# Radiomic combination of spatial and temporal features extracted from DCE-MRI for prostate cancer detection*

Catarina Dinis Fernandes, Massimo Mischi, *Senior Member, IEEE,* Hessel Wijkstra, *Senior Member IEEE*, Jelle O. Barentsz, Stijn W.T.P.J. Heijmink and Simona Turco, *Member, IEEE*

*Abstract*— **Multi-parametric MRI is part of the standard prostate cancer (PCa) diagnostic protocol. Recent imaging guidelines (PI-RADS v2) downgraded the value of Dynamic Contrast-Enhanced (DCE)-MRI in the diagnosis of PCa. A purely qualitative analysis of the DCE-MRI time series, as it is generally done by radiologists, might indeed overlook information on the microvascular architecture and function. In this study, we investigate the discriminative power of quantitative imaging features derived from texture and pharmacokinetic analysis of DCE-MRI. In 605 regions of interest (benign and malignant tissue) delineated in 80 patients, we found through independent cross-validation that a subset of quantitative spatial and temporal features extracted from DCE-MRI and incorporated in machine learning classifiers obtains a good diagnostic performance (AUC = 0.80-0.86) in distinguishing malignant from benign regions.**

*Clinical Relevance*— **These findings highlight the underlying potential of quantitative DCE-derived radiomic features in identifying PCa by MRI.**

## I. Introduction

Prostate cancer (PCa) is the most common cancer type in the male population in the western world [1], and the second cause of cancer-related death in men worldwide. Imaging is a crucial diagnostic element to detect and localize the disease and thus define the best treatment option. Multi-parametric Magnetic Resonance Imaging (mpMRI), comprising of anatomical (T2w) and functional images (Diffusion-Weighted Imaging, DWI and Dynamic Contrast Enhanced, DCE), is the standard diagnostic imaging modality. While DWI probes the diffusion of water molecules and provides a surrogate measure of cellular density, DCE dynamically monitors the distribution of a contrast agent within the prostate microvasculature and how it extravasates into the interstitium, thus providing an opportunity to assess microvascular perfusion and permeability. DCE-MRI is used to study cancer angiogenesis, i.e. the formation of a vascular network supporting tumor growth. In cancerous tissue, the vasculature is characterized by an irregular and inefficient structure, with higher tortuosity, permeability, and microvascular density in comparison to benign tissue [2]. On DCE-MRI, PCa is typically characterized by fast wash-in and wash-out of the contrast agent, whereas benign tissue shows a more gradual uptake and wash-out. A dynamic study of the signal enhancement in time, i.e. time-intensity curves (TIC), can in this way be used to distinguish malignant from benign tissue. The value of DCE-MRI for PCa detection has been debated [3,4], and the role of this imaging sequence in the most recent radiological guidelines (PI-RADS v2) is marginal [5]. However, traditional methods to assess DCE-MRI are limited to a qualitative assessment of the DCE time loop, which might be of challenging interpretation. On the other hand, quantitative assessment of DCE-MRI is possible and may provide a more effective way to evaluate DCE-MRI. This is frequently performed using pharmacokinetic (PK) analysis with the Tofts model (TM) [6]. An alternative method for PK analysis is dispersion MRI (MRDI), which has shown promising initial results for the PCa localization [7].

Radiomics is an imaging analysis approach in which a high number of quantitative features are extracted from medical images [8]. Radiomic analysis in the context of oncology is typically used as a method to noninvasively and quantitatively describe tumor phenotypes.

In this study, we investigate the diagnostic potential in identifying PCa using radiomic features obtained from texture analysis of three different DCE time points, in conjunction with DCE-MRI-derived PK imaging features from both TM and MRDI.

## II. Materials and Methods

### A. Data Acquisition

MRI data from patients treated at the Amsterdam University Medical Center (AMC), The Netherlands Cancer Institute (NKI), and Radboudumc (RUMC) were retrospectively collected after review-board approval. All patients had systematic biopsy-proven prostate cancer and underwent a clinical DCE-MRI examination prior to radical prostatectomy. Characteristics of the MRI examinations were institution-dependent. These are described in detail in a previously published article [9].

Histopathology analysis was performed on prostate specimens after surgical resection. Cancer areas were marked by a pathologist and, based on the pathology reports, the malignant and benign ROIs were delineated on MRI. Prostate

contours were manually drawn on each MRI slice by a technical expert.

