

Article

A Neoteric Feature Extraction Technique to Predict the Survival of Gastric Cancer Patients

Warid Islam ¹, Neman Abdoli ¹, Tasfiq E. Alam ², Meredith Jones ³, Bornface M. Mutembei ³, Feng Yan ³ and Qinggong Tang ^{3,*}

¹ School of Electrical and Computer Engineering, University of Oklahoma, Norman, OK 73019, USA; warid9666@ou.edu (W.I.); neman.abdoli@ou.edu (N.A.)

² School of Industrial and Systems Engineering, University of Oklahoma, Norman, OK 73019, USA; tasfiq@ou.edu

³ Stephenson School of Biomedical Engineering, University of Oklahoma, Norman, OK 73019, USA; meredith.jones@ou.edu (M.J.); bmutembei@ou.edu (B.M.M.); feng.yan@ou.edu (F.Y.)

* Correspondence: qtang@ou.edu

Abstract: Background: At the time of cancer diagnosis, it is crucial to accurately classify malignant gastric tumors and the possibility that patients will survive. Objective: This study aims to investigate the feasibility of identifying and applying a new feature extraction technique to predict the survival of gastric cancer patients. Methods: A retrospective dataset including the computed tomography (CT) images of 135 patients was assembled. Among them, 68 patients survived longer than three years. Several sets of radiomics features were extracted and were incorporated into a machine learning model, and their classification performance was characterized. To improve the classification performance, we further extracted another 27 texture and roughness parameters with 2484 superficial and spatial features to propose a new feature pool. This new feature set was added into the machine learning model and its performance was analyzed. To determine the best model for our experiment, Random Forest (RF) classifier, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Naïve Bayes (NB) (four of the most popular machine learning models) were utilized. The models were trained and tested using the five-fold cross-validation method. Results: Using the area under ROC curve (AUC) as an evaluation index, the model that was generated using the new feature pool yields $AUC = 0.98 \pm 0.01$, which was significantly higher than the models created using the traditional radiomics feature set ($p < 0.04$). RF classifier performed better than the other machine learning models. Conclusions: This study demonstrated that although radiomics features produced good classification performance, creating new feature sets significantly improved the model performance.

Keywords: gastric cancer; feature extraction; random forest (RF) classifier; radiomics features; support vector machine (SVM); K-Nearest Neighbors (KNN); Naïve Bayes (NB); area under ROC curve (AUC)



Citation: Islam, W.; Abdoli, N.; Alam, T.E.; Jones, M.; Mutembei, B.M.; Yan, F.; Tang, Q. A Neoteric Feature Extraction Technique to Predict the Survival of Gastric Cancer Patients. *Diagnostics* **2024**, *14*, 954. <https://doi.org/10.3390/diagnostics14090954>

Academic Editors: Najib Haboubi and Fabiano Bini

Received: 10 January 2024

Revised: 26 April 2024

Accepted: 28 April 2024

Published: 1 May 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

Gastric cancer (GC), one of the most prevalent malignant tumors, ranks as the second most common cause of cancer-related deaths worldwide [1]. Surgical procedures and subsequent chemotherapy are common therapeutic options. At the earliest possible stage of cancer diagnosis and prior to surgery, it is critical to accurately forecast the aggressiveness of malignant gastric tumors and the likelihood that patients will survive. Computed tomography (CT) is currently employed in clinical practice as the most common imaging modality to identify and diagnose gastric cancer [2]. We believe that CT scans can provide useful tumor phenotypic data related to tumor aggressiveness, which can be recognized and retrieved to assist in determining the chances of patients' overall survival (OS).

The American Joint Committee on Cancer (AJCC) staging system is a clinically used guideline for treatment allocation and prognostication on gastric cancer patients [3–5]. However, there may still be significant differences in the actual survival results among

patients with the same stage, suggesting that the current staging system is inadequate for individualized survival prediction [6]. Accurate survival prediction is difficult due to the high degree of heterogeneity in the pathophysiology of gastric cancer and the complicated etiologic and postoperative variables, but it is essential to improve patient treatment. Artificial intelligence (AI) has thus been extensively utilized in the healthcare sector recently to address this issue. AI refers to a machine's ability to replicate intelligent human behavior [7], and it is clear that a digital metamorphosis is taking place in healthcare as the use of AI grows [8,9]. AI has proven its potential of using large datasets to perform specific tasks such as image recognition. Machine learning is an application of AI which refers to an analytical technique in which models are trained on databases to generate predictions on new, unforeseen data [10]. By examining medical images, the field of radiomics, a distinct branch of AI, extracts quantitative features using data-characterization algorithms [11]. The well-established radiomics concept suggests that some radiomics features computed from radiographic images such as CT images strongly correlate with cancer prognosis, and that using computer-aided detection (CAD) schemes can more effectively extract clinically relevant radiomics features to create machine-learning models to predict cancer prognoses [12]. For instance, in our previous studies, we have developed CAD-supported machine learning models to predict the classification of breast cancer patients [13] and the risk of cancer recurrence in lung cancer patients after surgery [14].

Advanced proximal gastric cancer patients can be predicted to develop hilar lymph node metastases using radiomic characteristics in conjunction with other clinical indicators [15]. Shin et al. achieved an AUC of 0.714 by evaluating the performance of a radiomic signature-based model for predicting recurrence free survival (RFS) of locally advanced gastric cancer using preoperative contrast-enhanced CT [16]. Pham et al. uses radiomics features to predict the survival rate of gastric cancer patients and achieved an AUC of 0.72 by assuming that no additional clinical information, such as, age, gender, surgery history, and location of residence, is available, thus solely evaluating radiomic characteristics as a tool for predicting survival rate [17]. However, none of the earlier investigations compared the impact of various additional features with the radiomics features on the final classification performance. As a result, developing a new feature set may be useful for enhancing the performance of the machine learning model.

Medical images consist of many input features. It is pertinent to reduce the high dimensionality of the features such that only a small number of informative features are selected for classification [18], thereby improving the performance and reducing the complexity of the model.

The reduced feature sets can be fed into the machine learning model to predict the probability of gastric cancer patients' survival. Density-based methods are not effective for high dimensional data and usually require many training samples for a decent performance. This can be avoided by using kernels, as in Support Vector Machine (SVM)-based methods, or by artificially generated outliers [19]. One of the drawbacks of using this method is that sometimes excessive amounts of outliers are generated, which could lead to the inaccurate performance of the model. Randomized methods such as the RF classifier may be used to address this issue because they can both lower the quantity of outliers to generate and the size of the feature space in which the outliers are generated [20]. Aydadenta et al. applied machine learning techniques such as SVM, Artificial Neural Network (ANN), Naïve Bayes, k-Nearest Neighbor (kNN), and RF algorithm to image data such as colon, ovarian, central nervous, and lung to find out which algorithm works better [21]. It was found that the RF classifier outperforms SVM.

In this study, we extracted 27 texture and roughness parameters with 2484 superficial and spatial features to create a new feature pool. The unpaired t-student tests were used among the selected features and the features which are significantly different ($p < 0.05$) were selected. The selected features would then be fed in the RF classifier and its performance was further analyzed and compared with traditional radiomics features.

2. Materials and Dataset

2.1. Image Dataset

In this study, we assembled a retrospective image dataset from an existing database previously collected in our medical imaging laboratory, which has been used in our previous work [16]. A prior study provided the specifics of this image dataset, including patient demographics and clinical data as well as CT image scanning techniques and attributes [15]. In brief, this image dataset includes abdominal CT perfusion images acquired from 135 patients with pathologically confirmed gastric cancer. The patients were selected from the Fujian Medical University Union Hospital (Fuzhou, China) from January 2013 to December 2015 and January 2016 to December 2017, respectively. For each of these patients, three clinicopathological findings or markers obtained from the tumor resection specimen following surgery were also retrospectively recorded. The radiologists have determined and drawn a single region of interest (ROI) on the transverse imaging section of a tumor for each patient based on the response evaluation criteria in solid tumors (RECIST) standards currently utilized in clinical practice. This ROI shows the largest tumor size (or diameter). To pinpoint the locations of the gastric tumors, we used the radiologists' labeled CT image slices as ground truth. Other clinical diagnostic results and treatment outcomes are also recorded and collected in this dataset. We categorize these 135 patients into two groups based on the three-year survival statistics: Survival, and Non-Survival. According to patient follow-up information, 68 of these patients lived, while the other 67 did not. The dimensions of the images were 512×512 . The following steps are applied to develop and test the new CAD schemes based on a novel feature extraction technique to predict likelihood of patients' survival in this study.

