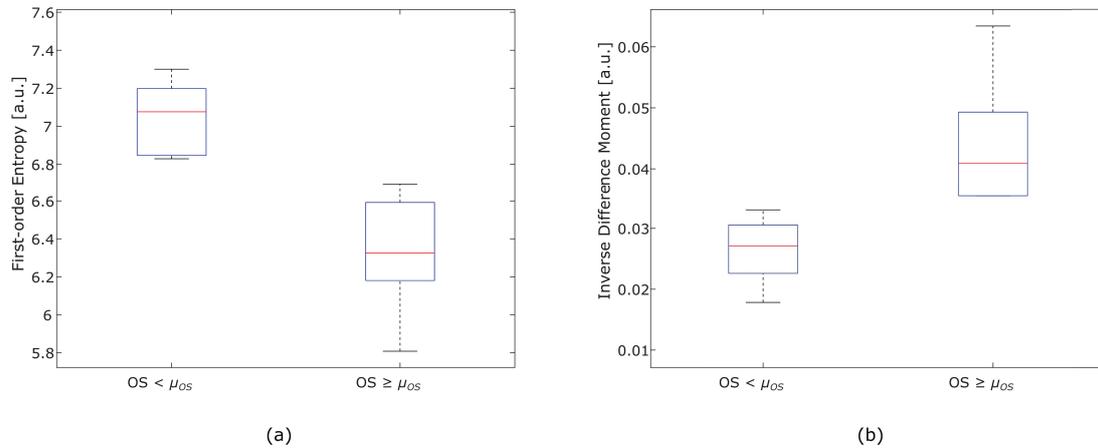


Fig. 5. Box plots showing first-order entropy (a) and inverse difference moment (b) derived from unenhanced CT images for $OS < \mu_{OS}$ (left) and $(OS \geq \mu_{OS})$ (right). The median is indicated as a red line in the boxes, while the vertical box size gives the interquartile range

REFERENCES

[1] Zhao Chen, Christine M Fillmore, Peter S Hammerman, Carla F Kim, and Kwok-Kin Wong. Non-small-cell lung cancers: a heterogeneous set of diseases. *Nat. Rev. Cancer*, 14(8):535–546, 2014.

[2] Martin Reck, David F Heigener, Tony Mok, Jean-Charles Soria, and Klaus F Rabe. Management of non-small-cell lung cancer: recent developments. *Lancet*, 382(9893):709–719, 2013.

[3] Marliese Alexander, Rory Wolfe, David Ball, Matthew Conron, Robert G Stirling, Benjamin Solomon, Michael MacManus, Ann Officer, Sameer Karnam, Kate Burbury, et al. Lung cancer prognostic index: a risk score to predict overall survival after the diagnosis of non-small-cell lung cancer. *British journal of cancer*, 117(5):744, 2017.

[4] Balaji Ganeshan, Elleny Panayiotou, Kate Burnand, Sabina Dizdarevic, and Ken Miles. Tumour heterogeneity in non-small cell lung carcinoma assessed by CT texture analysis: a potential marker of survival. *Eur. Radiol.*, 22(4):796–802, 2012.

[5] Stephen G Spiro, Nichole T Tanner, Gerard A Silvestri, Sam M Janes, Eric Lim, Johan F Vansteenkiste, and Robert Pirker. Lung cancer: progress in diagnosis, staging and therapy. *Respirology*, 15(1):44–50, 2010.

[6] Fergus Davnall, Connie SP Yip, Gunnar Ljungqvist, Mariyah Selmi, Francesca Ng, Bal Sanghera, Balaji Ganeshan, Kenneth A Miles, Gary J Cook, and Vicky Goh. Assessment of tumor heterogeneity: an emerging imaging tool for clinical practice? *Insights into imaging*, 3(6):573–589, 2012.

[7] Alessandro Bevilacqua and Serena Baiocco. Automatic classification of lung tumour heterogeneity according to a visual-based score system in dynamic contrast enhanced CT sequences. *Int. J. Mod. Phys. C*, pages 1650106–1–1650106–14, 2016.

[8] William D Figg and Judah Folkman. *Angiogenesis: an integrative approach from science to medicine*. Springer Science & Business Media, 2008.

[9] Gordon C Jayson, Robert Kerbel, Lee M Ellis, and Adrian L Harris. Antiangiogenic therapy in oncology: current status and future directions. *The Lancet*, 388(10043):518–529, 2016.

