



Article Machine Learning and Texture Analysis of [¹⁸F]FDG PET/CT Images for the Prediction of Distant Metastases in Non-Small-Cell Lung Cancer Patients

Armin Hakkak Moghadam Torbati ^{1,†}, Sara Pellegrino ^{1,†}, Rosa Fonti ¹, Rocco Morra ², Sabino De Placido ² and Silvana Del Vecchio ^{1,*}

- ¹ Department of Advanced Biomedical Sciences, University of Naples "Federico II", 80131 Naples, Italy; hakakarmin@gmail.com (A.H.M.T.); sara.pellegrino@unina.it (S.P.); rosa.fonti@unina.it (R.F.)
- ² Department of Clinical Medicine and Surgery, University of Naples "Federico II", 80131 Naples, Italy; rocco.mor4@gmail.com (R.M.); deplacid@unina.it (S.D.P.)
- * Correspondence: delvecc@unina.it; Tel.: +39-081-746-3307; Fax: +39-081-545-7081
- ⁺ These authors contributed equally to this work.

Abstract: The aim of our study was to predict the occurrence of distant metastases in non-small-cell lung cancer (NSCLC) patients using machine learning methods and texture analysis of ¹⁸F-labeled 2-deoxy-d-glucose Positron Emission Tomography/Computed Tomography {[¹⁸F]FDG PET/CT} images. In this retrospective and single-center study, we evaluated 79 patients with advanced NSCLC who had undergone [¹⁸F]FDG PET/CT scan at diagnosis before any therapy. Patients were divided into two independent training (n = 44) and final testing (n = 35) cohorts. Texture features of primary tumors and lymph node metastases were extracted from $[^{18}F]FDG$ PET/CT images using the LIFEx program. Six machine learning methods were applied to the training dataset using the entire panel of features. Dedicated selection methods were used to generate different combinations of five features. The performance of selected machine learning methods applied to the different combinations of features was determined using accuracy, the confusion matrix, receiver operating characteristic (ROC) curves, and area under the curve (AUC). A total of 104 and 78 lesions were analyzed in the training and final testing cohorts, respectively. The support vector machine (SVM) and decision tree methods showed the highest accuracy in the training cohort. Seven combinations of five features were obtained and introduced in the models and subsequently applied to the training and final testing cohorts using the SVM and decision tree. The accuracy and the AUC of the decision tree method were higher than those obtained with the SVM in the final testing cohort. The best combination of features included shape sphericity, gray level run length matrix_run length non-uniformity (GLRLM_RLNU), Total Lesion Glycolysis (TLG), Metabolic Tumor Volume (MTV), and shape compacity. The combination of these features with the decision tree method could predict the occurrence of distant metastases with an accuracy of 74.4% and an AUC of 0.63 in NSCLC patients.

Keywords: machine learning; texture features; [18F]FDG PET/CT; non-small-cell lung cancer; metastases

1. Introduction

Lung cancer is recognized as the leading cause of cancer death worldwide [1]. Each year, a total of 11.4% of all new cancer cases originate from the lungs, making it the second most common type of cancer [1]. At diagnosis, about 85% of patients have non-small-cell lung cancer (NSCLC), which, at histopathological examination, includes adenocarcinoma, squamous cell carcinoma, and large cell carcinoma [2]. Following diagnosis, patients undergo a complete work-up for the accurate staging of the disease that will guide the selection of the most appropriate treatment regimen for each patient [3]. Early-stage lung cancer patients are candidates for surgery or curative radiotherapy, whereas advanced-stage lung cancer patients will receive chemotherapy, chemo-radiotherapy, targeted therapy, and



Citation: Hakkak Moghadam Torbati, A.; Pellegrino, S.; Fonti, R.; Morra, R.; De Placido, S.; Del Vecchio, S. Machine Learning and Texture Analysis of [¹⁸F]FDG PET/CT Images for the Prediction of Distant Metastases in Non-Small-Cell Lung Cancer Patients. *Biomedicines* 2024, *12*, 472. https://doi.org/10.3390/ biomedicines12030472

Academic Editor: Jinghuei Lee

Received: 9 January 2024 Revised: 1 February 2024 Accepted: 16 February 2024 Published: 20 February 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/). immunotherapy [4–6]. Despite the transitory success of therapy, patients with lung cancer in stages I-III will often develop distant metastases that probably were already present at diagnosis in a subclinical phase and thus being undetectable at the staging work-up.

Lung cancer diagnosis, staging, evaluation of response to treatment, and follow up are based on multimodality imaging, including Computed Tomography (CT), Positron Emission Tomography/Computed Tomography (PET/CT), and Magnetic Resonance Imaging (MRI). The recent development of methods for the analysis of medical images based on artificial intelligence (AI) significantly improved the accuracy and efficiency of image analysis in several clinical contexts, such as lung cancer screening and diagnosis, nodule detection, molecular characterization, lung cancer staging, response to treatment, and prognosis [7,8]. In particular, AI-based algorithms including machine learning, deep learning, and radiomics have significantly enhanced clinical decision making based on individual patient's imaging, clinical, molecular, and pathological data. Machine learning methods include several algorithms that are capable of learning from data observation, thus obtaining inferences. Deep learning methods include algorithms based on artificial neural networks composed of layers of artificial neurons capable of extracting features from data and building complex models. Radiomics is based on the extraction and quantification of subvisual image features using data characterization algorithms.

