




Article

# CT Radiomics in Colorectal Cancer: Detection of KRAS Mutation Using Texture Analysis and Machine Learning

Víctor González-Castro <sup>1</sup>, Eva Cernadas <sup>2,\*</sup>, Emilio Huelga <sup>3</sup>, Manuel Fernández-Delgado <sup>2</sup>, Jacobo Porto <sup>3</sup>, José Ramón Antunez <sup>4</sup> and Miguel Souto-Bayarri <sup>3</sup>

<sup>1</sup> Department of Electrical, Systems and Automation Engineering, Universidad de León, Campus de Vegazana s/n, 24071 León, Spain; victor.gonzalez@unileon.es

<sup>2</sup> Centro Singular de Investigación en Tecnoloxías Intelixentes da USC (CiTIUS), Universidade de Santiago de Compostela, Rúa Xenaro de la Fuente Domínguez, 15782 Santiago de Compostela, Spain; manuel.fernandez.delgado@usc.es

<sup>3</sup> Department of Radiology, Complejo Hospitalario Universitario de Santiago, USC, 15706 Santiago de Compostela, Spain; Emilio.Huelga.Zapico@usc.es (E.H.); jacoporto@hotmail.com (J.P.); miguel.souto@usc.es (M.S.-B.)

<sup>4</sup> Department of Pathology, Complejo Hospitalario Universitario de Santiago, USC, 15706 Santiago de Compostela, Spain; Jose.Ramon.Antunez.Lopez@sergas.es

\* Correspondence: eva.cernadas@usc.es

Received: 6 July 2020; Accepted: 4 September 2020; Published: 7 September 2020



**Featured Application:** Detection of the KRAS mutation without an invasive technique (biopsy) may have an important role in the diagnosis, prognosis, treatment, and monitoring of the patients with colorectal cancer. Machine learning algorithms allow determination of the presence of the mutation from the analysis of the CT image, and prevent mistakes by biopsying only part of the tumour.

**Abstract:** In this work, by using descriptive techniques, the characteristics of the texture of the CT (computed tomography) image of patients with colorectal cancer were extracted and, subsequently, classified in KRAS+ or KRAS-. This was accomplished by using different classifiers, such as Support Vector Machine (SVM), Grading Boosting Machine (GBM), Neural Networks (NNET), and Random Forest (RF). Texture analysis can provide a quantitative assessment of tumour heterogeneity by analysing both the distribution and relationship between the pixels in the image. The objective of this research is to demonstrate that CT-based Radiomics can predict the presence of mutation in the KRAS gene in colorectal cancer. This is a retrospective study, with 47 patients from the University Hospital, with a confirmatory pathological analysis of KRAS mutation. The highest accuracy and kappa achieved were 83% and 64.7%, respectively, with a sensitivity of 88.9% and a specificity of 75.0%, achieved by the NNET classifier using the texture feature vectors combining wavelet transform and Haralick coefficients. The fact of being able to identify the genetic expression of a tumour without having to perform either a biopsy or a genetic test is a great advantage, because it prevents invasive procedures that involve complications and may present biases in the sample. As well, it leads towards a more personalized and effective treatment.

**Keywords:** KRAS mutation; colorectal cancer; texture analysis; wavelets; haralick texture descriptors; Support Vector Machine; Grading Boosting Machine; Neural Network; Random Forest

## 1. Introduction

Colorectal cancer is second in incidence in women after breast cancer, followed in third and fourth place by lung and cervix cancers, respectively. In men, lung cancer is the highest in incidence, with prostate cancer being the second, followed in third place by colorectal cancer. Despite the important advances in its treatment, colorectal cancer remains the second cause of death from cancer considering the general population, with an estimated mortality in both sexes worldwide in 2018 of 880,792 deaths [1,2]. In the United States, colorectal cancer is the third most common cancer diagnosed in both sexes, and The American Cancer Society estimates 104,610 new cases of colorectal cancer will occur in 2020 [3]. In Spain, the colorectal cancer was the cancer most frequently diagnosed in both sexes in 2019 (with 44,937 new cases), the second in men after prostate cancer, and the second in women after breast cancer [4].

There seems to be a process with multiple stages of evolution from the normal colonic mucosa to a potentially fatal carcinoma. Cells, which must have a certain genetic predisposition or suffer a series of genotoxic phenomena, are induced to proliferate, thus passing through a series of stages until they end up doing so in a completely uncontrolled manner. One of the proto-oncogenes that can be altered in this process is KRAS, which is involved in the pathway of transduction of growth signals and cell differentiation. It is estimated that 65% of sporadic colorectal carcinomas present activation by point mutations in a gene of the RAS family (mainly KRAS) [5]. The RAS gene is now part of the criteria of the new TNM classification of malignant tumours as a prognostic and predictive factor, since patients with early stage colorectal cancer (I and II) have a lower survival rate when related to the KRAS mutation. In patients with metastatic colorectal cancer, the median overall survival without treatment is approximately 6 months. With the use of current therapies, survival can be prolonged to a median of 30 months, with a response rate for some therapies of 50% [6]. The introduction of targeted therapies resulted in an increase in progression-free survival and also in overall survival in patients with metastatic colorectal cancer. Therefore, in addition to combined chemotherapy regimens, there are currently a series of agents that act against certain specific targets important in the pathological process of CRC (colorectal cancer). These therapies are monoclonal antibodies against EGFR, monoclonal antibodies against VEGF-A, and also fusion proteins against multiple molecules and growth factors. For the use of this targeted therapy, the existence of mutations in KRAS and NRAS must be previously verified. If these mutations are present, the tumour cell presents an activation of multiple signalling pathways that render useless the targeted treatment against those genes [7]. Therefore, it has been described that mutations in the KRAS proto-oncogene cause a poor response to anti-EGFR treatment in cases of metastatic colorectal cancer. The presence or absence of this mutation must be confirmed, because patients with colorectal cancer who present native KRAS can benefit from targeted therapy with anti-EGFR [8,9].

Radiological images contain information about the underlying pathophysiology that can be extracted through a quantitative analysis of the texture of the image [10,11]. With radiomics, analysing the relationship between the intensity/density of the pixels that make up the texture of the tumour image, we could obtain an identificative pattern of an underlying pathophysiological alteration, as has already been demonstrated in other tumours using other radiological techniques [12–17]. This information is different from that provided by laboratory tests and other pathological or genomic analysis. With these data, obtained from the texture analysis of an image, a common pattern can be detected in patients with a specific genomic mutation, such as the KRAS mutation in patients with colorectal cancer.

Trying to describe different textures of an object is extremely difficult for the human eye, even more so if that description is intended to be done on a CT image. However, by using computer programs it is possible to assign a quantitative value to a certain texture of a CT image, taking into account many characteristics that are invaluable to the human eye. If this quantitative analysis of the texture of a tumour with a certain pathophysiological singularity (mutation in the KRAS oncogene) is able to obtain a characteristic pattern, it will be possible to check whether other tumours have this

pattern (and therefore, the KRAS oncogene mutation). This is of utmost importance because the radiological data derive from the entire tumour, while the pathological data only derives from the sample obtained invasively. Therefore, radiomics can provide valuable information regarding the genomics of the tumour process, and this may have an important role in the diagnosis, prognosis, treatment, and monitoring of these patients [18].