[10] Balaji Ganeshan and Kenneth A Miles. Quantifying tumour heterogeneity with CT. *Cancer imaging*, 13(1):140, 2013.

[11] Hugo JW Aerts, Emmanuel Rios Velazquez, Ralph TH Leijenaar, Chintan Parmar, Patrick Grossmann, Sara Carvalho, Johan Bussink, René Monshouwer, Benjamin Haibe-Kains, Derek Rietveld, et al. Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.*, 5(4006):1–8, 2014.

[12] S Baiocco, A Bevilacqua, D Barone, A Delmonte, and G Gavelli. L’eterogeneità perfusionale: un biomarker prognostico per il cancro del polmone non a piccole cellule (NSCLC). *Il Giornale Italiano di Radiologia Medica*, 4(4):665–669, 2017.

[13] Lejla Alic, Wiro J Niessen, and Jifke F Veenland. Quantification of heterogeneity as a biomarker in tumor imaging: a systematic review. *PLoS one*, 9(10):e110300, 2014.

[14] RCJ Beckers, S Trebeschi, M Maas, RS Schnerr, JML Sijmons, GL Beets, JB Houwers, RGH Beets-Tan, and DMJ Lambregts. CT texture analysis in colorectal liver metastases and the surrounding liver parenchyma and its potential as an imaging biomarker of disease aggressiveness, response and survival. *European journal of radiology*, 102:15–21, 2018.

[15] Renaud Lopes and Nacim Betrouni. Fractal and multifractal analysis: a review. *Medical image analysis*, 13(4):634–649, 2009.

[16] Sugama Chicklore, Vicky Goh, Musib Siddique, Arunabha Roy, Paul K Marsden, and Gary JR Cook. Quantifying tumour heterogeneity in 18 F-FDG PET/CT imaging by texture analysis. *European journal of nuclear medicine and molecular imaging*, 40(1):133–140, 2013.

[17] Zhaohai Yang, Laura H Tang, and David S Klimstra. Effect of tumor heterogeneity on the assessment of ki67 labeling index in well-differentiated neuroendocrine tumors metastatic to the liver: implications for prognostic stratification. *The American journal of surgical pathology*, 35(6):853–860, 2011.

[18] B Ganeshan, K Skogen, I Pressney, D Coutroubis, and K Miles. Tumour heterogeneity in oesophageal cancer assessed by CT texture analysis: preliminary evidence of an association with tumour metabolism, stage, and survival. *Clinical radiology*, 67(2):157–164, 2012.

[19] Connie Yip, David Landau, Robert Kozarski, Balaji Ganeshan, Robert Thomas, Andriana Michaelidou, and Vicky Goh. Primary esophageal cancer: heterogeneity as potential prognostic biomarker in patients treated with definitive chemotherapy and radiation therapy. *Radiology*, 270(1):141–148, 2014.

[20] Philippe Lambin, Emmanuel Rios-Velazquez, Ralph Leijenaar, Sara Carvalho, Ruud GPM van Stiphout, Patrick Granton, Catharina ML Zegers, Robert Gillies, Ronald Boellard, André Dekker, and Aerts Hugo J. Radiomics: extracting more information from medical images using advanced feature analysis. *Eur. J. Cancer*, 48(4):441–446, 2012.

[21] Charles A Harlow and Sharon A Eisenbeis. The analysis of radiographic images. *IEEE Transactions on Computers*, 100(7):678–689, 1973.

[22] S Herlidou-Meme, JM Constans, B Carsin, D Olivie, PA Eliat, L Nadal-Desbarats, C Gondry, E Le Rumeur, I Idy-Peretti, and JD De Certaines. MRI texture analysis on texture test objects, normal brain and intracranial tumors. *Magnetic resonance imaging*, 21(9):989–993, 2003.