2.2. Image Preprocessing

The input image was filtered with a Gaussian filter to remove the unwanted noise, and histogram equalization was performed to improve the contrast of the image [8]. In addition, the tumor was segmented based on the annotations from radiologists. We placed a segmentation seed based on the center of the tumor in one CT slice. On this CT slice, the tumor was segmented throughout the center region of the tumor from the seed location using an adaptive multi-layer region growth method [22]. A threshold value for the first layer was determined using the digital value of the local minimum point found in the region growing algorithm's seed. This allowed for a partial correction of the Hurter Driffield (H-D) curve effect in the source films. Two additional threshold values were calculated for the second and third layers from the old threshold value and the contrast value. The detail of this process is described in [22]. We used MATLAB (Version: R2021a, Natick, MA, USA) to segment the tumor regions. Similarly, this algorithm has also been used for segmentation purposes in our previous studies [23,24]. The segmented tumors were visually reviewed and manually corrected in cases where the automatic segmented didn't produce desirable results. Figure 1 shows the illustration of the segmentation process. Some more illustrations of the original images and the segmentation process are mentioned in Figures S1 and S2 of the Supplementary Materials.

2.3. Feature Extraction

The pyradiomics platform is an open-source package developed to extract radiomics features from medical images, and the features extracted by this platform are in accordance with the feature definitions as described by the Imaging Biomarker Standardization Initiative (IBSI) [25]. In this research, we used pyradiomics platform [26] to extract radiomics features and calculate 73 features. The features are described below:

2.3.1. Gray-Level Co-Occurrence Matrix (GLCM)

GLCM is a well-known statistical technique for extracting texture information from gray level images [27]. It depicts how the different shades of gray are distributed spatially and how they interact in each area. The GLCM approach can be used to locate an ex-

tremely gray-level pixel [28]. Rashed et al. illustrates the GLCM features with descriptions and equations.

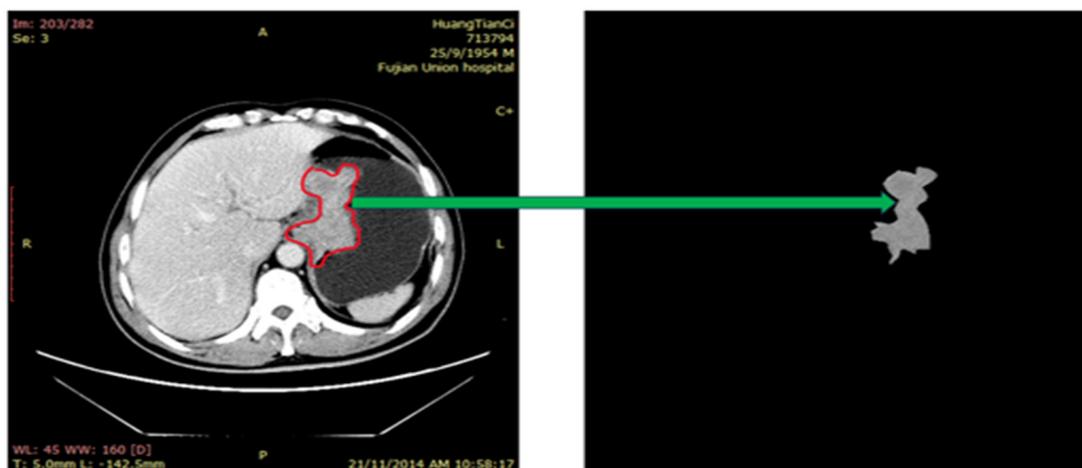


Figure 1. Example of illustrating segmented tumor regions on CT slice.

2.3.2. Gray-Level Run-Length Matrix (GLRLM)

GLRLM is a type of 2D histogram-like matrix that logs the occurrence of every possible combination of gray-level values and runs in a ROI for a specific direction. In the matrix, row and column keys are typically used to represent gray-level values and runs. The (i, j) -th item in the matrix indicates the number of pairings whose gray-level value is i and whose run length is j . Gray-level values and runs are often represented as row and column keys in the matrix [29,30]. Rashed et al. provide descriptions and equations for the GLRLM features [31].

2.3.3. Shape Features

Connected sections in binary images that are invariant to translation, rotation, and scaling are a hallmark of MI. The shapes of objects can be demonstrated using the moments. The moment-invariant features space has multiple dimensions, which invariance acknowledges. The central moments, which determine the seven shape qualities, are independent of the object's scale, orientation, or translation. Using the center of gravity, invariant moments at the origins of the central moments are translated [32]. Bharti et al. show the calculation of the moment invariance of order (p, q) [33]. In case of central invariance, the central moments of order (p, q) is calculated as shown in [34].

2.3.4. First Order Texture Features

Several first order statistical features were computed including Energy, Entropy, Interquartile Range, Kurtosis, Maximum, Mean Absolute Deviation, Mean, Minimum, Range, Robust Mean Absolute Deviation, Root Mean Square, Skewness, Total Energy, Uniformity, and Variance.

2.3.5. Gray Level Dependence Matrix (GLDM)

A Gray Level Dependence Matrix (GLDM) measures the dependencies between gray levels in an image. The detailed equation is described in this study [35].

During the feature extraction stage, we extracted a set of features from images that indicate significant information necessary for classification and diagnosis. The feature sets included texture-based features, shape-based features, and features calculated from the GLC in four directions (0, 45, 90, and 135), the GLRLM in four directions (0, 45, 90, and 135), and the GLDM, respectively. Thus, we obtained 16 features from the GLCM method, 16 features from the GLRLM method, 14 features from the GLDM method, 12 features from

the shape-based method, and 15 features from the first order texture-based features. A radiomics feature set was created using the features above.

Several other techniques like roughness [36], morphology [37], and multiscale [38] texture studies, can be utilized to extract surface and spatial properties from intensity images for pattern recognition and image classification, increasing the efficacy of structural image analysis. Characteristic distributions and patterns that are not immediately visible in the original intensity structures can be revealed using this spatial and superficial information from photos. Applications for spatial and surface feature analysis are numerous in satellite and geological image analysis, and are useful for deriving data on the distribution of regions, item identification, geological structures, and tectonics [39–41]. CT images combined with multiple texture analyses have been shown in some studies to be useful for tissue differentiation in diagnostics [42], retinal thickness classification [43], urinary bladder cancer recognition [44], ovarian tissue characterization [45], gastrointestinal tissues classification [46], oral carcinoma identification [47], and skin layer segmentation and surface quantification [48]. These benefits stem from the ability to expand image features through superficial and spatial parameters.

These additional features were extracted to find an effective feature pool. They are as follows:

- Statistical texture parameters: Histogram, Local Binary Descriptors (LBP), Histogram of Oriented Gradients (HOG), Gray Level Difference Statistics (GLDS), First-Order Statistics (FOS), Correlogram, Statistical Feature Matrices (SFM), and Gray Level Size Zone Matrix (GLSZM)
- One structural texture parameter: Shape-parameter
- One model-based texture parameter: Fractal Dimension Texture Analysis (FDTA)
- Six transform-based texture parameters: Gabor Pattern (GP), Wavelet Packets (WP), Discrete Wavelet Transform (DWT), Stroke Width Transform (SWT), Higher Order Spectra (HOS), and Laws Texture Energy (LTE)
- One model-based texture parameter: Amplitude Modulation-Frequency Modulation (AMFM)
- Two image moments texture parameters: Zero moments and Hu moments
- Two pattern spectrum and shape size texture parameters: Multilevel Binary Morphological Analysis (MultiBNA) and Gray Scale Morphological Analysis (GSMA)
- One thresholding adjacency statistics (TAS) texture parameter
- One multi-regional histogram (multiregional)
- One surface roughness parameter

The details of these parameters included: (1) Histogram (32 features), (2) Gray-Level Co-occurrence Matrix (GLCM, 28 features), (3) Local Binary Descriptors (LBP, 6 features), (4) Gray-Level Run Length Matrix (GLRLM, 10 features), (5) Histogram of Oriented Gradients (HOG, 4 features), (6) Gray Level Difference Statistics (GLDS, 5 features), (7) First-Order Statistics (FOS, 16 features), (8) Correlogram_ht (1024 features), (9) Correlogram_hd (1024 features), (10) Statistical Feature Matrices (SFM, 4 features), (11) Gray Level Size Zone Matrix (GLSZM, 14 features), (12) Shape-parameter (5 features), (13) Fractal Dimension Texture Analysis (FDTA, 6 features), (14) Gabor Pattern (GP, 4 features), (15) Wavelet Packets (WP, 126 features), (16) Discrete Wavelet Transform (DWT, 18 features), (17) Stroke Width Transform (SWT, 18 features), (18) Higher Order Spectra (HOS, 5 features), (19) Laws Texture Energy (LTE, 6 features), (20) Amplitude Modulation-Frequency Modulation (AMFM, 4 features), (21) Zero moments (25 features), (22) Hu moments (7 features), (23) Multilevel Binary Morphological Analysis (MultiBNA, 6 features), (24) Gray Scale Morphological Analysis (GSMA, 2 features), (25) thresholding adjacency statistics (TAS, 54 features), (26) multiregional (4 features), and (27) Roughness (26 features) parameters.