Texture analysis is an emerging radiomic tool for the evaluation and quantification of spatial signal variation in a segmented volume of medical images [9,10]. A combination of texture analysis and AI-based algorithms has been applied to evaluate tumor heterogeneity, overall and progression-free survival (OS and PFS), response to therapy, and biological and molecular characteristics of different types of tumors [11,12]. In particular, previous studies performed texture analysis in lung cancer patients and showed that several features, including the coefficient of variation, dissimilarity, coarseness, and entropy, were able to predict both PFS and OS in patients [13–18]. Other studies successfully used texture analysis and machine learning or deep learning methods to discriminate involved lymph nodes from reactive lymph nodes in lung cancer patients by analyzing and processing ¹⁸F-labeled 2-deoxy-d-glucose {[¹⁸F]FDG} PET/CT or CT alone [19,20]. Only a few studies reported the use of texture analysis to predict distant metastases in NSCLC patients and even fewer combined texture analysis and machine learning methods to identify patients at high risk of developing distant metastases [21–23]. In the present study, we compared six different machine learning methods and selected those that combined with texture analysis could predict the development of distant metastases in lung cancer patients, thus helping clinicians to select the appropriate treatment for patients at high risk of metastatic dissemination. To this end, we analyzed primary tumors and lymph nodes that were positive at [¹⁸F]FDG PET/CT scan using texture analysis and tested different machine learning methods in order to develop an effective classification model. By applying a radiomic approach, we aim to exploit all the information contained in the images to test whether they can predict the development of distant metastases.

2. Materials and Methods

2.1. Patients

We retrospectively evaluated 79 patients (54 men, 25 women) using the following inclusion criteria: histologically proven non-small-cell lung cancer; stage III and IV disease; whole-body [¹⁸F]FDG PET/CT scan performed at our institution before any therapy; and clinical and imaging follow up for at least 6 months. The exclusion criteria were prior lung or chest malignancy; prior chemotherapy or chest radiotherapy; no pathological diagnosis on primary lung lesion; missing imaging data for analysis; and missing clinical and imaging follow up for at least 6 months. Distant metastases were detected by [¹⁸F]FDG PET/CT scan, whole body contrast-enhanced CT, brain MRI, and in selected cases, MRI of different anatomical districts, and were subsequently confirmed at clinical and imaging follow up. This single institutional study was approved by the local ethics committee (Protocol No. 352/18), and all subjects signed an informed consent form.

We divided the overall population into two independent cohorts based on the time of imaging (before or after 2019). The first cohort included 44 patients and was used for training purposes. The second independent cohort included 35 patients and was used for final testing purposes. The clinical characteristics of all patients are summarized in Table 1. In the training dataset, there were 11 patients in stage III (5 IIIA, 4 IIIB, and 2 IIIC), while 33 patients were in stage IV (14 IVA and 19 IVB). There were 18 patients with adenocarcinoma, 12 with squamous cell carcinoma, 2 with large cell carcinoma, and 12 with NSCLC not otherwise specified (NOS). In the final testing cohort, there were 15 patients in stage III (7 IIIB and 8 stage IIIC), while 20 patients were in stage IV (4 IVA and 16 IVB); there were 20 patients with adenocarcinoma, 8 with squamous cell carcinoma, 1 with large cell carcinoma, and 6 with NSCLC not otherwise specified (NOS). The overall population was treated according to their stage, histology, molecular pathology, age, performance status, and comorbidities [4–6], as reported in Table 1.

Characteristic	Overall	Training Cohort	Final Testing Cohort	
Patients	79	44	35	
Age				
Mean \pm SD	65 ± 12	64 ± 13	67 ± 10	
Range	38–86	38–86	41–71	
Gender				
Male	54	29	25	
Female	25	15	10	
Histology				
Adenocarcinoma	38	18	20	
Squamous cell carcinoma	20	12	8	
Large cell carcinoma	3	2	1	
Not otherwise specified	18	12	6	
TNM stage				
Stage III	26	11	15	
Stage IV	53	33	20	
Treatment				
Chemotherapy	46	30	16	
Chemoradiotherapy	3	3		
Chemotherapy/Immunotherapy	15	3	12	
No cancer therapy	15	8	7	

Table 1. Clinical characteristics, histology, stage, and treatment of 79 patients with advanced NSCLC.

The presence of metastases was established at the time of $[^{18}F]FDG$ PET/CT in both the training and final testing cohorts.

2.2. [¹⁸F]FDG PET/CT Study

After fasting for 8 h, patients received 370 MBq of [18 F]FDG by intravenous injection and 60 min later underwent [18 F]FDG PET/CT scan using an Ingenuity TF scanner (Philips Healthcare, Best, The Netherlands). Before tracer injection, blood glucose level was measured in all patients and was <120 mg/dL. The parameters of the multidetector CT scan were 120 kV, 80 mAs, 0.8 s rotation time, and a pitch of 1.5; if needed, a fully diagnostic contrast-enhanced CT was performed. For PET scan acquisition, 3-dimensional mode, 3 min per bed position, and six to eight bed positions per patient were used. An ordered subset-expectation maximization algorithm was applied for image reconstruction. A filtered back projection of CT reconstructed images (Gaussian filter with 8 mm full width half maximum) was used for the attenuation correction of PET emission data. [18 F]FDG

PET/CT co-registered images in the transaxial, sagittal, and coronal planes were preliminarily visualized with Ingenuity TF software (IntelliSpace Portal V5.0, Philips Healthcare, Best, The Netherlands).