[23] Davide Prezzi, Aisha Khan, and Vicky Goh. Perfusion CT imaging

- of treatment response in oncology. *European journal of radiology*, 84(12):2380–2385, 2015.
- [24] G Petralia, L Bonello, S Viotti, L Preda, G d’Andrea, and M Bellomi. CT perfusion in oncology: how to do it. *Cancer Imaging*, 10(1):8–19, 2010.
- [25] Nunzia Tacelli, Teresa Santangelo, Arnaud Scherpereel, Alain Duhamel, Valérie Deken, Ernst Klotz, Alexis Cortot, Jean-Jacques Lafitte, Frédéric Wallyn, Jacques Remy, et al. Perfusion CT allows prediction of therapy response in non-small cell lung cancer treated with conventional and anti-angiogenic chemotherapy. *Eur. Radiol.*, 23(8):2127–2136, 2013.
- [26] Gaspard d’Assignies, Anne Couvelard, Stéphane Bahrami, Marie-Pierre Vullierme, Pascal Hammel, Olivia Hentic, Alain Sauvanet, Pierre Bedossa, Philippe Ruszniewski, and Valérie Vilgrain. Pancreatic endocrine tumors: tumor blood flow assessed with perfusion CT reflects angiogenesis and correlates with prognostic factors. *Radiology*, 250(2):407–416, 2009.
- [27] Avinash Kambadakone, Sam S Yoon, Tae-Min Kim, Daniel L Karl, Dan G Duda, Thomas F DeLaney, and Dushyant V Sahani. CT perfusion as an imaging biomarker in monitoring response to neoadjuvant bevacizumab and radiation in soft-tissue sarcomas: comparison with tumor morphology, circulating and tumor biomarkers, and gene expression. *Am. J. Roentgenol.*, 204(1):W11–8, 2015.
- [28] Alessandro Bevilacqua, Domenico Barone, Serena Baiocco, and Giampaolo Gavelli. A novel approach for semi-quantitative assessment of reliability of blood flow values in DCE-CT perfusion. *Biomed. Signal Process. Control*, 31:257–264, 2017.
- [29] Serena Baiocco, Domenico Barone, Giampaolo Gavelli, and Alessandro Bevilacqua. Analysis of CT Perfusion Blood Flow maps in Patients with Lung Cancer: Correlation with the Overall Survival. In *23rd Conference of Open Innovations Association (FRUCT)*, pages 55–61. IEEE, 2018.
- [30] Alessandro Bevilacqua, Domenico Barone, Silvia Malavasi, and Giampaolo Gavelli. Quantitative assessment of effects of motion compensation for liver and lung tumors in CT perfusion. *Acad. Radiol.*, 21(11):1416–1426, 2014.
- [31] Yoshiharu Ohno, Yasuko Fujisawa, Hisanobu Koyama, Yuji Kishida, Shinichiro Seki, Naoki Sugihara, and Takeshi Yoshikawa. Dynamic contrast-enhanced perfusion area-detector CT assessed with various mathematical models: Its capability for therapeutic outcome prediction for non-small cell lung cancer patients with chemoradiotherapy as compared with that of FDG-PET/CT. *Eur. J. Radiol.*, 86:83–91, 2017.
- [32] KA Miles and MR Griffiths. Perfusion CT: a worthwhile enhancement? *Br. J. Radiol.*, 76(904):220–231, 2003.
- [33] KA Miles. Perfusion CT for the assessment of tumour vascularity: which protocol? *Br. J. Radiol.*, 76(1):S36–42, 2003.
- [34] Gunnar Brix, Stefan Zwick, Jürgen Griebel, Christian Fink, and Fabian Kiessling. Estimation of tissue perfusion by dynamic contrast-enhanced imaging: simulation-based evaluation of the steepest slope method. *Eur. Radiol.*, 20(9):2166–2175, 2010.
- [35] Sylvain Goutelle, Michel Maurin, Florent Rougier, Xavier Barbaut, Laurent Bourguignon, Michel Ducher, and Pascal Maire. The Hill equation: a review of its capabilities in pharmacological modelling. *Fundam. Clin. Pharmacol.*, 22(6):633–648, 2008.
- [36] Alessandro Bevilacqua, Giampaolo Gavelli, Serena Baiocco, and Domenico Barone. CT Perfusion in Patients with Lung Cancer: Squamous Cell Carcinoma and Adenocarcinoma Show a Different Blood Flow. *BioMed Research International*, 2018, 2018.
- [37] Robert M Haralick, Karthikeyan Shanmugam, et al. Textural features for image classification. *IEEE Transactions on systems, man, and cybernetics*, (6):610–621, 1973.
- [38] E Roy Davies. *Machine vision: theory, algorithms, practicalities*. Elsevier, 2004.
- [39] Sergios Theodoridis and Konstantinos Koutroumbas. *Pattern Recognition, Second Edition*. Academic Press, Inc., Orlando, FL, USA, 2003.
- [40] Elisa Meldolesi, Nicola Dinapoli, Roberto Gatta, Andrea Damiani, Vincenzo Valentini, and Alessandra Farchione. How Can Radiomics Improve Clinical Choices? In *Multidisciplinary Management of Rectal Cancer*, pages 135–149. Springer, 2018.