

Article

Classification of Pancreatic Ductal Adenocarcinoma Using MALDI Mass Spectrometry Imaging Combined with Neural Networks

Frederic Kanter ¹, Jan Lellmann ^{1,*}, Herbert Thiele ², Steve Kalloger ³, David F. Schaeffer ^{3,4,5}, Axel Wellmann ⁶ and Oliver Klein ^{7,*}

¹ Institute of Mathematics and Image Computing, Universität zu Lübeck, 23562 Luebeck, Germany

² Fraunhofer Institute for Digital Medicine MEVIS, 23562 Luebeck, Germany

³ Department of Pathology and Laboratory Medicine, University of British Columbia, Vancouver, BC V6T 1Z4, Canada

⁴ Pancreas Centre BC, Vancouver, BC V5Z 1G1, Canada

⁵ Division of Anatomic Pathology, Vancouver General Hospital, Vancouver, BC V5Z 1M9, Canada

⁶ Institute of Pathology, Wittinger Strasse 14, 29223 Celle, Germany

⁷ BIH Center for Regenerative Therapies, Berlin Institute of Health at Charité-Universitätsmedizin Berlin, 13353 Berlin, Germany

* Correspondence: lellmann@mic.uni-luebeck.de (J.L.); oliver.klein@bih-charite.de (O.K.)

Simple Summary: Pancreatic ductal adenocarcinoma (PDAC) accounts for more than 90% of all pancreatic malignancies, and has a generally poor prognosis. Although it is the most common neoplastic disease of the pancreas, differential diagnosis is hindered by the lack of accurate and reliable diagnostic assays. Identification of molecular signatures for PDAC diagnosis offers a solution to improve the clinical and patient management. However, comprehensive omics profiling is time- and relatively cost-intensive and limited by tissue heterogeneity. Thus, it is not implemented in clinical routine. In this study, we investigate the feasibility of MALDI-MSI in combination with a neural-network-based analysis for accurate classification of pancreatic-ductal-adenocarcinoma patients. We provide evidence of the usefulness of this technology to support PDAC assessment which is promising for pathological aid.

Abstract: Despite numerous diagnostic and therapeutic advances, pancreatic ductal adenocarcinoma (PDAC) has a high mortality rate, and is the fourth leading cause of cancer death in developing countries. Besides its increasing prevalence, pancreatic malignancies are characterized by poor prognosis. Omics technologies have potential relevance for PDAC assessment but are time-intensive and relatively cost-intensive and limited by tissue heterogeneity. Matrix-assisted laser desorption/ionization mass spectrometry imaging (MALDI-MSI) can obtain spatially distinct peptide-signatures and enables tumor classification within a feasible time with relatively low cost. While MALDI-MSI data sets are inherently large, machine learning methods have the potential to greatly decrease processing time. We present a pilot study investigating the potential of MALDI-MSI in combination with neural networks, for classification of pancreatic ductal adenocarcinoma. Neural-network models were trained to distinguish between pancreatic ductal adenocarcinoma and other pancreatic cancer types. The proposed methods are able to correctly classify the PDAC types with an accuracy of up to 86% and a sensitivity of 82%. This study demonstrates that machine learning tools are able to identify different pancreatic carcinoma from complex MALDI data, enabling fast prediction of large data sets. Our results encourage a more frequent use of MALDI-MSI and machine learning in histopathological studies in the future.

