Computer Methods and Programs in Biomedicine 208 (2021) 106291

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

journal homepage: www.elsevier.com/locate/cmpb

Building robust pathology image analyses with uncertainty quantification

Jeremias Gomes[a], Jun Kong[b,e,f], Tahsin Kurc[c,d], Alba C.M.A. Melo[a], Renato Ferreira[g], Joel H. Saltz[c], George Teodoro[a,c,g,*]

[a] Department of Computer Science, University of Brasília, Brasília, Brazil
[b] Biomedical Informatics Department, Emory University, Atlanta, USA
[c] Biomedical Informatics Department, Stony Brook University, Stony Brook, USA
[d] Scientific Data Group, Oak Ridge National Laboratory, Oak Ridge, USA
[e] Department of Biomedical Engineering, Emory-Georgia Institute of Technology, Atlanta, USA
[f] Department of Mathematics and Statistics, Georgia State University, Atlanta, USA
[g] Department of Computer Science, Universidade Federal de Minas Gerais, Belo Horizonte, Brazil

ARTICLE INFO

Article history:
Received 12 March 2021
Accepted 9 July 2021

Keywords:
Whole slide image analysis
Uncertainty quantification
Sensitivity analysis
Microscopy
Survival analysis

ABSTRACT

Background and Objective: Computerized pathology image analysis is an important tool in research and clinical settings, which enables quantitative tissue characterization and can assist a pathologist's evaluation. The aim of our study is to systematically quantify and minimize uncertainty in output of computer based pathology image analysis. *Methods:* Uncertainty quantification (UQ) and sensitivity analysis (SA) methods, such as Variance-Based Decomposition (VBD) and Morris One-At-a-Time (MOAT), are employed to track and quantify uncertainty in a real-world application with large Whole Slide Imaging datasets - 943 Breast Invasive Carcinoma (BRCA) and 381 Lung Squamous Cell Carcinoma (LUSC) patients. Because these studies are compute intensive, high-performance computing systems and efficient UQ/SA methods were combined to provide efficient execution. UQ/SA has been able to highlight parameters of the application that impact the results, as well as nuclear features that carry most of the uncertainty. Using this information, we built a method for selecting stable features that minimize application output uncertainty. *Results:* The results show that input parameter variations significantly impact all stages (segmentation, feature computation, and survival analysis) of the use case application. We then identified and classified features according to their robustness to parameter variation, and using the proposed features selection strategy, for instance, patient grouping stability in survival analysis has been improved from in 17% and 34% for BRCA and LUSC, respectively. *Conclusions:* This strategy created more robust analyses, demonstrating that SA and UQ are important methods that may increase confidence digital pathology.

1. Introduction

Opportunities enabled by analysis of whole slide tissue images (WSIs) in digital pathology have motivated the development of computer aided diagnosis (CAD) systems [1,2]. These systems have been employed successfully to extract cellular/sub-cellular characteristics and their spatial organizations from WSIs and correlate them with molecular and clinical data [1–9]. Despite current success of digital pathology, there still are challenges in algorithm development, high computational costs, and storage requirements to enable their routinely use. Here, we look at an important component of algorithm development and evaluation: uncertainty quantification (UQ) and parameter sensitivity analysis (SA) of digital pathology image analysis algorithms with respect to their input parameters. UQ and SA are parameter study strategies that, respectively, (i) quantify changes in an application's output as input parameters are varied and (ii) examine and assess the contributions of different parameters to uncertainty in output. They can be used in the domain to guide parameter tuning, simplify application workflow, or to reduce overall application uncertainty.

UQ and SA are well-established mechanisms in other critical application domains, such as aerospace systems and risk management [10]. These studies are also considered essential to enable and improve confidence in machine-assisted medical decision making [11–17], but they are still not widely used in digital pathology. UQ and SA studies in pathology image analysis are challenging because of the high computation costs of analysis pipelines, the large

* Corresponding author.
E-mail address: teodoro@dcc.ufmg.br (G. Teodoro).

https://doi.org/10.1016/j.cmpb.2021.106291
0169-2607/

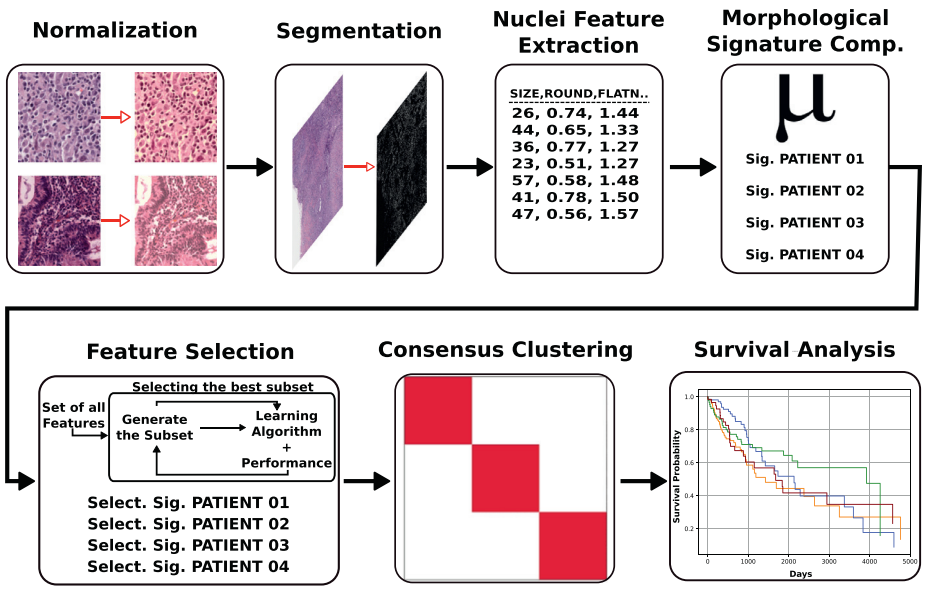


Fig. 1. Use-case Application. WSIs are processed to identify and compute nuclei features. Further, feature selection is applied and clustering is used to group patients by nuclei features (signatures), and survival analysis is carried out.

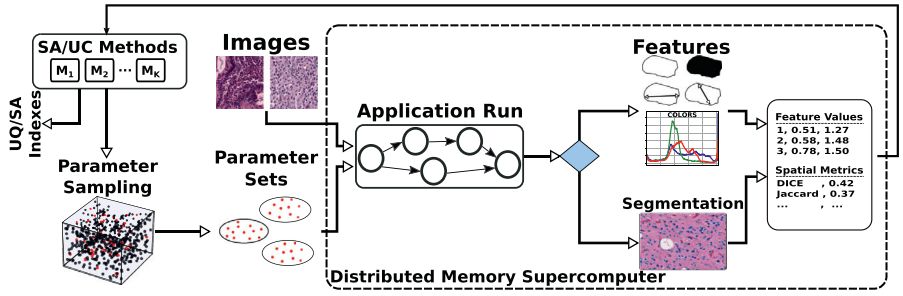


Fig. 2. Framework for performing UQ/SA studies in pathology image analysis applications.

number and size of images, and the high number of input parameters. Processing a single WSI may take hours on a workstation, and hundreds to thousands of parameters values would be evaluated depending on the UQ/SA method employed.