## B. Temporal analysis

DCE-MRI data were processed using Matlab (Matlab, version 2017a, MathWorks). TICs were extracted at each pixel within the prostate contours and were converted into concentration-time curves in accordance with the study by Schabel and Parker [10]. Information regarding the used contrast agent relaxivity times and values for prostate tissue relaxation can be found in [9]. Quantitative PK analysis of DCE-MRI was performed with the TM and MRDI.

### 1) Tofts Model

Using the TM [6], it is possible to estimate parameters related to tissue microvascular permeability such as the volume transfer constant between the intravascular and extravascular spaces ($K^{trans}$), the fractional extravascular volume ($v_e$), and the flux rate constant ($k_{ep} = K^{trans} / v_e$) characterizing leakage in the extravascular extracellular space. The one-compartment TM was applied as follows:

$$C_t(t) = K^{trans} C_p(t) * e^{-\frac{K^{trans}}{v_e}t}, \qquad (1)$$

where $C_t(t)$ is the tissue concentration at time $t$, $v_e$ is the extravascular volume fraction, and $C_p(t)$ is the blood plasma concentration at the capillary-tissue exchange site, substituted by an arterial input function (AIF) measured in an ROI in the iliac artery. Details on the extraction and usage of the AIF can be found in [9].

### 2) Dispersion MRI

In MRDI, the transport of the contrast agent within the vasculature can be described by a convective dispersion model [11]. A local solution of this model, the modified local density random walk, provides the intra-vascular contrast concentration, $C_i(t)$. As described in [9], this model circumvents the need for AIF estimation, and a reduced dispersion model can be used to calculate the pharmacokinetic parameters:

$$C_i(t) = \alpha * K^{trans} \sqrt{\frac{k_d}{2\pi (t-t_0)}} \, e^{\frac{-k_d(t-t_0-T_i)^2}{2(t-t_0)}} * e^{k_{ep}t}, \qquad (2)$$

where $T_i$ is the intravascular mean transit time between the injection and detection sites, $\alpha$ is the time integral of $C_i(t)$, $t_0$ is the theoretic injection time, and $k_d$ is the dispersion parameter. The model allows the estimation of the flux rate $k_{ep}$, a surrogate of microvascular leakage, and the dispersion parameter $k_d$, which is determined by the underlying microvascular architecture and used as a marker to assess cancer-induced vascular changes [2, 7, 12]. As a separate estimation of $K^{trans}$ is not possible, the parameter $A = \alpha*K^{trans}$ is in turn investigated.

## C. Spatial analysis

To summarize the spatial characteristics of benign and malignant tissue, the median value in the ROIs was computed for all quantitative PK maps obtained from both TM and MRDI. In addition, a radiomic analysis was used to describe the higher-order spatial characteristics of the delineated benign and malignant ROIs in different time points of the DCE series (Section *F. Feature extraction*). The extracted radiomic features encompassed global texture parameters (i.e. variance, skewness, and kurtosis), and gray-level matrix-derived characteristics (GLCM, GLRLM, GLSZM, and NGTDM) – Table I. A detailed explanation of the extracted features and their origin can be found in [13].

## F. Feature extraction

Matlab (Matlab, version 2020a, MathWorks) was used to extract imaging features. From MRDI, the median $k_d$, $k_{ep}$, A, $\mu$, and $t_0$ values were calculated for every ROI. The median $K^{trans}$, $k_{ep}$, and $v_e$ as extracted from the TM were also computed per ROI.

Per patient, three-time points in DCE TIC were defined as to capture: 1) the wash-in (DCE_wi), 2) the peak enhancement (DCE_pe) and 3) the wash-out (DCE_wo) of the contrast agent. Region-based 2D radiomic features were extracted for each DCE time point using the open-source radiomics toolbox published by Vallières et al. [13].

Voxels were analyzed at an isotropic grid of 1 mm. For the computation of global textural features, histograms with 100 bins were used for image discretization. Texture features were computed using a sliding window approach (window size=15 pixels, stride = 3), the Loyd quantization algorithm, and 64 gray-levels. A spatial convolution was applied to convert the obtained feature maps back to the images' original size. The final features provided to the classifiers consisted of the median value for the respective benign and malignant ROIs.