These additional features were used to create a new feature set. We used a total of 27 texture and roughness parameters to extract 2484 superficial and spatial features from each image. Figure 2 shows the schematic of our novel feature extraction method.

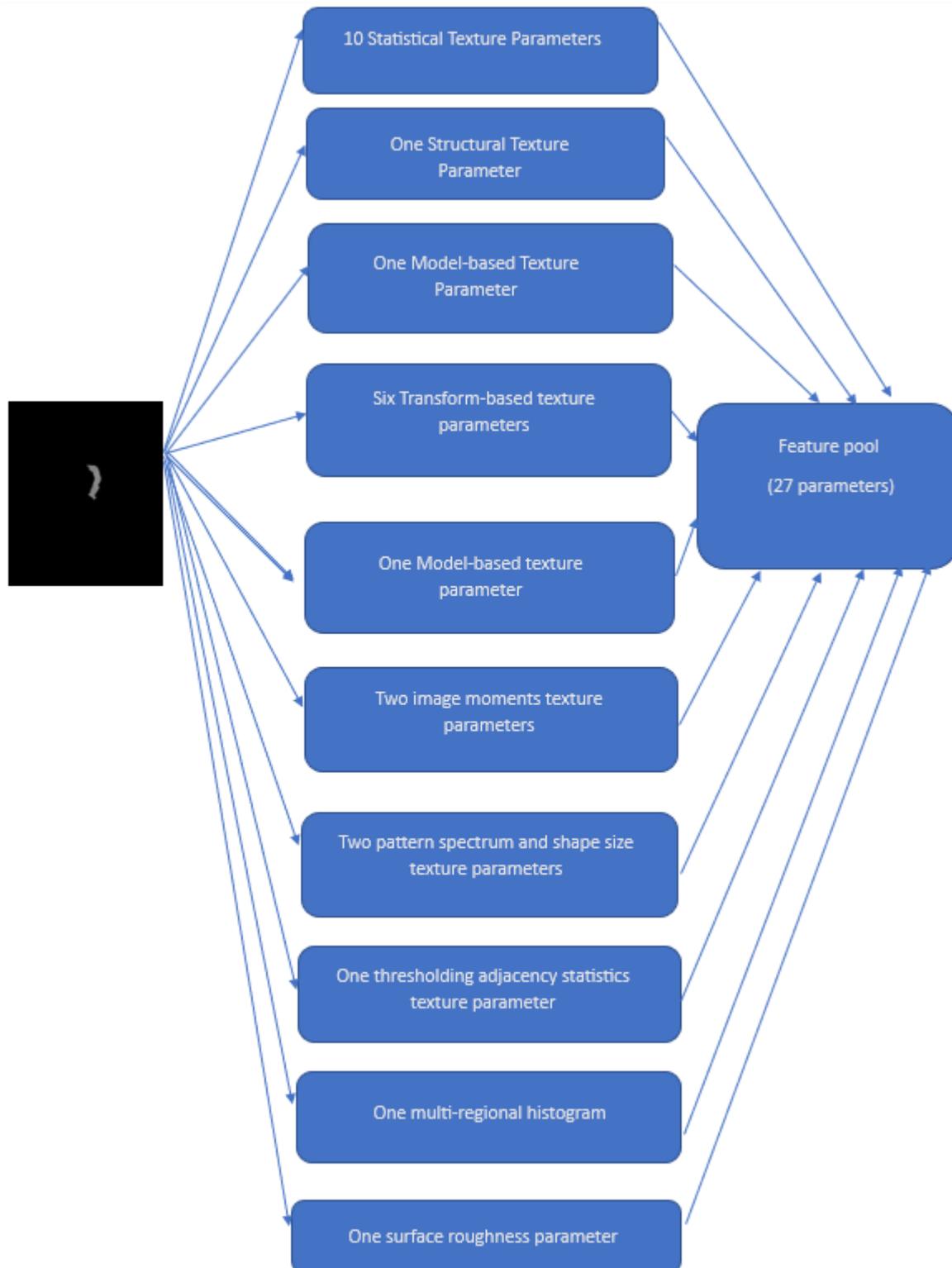


Figure 2. Flow diagram of feature extraction process.

2.4. Random Forest Classifier

The RF algorithm is an extension of the bagging method which uses feature randomness in addition to bagging to produce an uncorrelated forest of decision trees [49]. The three primary hyperparameters of random forest algorithms must be set prior to training. The number of trees, node size, and sampled feature counts are a few of them. Following that, regression or classification issues can be resolved using the random forest classifier.

Each decision tree in the ensemble that makes up the random forest method is built of a data sample taken from a training set with replacement known as the bootstrap sample. One-third of the training sample, or the out-of-bag (OOB) sample, is set aside as test data. The next step is feature bagging, which introduces yet another randomization event while increasing dataset variety and decreasing decision tree correlation. The method for making a prediction will differ depending on the type of problem. The individual decision trees will be averaged for the regression job, whereas for the classification task, the predicted class will be determined by the majority vote, or the most common categorical variable. The prediction is then finalized by cross-validation using the OOB sample. Figure 3 shows the schematic of the RF algorithm process. Some of the benefits of random forest classifier include:

- Reduced risk of overfitting, as the presence of robust numbers of decision trees reduces the overall variance and prediction error.
- Provides flexibility, as it can perform both regression and classification tasks with a high level of accuracy. The model is useful for estimating missing values due to the inclusion of feature bagging because it retains accuracy even when some of the data is missing.
- Easy to determine feature importance by evaluating the importance of a variable's contribution to the model is simple with RF.

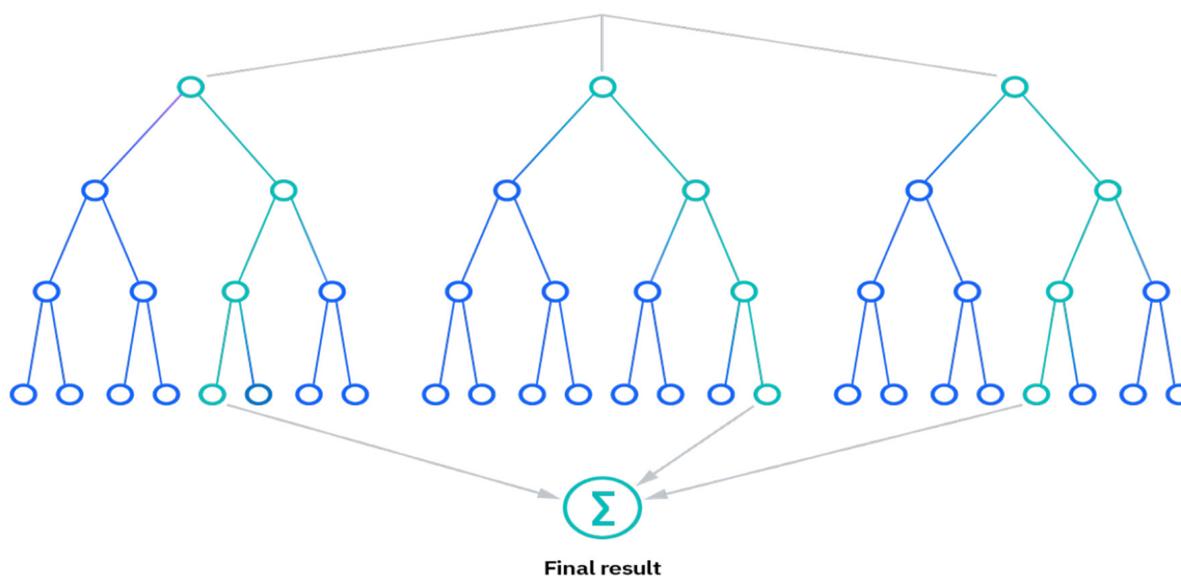


Figure 3. Illustration of RF Model.