The use of UQ and SA in digital pathology is addressed in this work with strategies to efficiently compute UQ/SA and with a novel approach to reduce application output uncertainty based on the UQ/SA results. We use a real-world correlative survival analysis application as use-case [18–20]. This application consists of several data processing steps, including normalization, object segmentation, feature computation (or extraction), computation of morphological signatures, and survival analysis (see Fig. 1). In our work a set of UQ/SA studies are carried out by systematically varying the input parameters of the segmentation phase and measuring the impact in the other phases of the application. The segmentation phase was chosen as the variable, because it is the most parameterized stage of the analysis application and it extracts crucial information from images used in the downstream analysis stages. We employ two segmentation algorithms that differ in the strategies to refine cell nuclei boundary and separate clumped objects (Figs. S1(a) and S1(b)). The contributions of our work can be summarized as follows:

- We propose an integrated platform (Fig. 2) for execution of UQ/SA studies in digital pathology, which implements incremental multi-phase analysis to quickly prune unimportant parameters and can take advantage of supercomputers to accelerate studies.
- The platform enables more sophisticated studies, such as the quantification of uncertainty propagation between the application stages. Using object/nuclei features commonly employed in the domain, our experiments show that features are subject to different levels of uncertainty. Additionally, the sets of features that are the most or least uncertain are consistent across different algorithms and tissue types.
- We propose a novel strategy which selects features to minimize uncertainty propagation. We evaluate the feature selection strategy in a survival analysis scenario with 943 Breast Invasive Carcinoma (BRCA) and 381 Lung Squamous Cell Carcinoma (LUSC) patients/WSIs. In the experiments we measured the impact of input parameter changes to patient clustering and Kaplan-Meier survival analysis [21] with and without the proposed feature selection method. The results shown that changes in input parameters have significant impact on the survival results without our approach. As the input parameters are varied, different groupings of patients (significant or not) are computed (Table 1). When our method is used, the differences between patient groupings are small (Table 2) for both cancer types. Also, the stability of the patient groups measured by the Fowlkes-Mallows Index (FMI) [22], as application parameters are varied, has been improved on average from 0.743(LUSC) and 0.793(BRCA) when our feature selection approach is not used to 0.867(LUSC) and 0.852(BRCA) with our uncertainty feature selection (higher value indicates more stable results).
- Our work is the first study to quantify uncertainty in a full application using large dataset and, more importantly, to show

Table 1
P-value of the Logrank Test for Normalized Cut feature selection. Cancer type, % of features used, and input parameter value are varied.

BRCA												
Patient	100%			50%			25%			10%		
Groups	5	46	80	5	46	80	5	46	80	5	46	80
1-2	0.237	0.489	0.087	0.430	0.569	0.246	0.346	0.667	0.050	0.479	0.327	0.201
1-3	0.410	0.718	0.322	0.566	0.729	0.609	0.543	0.748	0.256	0.553	0.480	0.269
2-3	0.941	0.903	0.748	0.716	0.955	0.615	0.611	0.866	0.586	0.879	0.919	0.940
						LUSC						
Patient		**100%**			**50%**			**25%**			**10%**	
Groups	**5**	**46**	**80**	**5**	**46**	**80**	**5**	**46**	**80**	**5**	**46**	**80**
1-2	0.327	0.152	0.032	0.123	0.084	0.124	0.187	0.115	0.058	0.163	0.222	0.068
1-3	0.516	0.183	0.144	0.411	0.233	0.148	0.196	0.124	0.120	0.426	0.457	0.166
2-3	0.582	0.956	0.986	0.459	0.544	0.944	0.980	0.996	0.890	0.506	0.617	0.641

Table 2
P-value of the Logrank Test with uncertainty based (CV) feature selection. Cancer type, % of features used, and input parameter value (G1) are varied.

BRCA												
Patient	100%			50%			25%			10%		
Groups	5	46	80	5	46	80	5	46	80	5	46	80
1-2	0.237	0.489	0.087	0.015	0.018	0.038	0.407	0.581	0.435	0.020	0.042	0.014
1-3	0.410	0.903	0.322	0.191	0.189	0.104	0.025	0.050	0.026	0.082	0.117	0.060
2-3	0.941	0.718	0.748	0.069	0.091	0.305	0.127	0.123	0.101	0.485	0.543	0.410
						LUSC						
Patient		**100%**			**50%**			**25%**			**10%**	
Groups	**5**	**46**	**80**	**5**	**46**	**80**	**5**	**46**	**80**	**5**	**46**	**80**
1-2	0.327	0.152	0.032	0.057	0.015	0.142	0.035	0.039	0.066	0.010	0.003	0.005
1-3	0.516	0.183	0.144	0.292	0.065	0.179	0.074	0.209	0.187	0.020	0.024	0.012
2-3	0.582	0.956	0.986	0.331	0.418	0.922	0.933	0.580	0.786	0.870	0.756	0.998

how to use uncertainty quantification to reduce uncertainty of a complex application. We believe these tools and methods are important for the development of robust pathology image analysis workflows.

2. Methods

2.1. Uncertainty quantification and sensitivity analysis

Uncertainty Quantification (UQ) and Sensitivity Analysis (SA) are closely related methods for studying uncertainty or variation on the output of an application – in this work we focus on output variations due to input parameter changes. UQ and SA may be categorized as local or global. The local methods investigate the effect of small perturbations in parameters near a parameter value of interest to the output. The global methods measure the impact of changes in the whole range of parameter values [23]. In our work we investigate and use the global methods.

Uncertainty Quantification characterizes the application output (Y) with metrics, such as, mean (μ), variance (σ^2), Skewness, Kurtosis, and output prediction intervals. These metrics are computed as application parameters are varied. The parameters variation may use several sampling approaches, described bellow in this section, which are also employed in SA. *Sensitivity Analysis* computes the impact of parameters to output variation. The most used SA methods include the Morris One-At-A-Time (MOAT) [24], importance metrics as Pearson's and Spearman's correlation coefficients [25], and the Variance-based Decomposition (VBD) [26]. These methods are supported in our work and are used in combination to reduce the overall computation cost of a study. Since they are increasingly costly, MOAT executes to remove non-important parameters, before other methods are applied.

MOAT computes the variation of Y as each parameter is modified alone [27]. The importance metrics includes Pearson's and partial correlation coefficients, and Sperman's rank correlation [25]. Differently from MOAT, they measure inter-parameter correlations. The variance-based sensitivity analysis computes Sobol sensitivity indices [28] that apportion the output variation among input parameters. The main effect (S_i) of the i-th parameter (X_i) is the output variation attributed to a single parameter, whereas the total-order effect (S_i^T) includes higher-order parameter interactions. The main effect is computed as follows: $S_i = V(E(Y|X_i))/V(Y)$, where V(Y) is the model unconditional variance, $V(E(Y|X_i))$ is the variance of all factors, except for X_i. The sum of S_i should not be greater than one. If it is one, no parameter interaction exists.

Parameter sampling is crucial in UQ/SA. It should be able to sample the parameter domain effectively and avoid bias with a small number of parameter values (and respective application runs). We support a number of stochastic sampling strategies through an integration with Dakota [29]: Latin hypercube, Monte Carlo, and quasi-Monte Carlo with Halton or Hammersley sequences that have been shown to perform well in a variety of problems [30].

2.2. Use-case correlative analysis application

We use a correlative analysis application, which consists of a series of image normalization, segmentation, nuclear feature extraction, feature selection and survival analysis, as our use case. It is an example of applications that extract and use Pathology imaging features to carry out correlative analyses, such as survival analysis or identification of significant gene expressions [18,31–35].

The workflow of the correlative analysis application is shown in Fig. 1. The application receives as input WSI tissue images that are partitioned into 4K×4K tiles for parallel processing. The appli-

cation processes all of the tiles to extract cell level characteristics. It effectively processes the entire WSI and does not require or use pre-selected structures or areas of interest. The application starts by computing color normalization to remove image color variations [36] using the image presented in Fig. S2 as a reference. It is worth mentioning that recent works have proposed new methods to compute normalization [37–40]. Another study has analyzed the impact of normalization on CNN-based classification tasks [41]. In our work we use the original version of the application normalization stage published in previous works [18–20] in order to evaluate the proposed uncertainty quantification strategies. In the future, we envision to evaluate variations in other application stages, including normalization, to understand their impacts to the overall uncertainty.

After the normalization step, the segmentation step is executed to delineate boundaries of cell nuclei. Segmentation is a commonly used and studied operation in microscopy image analysis. Thus, we employed two segmentation algorithms in order to better evaluate the impact of input parameter changes on extracted features and correlative analysis. The first algorithm uses morphological operations to identify nuclei and watershed to separate the clumped objects [19] (Fig. S1(a)). The second one employs level set in the nuclei boundary refinement along with mean-shift clustering to separate clumped objects [20] (Fig. S1(b)). The parameters used by both segmentation algorithms are shown in Table S1.