To identify a pattern in the images of patients with a mutation in the KRAS gene, it is necessary to extract from CT images characteristics of the texture which cannot be extracted by the visual analysis of the radiologists. Subsequently, using machine learning algorithms, e.g., Support Vector Machine (SVM), Grading Boosting Machine (GBM), Neural Networks (NNET), or Random Forest (RF), the molecular signature of colon cancers could be automatically classified into KRAS + or KRAS- based on those previously extracted texture characteristics. In this way, the existence of common texture patterns that are present in patients with colorectal cancer with mutated KRAS, but not in patients with native KRAS, will be demonstrated, which will allow determination of the presence of the mutation from the analysis of the CT image.

The Materials and Methods section details the selection of the patients, the description of the texture features and machine learning models, and the performance measures. The Results and Discussion section summarizes the results achieved. Finally, the main conclusions and future work are reported in the Conclusions section.

## 2. Objectives

The objective of this research is to use radiomics to detect the KRAS oncogene mutation in patients with colorectal cancer by analysing the texture of the images obtained by CT, thus reflecting that there is a relationship between the texture pattern of a radiological imaging test with the underlying genetic mutation. The detection of this KRAS mutation without an invasive technique may have an important role in the diagnosis, prognosis, treatment, and monitoring of the patients with colorectal cancer.

A secondary objective, but no less important, is to contribute to increasing the bibliography and the available experience regarding the possibilities of radiomics, a field that is booming and that requires the creation of common standards and protocols to establish the ideal methodology to get the most out of this promising field.

## 3. Materials and Methods

The first subsection describes the definition of the dataset used to evaluate the performance of the classification system of KRAS+ and KRAS- genetic disease from the TC images. This encloses the selection of patients and the annotation of the CT images. Texture analysis is a very important topic in medical image analysis [13,19]. Texture classification is concerned with the problem of assigning a sample image to one in the set of known texture classes (in our case, KRAS+ and KRAS-). The texture classification process normally encloses the following two steps: (1) computation of the texture features, which is described in the second subsection; and (2) design of machine learning models to predict the output class of an unknown input patient (third subsection). Finally, the fourth subsection describes the performance measurements used in this work.

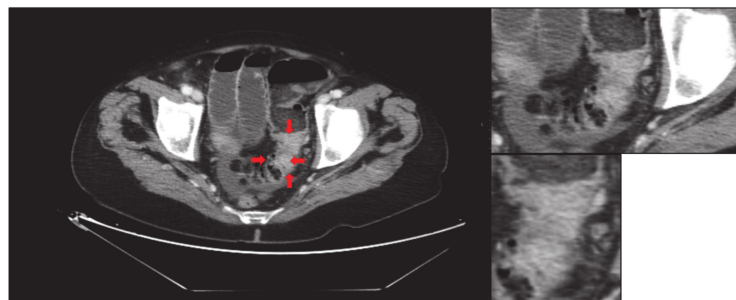
### 3.1. Patients Selection and Annotation of Regions of Interest (ROIs)

To carry out this cross-sectional study, the approval of the Research Ethics Committee was obtained. For this study, 47 patients were selected from the databases of the CHUS University hospital (Complejo Hospitalario Universitario de Santiago de Compostela). The inclusion criteria were as follows: (1) patients with colorectal cancer with determination of the KRAS mutation by biopsy between 2016 and 2017 (27 KRAS+ and 20 KRAS- patients, respectively); (2) selection of those patients whose intravenous contrast CT dated before any treatment; (3) obtaining the CT images with a slice thickness of 5 mm. Very thin slice thickness improves spatial resolution but decreases contrast. Therefore, 5 mm is the thickness we estimate best to determine the difference in texture of the tumour pixels. In addition,

the patients included in the study already had the CT done, since it is one of the first tests to be performed in tumour staging; the subsequent CTs could not be included in the study, since they are post-treatment and may bias its results.

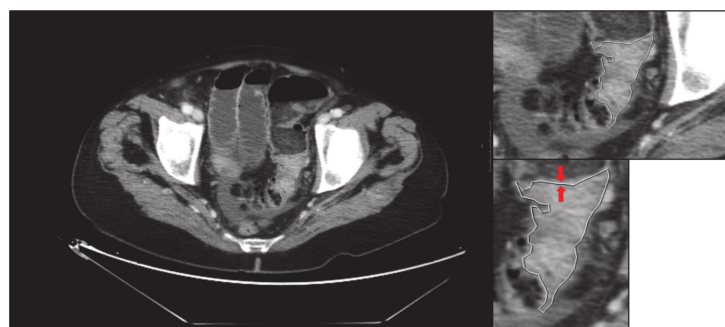
The CT images used must be selected to any treatment applied. Therefore, to obtain the images of each patient, we will resort to the CT performed at the time of diagnosis. From these images, a series of slices are obtained: three slices of the tumour are selected, i.e., the slice in which the tumour presents the largest diameter (central slice) and the slices immediately cranial and immediately caudal to that central slice. The selection of three slices for each tumour is made to encompass the largest possible tumour volume in the analysis, trying to approach a 3D analysis of it. With the three slices we make sure to select the largest amount of tumour that the imaging technique allows. We did not select more slices because in the analysis of some tumours, due to their size and disposition, it was not possible to analyse more than three slices, so it was finally decided to standardize the analysis of all patients' tumours to three slices. Then, with each image of the three slices, the tumour is manually delimited by an expert radiologist, performing the following segmentation sequence:

- First, an image of tumour is saved without limitation in original size as “large-original.tiff” (i.e., Figure 1). The sequence of obtaining the images is very important to speed up the process and ensure that the images have the same focus when they are enlarged. An expert radiologist searches the cut with the largest diameter of the tumour. All images must be saved in “.tiff” format, as we will explain later.



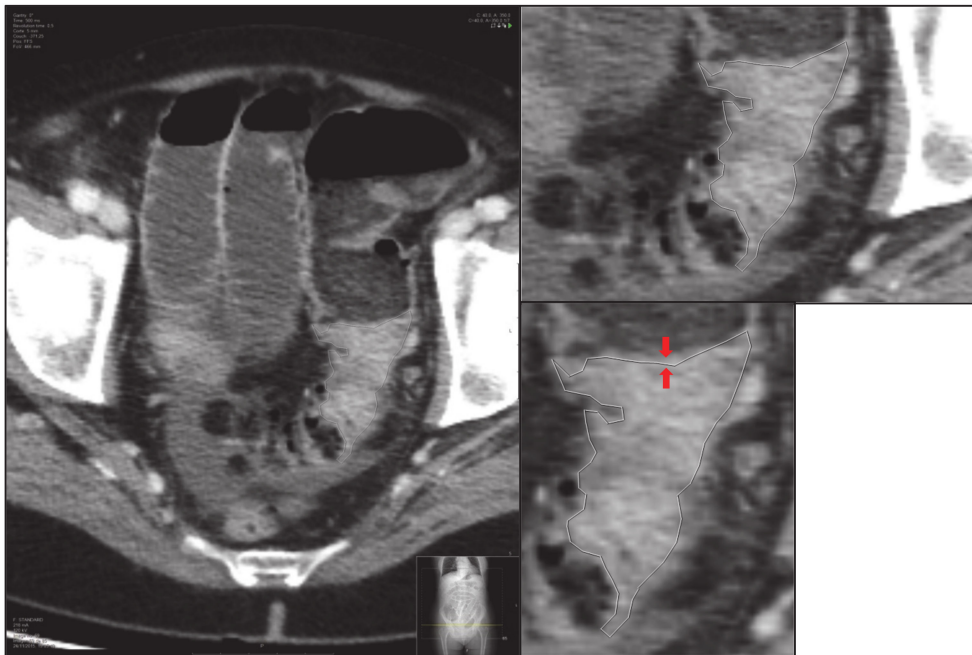
**Figure 1.** “Large-original” is the first image to obtain. This image belongs to a KRAS+ patient included in the study. The tumour under study is highlighted with red arrows and can also be seen in greater detail in the adjacent images.