2.5. Model Preparation

After processing images and computing features from all 135 images, we built machine learning models to predict the survival of the patients. RF classifier was used as the machine learning model. The parameters used in RF classifier are Number of Trees (`n_estimators`), Maximum Depth of Trees (`max_depth`), Minimum Samples Split (`min_samples_split`), Minimum Samples Leaf (`min_samples_leaf`), Maximum Features (`max_features`), Criterion (`criterion`), and Bootstrap Samples (`bootstrap`). We performed hyperparameter tuning separately when comparing different feature spaces. This technique involves grid search or randomized search to find the optimal combination of hyperparameters for each feature space. Second, we used the unpaired t-student tests to select the features which are significantly different ($p < 0.05$). Third, to increase the size and diversity of the training cases, as well as reduce potential bias in case partition and selection, we used the five-fold cross validation to train the classification model and evaluate its performance [50]. After training, for each testing image, the machine learning model will generate a classification

score to divide all the testing cases into two classes of Survival and Non-Survival. The score ranges from 0 to 1. From the numbers of testing cases in two classes, we computed and created confusion matrix by applying an operation threshold on the RF generated classification scores ($T = 0.5$). From the confusion matrices, the classification accuracy, sensitivity, specificity, and F1 score were computed and compared. Figure 4 shows the schematic of our model.

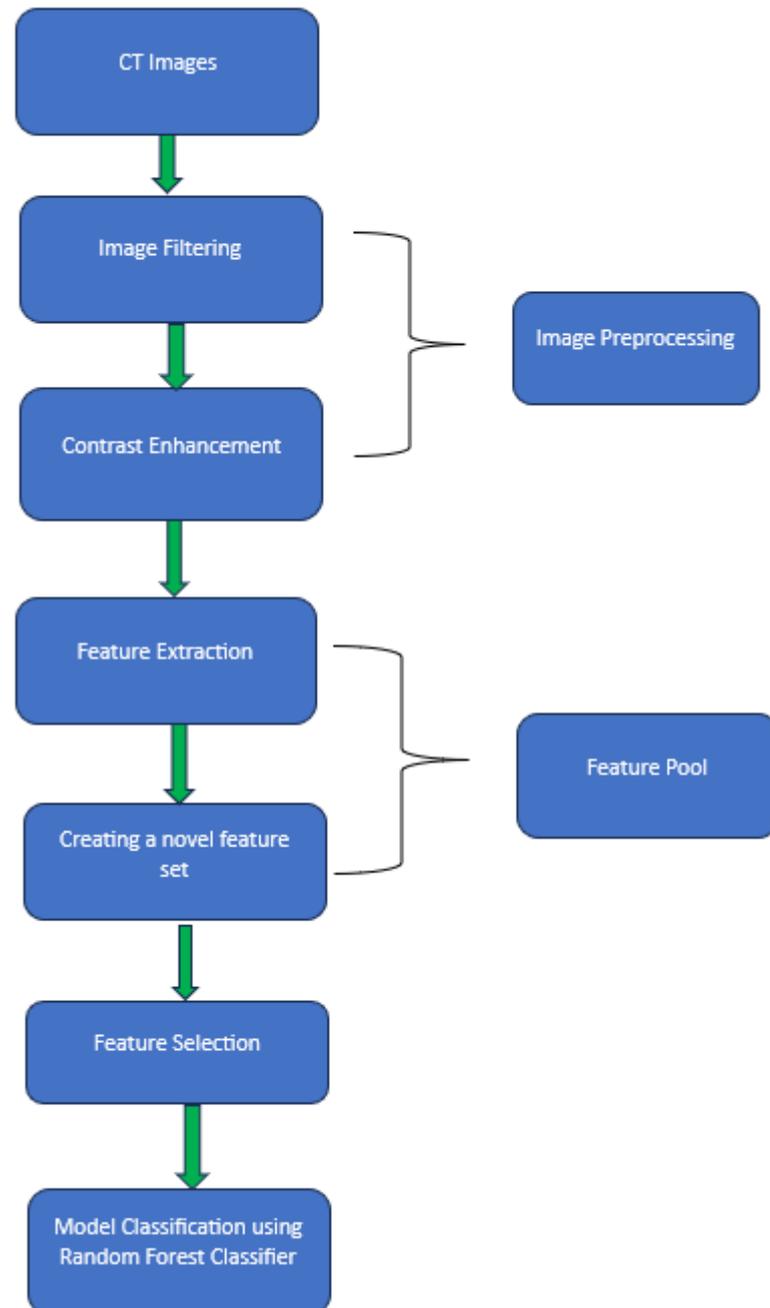


Figure 4. Schematic of machine learning model.

2.6. Performance Evaluation

We applied the following two statistical data analysis methods to evaluate and compare the gastric tumor classification performance of machine learning models based on different feature extraction techniques. First, we used a receiver operating characteristic (ROC) type data analysis method and the area under the ROC curve (AUC) as an evaluation index. Second, based on model-generated probability scores between two

output/classification nodes that predicted patient survival (Survival/Non-Survival), a testing image was assigned to one class of either Survival or Non-Survival, based on which probability score was greater between two output nodes of the model. Then, a confusion matrix was generated for each testing subset. Based on the confusion matrix, several additional evaluation indices were computed, which included the overall classification accuracy, sensitivity, specificity, and F1 score computed as follows to evaluate model performances used in a binary or practical decision environment:

$$Accuracy = (TP + TN) / (TP + TN + FP + FN)$$

$$Sensitivity = TP / (TP + FN)$$

$$Specificity = TN / (TN + FP)$$

$$F1\ Score = 2TP / (2TP + FP + FN)$$

where TP and TN represent the number of correctly classified Survival and Non-Survival cases, while FP and FN represent misclassified Survival and Non-Survival cases.

3. Results

We employed a total of 27 texture and roughness parameters to extract 2484 superficial and spatial features from each image. Within this pool of features, we applied the unpaired t-student tests among all the features. Figure 5 displays a heatmap of p -values less than 0.05 across all the features. Only the features that are significantly different ($p < 0.05$) were selected in our final model.

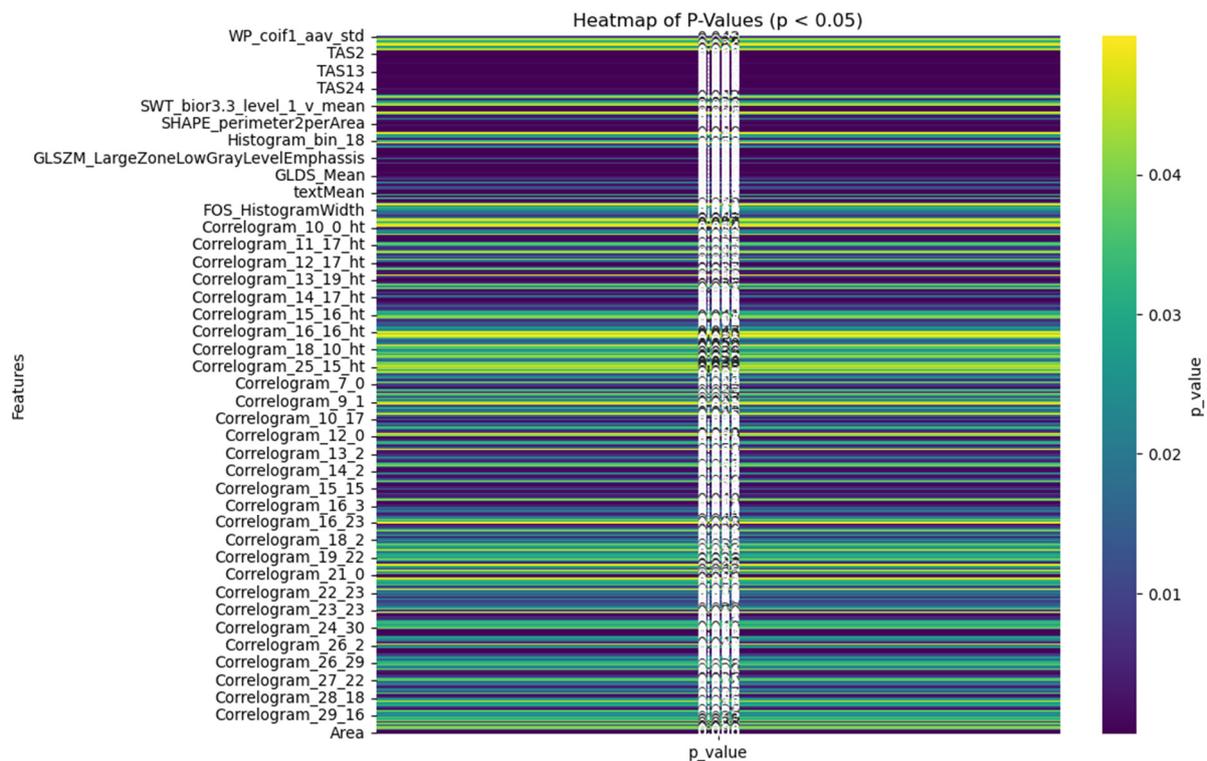


Figure 5. Heatmap of features with p -values less than 0.05.