In the feature extraction phase, each nucleus is described by a set of 51 features (17 quartile normalized features shown in Table S2) commonly used in the domain. They may be classified in morphometry, color intensity statistics, and moment features. The quartile normalization enables comparing different scales of feature distributions. Features of nuclei in a WSI/patient (one WSI per patient was used) are averaged to create a single feature vector called patient signature. Signatures are submitted to feature selection to remove redundant features. Multiple feature selection approaches were evaluated: Spectral Graph Theory (SPEC), Lasso, LFS-BSS (Localized Feature Selection, Based on Scattered Separability), Muti-Class Feature Selection (MCFS) [42], and Normalized Cut (NC). Our experimental evaluation with these feature selection methods shows that NC is the only strategy that leads to significant separation in survival analysis. Consequently, it is used in the feature selection steps of the use case application.

The signatures with remaining features are submitted to a consensus clustering [43] to separate patients in groups, which are analyzed for survival with a Kaplan-Meier estimator [21] defined as: $\hat{S}(t) = \prod_{t_i<t} \frac{n_i - d_i}{n_i}$. Here, d_i and n_i are, respectively, the number of deaths and patient with death risk at time t_i. The survival employed The Cancer Genome Atlas (TCGA) data with “days to death” for non-right-censored patients and “days to last follow-up” for right-censored patients. More details on the application are presented Section S1.

2.3. UQ and SA in pathology image analysis

This section describes our strategy to perform UQ/SA in pathology image analysis. The overall process is presented in Fig. 2. It starts with a scientist selecting the UQ/SA method, the parameters to be studied and their range values, and the WSIs to be used. The parameter sampling creates parameter sets to be evaluated. The application is executed and its outputs are used to compute the UQ or SA indices.

One might be interested in evaluating the stability of the output in different phases of the application or the uncertainty propagation. In our case, for instance, we may want to analyze the changes in output of the segmentation or the feature extraction. For the case of the segmentation, the output of the application is a mask with nuclei identified, but the UQ/SA methods expect a numeric value. Thus, we use a spatial engine to compute coefficients (Dice or Jaccard) that describe differences between the mask computed with each parameter set and a fixed reference mask generated using the application default parameters. When the target is to evaluate the features, the process is simpler and the actual feature value (or patient signature value for each feature) is passed directly to the UQ/SA method.

There are several challenges for UQ/SA use in practice within the pathology image analysis domain. The main problem arises from the high computational costs. In our use-case, for instance, the execution of a single parameter set may take several minutes in a single machine and one WSI. However, UQ/SA studies may use hundreds of WSIs and test thousands of parameters sets (application runs). This is addressed in our work with methodological and computational optimizations that include: (i) an incremental multi-phase analysis that to quickly prune unimportant parameters of the study and (ii) the utilization of distributed memory supercomputer machines to accelerate inevitable application runs. Further, we shown how UQ/SA can be used in practice in large-scale studies. This done by employing UQ/SA results to select features robust to parameter variations, which significantly reduce the overall application uncertainty (Section 2.3.2).

2.3.1. Incremental multi-phase parameter study

The incremental methodology intends to minimize computing resources required in a study. Our strategy uses (i) a combination of methods to reduce the number of application runs and (ii) domain specific knowledge to incrementally add application stages as unimportant parameters are pruned. The methods combination first studies parameters with MOAT to reveal unimportant ones. After that, more compute intensive methods, such as importance metrics and VBD, are used to refine the study and compute detailed indices. Being k the number of parameters, a MOAT study requires $n = r(k+1)$ application runs with r between 5 and 15, whereas VBD demands $n = r(k+2)$ executions for r in the range of 100 to 1000 [26]. Thus, the MOAT cost is negligible as compared to the VBD, while it can prune a significant number of unimportant parameters from VBD.

The incremental study consists in analyzing application stages incrementally. This also allows for early identification and pruning of unimportant parameters. For instance, in our use-case, as parameters of the segmentation are studied, their effects to uncertainty could be investigated by evaluating the survival analysis at the end of the workflow. However, this might be very expensive, because all stages of the application would have to be executed for each parameter set. Instead, with the incremental analysis, the segmentation output is first evaluated alone, and parameters that have little effect on the masks are pruned. The point here is that if the mask generated is not significantly modified by a parameter, the feature extracted should also not be affected by those unimportant parameters, and the same will propagate to other stages of the workflow. In our use-case, this incremental process follows with the analysis of the features extracted, before the entire application with the correlative survival analysis is used.

2.3.2. Using UQ to select features robust to parameter variation

Feature selection is an important step in data analysis that judges features according to their contributions to results. In most cases, not all features contribute significantly and may even add unnecessary noise. While feature selection has been shown to be relevant [18], its use has generally been restricted to the classic features selection analysis considering relevance, redundancy, etc. We propose a novel strategy which considers uncertainty of a feature in the selection process. The goal is to filter features that are more susceptible to variations (uncertainty) due to parameter value changes. The proposed strategy is unsupervised and uses as

input all features extracted in the application – features in our use case application are listed in the supplementary document in Table S2.

Our strategy employs the Coefficient of Variation (CV) computed during the features uncertainty quantification as an uncertainty metric. The strategy is called UQFC (Uncertainty Quantification based Feature Selection). Given the uncertainty expressed by CV, it will select the k most stable features to be included in the analysis. This process may be defined as follows. Given N parameter sets selected by the UQ sampling strategy $P_i = (p_{i1}, p_{i2}, \ldots, p_{id})$, where $i = 1, 2 \ldots, N$. Assuming the application output is a feature vector Y_i with s features, the CV or relative standard deviation of each feature j is defined as $CV(Y_j) = \frac{\sigma_j}{\mu_j}$. This coefficient is then used to rank features and the k ones with smallest CVs are selected. Please, notice that the σ_j and μ_j moments are outputted by the UQ methods, and other moments available could also be used. Also, our UQFC does not intend to replace existing feature selection approaches, but to complement them. Thus, it may be used to remove features that propagate uncertainty before other traditional feature selection strategies are applied.

2.3.3. Accelerating parameter studies

The studies executed in our proposed integrated systems were performed using distributed memory parallel supercomputers to speedup application runs. In our solution, whenever the application needs to be executed, we use our Region Templates (RT) framework [44]. This framework has been developed as an easy interface for deploying pathology image analysis applications on large clusters of machines. It allows for the user to express the computation as a set of stages, where each stage can be again decomposed into another workflow of fine-grain tasks. Stage instances can be schedule for execution into different nodes of the computing system. Stages communicate through read/write operations from/to a storage layer implemented by the RT. This simplifies the application development as users do not have to deal with message passing interfaces, which may be complex in some applications. This also enriches the system decision space, since RT controls the data placement and, as a consequence, can use that information to assign computing stage instances wisely and minimize data movements. For sake of parameter studies, this system may allow us to use several computing nodes transparently to accelerate our application runs.

3. Results

We used WSI images from the TCGA repository, including 943 cases for Breast Invasive Carcinoma (BRCA) and 381 cases of Lung Squamous Cell Carcinoma (LUSC). The WSI were partitioned in 4K×4K tiles for parallel computation using Region Templates. The list of the images is provided in Table S15. The clinical variables associated with the corresponding cases are summarized in Tables S13 and S14, and shown in Figs. S20 and S21. The experiments were executed on a distributed memory parallel machine where each node has two Xeon E52680 8-core Sandy Bridge processors. We executed the use-case application with the watershed and level set based segmentation methods. The parameters studied in the experiments are presented in Table S1. The parameter studies of the segmentation phase are presented in Section S2 for reference. Here, we focus on the study of the features extraction and full application with survival analysis.