- Secondly, an expert radiologist delimits the tumour, and an image of the delimited tumour is saved in its original size: “large-delimited.tiff” (i.e., Figure 2). The delimitation process is carried out manually by an expert, since due to the morphological heterogeneity of the tumour it is impossible to make an automatic delimitation.



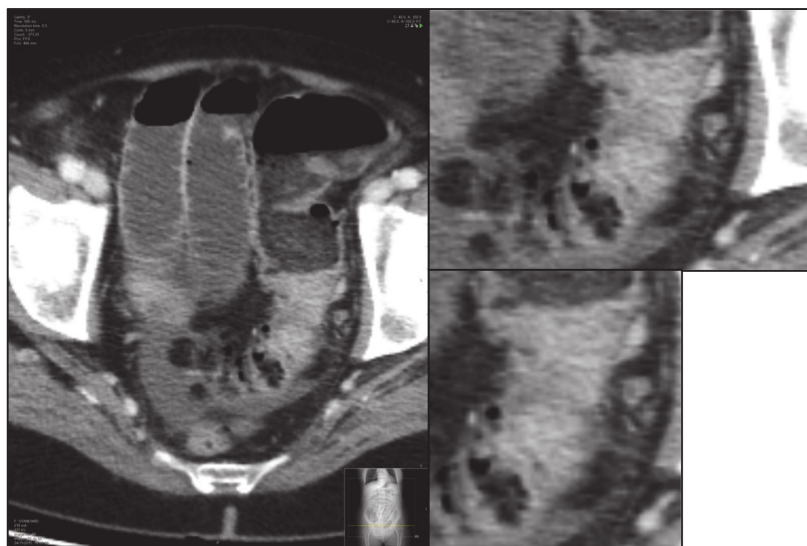
**Figure 2.** “Large-delimited” is the second image to be obtained. In the two adjacent small images, the delimited tumour can be seen in detail. The delimitation line is highlighted between the two red arrows in the most enlarged image.

- Third, the screen zoom is adjusted by enlarging the image and an image of the segmented and enlarged tumour is saved: “small-delimited.tiff” (i.e., Figure 3).



**Figure 3.** “Small-delimited”. To take this image, starting from the “large-delimited” image, the zoom is enlarged focusing on the delimited tumour. In the two adjacent small images, the delimited tumour can be seen in detail. In the enlarged image the delimitation is highlighted between the two red arrows.

- Finally, without modifying the zoom, the previous delimitation is eliminated, and one image of the tumour is saved without delimiting and enlarged: “small-original.tiff” (i.e., Figure 4).



**Figure 4.** “Small-original”. In the two adjacent small images, the tumour can be seen in detail. In this case, the images have not delimitation.

This process of obtaining the images is repeated with the slices immediately cranial and immediately caudal. Finally, a set of 12 images is obtained for each patient. As already mentioned, the order of obtaining the four images in each of the three slices is very important so that the enlarged

images have the same size and focus, thus ensuring that the enlarged images are in the same measure. In addition, this order of taking images also streamlines the process by being able to take them consecutively without almost having to make adjustments between shots.

### 3.2. Extraction of Texture Features

The texture is a visual pattern in the image, modelled by the computer as a vector of numerical values, which codifies the image structure. Many texture extraction techniques have been proposed in the literature and they have successfully applied to many applications [20]. Most of them are applied to rectangular or squared images. Nevertheless, in our application, the cancer tumours are irregular regions in the image. In a previous work, we have adapted some popular texture descriptors to their application on irregular patches in the image [21]. In this paper, we compare the performance of some descriptors of the two most popular families of texture descriptors: second order statistics of grey level pixels and spectral texture features. The family of statistical features encode the spatial distribution of grey levels in the image. We use the Local Binary Patterns (LBP), proposed by Ojala et al. in 2002 [22], and the Haralick coefficients, calculated from the grey level cooccurrence matrix [23]. The family of spectral features, like wavelets or Gabor filters, does decompositions of the signal or image in the frequency domain. Gabor filters uses the FFT (Fast Fourier Transform), which is a global technique, and they are not suitable to analyse irregular patches. Therefore, we only use the wavelet features.

Let  $G = \{0, \dots, N_g - 1\}$  be the set of  $N_g$  grey levels,  $S$  a finite subset of indexes specifying the region of interest (ROI) to be analysed (in our case, the tumour),  $I(x, y) \in G$  the grey level in the pixel  $(x, y) \in S$ . The Grey Level Cooccurrence Matrix ( $M$ ) of the image is a matrix of  $N_g \times N_g$  elements. The element  $M_{ij}$  represents the number of pairs of pixels with grey levels  $i$  and  $j$  that occur among the pixels belonging to the set  $S$  for a certain distance (i.e., scale)  $d$  and an orientation  $\theta$  from each other. For the construction of a rotation invariant cooccurrence matrix, we average the matrix to different orientations. The probability density function  $P_d(i, j), i, j \in G$  for each  $d$  is obtained dividing all elements of  $M_{ij}$  by the number of cooccurrences in  $M_{ij}$ . This probability matrix is normally characterized by the following features: energy ( $E$ ), correlation ( $Cor$ ), contrast ( $Con$ ), homogeneity ( $Ho$ ), and entropy ( $H$ ), which are defined as:

$$E = \sum_i \sum_j P_d(i, j)^2 \tag{1}$$

$$Cor = \frac{1}{\sigma_x \sigma_y} \sum_i \left[ \sum_j ij P_d(i, j) - \mu_x \mu_y \right] \tag{2}$$

$$Con = \sum_i \sum_j \frac{(i - j)^2 P_d(i, j)}{(N_g - 1)^2} \tag{3}$$

$$Ho = \sum_i \sum_j \frac{P_d(i, j)}{1 + |i - j|} \tag{4}$$

$$H = - \sum_i \sum_j P_d(i, j) \log(P_d(i, j)) \tag{5}$$

where  $\mu_x = \sum_i i \sum_j P_d(i, j)$ ,  $\mu_y = \sum_j j \sum_i P_d(i, j)$ ,  $\sigma_x = \sqrt{\sum_i (i - \mu_x)^2 \sum_j P_d(i, j)}$ , and  $\sigma_y = \sqrt{\sum_j (j - \mu_y)^2 \sum_i P_d(i, j)}$ . All the sums over  $i$  and  $j$  are for  $i, j = 0, \dots, N_g - 1$ . We compute eight features vector  $HF_{df}$  varying the distances used  $d = \{1, 2, 3, 123\}$  (distance one, two or three pixels and the concatenation of these three distances) and the number of features  $f = \{4, 5\}$ , feature vectors are constructed using the five above features (i.e., energy, correlation, contrast, homogeneity, and entropy) or excluding the entropy. So, the number of texture features is: 4 for  $HF_{d4}, d = \{1, 2, 3\}$  and 5 for  $HF_{d5}, d = \{1, 2, 3\}$ , and 12 and 15 features for  $HF_{123f}, f = 4$  and  $f = 5$ , respectively.