Keywords: MALDI mass spectrometry imaging; pancreatic ductal adenocarcinoma; neural-network models



Citation: Kanter, F.; Lellmann, J.; Thiele, H.; Kalloger, S.; Schaeffer, D.F.; Wellmann, A.; Klein, O. Classification of Pancreatic Ductal Adenocarcinoma Using MALDI Mass Spectrometry Imaging Combined with Neural Networks. *Cancers* **2023**, *15*, 686. <https://doi.org/10.3390/cancers15030686>

Academic Editors: Rita Casadonte, Kristina Schwamborn and Mark Kriegsmann

Received: 13 December 2022

Revised: 17 January 2023

Accepted: 20 January 2023

Published: 22 January 2023



Copyright: © 2023 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

Pancreatic ductal adenocarcinoma (PDAC) accounts for more than 90% of all pancreatic malignancies, and has generally a poor prognosis [1]. With a 5 year overall-survival of less than 8%, PDAC is the fourth most frequent cause of cancer-related deaths worldwide [2]. Projections indicate that the number of PDAC diagnoses as well as PDAC-related deaths will more than double in the next decade in the United States [3] and in European countries [4]. This is due to lifestyle habits such as alcohol and tobacco abuse, which are known to increase the risk of several other cancers and also seem to play a role in the development of PDAC [4–8]. Pancreatic ductal adenocarcinoma (PDAC) and ampullary carcinoma (AC) are gastrointestinal cancers with overlapping clinical symptoms [9]. Although several studies hypothesize that pathogeneses and molecular composition are different, the clinical regime and therapy remain similar [10]. Moreover, existing studies demonstrate that the 5 year overall-survival-rate of patients with PDAC is lower in comparison with patients with AC [11]. Pancreatic ductal adenocarcinoma of the pancreas and ampullary carcinoma arise in close proximity to each other [12]. As a result, differential diagnosis of PDAC remains clinically challenging. Differential diagnosis is hindered by the lack of accurate and reliable diagnostic assays. Identification of molecular signatures for PDAC diagnostics offers the possibility of improving the clinical and patient management. However, common tissue-based proteomic and genomic techniques require large amounts of homogenized tissue material, which does not enable a direct correlation of molecular alterations with tissue histology.

Matrix-assisted-laser-desorption/ionization (MALDI) imaging technology combines common mass spectrometry with histological approaches. This technique is suitable for analyzing molecules (e.g., metabolites, proteins, peptides, lipids and glycans) and their spatial distribution in a single tissue-section in an unsupervised and label-free manner [13–16]. MALDI mass spectrometry imaging (MALDI-MSI) enables the high-throughput determination of spatial molecular-signatures in a clinically acceptable period and at relatively low cost, in comparison to other omics technologies. This provides new capabilities to classify different patient subgroups and even supports prediction of disease progression and/or resistance development.

In previous studies, MALDI-MSI was applied to in situ proteomic analysis of pre-neoplastic lesions in pancreatic cancer in genetically engineered mice [17]. In the study, intraepithelial neoplasia (PanIN) and intraductal papillary mucinous neoplasm (IPMN) could be discriminated from normal pancreatic tissue and early pancreatic ductal adenocarcinoma. Further studies have shown differences in the chemical structure of phospholipids and their distribution patterns in human pancreatic islets with intra-islet spatial resolution using MALDI-MSI [18]. Besides the investigation of the underlying mechanism, several proof-of-concept studies demonstrate the potential of MALDI-MSI in combination with machine learning algorithms to identify peptide signatures of prognostic relevance in pancreatic cancer [19,20].

In recent years, neural networks gained great popularity in many machine learning tasks such as image and speech recognition [21], image segmentation [22], and various classification tasks. Convolutional neural networks especially outperform many classical machine-learning approaches [23]. Neural networks are able to approximate highly complex decision functions, through their layer structure. Each layer consists of so-called neurons, which themselves contain weights and a nonlinear gating function. During training, the weights are tuned to solve the task at hand, which typically goes hand in hand with intrinsic learning, a meaningful feature-representation of the given data [24]. In the context of MALDI-MSI, machine learning algorithms have been used less frequently. Most studies make use of statistical methods, such as hierarchical clustering [25] or linear discriminant analysis (LDA). The high-dimensional nature of MALDI-MSI spectra often prohibits applying machine learning directly to the raw spectrum. Thus, strategies for feature selection or dimension reduction are employed to enable machine learning tools to

be used. The most commonly used techniques are peak picking [26], principal component analysis (PCA), and non-negative matrix factorization (NMF) [27].