2.3. [¹⁸F]FDG PET/CT Image Analysis

PET/CT data, after transfer in DICOM format to a different workstation, were processed by the LIFEx program [24]. Areas of focal [¹⁸F]FDG uptake were considered positive if detected at least on 2 contiguous PET slices and corresponded to structural CT abnormalities. When multiple lymph nodes were coalescent, they were considered as a single lesion. Areas of physiological tracer uptake were excluded from the analysis. In agreement with previous studies [25–27], an automated contouring program was applied for drawing a tridimensional region around the target lesion using an absolute threshold for SUV at 2.5 (Figure 1), thus obtaining a Volume of Interest (VOI). The transfer of the VOI on the corresponding CT images confirmed the accuracy of lesion delimitation.



Figure 1. Representative images of $[^{18}F]$ FDG PET/CT scan in a patient with NSCLC. Representative images of $[^{18}F]$ FDG PET/CT scan in a patient with NSCLC: (**a**) transaxial $[^{18}F]$ FDG PET image; (**b**) co-registered PET/CT fusion image; (**c**) co-registered CT image. A tridimensional region of interest was drawn around the FDG avid tumor lesion (pink) using an automated contouring program setting an absolute threshold for SUV at 2.5 (**a**). Fusion and CT images confirmed the accuracy of lesion segmentation (**b**,**c**).

Texture features of the primary tumor and lymph node metastases were extracted using the LIFEx package (developed at CEA, Orsay, France, http://www.lifexsoft.org, accessed on 19 February 2024). VOIs that did not reach the minimum number of 64 voxels were excluded from the analysis to avoid the inaccurate quantification of texture features inside small lesions. Tonal discretization of the gray scale for PET images was adjusted using 64 gray levels with an absolute scale bound between 0 and 25 SUV. Texture features included conventional and histogram-based parameters, shape and size, and second- and high-order features, as shown in Supplementary Table S1 online. In particular, 39 features plus conventional [¹⁸F]FDG PET parameters SUVmax, SUVmean, SDmean, SUVmin, SUV-peak, MTV, and TLG were extracted from each lesion and then subjected to selection based on redundancy and importance in determining the clinical endpoint.

2.4. Selection of the Best Machine Learning Method

There are a large number of machine learning methods that can be used to train a model. Selecting the best method highly depends on the type of data and targets. Therefore, it is often needed to apply various machine learning methods to data and compare their performance. In this study, we compared the efficacy of 6 renowned methods, including the decision tree, linear discriminant analysis, the naïve Bayes classifier, the support vector machine (SVM), k-nearest neighbors, and the feedforward neural network [28–30].

Briefly, the decision tree is a supervised algorithm used for classification and regression. It has a hierarchical tree-like structure composed of a root node, branches, and leaf nodes. The predicted class is obtained by applying a top-down approach to the dataset from the root node to a leaf node [31]. Linear discriminant analysis is an algorithm that finds

linear discriminants that maximize the separation between classes in a dataset [32]. The naïve Bayes classifier is a supervised algorithm that computes the class of a given sample by maximizing its probability of belonging to that class. It assumes that predictors are independent of one another in each class [33]. The support vector machine is a supervised algorithm used for binary classification or regression prediction. It finds the boundary that best separates two classes while remaining at the maximum distance from the closest data points in each class. The term "support vectors" refers to the closest data points in each class [28]. K-nearest neighbor is an algorithm that identifies the k points closest to the unknown sample based on their distance. If k-nearest points belong to a given class, the assumption is that also the unknown sample belongs to the same class [34]. The feedforward neural network is a class of algorithms composed of layers of artificial neurons; basically, the input layer receives input data, the hidden layer extracts features from the input data, and the output layer produces the result. The flow of information is unidirectional from one layer to the next [35]. The selection of the tested classifiers was based on simplicity, interpretability of results, efficiency with a small dataset, computational time, and intensity. The experimental setting of each model is summarized as follows. In the decision tree model, the split criterion was the Gini's diversity index, and the maximum tree depth was set to 10 levels. Linear discriminant analysis was performed using singular value decomposition. The naïve Bayes classifier used the Gaussian model and assumed that each class is normally distributed. The support vector machine algorithm used a linear kernel. The k-nearest neighbor algorithm used a k parameter of 1 and Euclidean distance as the distance metric. The feedforward neural network model used an input layer of 10 neurons fully connected to a hidden layer of 10 neurons followed by an output layer, using sigmoid as the activation function.

The six methods were applied to the entire set of features in the training cohort of 104 lesions using MATLAB (version R2021a, Mathworks Inc., Natick, MA, USA), and the results were subjected to univariate analysis to select the best method for the prediction of distant metastases based on accuracy. A 5-fold cross-validation approach was used to train all six methods in the training cohort of 104 lesions. Among these lesions, 70% of the input data went into the training set, leaving 30% for the validation group.

2.5. Selection of Texture Features

Since many features are correlated, to avoid redundancy, we calculated the Pearson correlation coefficient of pairs of features. When the correlation coefficient of the two variables was >0.8, one of the two features was removed from the analysis. To select features on the basis of their importance in determining the clinical endpoint, we applied the least absolute shrinkage and selection operator algorithm (LASSO) in a MATLAB environment along with the Cox survival model in the training cohort. This algorithm includes several selection methods, such as Fscchi2, Fscmrmr, Fscnca, Fsrftest, Fsrnca, Fsulaplacian, and Relieff, which identify a subset of measured features highly correlated with the clinical endpoint. The first 5 features selected by each method were employed in the training model.

2.6. Selection and Evaluation of the Best Model

The selected machine learning methods were combined with each pool of selected features and accuracy, the confusion matrix, and ROC curves were obtained to identify the best model. Then this model was used in the final testing cohort of patients. The accuracy of the model was calculated for each combination of the selected features considering the presence or absence of metastases at the time of [¹⁸F]FDG PET/CT.