The original local binary pattern [22] extracts information locally from the texture of the images comparing the intensity of each pixel  $z_c = (x_c, y_c) \in S$  in the ROI with that of its neighbours. The neighbouring pixels  $P$  are angularly distributed over a circle with radius  $R$  centred in  $z_c \in S$ . The LBP signature assigns a binary label, i.e., 0 or 1, to each neighbour depending on whether the central pixel  $z_c$  has higher intensity value than the neighbouring pixels. Such an LBP is not robust against rotations. The rotationally invariant LBP for  $z_c$ , given a radius  $R$  and the number of involved neighbours  $P$ , is given by:

$$LBP_{ri}(z_c) = \min_{0 \leq n < P} \left\{ \sum_{i=0}^{P-1} u(I(z_i) - I(z_c)) 2^{(i+n)\%P} \right\}, \quad (6)$$

where  $z_i$  is the  $i$ -th neighbouring pixel,  $i = 0, \dots, P-1$ .  $I(z_i)$  is the grey level in pixels  $z_i$  and  $u(x) = 1$  for  $x \geq 0$  and  $u(x) = 0$ . The symbol  $\%$  denotes the remainder operation. The circular shift effect produced when the image is rotated is removed by finding the minimum value among all possible values. The original LBP texture descriptor yields a vector of  $2^P$  features, which is too large for higher values of  $P$  for many classifiers, especially if the number of input patterns is low. To overcome this drawback, we use the LBP uniform patterns, which refer to the patterns which have a limited number of transitions from 0 to 1 or vice versa, lower, or equal than two. In a circularly symmetric neighbour set of  $P$  pixels can occur  $P+1$  uniform binary patterns. In the uniform patterns, the number of 1's in the binary pattern is the label of the patterns, while the non-uniform patterns are labelled by  $P+1$ . The histogram of the patterns' labels accumulated over the set  $S$  is employed as texture feature vector. So, we compute three texture feature vectors  $LBP_{R_i}$  for each radius  $R_i = \{1, 2, 3\}$  using  $P = 8$ , and the feature vector  $LBP_{R_{123}}$  with 30 features for the concatenation of the previous three LBP vectors.

The discrete wavelet transform (DWT) representation can be used as spectral texture descriptors. The DWT decomposes the image into low-pass and high-pass frequency bands. Applying recursively, the filters to the bands produce the wavelet decomposition. Therefore, the wavelet transform involves filtering and subsampling. A more compact representation is normally derived from the bands as texture descriptors. Some possible measures to compute over the transform coefficients for each sub-band and decomposition level are mean, energy, entropy, and standard deviation. The measures are computed using the pixels  $(x, y) \in S$ . We construct eight feature vectors  $DWT_{sfm}$  varying the number of levels of decomposition  $s = \{2, 3\}$ , the number of filters  $f = \{LL, All\}$  (i.e., using only the low-pass sub-bands or using all sub-bands), and the number of measures  $m = \{EEMS, EMS\}$ , i.e., using the four measures EEMS (energy, entropy, mean, and standard deviation) or EMS, excluding the entropy.

Another popular family of texture features for classification is multi-scalar analysis, which encodes texture feature in different scales. This encoding can be done changing the size of neighbourhood in the statistical texture features, e.g., the LBP texture features or Haralick coefficients mentioned above or doing decompositions and down-sampling in the wavelet transforms, as the  $DWT_{sfm}$  texture feature vectors. We also achieved multi-scalar analysis combining subsampling using wavelets with statistical analysis. Specifically, we compute the cooccurrence matrix and Haralick features over a wavelet decomposition in the following ways: (1) texture feature vectors Wavelet Co-occurrence Features,  $WCF_{dm}$ , and (2) texture feature vectors Wavelet Decomposition Cooccurrence Features,  $WDC_{Ffm}$ . The features  $WCF_{dm}$  are calculated applying the wavelet low-pass (L) and high-pass (H) filtering to the original image  $I$  to produce the sub-bands LL, LH, HL, and HH. We calculate the cooccurrence matrix over the region of interest  $S$  (and scaled  $S$  for the sub-bands) for different distances  $d = \{1, 2, 3\}$  and calculating different sets of Haralick features for  $m = \{4, 5\}$ , depending on whether the feature entropy is included or not. This process develops texture feature vectors of 20 characteristics for  $WCF_{d4}$  [24] and 25 for  $WCF_{d5}$ , with  $d = \{1, 2, 3\}$ . The features  $WDC_{Ffm}$  are calculated applying three level wavelet decomposition levels to the original image  $I$ . We apply the decomposition only over the low-pass LL subbands and compute the cooccurrence matrix using a distance  $d = 1$  and Haralick coefficients  $m = \{4, 5\}$  using the LL subbands, i.e.,  $f = LL$ , or all subbands, i.e.,  $f = All$ . The number of features is 16 for  $WDC_{FLL4}$ , 20 for  $WDC_{FLL5}$ , 52 for  $WDC_{FAll4}$  and 65 for  $WDC_{FAll5}$ .

### 3.3. Machine Learning Models

We use supervised Machine Learning classifiers to predict the genetic disease of the patient (KRAS+ or KRAS-) from the CT sequence, represented as a texture feature vector for each slice. The classifier learns a function from examples (diagnosed patients using biopsy) that predicts the diagnosis of the patient from the CT sequence. The trained classifier can generalize its prediction to new patients that have not been seen during training. In the current study, we selected some of the classifiers identified as best-performing in our exhaustive comparison [25]. These classifiers are:

- Support Vector Machine (SVM), a state-of-the-art classifier that simultaneously minimizes the training error and the classification margin of a hyperplane (related to the ability of generalization) on a hidden space projected from the input space using a non-linear kernel. We used the *LibSVM* library (<https://www.csie.ntu.edu.tw/~cjlin/libsvm>) under the Octave interface. The hyper-parameters of the SVM are the inverse spread  $\gamma$  of the Radial Basis Function kernel, tuned with values  $2^{-10}, 2^{-8}, 2^{-6}, \dots, 2^{10}$ , and the regularization parameter  $C$ , tuned with values  $2^{-5}, 2^{-3}, 2^{-1}, \dots, 2^{10}$ .
- Gradient Boosting Machine (GBM) is an ensemble of classification trees trained using the functional gradient descent variant of the Boosting algorithm. We used the *gbm* package of the R statistical computing language (<http://www.r-project.org>), tuning the maximum depth of each tree (from 1 to 10) and the number of trees in the ensemble (from 50 to 500). The specific values of each parameter are generated using the utilities of the *caret* R package, being in general different for each combination of features.
- Multi-Layer Perceptron Neural Network (NNET) is the classical neural network, trained using the Back-Propagation algorithm implemented by the *nnet* package in R. The hyper-parameters, defined by the *caret* package, are the number of hidden neurons (tuned with values from 1 to 19) and the weight decay (10 values from 0.1 to 0.001).
- Random Forest (RF) is a bagging ensemble of random classification trees which is also considered a state-of-the-art classifier. Each node in the tree splits one input in two intervals, and each tree learns a different view of the classification problem. A datum is classified using a voting among all the trees in the forest. We used an RF of 500 trees, implemented by the *randomForest* R package, tuning the hyper-parameter *mtry* (number of variables that are randomly sampled as candidates in each split), with values between 2 and the number of inputs.
- Linear Discriminant Analysis (LDA), is a classical approach that performs a linear classification oriented to simultaneously maximize the inter-class variance and minimize the within-class variance. This linear mapping is defined by the eigenvectors associated to the largest eigenvalues of the inter-class and within-class covariance matrices. Usually, LDA is used for dimensionality reduction prior to classification with more powerful classifiers. We included it in our study to provide a baseline performance for the remaining classifiers. We used the *lda* function in the *MASS* R package.