In this proof-of-concept study, we investigate the feasibility of a neural network-based analysis of MALDI-MSI data for accurate classification of pancreatic ductal adenocarcinoma (Figure 1). We explicitly rely on the inherent feature-selection capability of neural networks, and evaluate the feasibility of feeding the full-scale spectral data to the classification methods in order to reduce the problem-specific modeling overhead as well as human interaction, and allow all available data to be used.

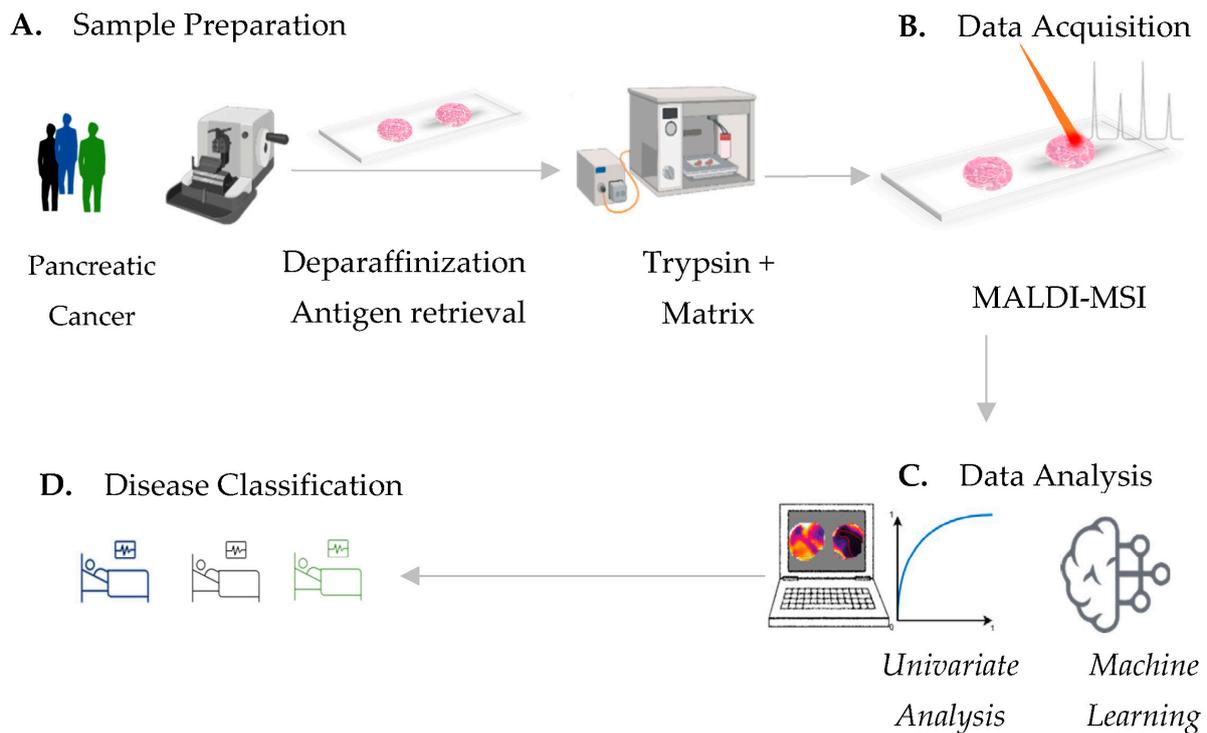


Figure 1. MALDI-MSI workflow for pancreatic-cancer-tissue assessment. (A) Pancreatic-cancer-tissue sample preparation includes deparaffinization, antigen retrieval followed by tryptic digestion and matrix application. (B) In situ determination of spatial mass to charge (m/z) signatures by means of MALDI mass spectrometry imaging. (C) Univariate testing of MALDI-MSI-derived data provides single discriminative m/z values, which can potentially identify distinct tissue types. Machine learning by neural-network strategies provides (D) an accurate classification of pancreatic-cancer types, which is potentially sufficient for clinical routine.