We exhibited the difference distribution of all the selected features. We used volcano plots to assess the significance and magnitude of the differences (Figure 6) as they are particularly effective for visualization. They plot the \log_2 fold change on the x-axis (which shows the magnitude of the change) against the $-\log_{10}$ of the p -value on the y-axis (which indicates

the statistical significance of the change). This helps to illustrate both a large magnitude of change and high statistical change. These are usually the most important features.

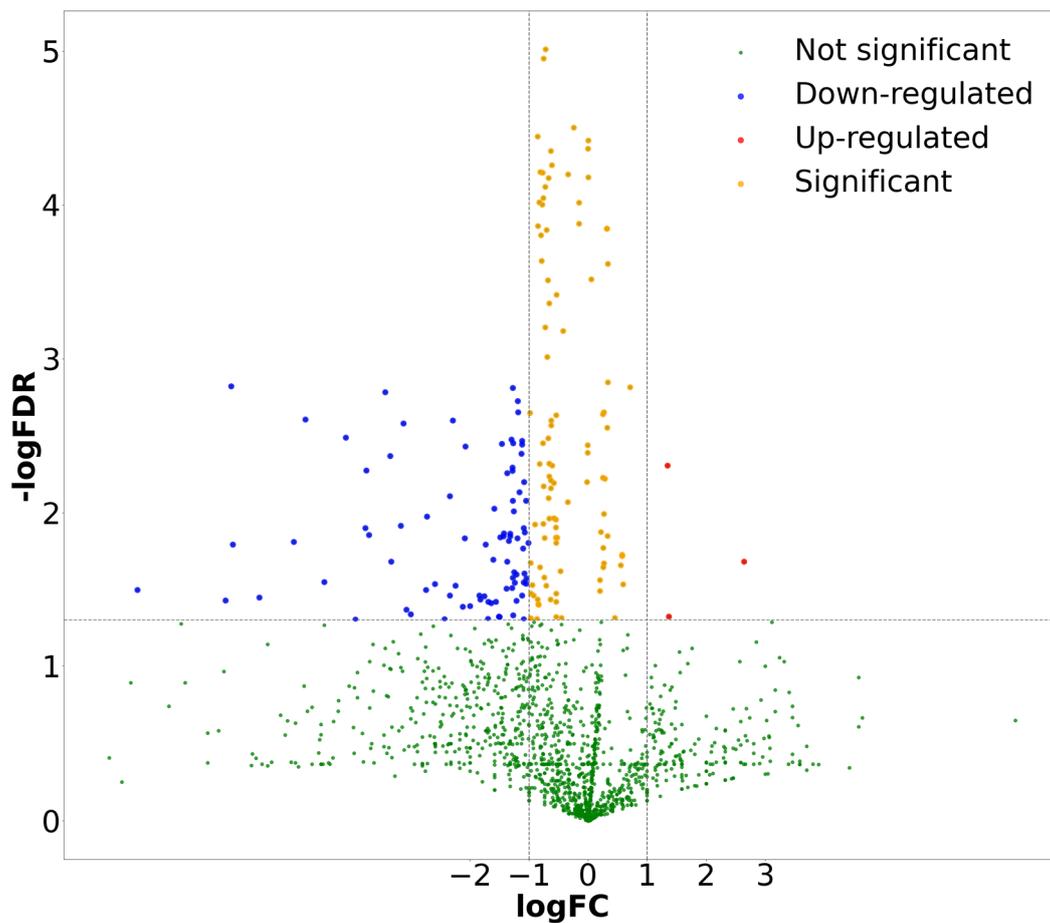


Figure 6. Volcano plot of difference among selected features.

The selected features were fed into the RF classifier and their classification performance was analyzed. Another machine learning model was built using the traditional radiomics feature set. The two models were analyzed by computing several evaluation metrics. The AUC value of the model that was built using the new feature set is 0.98 ± 0.01 which is significantly higher than the AUC value (0.72 ± 0.03) of the model built using the traditional radiomics feature set ($p < 0.05$). The results are illustrated in Figure 7.

Secondly, based on the above confusion matrices in Figure 8, the computed classification accuracy of the model built using the new feature set is 0.92, which is higher than the model built using the conventional radiomics feature set (0.64). Additionally, sensitivity, specificity, and the F-1 score of the model built using the new feature set was higher than the machine learning model built using the traditional radiomics feature set as shown in Table 1.

Table 1. Performance evaluation of the RF model using the two feature sets.

| Feature Set | AUC | Accuracy | Sensitivity | Specificity | F-1 Score |
|-------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| Radiomics | 0.72 ± 0.02 | 0.64 ± 0.04 | 0.64 ± 0.04 | 0.63 ± 0.02 | 0.64 ± 0.03 |
| New | 0.98 ± 0.01 | 0.92 ± 0.02 | 0.94 ± 0.03 | 0.90 ± 0.03 | 0.92 ± 0.04 |

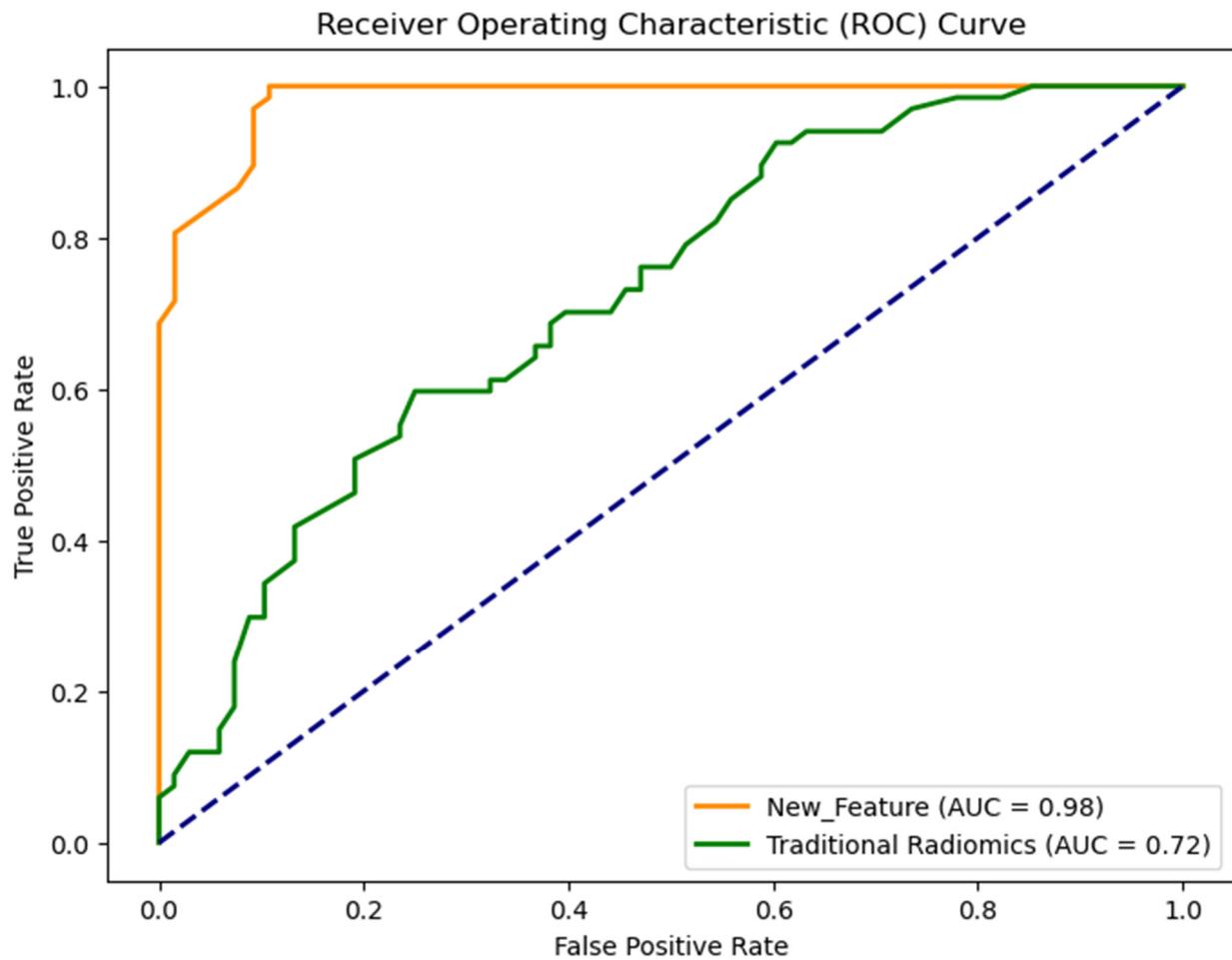


Figure 7. Receiver operating characteristic (ROC) curve of two models.

| | | Actual | |
|-----------|--------------|----------|--------------|
| | | Survival | Non-Survival |
| Predicted | Survival | 59 | 7 |
| | Non-Survival | 4 | 65 |

(a)

| | | Actual | |
|-----------|--------------|----------|--------------|
| | | Survival | Non-Survival |
| Predicted | Survival | 43 | 25 |
| | Non-Survival | 24 | 43 |

(b)

Figure 8. (a) Confusion matrix generated from model built using new feature set after five-fold cross-validation across entire dataset. (b) Confusion matrix generated from model built using traditional radiomics feature set after five-fold cross validation across entire dataset.