3.1. Features uncertainty quantification

This section performs an uncertainty quantification of features. As such, the normalization, segmentation, and feature computation stages are executed. The watershed and level set segmentation are evaluated with 300 parameter sets selected from the range in Table S1 with a quasi-Monte Carlo with Halton sequence. The output of the study includes the average, standard deviation, Skewness, Kurtosis, and 95% confidence interval for average and standard deviation of each feature (Table S2). The raw values for these analyses are presented in Tables S3–S6.

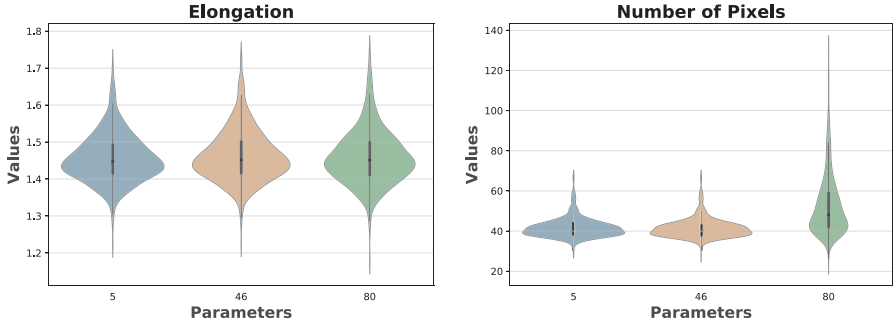
The results show that features are significantly impacted by parameters changes in both segmentation algorithms. For sake of better visualizing these variations, we employed the Coefficient of Variation (CV) that is commonly used as an uncertainty index [45]. The CV is defined as the standard deviation divided by the mean, and is presented for all features with the watershed based segmentation and BRCA in Fig. 3. Here, it is clear that the variation in features uncertainty as measured by CV is high. For instance, the number of pixels or area ($numbp\_*$) of objects is strongly affected by the parameter variations, while shape features such as elongation and flatness tend to be more stable. It is interesting that features have similar CVs regardless of the quartile in which they are measured (25%, 50% or 75%), showing that uncertainty is a characteristic of the features and not how they are probed (quartile) from their distribution.

We further analyzed the effect of uncertainty on feature distributions. This was performed by computing the features for all BRCA patients as the parameter (G1) that impacts the most the watershed segmentation (see Section S2) is varied. The selected features are elongation (`elong_q25`) and number of pixels (`numbp_q25`) that have, respectively, low and high CV values. As shown in Fig. 4, elongation remains very stable regardless of the G1 value used with very little changes to the overall distribution. The number of pixels varied significantly, and the changes were not a simple distribution shift but a shape change. These changes in `numbp_q25` were a consequence of the segmentation algorithm being impacted in its ability to separate clumped objects. These complex changes in the features space make it difficult for the remaining stages of the application to delivery the same stable and reliable results. Similar trends were observed in other features, which have their distributions presented in Figs. S10 and S11.

3.2. Uncertainty with different algorithms and tissue types

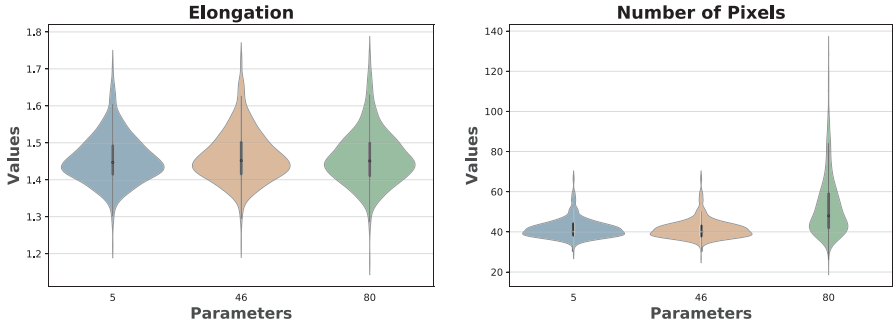
This section compares uncertainty in features as different tissue types (BRCA and LUSC) and segmentation algorithms (watershed and level set based) are used. Feature uncertainty values or CV for the two cancer types are presented in Fig. 5 with the watershed segmentation. It is possible to notice that the same set of features remains the least susceptible to parameter variation (smaller uncertainty) across tissue types. When looking at the topmost uncertain features, there is also a significant overlapping between features for both cancer types. For instance, when comparing the 10 features with smallest CV in both cases, 8 of them are found in both cancer types. This can be better visualized in Fig. S8. This result is interesting as it shows, for instance, that 4 features are ranked in all cases as the top 10 ones. Another relevant aspect are the higher overall feature CV values for BRCA as compare to LUSC. We have examined the images and noticed that the BRCA tissue had cell nuclei closer in the images. As a result, the variation of the parameters would impact more the separation of clumped nuclei in BRCA and, consequently, the size of the resulting segmented objects. An imaging showing a tissue of each type is presented in Fig. S3.

Further, Fig. 6 presents the CV for BRCA with both segmentation algorithms. As features uncertainty are compared between algorithms, it is clear that CV is higher for the level set as compared to the watershed based algorithm. The comparison of the most and least uncertain features, on the other hand, shows that both sets of features are similar between the segmentation strategies. This



(a) Elongation (elong_q25) (b) Number of Pixels (numbp_q25)

Fig. 3. Features Uncertainty using Coefficient of Variation Metric and Watershed segmentation.



(a) Elongation (elong_q25) (b) Number of Pixels (numbp_q25)

Fig. 4. Features values distribution as watershed segmentation G1 parameter is varied with the BRCA cancer images.

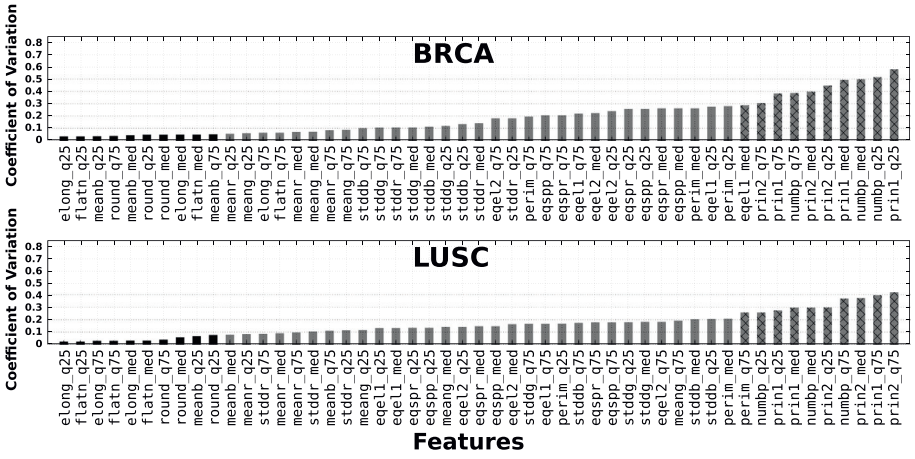


Fig. 5. Feature Uncertainty with Watershed Segmentation for BRCA and LUSC.

commonality may allow us to select a core set of features with reduced uncertainty, which in turn may lead to more stable and repeatable analyses.

3.3. Effect of feature uncertainty to correlative analysis

This section analyzes the impact of the uncertainty in the feature space on survival analysis. This is computed by separating patients into disjoint groups with a consensus clustering based on patients morphological signatures, as detailed in Section 2.2. The morphological signatures for all patients in the BRCA and LUSC cancer types were computed using the watershed and level set segmentation. This is a very compute demanding process because of the large size and number of WSIs employed. Thus, we varied only the parameter that impacts the most segmentation results for each algorithm: G1 and OTSU. Both parameters are evaluated using the default (middle value) and the minimum/maximum values as defined by the developers.