3. Results

Texture analysis was performed on [¹⁸F]FDG PET/CT scans of NSCLC patients. A total of 104 lesions (37 primary tumors and 67 lymph nodes) were analyzed in the training group, while 78 lesions (32 primary tumors and 46 lymph nodes) were analyzed in the final testing dataset. Texture analysis provided 39 features for each lesion plus conventional

[¹⁸F]FDG PET parameters SUVmax, SUVmean, standard deviation of SUVmean (SDmean), SUVmin, SUVpeak, Metabolic Tumor Volume (MTV), and Total Lesion Glycolysis (TLG).

Six methods were applied using the entire set of features, and univariate analysis was performed to find the best method for the prediction of metastases based on accuracy [36]. Table 2 shows the mean and standard deviation of the accuracy of each method for the training dataset. The SVM and decision tree methods had the best accuracy with 78.26% and 71.68%, respectively. So, these two methods were considered for further data processing.

MethodsMean Accuracy (%) \pm SDDecision tree 71.68 ± 0.89 Linear discriminant analysis 65.39 ± 1.04 Naïve Bayes classification 62.05 ± 0.98 Support vector machine 78.26 ± 0.98 K-nearest neighbor 63.39 ± 1.07 Feedforward neural network 66.09 ± 1.01

Table 2. Mean and standard deviation of the accuracy of six training machine learning methods.

Correlation among features was tested using the Pearson correlation coefficient, and 27 features (23 from texture analysis plus SUVmax, SUVmean, MTV, and TLG) remained after the exclusion of correlated features. Then, seven feature selection methods (LASSO) were applied to twenty-seven features to rank them in terms of importance and relative weight. The features extracted by each method are listed in Table 3.

Table 3. Features extracted by seven LASSO feature selection methods on the basis of their high correlation with the clinical endpoint.

Fscchi2	Fscmrmr	Fscnca	Fsrftest	Fsrnca	Fsulaplacian	Relieff
GLCM_ dissimilarity	GLRLM_RP	GLZLM_ LZHGE	GLCM_ dissimilarity	GLZLM_ LZHGE	MTV	Shape sphericity
GLCM_energy	TLG	GLRLM_ RLNU	GLCM_energy	GLRLM_ RLNU	SUVmean	GLRLM_ RLNU
GLCM_ homogeneity	HISTO_ kurtosis	CoV	GLCM_ homogeneity	CoV	TLG	TLG
TLG	CoV	GLRLM_ LRHGE	TLG	GLRLM_ LRHGE	CoV	MTV
GLZLM_ SZLGE	Shape sphericity		GLZLM_SZLGE		SUVmax	Shape compacity

The variables selected by Fscchi2 and Fsrftest and those selected by Fscnca and Fsrnca are similar since the methods have similar mathematic fundamentals. In the Fscnca and Fsrnca algorithms, the fifth feature had a slight weight compared to the previous four selected by the same method and was not included in further analysis.

In the subsequent stage, we tested the best combination of selected features and machine learning techniques. To this end, some models were designed using features in Table 3 as the input and the presence of distant metastases as the output. These models were trained with the SVM and decision tree methods as well. Accuracy, the confusion matrix, and ROC curves were obtained to identify the best model.

Although the SVM had good accuracy in the training dataset using different combinations of features (71.4–83.7%), its accuracy was very low in the final testing dataset (40–51.3%), as shown in Supplementary Tables S2 and S3 online. For example, one of the best performances for this method belonged to a model that combined five extracted features by the Fscchi2 algorithm as inputs. Despite the fact that this model achieved fairly good accuracy in training (83.7%), the accuracy in the final testing dataset was just 51.3%. In addition, the values of the AUC derived from the ROC curve for the training and final testing datasets were very low (training AUC = 0.43, final testing AUC = 0.39).

Tables 4 and 5 show that the decision tree method had great accuracy in both the training and final testing datasets using different combinations of features.

Table 4. The training accuracy of the model calculated for each combination of the selected features using the decision tree method in 104 lesions.

	Number of Combined Features					
LASSO method	First feature	Two first features	Three first features	Four first features	Five first features	
Fscchi2	74	75	76	82.7	79.8	
Fscmrmr	76	73.1	77.9	75	74	
Fscnca	77.9	78.8	72.1	71.2	-	
Fsrftest	74	75	76	82.7	79.8	
Fsrnca	77.9	78.8	72.1	71.2	-	
Fsulaplacian	75	71.2	72.1	71.2	73.1	
Relieff	76.9	76.9	76	77.9	76.9	

Table 5. Accuracy of the model calculated for each combination of the selected features using the decision tree method in the final testing dataset of 78 lesions.

	Number of Combined Features				
LASSO method	First feature	Two first features	Three first features	Four first features	Five first features
Fscchi2	73.1	71.8	69.2	70.5	69.2
Fscmrmr	66.7	67.9	60.3	65.4	65.4
Fscnca	71.8	75.5	70.5	69.2	-
Fsrftest	73.1	71.8	69.2	70.5	69.2
Fsrnca	71.8	75.5	70.5	69.2	-
Fsulaplacian	71.8	70.1	65.4	69.5	70.6
Relieff	69.2	73.1	73.1	71.9	74.4

Table 5 reports the accuracy of the decision tree method using different combinations of features from 1 to 5 in the final testing dataset. The best accuracy of the model using both Fscchi2 and Fsrftest was 73.1% by employing only the first feature of the combination, i.e., GLCM_dissimilarity. Using Fscnca and Fsrnca, the best accuracy was 75.5% by combining GLZLM_LZHGE with GLRLM_RLNU, the first two features of the selected combination. The best accuracy using Fscmrmr was 67.9% by combining GLRLM_RP with TLG, whereas with the Fsulaplacian combination, an accuracy of 71.8% was obtained employing only MTV. Finally, the best accuracy using Relieff was 74.4% with a combination of shape sphericity, GLRLM_RLNU, TLG, MTV, and shape compacity. To select the best model, we also consider the AUC and ROC curves, and Figure 2 shows the values of the AUC and accuracy for each model that were calculated based on the confusion matrix.