### 3.4. Performance Measures

A variant of the common cross-validation methodology has been used to develop the experiments to predict the genetic disease, which consists of: (1) the use of three data partitions (or sets), for training, validation and test, instead of classical cross-validation, which only uses training and test sets; and (2) the dataset is created leaving one patient out (i.e., the feature vectors that describe the tumour in the three slices of a patient) instead of leaving one slice out. The texture feature vector calculated for each image is used as input pattern for the classifier, so there are so many patterns as images. The values selected for the hyper-parameters of the classifiers are the ones that achieve the best performance over the validation set.

The classifier's performance is measured using the Cohen's kappa value (K), which measures the agreement between the true and predicted categories labels excluding the agreement by chance.



In addition, its value is always lower and, therefore, more conservative and with stronger statistical validity, than the balanced accuracy, Area Under the ROC Curve (AUC), or Matthews coefficient [26]. Let the true positive (TP) be the number of sick people correctly identified as sick, the false positive (FP) be the number of healthy people incorrectly identified as sick, the true negative (TN) be the number of healthy people correctly identified as healthy, and the false negative (FN) be the number of sick people incorrectly identified as healthy. The confusion matrix (Table 1) visualizes the classifier prediction (column) and the true class (row). The diagonal corresponds to the patients correctly classified. The kappa (in %) is defined for classification problems with two categories, in our case KRAS+ and KRAS-, as

$$K = 100 \frac{p_a - p_e}{s - p_e} \tag{7}$$

where  $p_a = \sum_{i=1}^C C_{ii}$ ,  $p_e = \frac{1}{N^2} \sum_{i=1}^C (\sum_{j=1}^C C_{ij})(\sum_{j=1}^C C_{ji})$  and  $s = \sum_{i=1}^C \sum_{j=1}^C C_{ij}$ , while  $C = 2$  is the number of classes,  $C_{ij}$  is the  $ij$ -th value in the confusion matrix, e.g.,  $C_{ij}$  is the number of patients of the  $i$ -th category for which the classifier predicts the  $j$ -th category. The sensitivity ( $Se$ ) refers to the classifier’s ability to correctly detect sick (KRAS+) patients while the specificity ( $Sp$ ) relates to the classifier’s ability to correctly reject healthy patients (KRAS-). They are defined by

$$Se(\%) = \frac{100TP}{TP + FN} \tag{8}$$

$$Sp(\%) = \frac{100TN}{TN + FP} \tag{9}$$

**Table 1.** Representation of a confusion matrix used to visualize the classifier performance. Each column in the matrix represents the instances in the predicted class while each row represents the instances in the true class.

		Predicted Class	
		KRAS+	KRAS-
True class	KRAS+	TP	FP
	KRAS-	FN	TN

## 4. Results and Discussion

### 4.1. Experimental Setup

This section describes the dataset used and the evaluation methodology used to obtain the results of the next section. Our data set contains 47 patients (27 KRAS+ and 20 KRAS-) from the CHUS University hospital in Santiago de Compostela. The inclusion criteria were as follows: (1) patients with colorectal cancer with determination of the KRAS mutation by biopsy between 2016 and 2017; (2) selection of those patients whose intravenous contrast CT dated before any treatment; (3) obtaining the CT images with a thickness of slice of 5 mm. From each patient, an experienced radiologist selects the slice with the largest transverse diameter of the tumour (central slice) and the slices immediately above and below it. From these three slices, the radiologist takes four images of each of the slices, two of which are manually outlined. The total number of images (slices) is 141 (47 patients multiplied by 3 cuts per patient).

As mentioned, we use the LOPO (leave-one-patient-out) cross-validation approach. To develop the experiment,  $N$  partitions were created, being  $N$  the number of patients, each including a training, a validation and a test set. For each partition  $i = 1, \dots, N$ , the three slices belonging to the  $i$ -th patient are used as test set, and the remaining patients (three slices per patient) are distributed into the training and validation sets (50% of patients each one). All the inputs, that correspond to the different texture feature vectors extracted from the image, are pre-processed to have zero mean and standard deviation

one. For each partition,  $i = 1, \dots, N$ , the classifier is trained on the training set for each combination of hyper-parameter values and tested on its corresponding validation set. The combination that achieves the highest kappa on the validation set is selected. Then, the classifier is trained using jointly the training and validation sets, using the selected hyper-parameter values, and tested on the test set, composed by the three images (slices) of the patient that has been left out. The classifier prediction for this patient is the class that has been most voted over the three slices. This process is repeated for the  $N$  partitions to achieve the predicted class (KRAS+ or KRAS-) for each patient. Once this prediction is known, the kappa statistic of the classifier is calculated to measure its agreement with the true class labels.

#### 4.2. Results

We develop experiments using the five classifiers over the following 30 texture feature vectors: (1) Haralick's features: the eight vectors  $Hf_{df}$  varying the distance  $d = \{1, 2, 3, 123\}$  and the number of Haralick's coefficients  $f = \{4, 5\}$ ; (2) LBP features: the four vectors  $LPB_{R_i}$  where  $R_i = \{R_1, R_2, R_3, R_{123}\}$  containing uniform LBP descriptors using eight neighbours and radius of 1, 2, and 3 for  $R_1$ ,  $R_2$  and  $R_3$  and the concatenation of the previous three vectors; (3) Wavelet features: the eight vectors ( $DWT_{sfm}$ ) where  $s = \{2, 3\}$  means the levels of decomposition,  $f = \{LL, All\}$  means if the decomposition was applied to the low band or over all subbands, respectively, and  $m = \{EEMS, EMS\}$  means if the entropy is considered (EEMS) or not (EMS); (4) Wavelet + cooccurrence features: the six vectors ( $WCF_{dm}$ ) varying the distance  $d = \{1, 2, 3\}$  to calculate the cooccurrences matrices and the number of Haralick's coefficients  $m = \{4, 5\}$ ; and (5) Wavelet + cooccurrence features: the four vectors  $WDCF_{fm}$  varying the number of Haralick's coefficients  $m = \{4, 5\}$  and the type of decomposition  $f = \{LL, All\}$  using a distance  $d = 1$ . We developed a total of 150 experiments (30 feature vectors times 5 classifiers).

Table 2 show the highest kappa achieved by each family of texture features for all the classifiers tested. The best result (kappa = 64.7%) is achieved by the texture vectors  $WCF_{dm}$  and the classifier NNET. This set of texture vectors achieved the best performance for classifiers SVM (59.9%) and LDA (50%), while GBM achieved the lowest result (11.8%). In general, the combination of wavelets with second-order statistical characteristics like Haralick's coefficients provided the best results for all the classifiers. The classifier GBM is the best for all the statistical texture features: the local binary pattern (vectors  $LPB_{R_i}$ ), achieving a kappa = 25.5%, and the textures vectors ( $HF_{df}$ ), with kappa = 17.8%. The wavelet texture features (vector  $DWT_{sfm}$ ) achieved the best result (34.3%) for the NNET classifier. But the results provided by the combination of wavelet and cooccurrence features significantly overcame the other texture features for all the classifiers. Therefore, the texture features vectors  $WCF_{dm}$  are clearly superior to the others texture descriptors. In relation of the classifier, the NNET achieved the best results for two texture families, while GBM, SVM, and RF are only the bests in one texture feature. The LDA is clearly the worst classifier, as expected, and the other classifiers showed its ability to overcome its baseline performance.