We added the new feature pool to the other machine learning models such as SVM, KNN, and Naïve Bayes, and discovered that the RF model performs better than the other models. Table 2 displays the results we obtained.

Table 2. Performance comparison of two machine learning models using new feature set.

| Machine Learning Model | Accuracy | AUC |
|--------------------------|----------|------|
| SVM | 0.90 | 0.93 |
| Random Forest Classifier | 0.92 | 0.98 |
| KNN | 0.91 | 0.93 |
| Naïve Bayes | 0.91 | 0.94 |

4. Discussion

Different types of machine learning technologies have been extensively used to aid in the diagnosis of gastric cancer and prediction of patients' prognosis or survival [51–57]. Feature extraction is an important part of the classification process as it enables the model to classify the images based on their distinct features. In this study, we investigated and evaluated a novel method to develop and optimize CAD schemes in the diagnosis of gastric cancer patients. The study has several unique characteristics and new observations. First, conventional CAD schemes train machine learning classifiers using radiomics feature sets. Wang et al. used similar images from 515 patients to develop a CT-based radiomics nomogram for identification of No.10 LNs status in advanced proximal gastric cancer (APGC) patients. They computed 352 image features from each segmented image. However, the performance of these models is still not impressive. In this study, we computed several other features. We tested a novel subset of features that yielded much higher classification performance in comparison to the traditional radiomics features as shown in the studies by Shin et al. and Pham et al. Texture and roughness feature extractions are widely used in remote sensing and materials to describe the intrinsic homogeneity/heterogeneity, scales, regularity, directionality, and isotropy/anisotropy. These features can extract specific structure information for image segmentation, classification, pattern recognition, and profile analysis. Texture and roughness features allow us to extract the intrinsic distribution of internal tissue structures. In the pursuit of extracting image features, various texture extraction methods have emerged, encompassing statistical, structural, model-based, and transform-based approaches within the machine vision field. These approaches focus on initial units, spatial layout, physical model designs, and transformation functions. Each method employs distinct mathematical models for extracting image features. Currently, there are 26 classic texture methods specifically developed for texture feature extraction [58]. Additionally, roughness, a parameter comprising 26 features, is employed to extract spatial information from the material surface [59]. These selected 27 texture and roughness parameters are broadly utilized in the analysis of remote sensing and material processing. Therefore, we applied these 27 parameters that can extract 2484 surficial and spatial features in total in our CT images. This indicated the importance of research effort to continuously search for novel and more effective model finetuning and optimization methods in the development of machine learning-based CAD schemes with respect to medical images. To the best of our knowledge, no similar CAD study has been reported in earlier literature.

Secondly, one of the popular machine learning models for classification is the Support Vector Machine (SVM). However, SVM, might not be efficient for all datasets. Our study showed that RF classifier performs better than SVM for our dataset. We also applied the K-Nearest Neighbors (KNN) and the Naïve Bayes model to our dataset and found out that the RF classifier outperformed all the other models.

Third, we did not have a large dataset for our experiment. If we used the traditional train-test split, we would face a smaller test set which could affect the result. To alleviate this problem, we used the five-fold cross validation technique so that we were able to make predictions on all our data. To assess the scientific rigor or robustness of our model, we repeated the experiment five times. The analysis results showed small standard deviations (i.e., <0.02), which demonstrated the robustness of the model.

Therefore, the study outcomes validate our hypothesis that using the new feature set yields better classification performance than the traditional radiomics features.

Despite the encouraging results or observations made in this study, we also recognize that the study has some limitations. We used a small dataset for our experiment ($n = 135$). To assess the experiment's viability, a bigger dataset might be used. In addition, our study results also provide a new scientific rationale and background for us to develop deep learning model-based CAD schemes further optimally [36]. Instead of using handcrafted features, we could use deep learning to extract additional features. We could fuse the traditional handcrafted and the deep learning features and create a new feature pool. This new feature subset could be used to build a new machine learning model and its performance could be analyzed. Autoencoder neural network has been recently used extensively in medical image processing [60–62]. We plan to apply these networks to our dataset to see if we can come up with a better performance. The inference time on NVIDIA RTX 4090 GPUs was 0.19 s for this entire dataset. This timing could be reduced if multiple GPUs are available, and we plan to study them in our future work. Thus, the above limitations need to be further investigated in our future studies. However, we think that this study is reliable and that it successfully tests or validates the hypothesis that using the novel feature set extracted earlier to create a new feature pool would improve the performance of the machine learning model.

5. Conclusions

In this study, we proposed and investigated the feasibility of a new feature pool based on a new subset of features (27 texture and roughness parameters). The study's findings showed that the new feature set outperformed the traditional radiomics feature set in terms of performance. The robustness of the method was also tested and demonstrated by repeating the experiment five times. This provides a solid platform to further perform this experiment on a larger dataset and combine it with deep learning to produce better results.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/diagnostics14090954/s1>, Figure S1: (a–t) Examples of patients with gastric cancer; Figure S2: (a–j) Examples of illustrating the segmented tumor regions on one CT slice.

Author Contributions: Conceptualization, W.I. and Q.T.; Methodology, B.M.M.; Software, W.I., N.A. and B.M.M.; Validation, T.E.A.; Formal analysis, W.I., M.J. and F.Y.; Investigation, W.I. and Q.T.; Writing—original draft, W.I.; Writing—review & editing, W.I. and Q.T.; Visualization, W.I., T.E.A. and F.Y.; Supervision, Q.T.; Project administration, Q.T.; Funding acquisition, Q.T. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by grants from the Stephenson Cancer Center (3P30CA225520), Institutional Research Grant number 134128-IRG-19-142-01 from the American Cancer Society, National Science Foundation (OIA-2132161, 2238648), National Institute of Health (R01DK133717), Oklahoma Shared Clinical and Translational Resources (NIGMS U54GM104938), Oklahoma Center for the Advancement of Science and Technology (HR23-071), the medical imaging COBRE (P20 GM135009), and the Midwest Biomedical Accelerator Consortium (MBArc, 1U01HL152410), an NIH Research Evaluation and Commercialization Hub (REACH). And The APC was funded by the OU Libraries' Open Access Fund.

Institutional Review Board Statement: This study does not acquire new image data from the hospitals. It uses the existing de-identified image data selected from a pre-established retrospective database in our medical imaging laboratory.

Informed Consent Statement: This study uses the existing image data from a de-identified retrospective gastric cancer patient database, and no informed consent was required or obtained for this study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author (qtang@ou.edu).

Conflicts of Interest: The authors declare no conflicts of interest related to this article.