When we took into account the AUC and accuracy of each most accurate selection method, we found that the best AUC and accuracy values were obtained using the Relieff method. Therefore, the combination of shape sphericity, GLRLM_RLNU, TLG, MTV, and shape compacity is able to predict the presence of distant metastases with an AUC of 0.63 and an accuracy of 74.4%. Figure 3 shows the AUC value and ROC curve for the combination of five features of the Relieff model.



Figure 2. AUC and accuracy of models applied to the final testing dataset of 78 lesions using the decision tree method and different combinations of features.



Figure 3. Graph of the ROC curve showing the performance of the model combining the decision tree and Relieff methods for the prediction of distant metastases in the final testing dataset of 78 lesions. The combination of shape sphericity, GLRLM_RLNU, TLG, MTV, and shape compacity is able to predict the presence of distant metastases with an AUC of 0.63.

4. Discussion

In the present study, we compared six machine learning methods to select the best training model for the prediction of distant metastases in NSCLC patients and found the

highest accuracy using the SVM and decision tree methods on the whole set of variables. Then, by applying seven feature selection methods, we identified the best five features that were predictive of metastases. The combination of the selected machine learning methods with each group of five selected features was then tested in the training cohort of patients and evaluated in the final testing cohort. We found that by feeding the decision tree method with the combination of shape sphericity, GLRLM_RLNU, TLG, MTV, and shape compacity selected by Relieff, we were able to predict the presence of distant metastases with an AUC of 0.63 and an accuracy of 74.4% in the final testing cohort of patients. Therefore, texture analysis of [¹⁸F]FDG-positive lesions and the decision tree method can identify patients with a high risk of distant metastases that can be candidates for more aggressive treatments. The novelty of our study, rather than being the use of advanced AI-based algorithms, resides in the fact that a systematic evaluation of classical machine learning methods and different feature combinations even in a small cohort of patients can provide meaningful information about patient's risk of distant metastases that can affect subsequent therapeutic choices. Another important consideration that we can draw from our results is that both GLCM_dissimilarity and MTV are very powerful predictive features since by feeding alone the decision tree model they could identify patients with distant metastases. In relation to the SVM method, while it showed good accuracy in the training group of patients using different combinations of features, it failed to predict the development of distant metastases in the final testing cohort.

Our study enrolled patients with advanced NSCLC, and most of them already had distant metastases at the time of [¹⁸F]FDG PET/CT examination. This allowed us to rapidly identify the features strictly correlated with the presence of distant metastases and to include them in the training model. However, the evaluation of the model in the final testing dataset, showing good accuracy, may allow us to extend the model to the early stages of NSCLC. In our study, texture analysis was performed on both primary tumors and involved lymph nodes providing a large range of feature values that reflected the heterogeneity of [¹⁸F]FDG uptake in lesions at different steps of disease progression. Further studies are needed to test whether combining information on the heterogeneity of primary lesions and involved lymph nodes may improve the identification of patients with early-stage NSCLC at risk of distant metastases formation.

Nevertheless, our study has several limitations, including its retrospective design, the relatively limited number of patients, a moderate cohort imbalance in relation to the percentage of stage III patients in each cohort, and the absence of external validation. Therefore, our findings may require confirmation in a larger prospective study using stringent criteria for patient enrollment, including external validation in different centers. Furthermore, the limited size of training and testing datasets prevented the application of deep learning methods. Although we are aware that these methods are powerful tools to perform even complicated tasks, we believe that a simple approach using classical machine learning methods may provide key information to build more sophisticated models with a higher clinical impact.

Previous studies used machine learning, deep learning, and radiomics for the analysis of [¹⁸F]FDG PET/CT images of lung cancer patients in different clinical contexts. They reported AUC values up to 0.97 in the diagnosis of lung cancer using deep learning algorithms [7,22]. In the prediction of treatment response and prognosis, AUC values up to 0.95 were achieved using machine learning algorithms [7,22]. For staging purposes, lymph node involvement was predicted by both machine learning and deep learning models, which often included radiomic features and clinical data with AUC values up to 0.94 [19,22,37–40]. Only a few studies used AI-based algorithms for the prediction of distant metastases in lung cancer patients [21–23]. By applying a CNN, an accuracy of $63 \pm 5\%$ was achieved in a study that included 264 patients with different stages of the disease and 102 out of 264 with distant metastases [23]. In agreement with these findings, using the decision tree method in the final testing cohort, we found an accuracy of 74.4% in the prediction of distant metastases. Deep learning was also applied to predict occult

nodal metastases in [¹⁸F]FDG PET/CT images of patients with NSCLC [41]. In this study, 1911 patients were included in the training dataset, 355 patients for external validation and 999 patients for prospective validation. Deep learning applied to a combination of imaging features extracted from PET and CT images of primary tumors could predict occult metastases in regional lymph nodes with an AUC ranging between 0.875 and 0.958 in the different cohorts. These findings are very promising from the perspective of applying deep learning methods to the prediction of occult distant metastases in the early stages of the disease.