TABLE I. EXTRACTED IMAGING FEATURES.

| Imaging | Feature type | Feature name |
|---|---|---|
| **Tofts and MRDI** | Median value | $k_d$, $k_{ep}$, $A$, $T_i$, $t_0$, $K^{trans}$, $k_{ep}$ Tofts, $v_e$ Tofts |
| **DCE time point t (DCE_wi, DCE_pe, and DCE_wo)** | Global texture | Variance, Skewness Kurtosis |
| | GLCM matrix | Energy, Contrast, Entropy, Homogeneity, Correlation, Sum average, Variance, Dissimilarity, Auto Correlation |
| | GLRLM matrix | SRE, LRE, GLN, RLN, RP, LGRE, HGRE, SRLGE, SRHGE, LRLGE, LRHGE, GLV, RLV |
| | GLSZM matrix | SZE, LZE, GLN, ZSN, ZP, LGZE, HGZE, SZLGE, SZHGE, LZLGE, LZHGE, GLV, ZSV |
| | NGTDM matrix | Coarseness, Contrast, Busyness, Complexity, Strength |

## G. Feature selection and model definition

Python 3.8.3 with the scikit-learn machine learning toolbox was used to implement feature selection and model optimization. All features were scaled by subtracting the median and scaling the data according to the interquartile range. This method provides increased robustness against outliers and is implemented as a 'robust scaler' part of scikit-learn.

Feature selection is a crucial step to address the curse of dimensionality. Firstly, the correlation between features was computed and for pairs with a correlation higher than 0.9 one of the features was removed. In this study, we evaluate different feature selection techniques: an unsupervised method - principal component analysis (PCA), a wrapper method - recursive feature elimination (RFE), and a multivariate filter method - minimum redundancy maximum relevance algorithm (mRMR) [14]. With PCA, the optimal number of features to be selected was defined as those needed to explain 95% of the variability in the data. The selection criterion for RFE and mRMR was accuracy, as assessed with a support vector classifier (SVC) with a linear kernel.

The models under study were a logistic regression (LG), a K-Nearest Neighbors (KNN), and a support vector machine (SVM). These algorithms were chosen as they are widely used, simple to implement, and computationally efficient. The KNN and SVM classifiers are sensitive to the choice of hyperparameters, for which optimization was performed in the training set. The parameters under study were the number of neighbors, the leaf size, and the $p$ (Manhattan distance or Euclidean distance) for the KNN; and kernel type, C, and gamma (interconnected regularization parameters) values for the SVM. For the KNN the leaf size was varied between 1 and 50, the number of neighbors between 1 and 30, and p was either 1 or 2. For the SVM, the kernel type was linear, RBF, or poly, the studied C values were 0.1, 1.5, 10 and the gamma values were 1, 0.1, 0.01, 0.001, 0.0001.

## I. Statistical analysis

Independent cross-validation (CV) was implemented in a nested fashion. The external cross-validation loop consisted of a stratified shuffle with 10 splits, enforcing a data split of 80% for training and 20% for testing. The external cross-validation process was repeated 10 times. Nested 10-fold cross-validation schemes on the training data were implemented for RFE and mRMR feature selection while performing a random grid search to optimize the classifier's hyperparameters. The number of features chosen by each feature selection method in every iteration of the cross-validation scheme was recorded. With this information, it was possible to retrieve the prevalence of each choice. To assess which features were the most relevant in distinguishing benign from malignant samples, the RFE and mRMR were applied to the full dataset.

Model performance was assessed by the area under the curve (AUC). AUC values were obtained on the test dataset after feature selection, using the best model hyperparameters (which were recorded for each iteration of the CV), and were used to assess the performance of the different combinations between classifiers and feature selection methods.

## III. RESULTS

Imaging data from 80 patients, encompassing 410 benign and 195 malignant ROIs, were included in the study. For each patient, a total of 137 imaging features were available for further analysis. The training set consisted of 328 positive and 156 negative ROIs.

## A. Feature selection

From the initial 137, a total of 46 features remained after removing those with a correlation higher than 0.9. Fig.1 illustrates the prevalence of the optimal number of features for the different feature selection methods as compiled through all CV and repetitions. With PCA feature selection, the most commonly selected number of features was 9 (41%) and 10 (34%). As for RFE, the number of features oscillated from 6-46; the highest prevalence was for 14 (9%). At last, using mRMR, the number of features oscillated between 6-46, with the highest prevalence for 22 and 26 (both with 6%), and 21 and 24 (7%) features.
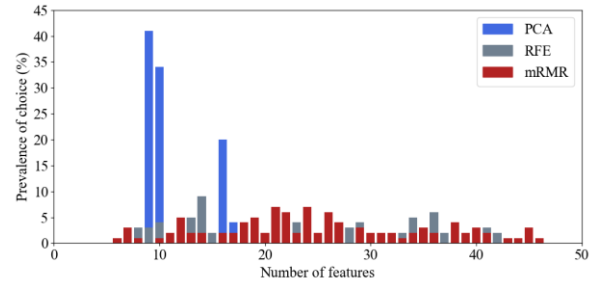


Figure 1. Prevalence of the optimal number of features as decided by the PCA, RFE, and mRMR feature selection methods.