**Table 2.** Highest kappa (in %) achieved by each family of texture descriptors for all classifiers tested (in bold letters the best result achieved by each texture vector).

Texture Features	LDA	GBM	NNET	RF	SVM
$Hf_{df}$	3.0	<b>17.8</b>	17.4	13.0	13.8
$LPB_{R_i}$	15.7	25.5	6.8	<b>34.3</b>	10.4
$DWT_{sfm}$	3.9	6.8	<b>34.3</b>	11.8	26.5
$WCF_{dm}$	50.0	<b>11.8</b>	<b>64.7</b>	30.3	59.9
$WDCF_{fm}$	21.7	18.5	28.5	32.5	<b>33.4</b>

Table 3 shows the highest kappa achieved by each family of texture features using the best configuration and classifier. The highest kappa (64.7%) was achieved by combining the wavelet decomposition and the Haralick's coefficients energy, correlation, contrast, homogeneity, and entropy

using a distance  $d = 1$  pixels and  $m = 5$  coefficients (vector  $WCF_{15}$ ). Table 4 reports the confusion matrix for the best performance in Table 3 (texture vector  $WCF_{15}$  and classifier NNET), with accuracy of 83%, sensitivity of 88.9% and specificity of 75.0%.

**Table 3.** Highest kappa (in %) achieved by each family of texture descriptors, with the parameters' values and classifier for which such kappa was achieved.

Texture Features	Configuration	Classifier	Kappa (%)
Hf <sub>df</sub>	$d = 2, f = 5$	GBM	17.8
LPB <sub>Ri</sub>	$i = 3$	RF	34.3
DWT <sub>sfm</sub>	$s = 3, f = All,$ $m = EEMS$	NNET	34.3
WCF <sub>dm</sub>	$d = 1, m = 5$	NNET	<b>64.7</b>
WDCF <sub>fm</sub>	$m = 4, f = All$	SVM	33.4

**Table 4.** Confusion matrix achieved with the texture descriptors and classifier that yielded the highest kappa (i.e., texture vector  $WCF_{15}$  and classifier NNET). The percentage is shown in parenthesis.

		Predicted Class	
		KRAS+	KRAS-
True class	KRAS+	24 (51.1%)	5 (10.6%)
	KRAS-	3 (6.4%)	15 (31.9%)

#### 4.3. Discussion

Our study demonstrated that CT based radiomics can predict the presence of the KRAS mutation in patients with colorectal cancer. In this study, we selected the subset or radiomic features based on the texture of the CT in the cancerous region, which we considered more useful for this purpose. Nonetheless, standard intensity features of the original image (i.e., mean, energy, entropy, and variance) were also considered in the feature vector DWT. In this study, the geometrical shape descriptors were discarded in favour of texture features, although their relevance may be explored in future works. There is a substantial agreement between the data obtained from an imaging technique and an underlying genetic alteration, as reflected by the highest kappa achieved (64.7%). Texture descriptors and automatic classifiers allowed to identify, only from the data obtained from a CT image of the tumour, a characteristic pattern that is different between a colorectal cancer with mutated KRAS and a colorectal cancer without such mutation, with a sensitivity of 88.9%, a specificity of 75.0% and an accuracy of 83%.

Our results are in good agreement with those from other investigators. From their publications, Yang et al. closely resembles our study, and obtained a lower sensitivity (75%) and a little higher specificity (83%), with a slightly higher number of patients included in the study (61 in the main cohort) [27]. More recently, He et al. [28] presented a method based on a Deep Learning, i.e., a ResNet network, to predict KRAS mutation in colorectal cancer. The best result they report is an AUC of 0.93, although the sensitivity and specificity (i.e., 0.59 and 1, respectively) are imbalanced. However, the patients' cohort (i.e., 117 for the training set) is not too large for training a deep learning model, which are known to need lots of training data to be adjusted [29]. Our work deals with this problem using traditional machine learning techniques (which can be trained reliably with much less samples) and handcrafted features, extracted only from the tumour area, thus avoiding pixels in the non-cancerous tissue that may distort the analysis. However, other studies consulted, such as Digumarthy et al., 2019 [30], used 36 patients, even lower than the current study.

Other similar studies with different radiological techniques also show that there is a statistically significant agreement between the radiological characteristics and the mutational state of the KRAS oncogene. Eun Oh et al. studied the applications of Magnetic Resonance Imaging (MRI) in the detection of the KRAS mutation in rectal cancer, and their results also reflect lower sensitivity (84%)

and accuracy (81.7%), but higher specificity (80%), in their case with a number of patients somewhat higher than ours (60) [31]. This study also refers to the comparison between the texture analysis by magnetic resonance and by CT. It concludes that the impact of image noise may be less in magnetic resonance imaging, and the texture analysis in CT may be artifacted due to its lower contrast resolution. However, in the process of diagnosis and staging of colorectal cancer, CT is indicated in all patients with both colon cancer and rectal cancer, while the main indication for magnetic resonance imaging is located in the staging of patients with rectal cancers; therefore, its global impact in the group of patients with colorectal cancer is less than CT. Furthermore, this study only refers to T2-weighted (T2w) MRI leaving the door open for further research including other magnetic resonance imaging sequences. In other study, Cui et al. also presented a T2w MRI-based radiomics signature for predicting KRAS mutation in rectal cancer [32]. The evaluation on an external validation dataset ( $n = 86$ ) yielded an accuracy, sensitivity and specificity lower than ours (i.e., 69.8%, 71.1%, and 68.8%, respectively). Besides, Meng et al. developed and validated radiomic models for assessing biological characteristics of rectal cancer based on multiparametric magnetic resonance imaging (MP-MRI), one of which is the KRAS-2 gene mutation status [33]. They reported an accuracy of 61.6%, a sensitivity of 58.1%, and a specificity of 64.3% on a cohort of 99 patients. This study acknowledges that, even though MP-MRI can provide more valuable data for radiomics through high-throughput extraction of quantitative image features, MRI signal is very sensitive to many parameters, such as the sequence, the strength and uniformity of the magnetic field or the normalization, and regularization methods, which may influence the classification and complicate radiomics.

Our results indicate the importance that these techniques may have the future in the medical decision making, together with the possibilities in terms of monitoring and early detection of recurrence that are reflected in the results of studies, such as Zhou et al., 2017 [34] and Bianconi et al., 2020 [35]. Therefore, it is essential to establish a common methodological criterion to obtain the best results that are possible and bring radiomics closer to the daily hospital clinic [35]. Regarding the best observed parameters, in our case, the texture features combining wavelet and Haralick's coefficients energy, correlation, contrast, homogeneity, and entropy ( $WCF_{dm}$  vector with  $d = 1$  pixels and  $m = 5$ ) are clearly superior than the other texture descriptors. Regarding to the classifier, the neural network (NNET) achieved the best results for two texture families.

As for the manual delimitation of the tumour by the radiologist, the automatic delimitation currently available for the segmentation of an organ or bone structure cannot be performed in a cancer because of the morphological heterogeneity of the tumour; this makes the radiologist's intervention in this process to be crucial. In any case, it is estimated that the differences between the segmentation performed by the radiologist and the segmentation that would be performed would not automatically be significantly different. In this study, a radiologist specialized in abdominal-pelvic CT analysis, with years of experience in the technique, was in charge of delimiting the tumour. Tumour delimitation lacks additional complication than being able to see the difference in texture between the tumour lesion and the wall of the abdominal viscera. Between two radiologists skilled in the art and with years of experience, the difference between the delimitations is not significant, as we verified when contrasting several delimitations made with other specialists in the field.