References

1. Bray, F.; Ferlay, J.; Soerjomataram, I.; Siegel, R.L.; Torre, L.A.; Jemal, A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancer in 185 countries. *CA Cancer J. Clin.* **2018**, *68*, 394–424. [[CrossRef](#)] [[PubMed](#)]
2. Ma, D.; Zhang, Y.; Shao, X.; Wu, C.; Wu, J. PET/CT for predicting occult lymph node metastasis in gastric cancer. *Curr. Oncol.* **2022**, *29*, 6523–6539. [[CrossRef](#)]
3. He, X.; Wu, W.; Lin, Z.; Ding, Y.; Si, J.; Sun, L.-M. Validation of the American Joint Committee on Cancer (AJCC) stage system for gastric cancer patients: A population-based analysis. *Gastric Cancer* **2018**, *21*, 391–400. [[CrossRef](#)] [[PubMed](#)]
4. In, H.; Solsky, I.; Palis, B.; Langdon-Embry, M.; Ajani, J.; Sano, T. Validation of the 8th edition of the AJCC TNM staging system for gastric cancer using the National Cancer Database. *Ann. Surg. Oncol.* **2017**, *24*, 3683–3691. [[CrossRef](#)] [[PubMed](#)]
5. Son, T.; Sun, J.; Choi, S.; Cho, M.; Kwon, I.G.; Kim, H.-I.; Hyung, W.J. Multi-institutional validation of the 8th AJCC TNM staging system for gastric cancer: Analysis of survival data from high-volume eastern centers and the SEER database. *J. Surg. Oncol.* **2019**, *120*, 676–684. [[CrossRef](#)] [[PubMed](#)]
6. Wu, Y.; Grabsch, H.; Ivanova, T.; Tan, I.B.; Murray, J.; Ooi, C.H.; Tan, P. Comprehensive genomic meta-analysis identifies intra-tumoural stroma as a predictor of survival in patients with gastric cancer. *Gut* **2013**, *62*, 1100–1111. [[CrossRef](#)] [[PubMed](#)]
7. Davenport, T.; Kalakota, R. The potential for artificial intelligence in healthcare. *Future Healthc. J.* **2019**, *6*, 94–98. [[CrossRef](#)] [[PubMed](#)]
8. Briganti, G.; Le Moine, O. Artificial intelligence in medicine: Today and tomorrow. *Front. Med.* **2020**, *7*, 27. [[CrossRef](#)] [[PubMed](#)]
9. Hao, D.; Li, Q.; Feng, Q.X.; Qi, L.; Liu, X.S.; Arefan, D.; Wu, S. SurvivalCNN: A deep learning-based method for gastric cancer survival prediction using radiological imaging data and clinicopathological variables. *Artif. Intell. Med.* **2022**, *134*, 102424. [[CrossRef](#)]
10. Bėbas, E.; Borowska, M.; Derlatka, M.; Oczeretko, E.; Hładuński, M.; Szumowski, P.; Mojsak, M. Machine-learning-based classification of the histological subtype of non-small-cell lung cancer using MRI texture analysis. *Biomed. Signal Process. Control.* **2021**, *66*, 102446. [[CrossRef](#)]
11. Mintz, Y.; Brodie, R. Introduction to artificial intelligence in medicine. *Minim. Invasive Ther. Allied Technol.* **2019**, *28*, 73–81. [[CrossRef](#)] [[PubMed](#)]
12. Aerts, H.J.; Velazquez, E.R.; Leijenaar, R.T.; Parmar, C.; Grossmann, P.; Carvalho, S.; Lambin, P. Decoding tumor phenotype by noninvasive imaging using a quantitative radiomics approach. *Nat. Commun.* **2014**, *5*, 4006. [[CrossRef](#)] [[PubMed](#)]
13. Islam, W.; Danala, G.; Pham, H.; Zheng, B. Improving the performance of computer-aided classification of breast lesions using a new feature fusion method. *Proc. SPIE* **2022**, *12033*, 84–91.
14. Emaminejad, N. Fusion of quantitative image features and genomic biomarkers to improve prognosis assessment of early stage lung cancer patients. *IEEE Trans. Biomed. Eng.* **2016**, *63*, 1034–1043. [[CrossRef](#)] [[PubMed](#)]
15. Wang, L.; Gong, J.; Huang, X.; Lin, G.; Zheng, B.; Chen, J.; Lin, W. CT-based radiomics nomogram for preoperative prediction of No. 10 lymph nodes metastasis in advanced proximal gastric cancer. *Eur. J. Surg. Oncol.* **2021**, *47*, 1458–1465. [[CrossRef](#)] [[PubMed](#)]
16. Shin, J.; Lim, J.S.; Huh, Y.M.; Kim, J.H.; Hyung, W.J.; Chung, J.J.; Kim, S. A radiomics-based model for predicting prognosis of locally advanced gastric cancer in the preoperative setting. *Sci. Rep.* **2021**, *11*, 1879. [[CrossRef](#)]
17. Pham, H.; Jones, M.A.; Gai, T.; Islam, W.; Danala, G.; Jo, J.; Zheng, B. Identifying an optimal machine learning generated image marker to predict survival of gastric cancer patients. In Proceedings of the Medical Imaging 2022: Computer-Aided Diagnosis, San Diego, CA, USA, 4 April 2022; Volume 12033, pp. 408–414.
18. Ge, G.; Zhang, J. Feature selection methods and predictive models in CT lung cancer radiomics. *J. Appl. Clin. Med. Phys.* **2023**, *24*, e13869. [[CrossRef](#)] [[PubMed](#)]
19. Hoi, H.; Jin, R.; Zhu, J.; Lyu, R. Semisupervised SVM batch mode active learning with applications to image retrieval. *ACM Trans. Inf. Syst.* **2009**, *27*, 1–16. [[CrossRef](#)]
20. Désir, C.; Bernard, S.; Petitjean, C.; Heutte, L. A random forest based approach for one class classification in medical imaging. In Proceedings of the Machine Learning in Medical Imaging: Third International Workshop, MLMI 2012, Held in Conjunction with MICCAI 2012, Nice, France, 1 October 2012; Revised Selected Papers 3. Springer: Berlin/Heidelberg, Germany, 2012; pp. 250–257.
21. Aydadenta, H. On the classification techniques in data mining for microarray data classification. *J. Phys. Conf. Ser.* **2018**, *971*, 012004. [[CrossRef](#)]
22. Zheng, B.; Chang, Y.; Gur, D. Computerized detection of masses in digitized mammograms using single-image segmentation and a multilayer topographic feature analysis. *Acad. Radiol.* **1995**, *2*, 959–966. [[CrossRef](#)]
23. Danala, G.; Patel, B.; Aghaei, F.; Heidari, M.; Li, J.; Wu, T.; Zheng, B. Classification of breast masses using a computer-aided diagnosis scheme of contrast enhanced digital mammograms. *Ann. Biomed. Eng.* **2018**, *46*, 1419–1431. [[CrossRef](#)] [[PubMed](#)]
24. Mirniaharikandehi, S.; Heidari, M.; Danala, G.; Lakshmivarahan, S.; Zheng, B. Applying a random projection algorithm to optimize machine learning model for predicting peritoneal metastasis in gastric cancer patients using CT images. *Comput. Methods Programs Biomed.* **2021**, *200*, 105937. [[CrossRef](#)] [[PubMed](#)]
25. Zwanenburg, A.; Vallières, M.; Abdalah, M.A.; Aerts, H.J.; Andrearczyk, V.; Apte, A.; Löck, S. The image biomarker standardization initiative: Standardized quantitative radiomics for high-throughput image-based phenotyping. *Radiology* **2020**, *295*, 328–338. [[CrossRef](#)] [[PubMed](#)]