The most commonly selected top 5 features for the RFE and mRMR methods are presented in Table II. PCA is not covered in this table as due to the algorithm's nature, the features lose their original meaning.

TABLE II.    FEATURES RANKED FOR THEIR RELEVANCE.

| Top ranking features | Feature selection methods | |
|---|---|---|
| | **RFE** *Training CV (prevalence)* | **mRMR** *Training CV (prevalence)* |
| 1 | MRDI median $k_d$ (100%) | MRDI median $k_d$ (100%) |
| 2 | TM median $v_e$ (47%) | MRDI median $A$ (84%) |
| 3 | *DCE_wi*, GLCM, Sum average (45%) | TM median $k_{ep}$ (33%) |
| 4 | TM median $k_{ep}$ (88%) | MRDI median $t_0$ (23%) |
| 5 | *DCE_wi*, GLSZM, HGZE (17%) | TM median $K^{trans}$ (22%) |

## B. Classifiers

The optimal SVM hyperparameters were an RBF kernel (93%), C = 10 (89%) and gamma = 0.01 (86%). Regarding KNN, the most prevalently chosen hyperparameters were 5 (85%) neighbors, a leaf size of 14 (85%) and a p = 2 (95%).

## C. Overall performance results

The median (range) AUC values as obtained on the test dataset for each feature selection method and classifier over all cross-validations and repetitions are reported in Table III.

Two combinations of feature selection methods and classifiers obtained the best performance with a median AUC = 0.86. These are highlighted in bold in Table III. Fig. 2 illustrates a malignancy probability map obtained using RFE feature selection and a KNN model trained on a subset of the original dataset (without this specific patient) and tested on this example case.

TABLE III. MEDIAN AUC VALUES OBTAINED FOR THE DIFFERENT COMBINATIONS OF FEATURE SELECTION AND CLASSIFIERS.

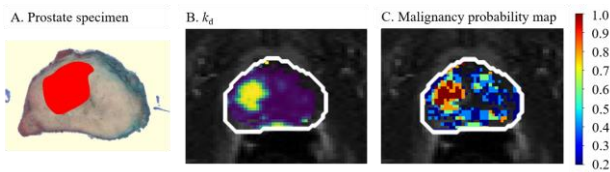| Feature selection | Classifier | AUC median (range) |
|---|---|---|
| *PCA feature selection* | KNN | 0.85 (0.79 – 0.92) |
| | SVM | 0.80 (0.69 - 0.92) |
| | LG | 0.80 (0.69 – 0.87) |
| *RFE feature selection* | KNN | **0.86 (0.79 – 0.93)** |
| | SVM | 0.82 (0.69 – 0.91) |
| | LG | 0.81 (0.71 - 0.90) |
| *mRMR feature selection* | KNN | **0.86 (0.79 – 0.93)** |
| | SVM | 0.81 (0.69 – 0.91) |
| | LG | 0.81 (0.70 – 0.89) |



Figure 2. A. Prostate specimen with histopathological tumor delineation (in red), B. $k_d$ dispersion map, and C. malignancy probability map for an example patient.

## IV. DISCUSSION

The obtained results highlight the value of quantitative DCE-MRI-derived parameters in PCa detection. With a median AUC ranging from 0.80 to 0.86, the proposed radiomic approach proved to have an excellent diagnostic potential [15]. This work combines both spatial as well as temporal information by incorporating the texture analysis of 3 timeframes from the DCE-MRI sequence, as well as the pixel-based pharmacokinetic analysis of the full concentration-time curves. The MRDI median $k_d$ dispersion feature was consistently chosen as the most relevant in distinguishing benign and malignant ROIs. Additionally, median $v_e$, $k_{ep,}$ and $K^{\text{trans}}$ as obtained through the TM, MRDI median A and $t_0$, and spatial characteristics extracted from the first DCE time point (wash-in) regularly appeared as the highest-ranking features. The variability observed in the reported AUC results highlights the fact that the available dataset is small in comparison to the number of extracted features. This is a common limitation in studies that rely on (labeled) medical data, and ours is not an exception. This AUC variability is the largest for SVM, potentially due to its higher complexity. Our feature selection methods suggest the use of a 6-30 subset of features leads to optimal performance, which fulfills a typical rule of thumb recommending a minimum of 10 outcome events per predictor variable [16]. The extension of this work encompasses evaluating the diagnostic potential of DCE-derived features in comparison and combination with T2w and ADC imaging features and improving the machine learning framework.