Regarding the clinical implications, this technique does not aspire to replace the biopsy, but rather to combine it, both to provide valuable data for decision-making regarding other possible diagnostic processes and to increase the information available to decide on the possible alternatives for treatment, establish the prognosis, or follow up on patients with colorectal cancer [35].

As we have previously pointed out, it has been described that the mutation in the KRAS oncogene must be confirmed in patients with colorectal cancer, since its presence determines a worse response to anti-EGFR treatments. The radiological images prior to any treatment on the tumour can be of great help in the treatment of patients with colorectal cancer, since from the first moment of diagnosis they can predict the presence or absence of the mutation, and therefore they can focus treatment from the beginning, without the need to wait for a confirmatory biopsy or pathological analysis of the

surgical specimen (in case the first treatment performed is resective surgery). This determination of the presence of the KRAS oncogene in a bloodless manner through the use of imaging techniques, can lead to a reduction in the time limits in the process of diagnosis and treatment of patients with colorectal cancer.

In addition, as pointed out in this study, the analysis using imaging techniques is intended to be an analysis of the entire tumour piece, which is an advantage over the analysis of a biopsy, which only covers a small part of the tumour. This is of paramount importance in patients who are not undergoing initial resective treatment and who require a histological analysis of the tumour to determine the presence or absence of the mutation.

On the other hand, the results obtained in the studies carried out are still not as categorical as the ones provided by histological methods, therefore, this technique is not yet in a position to compete with the gold standard for this indication. However, pending a development that will surely come with the frenetic advance in this field, imaging techniques can provide a valuable additional information, such as determining the areas of the tumour most likely to be biopsied for being suspicious of mutation in KRAS, since we know that this mutation is not distributed homogeneously by all the tumour cells. Another possible use of it is in the follow-up of patients already treated, trying to make a diagnosis as early as possible of a tumour recurrence or of a non-response to primary treatment.

The goal of the current paper was to present a viability analysis to test if texture features and machine learning techniques were able to predict the genetic disorder KRAS from CT images. The obtained results are encouraging and even better than similar studies [27–31]. Nonetheless, a possible limitation is the low sample size our study has been conducted with (i.e., 47 patients). Therefore, it is necessary to carry out further validation with more patients in future studies. Moreover, in future works, we will tackle the compliance in implementation of the features used in this study with the image biomarker standardisation initiative (IBSI) [36], and even a comparison with them.

**Author Contributions:** Conceptualization, M.S.-B. and E.C.; data curation, E.H., J.P. and J.R.A.; methodology, E.C., V.G.-C. and M.S.-B.; software, V.G.-C., E.C. and M.F.-D.; validation, V.G.-C., E.C. and M.F.-D.; formal analysis, M.F.-D., E.H., J.R.A. and J.P.; investigation, V.G.-C. and E.H.; resources, M.S.-B., J.R.A. and E.H.; writing—original draft preparation, E.C. and M.S.-B.; writing—review and editing, V.G.-C.; visualization, J.P.; supervision, E.C. and M.S.-B.; project administration, M.S.-B. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Acknowledgments:** This work has received financial support from the Xunta de Galicia (Centro singular de investigación de Galicia, accreditation 2020–2023) and the European Union (European Regional Development Fund—ERDF), Project MTM2016-76969-P.

**Conflicts of Interest:** The authors declare no conflict of interest.

## References

1. Observatorio del Cáncer de la AECC. *Incidencia y Mortalidad de Cáncer Colorrectal en España en la Población Entre 50 y 69 Años*; AECC: Madrid, Spain, 2018. Available online: <https://www.aecc.es/es> (accessed on 1 July 2019).
2. World Health Organization. Lyon: International Agency for Research on Cancer. Available online: <http://gco.iarc.fr> (accessed on 28 June 2019).
3. American Cancer Society. Sitio Web de la American Cancer Society Atlanta: Equipo de Redactores y Editores de la American Cancer Society. Available online: <https://www.cancer.org/es/cancer/cancer-decolon-o-recto/tratamiento/terapia-dirigida.html> (accessed on 15 February 2018).
4. Sociedad Española de Oncología Médica. Madrid: Fundación SEOM 2019. Available online: <https://seom.org> (accessed on 28 June 2020).
5. Feldman, M.; Friedman, L.; Brandt, L.J.; Sleissenger, M. *Enfermedades Gastrointestinales y Hepáticas*, 7th ed.; Luna, G.J., Ed.; Editorial Medica Panamericana: Madrid, Spain, 2004.