26. Van Griethuysen, J.J.; Fedorov, A.; Parmar, C.; Hosny, A.; Aucoin, N.; Narayan, V.; Aerts, H.J. Computational radiomics system to decode the radiographic phenotype. *Cancer Res.* **2017**, *77*, e104–e107. [[CrossRef](#)]
27. Park, Y.; Guldmann, J.M. Measuring continuous landscape patterns with Gray-Level Co-Occurrence Matrix (GLCM) indices: An alternative to patch metrics? *Ecol. Indic.* **2020**, *109*, 105802. [[CrossRef](#)]
28. Zhou, J.; Yang, M. Bone region segmentation in medical images based on improved watershed algorithm. *Comput. Intell. Neurosci.* **2022**, *2022*, 3975853. [[CrossRef](#)]
29. Venkatesh, U.; Balachander, B. Analysis of Textural Variations in Cerebellum in Brain to Identify Alzheimers by using Haralicks in Comparison with Gray Level Co-occurrence Matrix (GLRLM). In Proceedings of the 2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM), Gautam Buddha Nagar, India, 23–25 February 2022; Volume 2, pp. 549–556.
30. Chandraprabha, K.; Akila, S. Texture feature extraction for batik images using glcm and glrlm with neural network classification. *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.* **2019**, *5*, 6–15. [[CrossRef](#)]
31. Rashed, B.M.; Popescu, N. Performance Investigation for Medical Image Evaluation and Diagnosis Using Machine-Learning and Deep-Learning Techniques. *Computation* **2023**, *11*, 63. [[CrossRef](#)]
32. Khan, S.; Kaklis, P.; Serani, A.; Diez, M.; Kostas, K. Shape-supervised dimension reduction: Extracting geometry and physics associated features with geometric moments. *Comput. -Aided Des.* **2022**, *150*, 103327. [[CrossRef](#)]
33. Bharti, R.; Khamparia, A.; Shabaz, M.; Dhiman, G.; Pande, S.; Singh, P. Prediction of heart disease using a combination of machine learning and deep learning. *Comput. Intell. Neurosci.* **2021**, *2021*, 8387680. [[CrossRef](#)]
34. Zhang, H.; Hung, C.L.; Min, G.; Guo, J.P.; Liu, M.; Hu, X. GPU-accelerated GLRLM algorithm for feature extraction of MRI. *Sci. Rep.* **2019**, *9*, 10883. [[CrossRef](#)] [[PubMed](#)]
35. Kim, J.K.; Park, H.W. Statistical textural features for detection of microcalcifications in digitized mammograms. *IEEE Trans. Med. Imaging* **1999**, *18*, 231–238. [[PubMed](#)]
36. Strano, G.; Hao, L.; Everson, R.M.; Evans, K.E. Surface roughness analysis, modelling and prediction in selective laser melting. *J. Mater. Process. Technol.* **2013**, *213*, 589–597. [[CrossRef](#)]
37. Zhang, J.; Hu, J. Image segmentation based on 2D Otsu method with histogram analysis. In Proceedings of the 2008 International Conference on Computer Science and Software Engineering, Wuhan, China, 12–14 December 2008; Volume 6, pp. 105–108.
38. Haralick, R.M.; Sternberg, S.R.; Zhuang, X. Image analysis using mathematical morphology. *IEEE Trans. Pattern Anal. Mach. Intell.* **1987**, *PAMI-9*, 532–550.
39. Dierking, W. Quantitative roughness characterization of geological surfaces and implications for radar signature analysis. *IEEE Trans. Geosci. Remote Sens.* **1999**, *37*, 2397–2412. [[CrossRef](#)]
40. Singh, A.; Gaurav, K.; Rai, A.K.; Beg, Z. Machine learning to estimate surface roughness from satellite images. *Remote Sens.* **2021**, *13*, 3794. [[CrossRef](#)]
41. Celik, T. Multiscale change detection in multitemporal satellite images. *IEEE Geosci. Remote Sens. Lett.* **2009**, *6*, 820–824. [[CrossRef](#)]
42. Lubner, M.G.; Smith, A.D.; Sandrasegaran, K.; Sahani, D.V.; Pickhardt, P.J. CT texture analysis: Definitions, applications, biologic correlates, and challenges. *Radiographics* **2017**, *37*, 1483–1503. [[CrossRef](#)]
43. Sánchez, C.I.; Niemeijer, M.; Išgum, I.; Dumitrescu, A.; Suttorp-Schulten, M.S.; Abramoff, M.D.; van Ginneken, B. Contextual computer-aided detection: Improving bright lesion detection in retinal images and coronary calcification identification in CT scans. *Med. Image Anal.* **2012**, *16*, 50–62. [[CrossRef](#)] [[PubMed](#)]
44. Fan, T.W.; Malhi, H.; Varghese, B.; Cen, S.; Hwang, D.; Aron, M.; Duddalwar, V. Computed tomography-based texture analysis of bladder cancer: Differentiating urothelial carcinoma from micropapillary carcinoma. *Abdom. Radiol.* **2019**, *44*, 201–208. [[CrossRef](#)]
45. He, Z.; Chen, J.; Yang, F.; Pan, X.; Liu, C. Computed tomography-based texture assessment for the differentiation of benign, borderline, and early-stage malignant ovarian neoplasms. *J. Int. Med. Res.* **2023**, *51*, 03000605221150139. [[CrossRef](#)]
46. Ba-Ssalamah, A.; Muin, D.; Scherthaner, R.; Kulinna-Cosentini, C.; Bastati, N.; Stift, J.; Mayerhoefer, M.E. Texture-based classification of different gastric tumors at contrast-enhanced CT. *Eur. J. Radiol.* **2013**, *82*, e537–e543. [[CrossRef](#)] [[PubMed](#)]
47. Raja, J.V.; Khan, M.; Ramachandra, V.K.; Al-Kadi, O. Texture analysis of CT images in the characterization of oral cancers involving buccal mucosa. *Dentomaxillofacial Radiol.* **2012**, *41*, 475–480. [[CrossRef](#)]
48. Bonnin, A.; Durot, C.; Barat, M.; Djelouah, M.; Grange, F.; Mulé, S.; Hoeffel, C. CT texture analysis as a predictor of favorable response to anti-PD1 monoclonal antibodies in metastatic skin melanoma. *Diagn. Interv. Imaging* **2022**, *103*, 97–102. [[CrossRef](#)] [[PubMed](#)]
49. IBM. What Is Random Forest? Available online: <https://www.ibm.com/topics/random-forest> (accessed on 9 January 2024).
50. Khandelwal, R. (2018, November 2). K Fold and Other Cross-Validation Techniques. Available online: <https://medium.datadriveninvestor.com/k-fold-and-other-cross-validation-techniques-6c03a2563f1e> (accessed on 17 July 2023).
51. Liu, M.M.; Wen, L.; Liu, Y.J.; Cai, Q.; Li, L.T.; Cai, Y.M. Application of data mining methods to improve screening for the risk of early gastric cancer. *BMC Med. Inform. Decis. Mak.* **2018**, *18*, 23–32. [[CrossRef](#)] [[PubMed](#)]
52. Cai, Q.; Zhu, C.; Yuan, Y.; Feng, Q.; Feng, Y.; Hao, Y.; Li, Z. Development and validation of a prediction rule for estimating gastric cancer risk in the Chinese high-risk population: A nationwide multicentre study. *Gut* **2019**, *68*, 1576–1587. [[CrossRef](#)] [[PubMed](#)]
53. Safdari, R.; Arpanahi, H.K.; Langarizadeh, M.; Ghazisaiedi, M.; Dargahi, H.; Zendehtdel, K. Design a fuzzy rule-based expert system to aid earlier diagnosis of gastric cancer. *Acta Inform. Medica* **2018**, *26*, 19. [[CrossRef](#)] [[PubMed](#)]

54. Su, Y.; Shen, J.; Qian, H.; Ma, H.; Ji, J.; Ma, H.; Shou, C. Diagnosis of gastric cancer using decision tree classification of mass spectral data. *Cancer Sci.* **2007**, *98*, 37–43. [[CrossRef](#)]
55. Brindha, S.K.; Kumar, N.S.; Chenkual, S.; Lalruatfela, S.T.; Zomuana, T.; Ralte, Z.; Nath, P. Data mining for early gastric cancer etiological factors from diet-lifestyle characteristics. In Proceedings of the 2020 4th International Conference on Intelligent Computing and Control Systems (ICICCS), Madurai, India, 13–15 May 2020; pp. 65–73.
56. Mortezaagholi, A.; Khosravizadeh, O.; Menhaj, M.B.; Shafigh, Y.; Kalhor, R. Make intelligent of gastric cancer diagnosis error in Qazvin’s medical centers: Using data mining method. *Asian Pac. J. Cancer Prev. APJCP* **2019**, *20*, 2607. [[CrossRef](#)]
57. Taninaga, J.; Nishiyama, Y.; Fujibayashi, K.; Gunji, T.; Sasabe, N.; Iijima, K.; Naito, T. Prediction of future gastric cancer risk using a machine learning algorithm and comprehensive medical check-up data: A case-control study. *Sci. Rep.* **2019**, *9*, 12384. [[CrossRef](#)]
58. Islam, W.; Jones, M.; Faiz, R.; Sadeghipour, N.; Qiu, Y.; Zheng, B. Improving performance of breast lesion classification using a ResNet50 model optimized with a novel attention mechanism. *Tomography* **2022**, *8*, 2411–2425. [[CrossRef](#)]
59. Armi, L.; Fekri-Ershad, S. Texture image analysis and texture classification methods-A review. *arXiv* **2019**, arXiv:1904.06554.
60. Zhao, X.; Xia, X.; Wang, X.; Bai, M.; Zhan, D.; Shu, K. Deep learning-based protein features predict overall survival and chemotherapy benefit in gastric cancer. *Front. Oncol.* **2022**, *12*, 847706. [[CrossRef](#)] [[PubMed](#)]
61. Yeganeh, A.; Johannssen, A.; Chukhrova, N.; Erfanian, M.; Azarpazhooh, M.R.; Morovatdar, N. A monitoring framework for health care processes using Generalized Additive Models and Auto-Encoders. *Artif. Intell. Med.* **2023**, *146*, 102689. [[CrossRef](#)] [[PubMed](#)]
62. Aslam, M.A.; Xue, C.; Chen, Y.; Zhang, A.; Liu, M.; Wang, K.; Cui, D. Breath analysis based early gastric cancer classification from deep stacked sparse autoencoder neural network. *Sci. Rep.* **2021**, *11*, 4014. [[CrossRef](#)]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.