The results for the Kaplan-Meier survival estimations are presented in Fig. 7 for the watershed and LUSC, whereas the survival curves for the other combinations of segmentation algorithm and cancer types are presented the supplementary document in Figs. S12 to S19. Fig. 7 shows that survival curves for different G1 values, as the percentage of features (chosen by the normalized cut) used in the clustering phase is varied. The experiment changed the number of patient groups from 2 to 6, and the results for 3 groups (clusters) are shown as they led to overall best separation. First, for

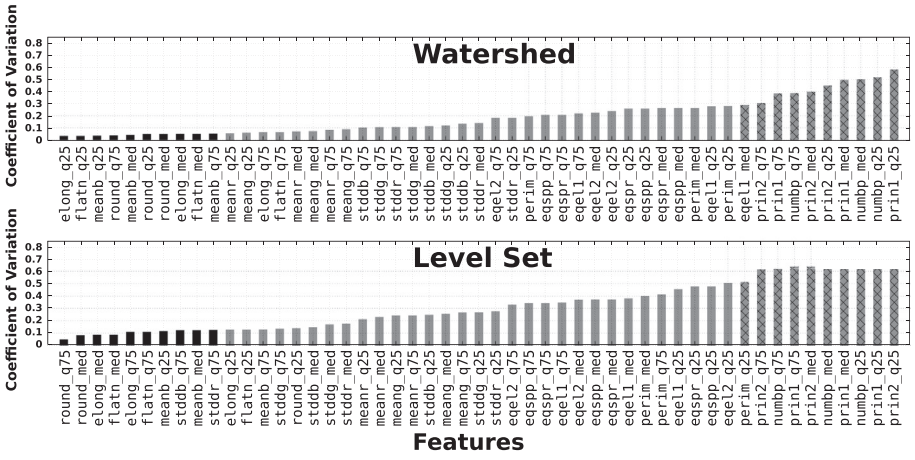


Fig. 6. Feature Uncertainty for Level-set and Watershed based segmentation with BRCA.

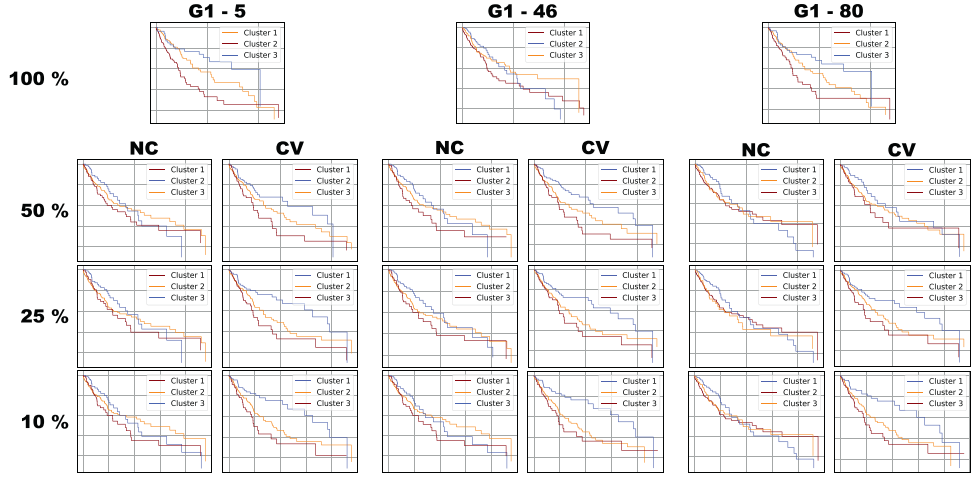


Fig. 7. Survival Analysis LUSC with Watershed algorithm using Normalized Cut (NC) and Uncertainty based (CV) feature selection methods.

each fixed percentage of features used with NC, it can be noticed that a significant variation exists in the shape of the Kaplan-Meier survival curves as G1 is changed. This highlights that uncertainty in features propagate to the application correlative results.

In order to systematically evaluate the significance of patient clusters, the logrank test was employed to compare pairs of survival curves. The results are present in Table 1 for the BRCA and LUSC cancer types with the watershed algorithm (See Table S9 for results with the level set segmentation). We considered to be significantly different pairs of patient groups in which the logrank p-value ≤ 0.05. The table cells/experiment configurations to meet this requirement are highlighted in green. As may be observed, there is one pair of curves or patient groups that are significantly different in each cancer type. However, separation significance does not hold as the parameter value and the percentage of features used vary. For instance, BRCA with 25% of the features was significantly different for a pair of patient groups, but it occurs for a specific parameter value (G1=80). Similar results are observed for the LUSC cancer type. This raises concerns about the robustness of the correlative results.

3.4. Impact of uncertainty-driven feature selection

This section evaluates the effectiveness of our feature selection to minimize the application uncertainty. For that sake, similarly to the previous section, we computed the survival analysis using watershed segmentation with BRCA and LUSC. In both cases, the input parameter that most impacts the segmentation is varied. The results with the level set segmentation are provided in the supplementary document in Table S10.

The p-values for the logrank tests of the survival estimations with our proposed features selection are presented in Table 2. Significantly different survival groups are highlighted in green. In our feature selection strategy, as the percentage of features used is reduced, the results tend to become more stable regardless of the parameter value used. For instance, for BRCA (Tables 2) with 25% of the features, the variation in the significance among parameter values is small. In the other words, the application output correlative results tend to be the same (significant or not) regardless of the parameter value used in the segmentation. For the same configuration (BRCA and 25%), the patient groups with NC shown in Table 1 may or may not be significantly different. This behavior is observed for other parameter values, percentage of features, and cancer types. The features selected by our CV and Normalized Cut are presented in Tables S7 and S8. As shown, feature selected or prioritized by the NC, such as principal moments and number of pixels, are in those identified as the most uncertain.

A direct measure of patient clustering changes with NC and our CV feature selection is presented in Table 3. The metric used is the Fowlkes-Mallows Index (FMI) [22] that combines the true positive and positive predictive rates to compare two clustering results. Higher values (close to 1) mean more similar clustering results. When 10% of the features are selected by our method, the average FMI values for LUSC and BRCA under parameter variations are, re-

Table 3
Fowlkes-Mallows Index comparing pairs of patient clustering results with the watershed algorithm as parameter value G1 is varied. Results are presented for NC and our CV approach with different % of features used.

NORMALIZED CUT								
	BRCA				**LUSC**			
	% of features				**% of features**			
Pair / %	**100**	**50**	**25**	**10**	**100**	**50**	**25**	**10**
5-46	0.964	0.944	0.947	0.964	0.649	0.849	0.918	0.954
5-80	0.761	0.700	0.682	0.713	0.551	0.728	0.738	0.643
46-80	0.751	0.701	0.680	0.704	0.643	0.658	0.681	0.633
COEFFICIENT OF VARIATION								
	BRCA				**LUSC**			
	% of features				**% of features**			
Pair / %	**100**	**50**	**25**	**10**	**100**	**50**	**25**	**10**
5-46	0.946	0.926	0.933	0.905	0.649	0.851	0.854	0.933
5-80	0.761	0.844	0.845	0.831	0.551	0.564	0.817	0.819
46-80	0.751	0.835	0.831	0.825	0.643	0.598	0.887	0.850

Table 4
Fowlkes-Mallows Index in a cross tissue analysis in which features independently selected by our CV method for LUSC are used in BRCA and vice versa. The watershed algorithm is used with the G1 parameter variation.

	BRCA				LUSC			
Param Value	% of features				% of features			
Pair (G1)	100	50	25	10	100	50	25	10
5-46	0.964	0.873	0.943	0.945	0.649	0.959	0.878	0.889
5-80	0.761	0.823	0.781	0.857	0.551	0.754	0.852	0.820
46-80	0.751	0.782	0.783	0.844	0.643	0.758	0.891	0.825

spectively, 0.867 and 0.853, while these values are 0.743 and 0.793 when our approach is not used (NC). Further, for the case with G1 varying from 5-80, our CV approach has improved patient grouping stability (FMI) in up to 17% and 27% for BRCA and LUSC. As may be observed, for both cancer types, there is a smaller change in clustering when comparing results of G1=5 to G1=46 (5-46) in all cases. However, for parameter pairs (5-80 or 46-80), using a subset of features selected with CV tends to increase the FMI values for both cancer types, whereas it does not happen with the NC feature selection. This corroborates with the survival analysis results.