Therapeutic regimens in NSCLC patients are currently designed on the basis of the histology, stage, and molecular characterization of oncogene drivers and immune checkpoint proteins. The addition of features expressing the heterogeneity of glycolytic phenotypes in primary tumors and involved lymph nodes may add subclinical information on the metastatic potential of early-stage NSCLC in individual patients. Analysis of combined clinical data and texture features from imaging studies by machine learning or deep learning methods may further improve the adaptation of therapy to individual patients.

In conclusion, our study reports an effective classification model based on the machine learning method and texture analysis of [¹⁸F]FDG PET/CT images that can be clinically employed to identify patients with NSCLC at risk of developing distant metastases.

Supplementary Materials: The following supporting information can be downloaded at https://www. mdpi.com/article/10.3390/biomedicines12030472/s1, Table S1: Conventional and texture features derived from [¹⁸F]FDG PET/CT images using LIFEx software; Table S2: Accuracy of models using the SVM method in the training dataset; Table S3: Accuracy of models using the SVM method in the final testing dataset.

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

Funding: This work was supported by the Associazione Italiana per la Ricerca sul Cancro (AIRC), Grant Number IG 2021—ID, 25945 project—P.I. Silvana Del Vecchio and by MUR PRIN 2022 Project No. 2022F5N25M, CUP E53D23001290006, under the National Recovery and Resilience Plan (NRRP), Mission 4 Component C2 Investment 1.1—MUR call No. 104 on 2 February 2022, funded by the European Union—NextGenerationEU.

Institutional Review Board Statement: The study was conducted in accordance with the Declaration of Helsinki and was approved by the Ethics Committee of the University of Naples "Federico II" (16/01/2019; Protocol No. 352/18).

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available upon request from the corresponding author.

Acknowledgments: Armin Hakkak Moghadam Torbati performed this study as a part of his PhD program at the University of Naples "Federico II".

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. CA Cancer J. Clin. 2021, 71, 209–249. [CrossRef]
- Herbst, R.S.; Morgensztern, D.; Boshof, C. The biology and management of non-small cell lung cancer. *Nature* 2018, 553, 446–454. [CrossRef] [PubMed]
- 3. Jones, C.M.; Brunelli, A.; Callister, M.E.; Franks, K.N. Multimodality treatment of advanced non-small cell lung cancer: Where are we with the evidence? *Curr. Surg. Rep.* **2018**, *6*, 5. [CrossRef] [PubMed]