6. Vergara, E.; Alvis, N.; Suárez, A. ¿Existen ventajas clínicas al evaluar el estado de los genes KRAS, NRAS, BRAFF, PIK3CA, PTEN y HER2 en pacientes con cáncer colorrectal? *Rev. Colomb. Cir.* **2017**, *32*, 45–55. [[CrossRef](#)]
7. Rafiyath, S.; Braden, C.D.; Espat, N.J.; Rose, L.J. Colon Cancer Staging. Available online: <https://emedicine.medscape.com/article/2006674-overview> (accessed on 28 June 2019).
8. Ochoa-Carrillo, F.J.; de la Astudillo, V.H.; Alvarado-Cabrero, I.; Ruiz-García, E.; Torrecillas-Torres, L.; Ruiz-García, A. Cáncer colorrectal metastásico, hacia un tratamiento personalizado. *Gac. Mex. Onc.* **2014**, *13*, 39–46.
9. Valtorta, E.; Misale, S.; Sartore-Bianche, A.; Nagtegaal, I.D.; Paraf, F.; Lauricella, C.; Dimartino, V.; Hobor, S.; Jacobs, B.; Ercolani, C.; et al. KRAS gene amplification in colorectal cancer and impact on response to EGFR-targeted therapy. *Int. J. Cancer* **2013**, *133*, 1259–1265. [[CrossRef](#)]
10. Gillies, R.J.; Kinahan, P.E.; Hricak, H. Radiomics: Images are more than pictures, they are data. *Radiology* **2016**, *278*, 563–577. [[CrossRef](#)]
11. Aerts, H.; Van Griethuysen, J.; Hosny, A.; Fedonov, A.; Aucoin, N.; Fillion-Robin, J.C.; Pieper, S. *Radiomics*; Harvard Medical School: Harvard, UK, 2017; Available online: <https://www.radiomics.io/index.html> (accessed on 28 June 2019).
12. Wibmer, A.; Hricak, H.; Gondo, T.; Matsumoto, K.; Veeraraghavan, H.; Fehr, D.; Zheng, J.; Goldman, D.; Moskowicz, C.; Fine, S.; et al. Haralick texture analysis of prostate MRI: Utility for differentiating non-cancerous prostate from prostate cancer and differentiating prostate cancers with different Gleason scores. *Eur. Radiol.* **2015**, *25*, 2840–2850. [[CrossRef](#)]
13. Ortiz-Ramón, R.; Hernández, M.C.V.; González-Castro, V.; Makin, A.; Armitage, P.; Aribisala, B.; Bastin, M.; Deary, I.; Wardlaw, J.; Moratal, D. Identification of the presence of ischaemic stroke lesions by means of texture analysis on brain magnetic resonance images. *Comput. Med. Imaging Graph.* **2019**, *74*, 12–24. [[CrossRef](#)]
14. Yip, S.S.; Kim, J.; Coroller, T.P.; Parmar, C.; Velazquez, E.R.; Huynh, E.; Mak, R.H.; Aerts, H.J. Associations between somatic mutations and metabolic imaging phenotypes in nonsmall cell lung cancer. *J. Nuc. Med.* **2017**, *58*, 569–576. [[CrossRef](#)]
15. Huang, Y.; He, L.; Dong, D.; Yang, C.; Liang, C.; Chen, X.; Ma, Z.; Huang, X.; Yao, S.; Changhong, L.; et al. Individualized prediction of perineural invasion in colorectal cancer: Development and validation of radiomics prediction model. *Chin. J. Cancer Res.* **2018**, *30*, 40–50. [[CrossRef](#)] [[PubMed](#)]
16. Bodalal, Z.; Trebeschi, S.; Nguyen-Kim, T.D.L.; Schats, W.; Beets-Tan, R. Radiogenomics: Bridging imaging and genomics. *Abdom. Radiol.* **2019**, *44*, 1960–1984. [[CrossRef](#)]
17. Rahmim, A.; Bak-Fredslund, K.; Ashrafinia, S.; Lu, L.; Schmidlein, C.R.; Subramaniam, M.; Morsing, A.; Keiding, S.; Horsager, J.; Munk, O.L. Prognostic modeling for patients with colorectal liver metastases incorporating FDG PET radiomic features. *Eur. J. Radiol.* **2019**, *113*, 101–109. [[CrossRef](#)] [[PubMed](#)]
18. García-Figueras, R.; Baleato, G.S.; Padhani, A.R.; Luna-Alcalá, A.; Marhuenda, A.; Vilanova, J.C.; Osorio-Vázquez, I.; Martínez-de-Alegría, A.; Gómez-Caamaño, A. Advanced imaging techniques in evaluation of colorectal cancer. *Radiographics* **2018**, *38*, 740–765. [[CrossRef](#)] [[PubMed](#)]
19. Hernández, M.C.V.; González-Castro, V.; Chappell, F.; Sakka, E.; Makin, S.; Armitage, P.; Nailon, W.; Wardlaw, J. Application of texture analysis to study small vessel disease and blood–brain barrier integrity. *Front. Neurol.* **2017**, *8*, 327. [[CrossRef](#)]
20. Cernadas, E.; Fernández-Delgado, M.; González-Rufino, E.; Carrión, P. Influence of normalization and color space to color texture classification. *Pattern Recognit.* **2017**, *61*, 120–138. [[CrossRef](#)]
21. González-Rufino, E.; Carrión, P.; Cernadas, E.; Fernández-Delgado, M.; Domínguez-Petit, R. Exhaustive comparison of colour texture features and classification methods to discriminate cells categories in histological images of fish ovary. *Pattern Recognit.* **2013**, *46*, 2391–2407. [[CrossRef](#)]
22. Ojala, T.; Piatikäinen, M.; Mäenpää, T. Multiresolution grey-scale and rotation invariant texture classification with local binary pattern. *IEEE Trans. Pattern Anal. Mach. Intell.* **2002**, *24*, 971–987. [[CrossRef](#)]
23. Haralick, R.M.; Shanmugam, K.; Dinstein, I. Textural features for image classification. *IEEE Trans. Syst. Man. Cybern.* **1973**, *3*, 610–621. [[CrossRef](#)]
24. Alegre, E.; González-Castro, V.; Alaiz-Rodríguez, R.; García-Ordás, M.T. Texture and moments-based classification of the acrosome integrity of boar spermatozoa images. *Comput. Meth. Prog. Bio.* **2012**, *108*, 873–881.

25. Fernández-Delgado, M.; Cernadas, E.; Barro, S.; Amorim, D. Do we need hundreds of classifiers to solve real world classification problems? *J. Mach. Learn. Res.* **2014**, *15*, 3133–3181.
26. McHugh, M.L. Interrater reliability: The kappa statistic. *Biochem. Medica.* **2012**, *22*, 276–282. [[CrossRef](#)]
27. Yang, L.; Dong, D.; Fang, M.; Zhu, Y.; Zang, Y.; Liu, Z.; Zhang, H.; Ying, J.; Zhao, X.; Tian, J. Can CT-based radiomics signature predict KRAS/NRAS/BRAF mutations in colorectal cancer? *Eur. Radiol.* **2018**, *28*, 2058–2067. [[CrossRef](#)]
28. He, K.; Liu, X.; Li, M.; Li, X.; Yang, H.; Zhang, H. Noninvasive KRAS mutation estimation in colorectal cancer using a deep learning method based on CT imaging. *BMC Med. Imaging* **2020**, *20*, 59. [[CrossRef](#)]
29. Rubin, M.; Stein, O.; Turko, N.; Nygate, Y.; Roitshtain, D.; Karako, L.; Barnea, I.; Giryes, R.; Shaked, N. TOP-GAN: Stain-free cancer cell classification using deep learning with a small training set. *Med. Image Anal.* **2019**, *57*, 176–185. [[CrossRef](#)] [[PubMed](#)]
30. Digumarthy, S.R.; Padole, A.M.; Rastogi, S.; Price, M.; Mooradian, J.; Sequist, L.V.; Kaldra, M.K. Predicting malignant potential of subsolid nodules: Can radiomics preempt longitudinal follow up CT? *Cancer Imaging* **2019**, *19*, 36. [[CrossRef](#)] [[PubMed](#)]
31. Oh, J.E.; Kim, M.J.; Lee, J.; Hur, B.Y.; Kim, B.; Kim, D.Y.; Baek, J.Y.; Chang, H.J.; Park, S.C.; Oh, J.H.; et al. Magnetic resonance-based texture analysis differentiating KRAS mutation status in rectal cancer. *Cancer Res. Treat.* **2020**, *52*, 51–59. [[CrossRef](#)]
32. Cui, Y.; Liu, H.; Ren, J.; Du, X.; Xin, L.; Li, D.; Yang, X.; Wang, D. Development and validation of a MRI-based radiomics signature for prediction of KRAS mutation in rectal cancer. *Eur. Radiol.* **2020**, *30*, 1948–1958. [[CrossRef](#)]
33. Meng, X.; Xia, W.; Xie, P.; Zhang, R.; Li, W.; Wang, M.; Xiong, F.; Liu, Y.; Fan, X.; Xie, Y.; et al. Preoperative radiomic signature based on multiparametric magnetic resonance imaging for noninvasive evaluation of biological characteristics in rectal cancer. *Eur. Radiol.* **2019**, *29*, 3200–3209. [[CrossRef](#)]
34. Zhou, Y.; He, L.; Huang, Y.; Chen, S.; Wu, P.; Ye, W.; Liu, Z.; Liang, C. CT-based radiomics signature: A potential biomarker for preoperative prediction of early recurrence in hepatocellular carcinoma. *Abdom. Radiol.* **2017**, *42*, 1695–1704. [[CrossRef](#)] [[PubMed](#)]
35. Bianconi, F.; Palumbo, I.; Spanu, A.; Nuvoli, S.; Fravolini, M.L.; Palumbo, B. PET/CT radiomics in lung cancer: An overview. *Appl. Sci.* **2020**, *10*, 1718. [[CrossRef](#)]
36. Zwanenburg, A.; Vallières, M.; Abdalah, M.A.; Aerts, H.J.W.L.; Andrearczyk, V.; Apte, A.; Echegaray, S. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* **2020**, *295*, 328–338. [[CrossRef](#)]