We carried out another experiment to analyze the clustering stability, referred to as cross tissue analysis. In this case, features independently selected in the analysis of LUSC are used in the analysis of BRCA and vice versa. This cross-tissue analysis allows us to understand how tissue specific feature sets would perform in a different tissue and patient set. The evaluation uses the FMI metric as in the previous experiment. The results are shown in Table 4. The average FMI values for BRCA and LUSC with 10% of the features in the cross tissue analysis are, respectively, 0.882 and 0.844, which are at a similar level with the case in which features were selected and used in the same tissue type (Table 3). This is a consequence of the high similarity among features selected in both cancer types, as previously presented in Figs. 5 and S8.

3.5. Efficient execution of the parameter studies analysis

We evaluated the efficient execution of the parameter studies using our target application in a distributed memory machine. In this case, the application workflow consisting of normalization, segmentation, and feature extraction was executed in our parallel system. The other stages are not compute demanding and do not require distributed memory parallelization, since they execute on patient basis morphological signatures. Here we use the large 943 Breast Invasive Carcinoma (BRCA) WSIs dataset with the watershed based algorithm in the segmentation. The benefits for the LUSC or level set algorithm in terms of scalability are similar to those presented below.

The performance of our distributed memory execution on top of RT in a SA parameter study is presented in Fig. 8. As shown, the execution time is very high and indeed represents a limiting factor for the use of parameter studies in several cases. However, our system has been able to achieve near to linear scalability with the number of machines used, leading to a parallel efficiency of about 0.945 with 256 nodes. This computing power is a key aspect to enable the use of large datasets and with complex methods.

We also evaluated the impact of performing parameter studies in an incremental fashion. We have compared the execution times of executing a VBD using all 15 watershed segmentation parameters vs. using MOAT before VBD to prune unimportant parameters. In our use-case we have been able to reduce the number of parameters with MOAT from 15 to 8 for the watershed based segmentation. The MOAT required $n = r(k+1) = 240$ runs where $k = 15$ (number of parameters) and $r = 15$, removing 7 parameters from further analysis with the VBD. VBD requires $n = r(k+2)$ runs with a larger value of n, which was set to 200 in our case to limit computation time. Thus, the number of runs in the incremental case was reduced to 2,000 (VBD with 8 parameters) + 240 (MOAT) vs. 3,400 when VBD is executed directly with all 15 parameters, which resulted in saving of about 35% of the execution time.

4. Discussion

A growing number of projects have developed and employed pathology image analysis methods for basic, translational and clinical research [1–9,46,47]. Nevertheless, there has been relatively limited effort on systematic evaluation and quantification of sensitivity of analysis results of input parameters. Our work shows that typical imaging features computed and used in pathology image analyses are sensitive to input parameter changes. It also shows that uncertainty in computed features can propagate to downstream correlative analyses. For example, clustering of patients into

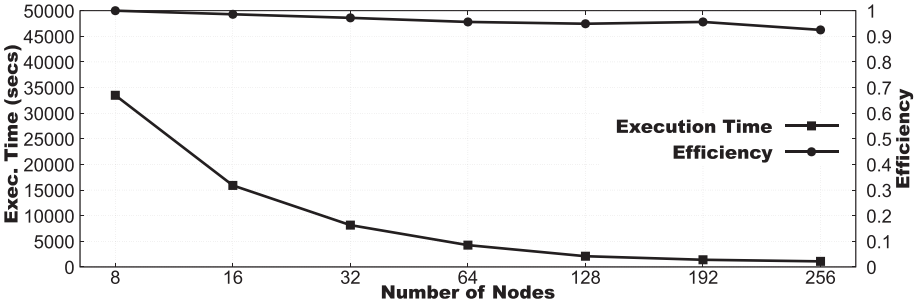


Fig. 8. Scalability of the distribute memory execution.

groups using a robust consensus clustering method may not be able to smooth variations coming from feature computation steps. As a result, survival analyses carried out with those features are strongly affected.

Parameter studies have been used in other domains and in biomedical applications [12,13,17,48–51] that are less compute demanding. Our earlier work [52] carried out sensitivity analyses and uncertainty quantification in a single stage of an analysis pipeline and focused on evaluating Surrogate Models as a tool to reduce the computation costs of these studies. In this work we have extended our earlier work and other related work with end-to-end application parameter studies, including analysis of features uncertainty. We also proposed an uncertainty driven feature selection methodology. This strategy enables the identification of features that contribute most to uncertainty propagation as well as the core set of features that are most stable for a given algorithm and tissue type. We have then demonstrated in large-scale survival analyses with 943 BRCA and 381 LUSC patients/WSIs that our approach can successfully identify robust features that minimize uncertainty in the overall analysis workflow. Another contribution of our work is that it demonstrates the use of UQ to reduce uncertainty in an end-to-end complex application.

We have shown that the uncertainty of features vary according to the segmentation workflow. However, their ranking remains stable or the most/least uncertain features are similar in both workflows (Fig. 6). The features extracted from the level set segmentation have a higher uncertainty, which is a consequence of the segmentation having a higher variation in the results as the input parameters are changed. This is presented in Fig. 9 using one of the 15 images with sizes ranging from 459 $\times$ 392 pixels to 1032 $\times$ 808 pixels manually annotated by a pathologist in our group. The variations in segmentation results for other images were similar. Each of the top 10 most uncertain features has an uncertainty value about 10$\times$ higher than those in the 10 least uncertain ones. Thus, there is a consistent set of features with far lower uncertainty, which may enable minimizing the overall uncertainty propagation.

The impact of the parameter variations to the entire application was evaluated by varying the parameter of highest influence in the segmentation (Section S2). It has shown that traditional feature selection algorithm (e.g., NC) failed in avoiding uncertainty propagation to the survival analysis. Consequently, Kaplan-Meier survival curves and the significance of the patient groups separation are strongly affected by parameter changes (Fig. 7). The use of our UQ based feature selection method (CV) is able to minimize uncertainty propagation. Conclusions drawn from the application when our method is applied are more stable. For instance as shown in Table 2, the variation on the *p*-value of the logrank test to measure significance of the curves (patient groups) separations are small when parameters are changed. Thus, regardless of the input parameter value used, the results and conclusions remain stable. We provided the full set of features ranked by uncertainty in each case in Tables S7 and S8. While we have not evaluated the methods

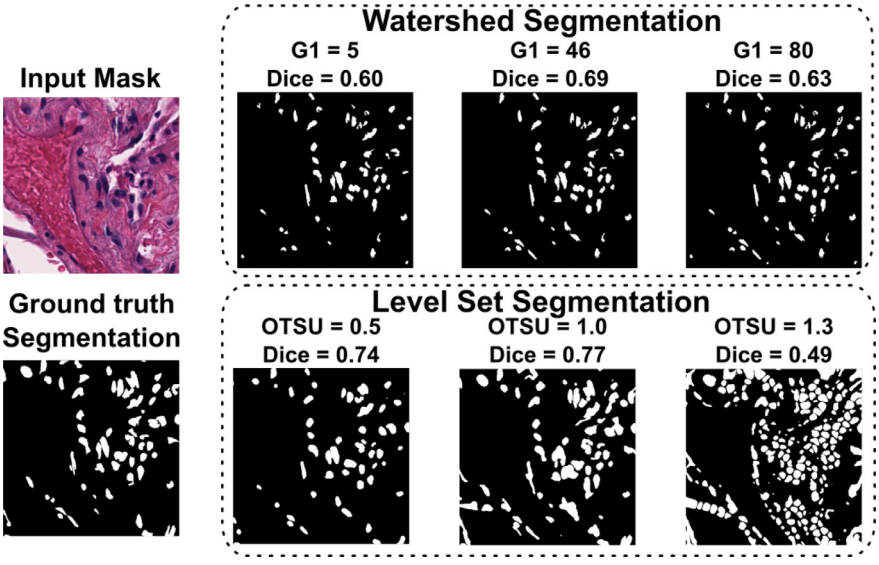


Fig. 9. Impact changing the most important parameter (identified by the UQ study) to the masks generated in each segmentation workflow.

proposed here with other applications, we believe our work can serve as a initial guide to understanding uncertainty in other related image analysis tasks. Previous works [53,54] have used Symmetric Uncertainty in feature selection. In that context, however, uncertainty is used to measure mutual information between features and target or pairs of features. Our work employs uncertainty in an orthogonal way. Here, uncertainty measures the impact of input parameter variations to changes in output, and we use that information to minimize uncertainty propagation in the application. Our method can be used as a filtering step before other feature selection strategies are used, including the symmetric uncertainty based ones.