- Postmus, P.E.; Kerr, K.M.; Oudkerk, M.; Senan, S.; Waller, D.A.; Vansteenkiste, J.; Escriu, C.; Peters, S.; on behalf of the ESMO Guidelines Committee. Early and locally advanced non-small-cell lung cancer (NSCLC): ESMO Clinical Practice Guidelines for diagnosis, treatment and follow-up. *Ann. Oncol.* 2017, 28, iv1–iv21. [CrossRef] [PubMed]
- Hendriks, L.E.; Kerr, K.M.; Menis, J.; Mok, T.S.; Nestle, U.; Passaro, A.; Peters, S.; Planchard, D.; Smit, E.F.; Solomon, B.J.; et al. Non-oncogene-addicted metastatic non-small-cell lung cancer: ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann. Oncol.* 2023, 34, 358–376. [CrossRef] [PubMed]
- Hendriks, L.E.; Kerr, K.M.; Menis, J.; Mok, T.S.; Nestle, U.; Passaro, A.; Peters, S.; Planchard, D.; Smit, E.F.; Solomon, B.J.; et al. Oncogene-addicted metastatic non-small cell lung cancer:ESMO Clinical Practice Guideline for diagnosis, treatment and follow-up. *Ann. Oncol.* 2023, 34, 339–357. [CrossRef] [PubMed]
- Gandhi, Z.; Gurram, P.; Amgai, B.; Lekkala, S.P.; Lokhandwala, A.; Manne, S.; Mohammed, A.; Koshiya, H.; Dewaswala, N.; Desai, R.; et al. Artificial Intelligence and Lung Cancer: Impact on Improving Patient Outcomes. *Cancers* 2023, 15, 5236. [CrossRef] [PubMed]
- Huang, S.; Yang, J.; Shen, N.; Xu, Q.; Zhao, Q. Artificial intelligence in lung cancer diagnosis and prognosis: Current application and future perspective. *Semin. Cancer Biol.* 2023, *89*, 30–37. [CrossRef] [PubMed]
- 9. Phillips, I.; Ajaz, M.; Ezhil, V.; Prakash, V.; Alobaidli, S.; McQuaid, S.J.; South, C.; ScuffhAm, J.; Nisbet, A.; Evans, P. Clinical applications of textural analysis in non-small cell lung cancer. *BJR* **2018**, *91*, 20170267. [CrossRef]
- 10. Rogers, W.; Seetha, S.T.; Refaee, T.A.G.; Lieverse, R.I.Y.; Granzier, R.W.Y.; Ibrahim, A.; Keek, S.; Sandaleanu, S.; Primakov, S.; Beuque, M.P.L.; et al. Radiomics: From qualitative to quantitative imaging. *BJR* **2020**, *93*, 20190948. [CrossRef]
- 11. Chiu, H.Y.; Chao, H.S.; Chen, Y.M. Application of Artificial Intelligence in Lung Cancer. Cancers 2022, 14, 1370. [CrossRef]
- Forghani, R.; Savadjiev, P.; Chatterjee, A.; Muthukrishnan, N.; Reinhold, C.; Forghani, B. Radiomics and Artificial Intelligence for Biomarker and Prediction Model Development in Oncology. *Comput. Struct. Biotechnol. J.* 2019, 17, 995–1008. [CrossRef] [PubMed]
- Lovinfosse, P.; Hatt, M.; Visvikis, D.; Hustinx, R. Heterogeneity analysis of ¹⁸F-FDG PET imaging in oncology: Clinical indications and perspectives. *Clin. Transl. Imaging* 2018, *6*, 393–410. [CrossRef]
- Pellegrino, S.; Fonti, R.; Moghadam, H.; Torbati, A.; Bologna, R.; Morra, R.; Damiano, V.; Matano, E.; De Placido, S.; Del Vecchio, S. Heterogeneity of glycolytic phenotype determined by ¹⁸F-FDG PET/CT using coefficient of variation in patients with advanced non-small cell lung cancer. *Diagnostics* 2023, *13*, 2448. [CrossRef] [PubMed]
- Cook, G.J.; Yip, C.; Siddique, M.; Goh, V.; Chicklore, S.; Roy, A.; Marsden, P.; Ahmad, S.; Landau, D. Are pretreatment ¹⁸F-FDG PET tumor textural features in non-small cell lung cancer associated with response and survival after chemoradiotherapy? J. Nucl. Med. 2013, 54, 19–26. [CrossRef] [PubMed]
- Lovinfosse, P.; Janvary, Z.L.; Coucke, P.; Jodogne, S.; Bernard, C.; Hatt, M.; Visvikis, D.; Jansen, N.; Duysinx, B.; Hustinx, R. FDG PET/CT texture analysis for predicting the outcome of lung cancer treated by stereotactic body radiation therapy. *Eur. J. Nucl. Med. Mol. Imaging* 2016, 43, 1453–1460. [CrossRef]
- Park, S.; Ha, S.; Lee, S.H.; Paeng, J.C.; Keam, B.; Kim, T.M.; Dong-Wan Kim, D.W.; Heo, D.S. Intratumoral heterogeneity characterized by pretreatment PET in non-small cell lung cancer patients predicts progression-free survival on EGFR tyrosine kinase inhibitor. *PLoS ONE* 2018, *13*, e0189766. [CrossRef] [PubMed]
- Tixier, F.; Hatt, M.; Valla, C.; Fleury, V.; Lamour, C.; Ezzouhri, S.; Ingrand, P.; Perdrisot, R.; Visvikis, D.; Le Rest, C.C. Visual versus quantitative assessment of intratumor ¹⁸F-FDG PET uptake heterogeneity: Prognostic value in non-small cell lung cancer. *J. Nucl. Med.* 2014, 55, 1235–1241. [CrossRef] [PubMed]
- Alves, A.F.; Souza, S.A.; Ruiz, R.L.; Reis, T.A.; Ximenes, A.M.; Hasimoto, E.N.; Lima, R.P.S.; Miranda, J.R.A.; Pina, D.R. Combining machine learning and texture analysis to differentiate mediastinal lymph nodes in lung cancer patients. *Phys. Eng. Sci. Med.* 2021, 44, 387–394. [CrossRef]
- Gao, X.; Chu, C.; Li, Y.; Lu, P.; Wang, W.; Liu, W.; Yu, L. The method and efficacy of support vector machine classifiers based on texture features and multi-resolution histogram from ¹⁸F-FDG PET-CT images for the evaluation of mediastinal lymph nodes in patients with lung cancer. *Eur. J. Radiol.* 2015, *84*, 312–317. [CrossRef]
- Wu, J.; Aguilera, T.; Shultz, D.; Gudur, M.; Rubin, D.L.; Loo, B.W., Jr.; Diehn, M.; Li, R. Early-stage non-small cell lung cancer: Quantitative imaging characteristics of (18)F fluorodeoxyglucose PET/CT allow prediction of distant metastasis. *Radiology* 2016, 281, 270–278. [CrossRef]
- Manafi-Farid, R.; Askari, E.; Shiri, I.; Pirich, C.; Asadi, M.; Khater, M.; Zaidi, H.; Beheshti, M. [¹⁸F]FDG-PET/CT Radiomics and Artificial Intelligence in Lung Cancer: Technical Aspects and Potential Clinical Applications. *Semin. Nucl. Med.* 2022, *52*, 759–780. [CrossRef]
- Tau, N.; Stundzia, A.; Yasufuku, K.; Hussey, D.; Metser, U. Convolutional neural networks in predicting nodal and distant metastatic potential of newly diagnosed non-small cell lung cancer on FDG PET images. *AJR Am. J. Roentgenol.* 2020, 215, 192–197. [CrossRef] [PubMed]
- Nioche, C.; Orlhac, F.; Boughdad, S.; Reuzé, S.; Goya-Outi, J.; Robert, C.; Pellot-Barakat, C.; Soussan, M.; Frouin, F.; Buvat, I. LIFEx: A freeware for radiomic feature calculation in multimodality imaging to accelerate advances in the characterization of tumor heterogeneity. *Cancer Res.* 2018, 78, 4786–4789. [CrossRef]