In the radiological PI-RADS v2 guidelines used in clinical practice for PCa diagnosis, DCE-MRI is exclusively evaluated in a qualitative manner. In this way, the rich spatial and temporal DCE dataset is reduced to a few characteristics of the TIC; as such, the wealth of information present in the DCE loops is not fully exploited. In this work, we offer evidence that DCE parameters, namely those derived from MRDI as well as spatial features from specific DCE-MRI timepoints, have potential for PCa diagnosis and should be the target of further investigation.

## V. CONCLUSIONS

Quantitative DCE-MRI-derived imaging features have diagnostic potential in PCa. This work suggests that the value of DCE-MRI, namely with the use of quantitative parameters derived from pharmacokinetic analysis, should be revisited in future prostate cancer diagnostic radiology guidelines.

REFERENCES

[1] L. A. Torre et al, "Global cancer incidence and mortality rates and trends: an update," *Cancer Epidemiol Biomarkers Prev*, vol. 25 (1), 2016, pp. 16–27.

[2] S. Turco et al, "Mathematical models of contrast transport kinetics for cancer diagnostic imaging: a review," *IEEE Rev Biomed Eng.*, vol. 9, 2016, pp. 121–47.

[3] M. de Rooij et al, "Accuracy of multiparametric MRI for prostate cancer detection: a meta-analysis," *AJR Am J Roentgenol*, vol. 202 (2), 2014, pp. 343–51.

[4] T. Auer et al, "Performance of PI-RADS version 1 versus version 2 regarding the relation with histopathological results," *World J Urol*, vol. 35 (5), 2016, pp. 1–7.

[5] J. C. Weinreb et al, "PI- RADS Prostate Imaging - Reporting and Data System: 2015, version 2," *Eur Urol,* vol.69 (1), 2016; pp.16–40.

[6] P.S. Tofts et al. "Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer: standardized quantities and symbols," *J Magn Reson Imaging*, vol.10 (3), 1999, pp. 223–32.

[7] M. Mischi et al, "Magnetic resonance dispersion imaging for localization of angiogenesis and cancer growth," *Invet Radiol,* vol.49 (8), 2014, pp. 561–69.

[8] R. J. Gilies et al, "Radiomics: Images are more than pictures, they are data," *Radiology*, vol. 278(2), 2016, pp. 563-77.

[9] S. Turco et al, "Evaluation of Dispersion MRI for Improved Prostate Cancer Diagnosis in a Multicenter Study," *AJR Am J Roentgenol*, vol. 211 (5), 2018, pp. 242-51.

[10] M. C. Schabel and D. L. Parker, "Uncertainty and bias in contrast concentration measurements using spoiled gradient echo pulse sequences," *Phys Med Bio*, vol. 53 (9), 2008, pp. 2345–73.

[11] G. Taylor, "Dispersion of soluble matter in solvent flowing slowly through a tube," *Royal Society*, 1953, vol 219 (1137).

[12] M.P. Kuenen et al, "Contrast-ultrasound diffusion imaging for localization of prostate cancer," *IEEE Trans Med Imaging,* vol. 30 (8), 2011, pp. 1493–502.

[13] M. Vallières et al, "A radiomics model from joint FDG-PET and MRI texture features for the prediction of lung metastases in soft-tissue sarcomas of the extremities," *Phys Med Bio*, vol. 60 (14), 2015, pp. 5471-96.

[14] H. Peng et al, "Feature selection based on mutual information: criteria of max-dependency, max-relevance, and min-redundancy," *IEEE Trans Pattern Anal Mach Intell*, vol. 27(8), 2005, pp.1226-38.

[15] D. W. Hosmer et al, "Applied Logistic Regression*,"*, NY: John Wiley and Sons, 2000, 2$^{nd}$ ed, ch. 5, pp. 160-4.

[16] S. J. Raudys et al, "Small sample size effects in statistical pattern recognition: recommendations for practitioners,*" IEEE Trans Pattern Analysis and Machine Intelligence*, vol. 13 (3), 1991, pp. 252–64.