A limitation of our work is that while we have been able to minimize our target application uncertainty with the feature selection, this method may affect differently other feature based correlative or classification tasks. Thus, we want to evaluate our methods with other applications and tasks, as well as in other cancer types and non-oncology related analysis to better understand how the proposed method will generalize. Although the features employed in this work are commonly used in the domain, and we did not intend to do an exhaustive feature study, increasing the number of features evaluated would benefit application developers. Finally, our studies have focused on hand-tuned features that are still very important in the domain, but the increasing use of deep features makes them one of our targets for future studies.

Declaration of Competing Interest

Authors declare that they have no conflict of interest.

Acknowledgment

This work was supported in part by U24CA180924, U24CA215109, 1UG3CA225021 from the NCI, R01LM011119-01 and R01LM009239 from the NLM, CNPq, and NIH K25CA181503. This research used resources of the XSEDE Science Gateways program under grant TG-ASC130023.

Supplementary material

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.cmpb.2021.106291.

References

[1] M.N. Gurcan, L. Boucheron, A. Can, A. Madabhushi, N. Rajpoot, B. Yener, Histopathological image analysis: a review, IEEE Rev. Biomed. Eng. 2 (2009) 147–171.
[2] D. Komura, S. Ishikawa, Machine learning methods for histopathological image analysis, Comput. Struct. Biotechnol. J. 16 (2018) 34–42.
[3] A.H. Beck, A.R. Sangoi, S. Leung, R.J. Marinelli, T.O. Nielsen, M.J. van de Vijver, R.B. West, M. van de Rijn, D. Koller, Systematic analysis of breast cancer morphology uncovers stromal features associated with survival, Sci. Transl. Med. 3 (108) (2011) 108–113.
[4] L.A. Cooper, D.A. Gutman, C. Chisolm, C. Appin, J. Kong, Y. Rong, T. Kurc, E.G.V. Meir, J.H. Saltz, C.S. Moreno, D.J. Brat, The tumor microenvironment strongly impacts master transcriptional regulators and gene expression class of glioblastoma, Am. J. Pathol. 180 (5) (2012) 2108–2119.
[5] K.-H. Yu, C. Zhang, G.J. Berry, R.B. Altman, C. Ré, D.L. Rubin, M. Snyder, Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features, Nat. Commun. 7 (2016) 12474.
[6] M. Veta, R. Kornegoor, A. Huisman, A.H.J. Verschuur-Maes, M.A. Viergever, J.P.W. Pluim, P.J. van Diest, Prognostic value of automatically extracted nuclear morphometric features in whole slide images of male breast cancer, Mod. Pathol. 25 (2012) 1559–1565.
[7] J.L. Carstens, P. Correa de Sampaio, D. Yang, S. Barua, H. Wang, A. Rao, J.P. Allison, V.S. LeBleu, R. Kalluri, Spatial computation of intratumoral T cells correlates with survival of patients with pancreatic cancer, Nat. Commun. 8 (2017) 13.
[8] M. Peikari, S. Salama, S. Nofech-Mozes, A.L. Martel, A cluster-then-label semi–supervised learning approach for pathology image classification, Sci. Rep. 8 (1) (2018) 1–13.
[9] M. Zhou, J. Scott, B. Chaudhury, L. Hall, D. Goldgof, K. Yeom, M. Iv, Y. Ou, J. Kalpathy-Cramer, S. Napel, R. Gillies, O. Gevaert, R. Gatenby, Radiomics in brain tumor: image assessment, quantitative feature descriptors, and machine-learning approaches, Am. J. Neuroradiol. 39 (2) (2018) 208–216.
[10] D. Bose, M.J. Wright, G.E. Palmer, Uncertainty analysis of laminar aeroheating predictions for Mars entries, J. Thermophys. Heat Transf. 20 (4) (2006) 652–662.
[11] E. Begoli, T. Bhattacharya, D. Kusnezov, The need for uncertainty quantification in machine-assisted medical decision making, Nat. Mach. Intell. 1 (1) (2019) 20–23.
[12] G. Ninos, V. Bartzis, N. Merlemis, I. Sarris, Uncertainty quantification implementations in human hemodynamic flows, Comput. Methods Programs Biomed. 203 (2021) 106021.
[13] R. Piemjaiswang, Y. Ding, Y. Feng, P. Piumsomboon, B. Chalermsinsuwan, Effect of transport parameters on atherosclerotic lesion growth: a parameter sensitivity analysis, Comput. Methods Programs Biomed. 199 (2021) 105904.
[14] M. Ghallab, Responsible AI: requirements and challenges, AI Perspect. 1 (1) (2019) 1–7.
[15] V.G. Eck, W.P. Donders, J. Sturdy, J. Feinberg, T. Delhaas, L.R. Hellevik, W. Huberts, A guide to uncertainty quantification and sensitivity analysis for cardiovascular applications, Int. J. Numer. MethodsBiomed. Eng. 32 (8) (2016) e02755.
[16] O. Eriksson, A. Jauhiainen, S. Maad Sasane, A. Kramer, A.G. Nair, C. Sartorius, J. Hellgren Kotaleski, Uncertainty quantification, propagation and characterization by Bayesian analysis combined with global sensitivity analysis applied to dynamical intracellular pathway models, Bioinformatics 35 (2) (2019) 284–292.
[17] P. Tar, N. Thacker, S. Deepaisarn, J. OConnor, A. McMahon, A reformulation of pLSA for uncertainty estimation and hypothesis testing in bio-imaging, Bioinformatics 36 (13) (2020) 4080–4087.
[18] L.A.D. Cooper, J. Kong, D.A. Gutman, F. Wang, J. Gao, C. Appin, S. Cholleti, T. Pan, A. Sharma, L. Scarpace, T. Mikkelsen, T. Kurc, C.S. Moreno, D.J. Brat, J.H. Saltz, Integrated morphologic analysis for the identification and characterization of disease subtypes, J. Am. Med. Inf.Assoc. 19 (2) (2012) 317–323.
[19] J. Kong, L.A.D. Cooper, F. Wang, J. Gao, G. Teodoro, L. Scarpace, T. Mikkelsen, M.J. Schniederjan, C.S. Moreno, J.H. Saltz, D.J. Brat, Machine-based morphologic analysis of glioblastoma using whole-slide pathology images uncovers clinically relevant molecular correlates, PLOS ONE 8 (11) (2013) 1–17.
[20] Y. Gao, V. Ratner, L. Zhu, T. Diprima, T. Kurc, A. Tannenbaum, J. Saltz, Hierarchical nucleus segmentation in digital pathology images, in: M.N. Gurcan, A. Madabhushi (Eds.), Medical Imaging 2016: Digital Pathology, vol. 9791, SPIE, International Society for Optics and Photonics, 2016, pp. 304–309.
[21] E.L. Kaplan, P. Meier, Nonparametric estimation from incomplete observations, J. Am. Stat. Assoc. 53 (1958) 457–481.
[22] E.B. Fowlkes, C.L. Mallows, A method for comparing two hierarchical clusterings, J. Am. Stat. Assoc. 78 (383) (1983) 553–569.
[23] A. Saltelli, Making best use of model evaluations to compute sensitivity indices, Comput. Phys. Commun. 145 (2) (2002) 280–297.
[24] M.D. Morris, Factorial sampling plans for preliminary computational experiments, Technometrics 33 (2) (1991) 161–174.
[25] A. Saltelli, S. Tarantola, F. Campolongo, M. Ratto, Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models, Wiley, 2004.
[26] V.G. Weirs, J.R. Kamm, L.P. Swiler, S. Tarantola, M. Ratto, B.M. Adams, W.J. Rider, M.S. Eldred, Sensitivity analysis techniques applied to a system of hyperbolic conservation laws, Reliab. Eng. Syst. Saf. 107 (2012) 157–170.
[27] F. Campolongo, J. Cariboni, A. Saltelli, An effective screening design for sensitivity analysis of large models, Environ. Modell. Softw. 22 (10) (2007) 1509–1518.
[28] I. Sobol, Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates, Math. Comput. Simul. 55 (1–3) (2001) 271–280.
[29] M.S. Eldred, K.R. Dalbey, W.J. Bohnhoff, B.M. Adams, L.P. Swiler, P.D. Hough, D.M. Gay, J.P. Eddy, K.H. Haskell, DAKOTA, A Multilevel Parallel Object-Oriented Framework for Design Optimization, Parameter Estimation, Uncertainty Quantification, and Sensitivity Analysis: Version 5.0 User's Manual, Tech. Rep. SAND2014-4633, Sandia National Laboratories, 2019.
[30] M.D. McKay, R.J. Beckman, W.J. Conover, A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, Technometrics 21 (2) (1979) 239–245.
[31] L.A. Cooper, J. Kong, D.A. Gutman, W.D. Dunn, M. Nalisnik, D.J. Brat, Novel genotype-phenotype associations in human cancers enabled by advanced molecular platforms and computational analysis of whole slide images, Lab. Invest. 95 (4) (2015) 366–376.
[32] X. Wang, A. Janowczyk, Y. Zhou, R. Thawani, P. Fu, K.A. Schalper, V. Velcheti, A. Madabhushi, Prediction of recurrence in early stage non-small cell lung cancer using computer extracted nuclear features from digital H&E images, Sci. Rep. 7 (1) (2017) 13543.
[33] M.-Y. Ji, L. Yuan, X.-D. Jiang, Z. Zeng, N. Zhan, P.-X. Huang, C. Lu, W.-G. Dong, Nuclear shape, architecture and orientation features from H&E images are able to predict recurrence in node-negative gastric adenocarcinoma, J. Transl. Med. 17 (1) (2019) 1–12.
[34] X. Luo, X. Zang, L. Yang, J. Huang, F. Liang, J. Rodriguez-Canales, I.I. Wistuba, A. Gazdar, Y. Xie, G. Xiao, Comprehensive computational pathological image analysis predicts lung cancer prognosis, J. Thorac. Oncol. 12 (3) (2017) 501–509.