- Pellegrino, S.; Fonti, R.; Mazziotti, E.; Piccin, L.; Mozzillo, E.; Damiano, V.; Matano, E.; De Placido, S.; Del Vecchio, S. Total metabolic tumor volume by ¹⁸F-FDG PET/CT for the prediction of outcome in patients with non-small cell lung cancer. *Ann. Nucl. Med.* **2019**, *33*, 937–944. [CrossRef] [PubMed]
- Im, H.J.; Pak, K.; Cheon, G.J.; Kang, K.W.; Kim, S.J.; Kim, I.J.; Chung, J.K.; Kim, E.E.; Lee, D.S. Prognostic value of volumetric parameters of (18)F-FDG PET in non-small-cell lung cancer: A meta-analysis. *Eur. J. Nucl. Med. Mol. Imaging* 2015, 42, 241–251. [CrossRef] [PubMed]
- 27. Pellegrino, S.; Fonti, R.; Pulcrano, A.; Del Vecchio, S. PET-based volumetric biomarkers for risk stratification of non-small cell lung cancer patients. *Diagnostics* **2021**, *11*, 210. [CrossRef]
- Uribe, C.F.; Mathotaarachchi, S.; Gaudet, V.; Smith, K.C.; Rosa-Neto, P.; Benard, F.; Black, S.E.; Zukotynski, K. Machine Learning in Nuclear Medicine: Part 1—Introduction. J. Nucl. Med. 2019, 60, 451–458. [CrossRef]
- 29. Zukotynski, K.; Gaudet, V.; Uribe, C.F.; Mathotaarachchi, S.; Smith, K.C.; Rosa-Neto, P.; Benard, F.; Black, S.E. Machine Learning in Nuclear Medicine: Part 2—Neural Networks and Clinical Aspects. J. Nucl. Med. 2021, 62, 22–29. [CrossRef]
- Lambin, P.; Leijenaar, R.T.H.; Deist, T.M.; Peerlings, J.; de Jong, E.E.C.; van Timmeren, J.; Sanduleanu, S.; Larue, R.T.H.M.; Even, A.J.G.; Jochems, A.; et al. Radiomics: The bridge between medical imaging and personalized medicine. *Nat. Rev. Clin. Oncol.* 2017, 14, 749–762. [CrossRef]
- Gupta, B.; Rawat, A.; Jain, A.; Arora, A.; Dhami, N. Analysis of Various Decision Tree Algorithms for Classification in Data Mining. Int. J. Comput. Appl. 2017, 163, 15–19. [CrossRef]
- 32. Gardner-Lubbe, S. Linear discriminant analysis for multiple functional data analysis. J. Appl. Stat. 2021, 48, 1917–1933. [CrossRef]
- 33. Lowd, D.; Domingos, P. Naive Bayes models for probability estimation. In Proceedings of the ICML '05: 22nd International Conference on Machine Learning, Bonn, Germany, 7–11 August 2005; pp. 529–536. [CrossRef]
- 34. Uddin, S.; Haque, I.; Lu, H.; Moni, M.A.; Gide, E. Comparative performance analysis of K-nearest neighbour (KNN) algorithm and its different variants for disease prediction. *Sci. Rep.* **2022**, *12*, 6256. [CrossRef]
- 35. Teso-Fz-Betono, A.; Zulueta, E.; Cabezas-Olivenza, M.; Teso-Fz-Betono, A.; Fernandez-Gamiz, U. A Study of Learning Issues in Feedforward Neural Networks. *Mathematics* 2022, 10, 3206. [CrossRef]
- Sarkar, J.P.; Saha, I.; Sarkar, A.; Maulik, U. Machine learning integrated ensemble of feature selection methods followed by survival analysis for predicting breast cancer subtype specific miRNA biomarkers. *Comput. Biol. Med.* 2021, 131, 104244. [CrossRef] [PubMed]
- 37. Chang, C.; Ruan, M.; Lei, B.; Yu, H.; Zhao, W.; Ge, Y.; Duan, S.; Teng, W.; Wu, Q.; Qian, X.; et al. Development of a PET/CT molecular radiomics-clinical model to predict thoracic lymph node metastasis of invasive lung adenocarcinoma ≤ 3 cm in diameter. *EJNMMI Res.* 2022, *12*, 23. [CrossRef] [PubMed]
- Yoo, J.; Cheon, M.; Park, Y.J.; Hyun, S.H.; Zo, J.I.; Um, S.W.; Won, H.H.; Lee, K.H.; Kim, B.T.; Choi, J.Y. Machine learning-based diagnostic method of pre-therapeutic ¹⁸F-FDG PET/CT for evaluating mediastinal lymph nodes in non-small cell lung cancer. *Eur. Radiol.* 2021, 31, 4184–4194. [CrossRef]
- Wang, H.; Zhou, Z.; Li, Y.; Chen, Z.; Lu, P.; Wang, W.; Liu, W.; Yu, L. Comparison of machine learning methods for classifying mediastinal lymph node metastasis of non-small cell lung cancer from ¹⁸F-FDG PET/CT images. *EJNMMI Res.* 2017, 7, 11. [CrossRef]
- Guberina, M.; Herrmann, K.; Pöttgen, C.; Guberina, N.; Hautzel, H.; Gauler, T.; Ploenes, T.; Umutlu, L.; Wetter, A.; Theegarten, D.; et al. Prediction of malignant lymph nodes in NSCLC by machine learning classifiers using EBUS TBNA and PET/CT. *Sci. Rep.* 2022, 12, 17511. [CrossRef]
- 41. Zhong, Y.; Cai, C.; Chen, T.; Gui, H.; Deng, J.; Yang, M.; Yu, B.; Song, Y.; Wang, T.; Sun, X.; et al. PET/CT based cross-modal deep learning signature to predict occult nodal metastasis in lung cancer. *Nat. Commun.* **2023**, *14*, 7513. [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.