[35] J. Cheng, J. Zhang, Y. Han, X. Wang, X. Ye, Y. Meng, A. Parwani, Z. Han, Q. Feng, K. Huang, Integrative analysis of histopathological images and genomic data predicts clear cell renal cell carcinoma prognosis, Cancer Res. 77 (21) (2017) e91–e100.
[36] E. Reinhard, M. Adhikhmin, B. Gooch, P. Shirley, Color transfer between images, IEEE Comput. Graph. Appl. 21 (5) (2001) 34–41.
[37] A. Vahadane, T. Peng, A. Sethi, S. Albarqouni, L. Wang, M. Baust, K. Steiger, A.M. Schlitter, I. Esposito, N. Navab, Structure-preserving color normalization and sparse stain separation for histological images, IEEE Trans. Med. Imaging 35 (8) (2016) 1962–1971.
[38] M. Salvi, N. Michielli, F. Molinari, Stain color adaptive normalization (SCAN) algorithm: separation and standardization of histological stains in digital pathology, Comput. Methods Programs Biomed. 193 (2020) 105506.
[39] S. Vijh, M. Saraswat, S. Kumar, A new complete color normalization method for H&E stained histopathological images, Appl. Intell. (2021) 1–14.
[40] T.A.A. Tosta, P.R. de Faria, J.P.S. Servato, L.A. Neves, G.F. Roberto, A.S. Martins, M.Z. do Nascimento, Unsupervised method for normalization of hematoxylin-eosin stain in histological images, Comput. Med. Imaging Graph. 77 (2019) 101646.
[41] M. Hermsen, T. de Bel, M. den Boer, E.J. Steenbergen, J. Kers, S. Florquin, J.J.T.H. Roelofs, M.D. Stegall, M.P. Alexander, B.H. Smith, B. Smeets, L.B. Hilbrands, J.A.W.M. van der Laak, Deep learning–based histopathologic assessment of kidney tissue, J. Am. Soc. Nephrol. 30 (10) (2019) 1968–1979.
[42] S. Alelyani, J. Tang, H. Liu, Feature selection for clustering: a review, in: Data Clustering, Chapman and Hall/CRC, 2018, pp. 29–60.
[43] S. Monti, P. Tamayo, J. Mesirov, T. Golub, Consensus clustering: a resampling-based method for class discovery and visualization of gene expression microarray data, Mach. Learn. 52 (1–2) (2003) 91–118.
[44] G. Teodoro, T. Pan, T. Kurc, J. Kong, L. Cooper, S. Klasky, J. Saltz, Region templates: Data representation and management for high-throughput image analysis, Parallel Comput. 40 (10) (2014) 589–610.
[45] G.H. White, Basics of estimating measurement uncertainty, Clin. Biochem. Rev. 29 (Suppl 1) (2008) S53.
[46] T.J. Fuchs, P.J. Wild, H. Moch, J.M. Buhmann, Computational pathology analysis of tissue microarrays predicts survival of renal clear cell carcinoma patients, in: Medical Image Computing and Computer-Assisted Intervention – MICCAI 2008, Berlin, Heidelberg, 2008, pp. 1–8.
[47] P. Mobadersany, S. Yousefi, M. Amgad, D.A. Gutman, J.S. Barnholtz-Sloan, J.E. Velázquez Vega, D.J. Brat, L.A.D. Cooper, Predicting cancer outcomes from histology and genomics using convolutional networks, Proc. Natl. Acad. Sci. 115 (13) (2018) E2970–E2979.
[48] C.-Y. Huang, P.-A. Nguyen, H.-C. Yang, M.M. Islam, C.-W. Liang, F.-P. Lee, Y.-C.J. Li, A probabilistic model for reducing medication errors: a sensitivity analysis using electronic health records data, Comput. Methods Programs Biomed. 170 (2019) 31–38.
[49] N. Gentilal, P.C. Miranda, Heat transfer during TTFields treatment: Influence of the uncertainty of the electric and thermal parameters on the predicted temperature distribution, Comput. Methods Programs Biomed. 196 (2020) 105706.
[50] T. Wang, F. Liang, Z. Zhou, X. Qi, Global sensitivity analysis of hepatic venous pressure gradient (HVPG) measurement with a stochastic computational model of the hepatic circulation, Comput. Biol. Med. 97 (2018) 124–136.
[51] B.M. Johnston, P.R. Johnston, Sensitivity analysis of ST-segment epicardial potentials arising from changes in ischaemic region conductivities in early and late stage ischaemia, Comput. Biol. Med. (2018) 288–299.
[52] J. Gomes, W. Barreiros Jr, T. Kurc, A.C. Melo, J. Kong, J.H. Saltz, G. Teodoro, Sensitivity analysis in digital pathology: handling large number of parameters with compute expensive workflows, Comput. Biol. Med. 108 (2019) 371–381.
[53] Y. Piao, M. Piao, K. Park, K.H. Ryu, An ensemble correlation-based gene selection algorithm for cancer classification with gene expression data, Bioinformatics 28 (24) (2012) 3306–3315.
[54] L. Yu, H. Liu, Efficient feature selection via analysis of relevance and redundancy, J. Mach. Learn. Res. 5 (2004) 1205–1224.