2. Materials and Methods

2.1. Patient and Sample Cohort

Tissue microarrays (TMAs) from formalin-fixed paraffin-embedded tissue from patients diagnosed with exocrine pancreatic cancer, in particular pancreatic ductal adenocarcinoma (PDAC) and ampullary carcinoma (AC) were prepared at the University of British Columbia Research (Table 1). The use of these samples is covered by ethical approval from the University of British Columbia Research Ethics Board H22-00073. Besides PDAC and AC, the investigated cohort included tissue material of other pancreatic cancer types, which are: acinar cell carcinoma, carcinoma NOS, benign chronic pancreatitis, intraductal papillary-mucinous carcinoma-invasive, intraductal papillary-mucinous carcinoma-noninvasive, mucinous cystic neoplasm-noninvasive, mucinous noncystic carcinoma, neuroendocrine tumor, pseudo-papillary tumor, serous cystadenoma, and signet-ring cell carcinoma.

Table 1. Clinicopathological characteristics of our patient cohort.

	Study Group	PDAC	AC	Other PC
Total <i>n</i>	446	260	103	83
Age at Surgery				
>60 years	311 (70%)	187 (72%)	74 (72%)	50 (60%)
<60 years	135 (30%)	73 (28%)	29 (28%)	33 (40%)
Sex				
female	211 (47%)	118 (45%)	45 (44%)	48 (58%)
male	235 (53%)	142 (55%)	58 (56%)	35 (42%)
Regional Lymph Nodes				
pN0: No regional lymph-node-metastasis	115 (26%)	64 (25%)	50 (49%)	1 (1%)
pN1: Regional lymph-node-metastasis	241 (54%)	180 (69%)	51 (50%)	10 (12%)
pNx: Cannot be assessed	90(20%)	16 (6%)	2(2%)	72 (87%)
Histologic Grade				
G1: Well differentiated	4 (1%)	3 (1%)	1(1%)	-
G2: Moderately differentiated	287 (64%)	181 (70%)	96 (93%)	10 (12%)
G3: Poorly differentiated	72 (16%)	65 (25%)	6 (6%)	1 (1%)
not assessed	94 (21%)	11 (4%)	-	72 (87%)

2.2. MALDI-MSI

Formalin-fixed paraffin-embedded (FFPE) tissue sections (tissue microarrays) were prepared as described before [28]. Briefly, 6- μ m thick tissue-sections were mounted onto conductive glass slides, coated with indium tin oxide (Bruker Daltonik GmbH, Bremen, Germany). Sections were preheated to 80 °C for 15 min, followed by paraffin removal, and heat-induced antigen retrieval. Trypsin solution (20 μ g Modified Porcine Trypsin in 800 μ L digestion buffer (20 mM ammonium bicarbonate with 0.01% glycerol) was applied by an automated spraying device (HTX TM-Sprayer, HTX Technologies LLC, ERC GmbH, Riemerling, Germany), at 30 °C. The tryptic digest was performed in a humidity chamber for 2 h, at 50 °C. Matrix solution (7 g/L *a*-cyano-4-hydroxycinnamic acid in 70% acetonitrile and 1% trifluoroacetic acid, at 75 °C) was applied using an HTX TM sprayer. A Rapiflex MALDI Tissue typer with flexImaging 5.1 and flexControl 3.0 software (Bruker Daltonik GmbH, Bremen, Germany) was used in positive-ion reflector mode, detection range of 800–3200 *m/z*, 500 laser-shots per spot, a sampling rate of 1.25 GS/s (gigasamples per second) and a raster width of 50 μ m for MALDI-MSI data acquisition. External calibration was carried out using a peptide calibration standard (Bruker Daltonik GmbH). After matrix removal, TMA sections were stained with hematoxylin and eosin, for histology. Tumor regions were digitally annotated in the QuPath open-source software by a pathologist, and transferred into SCiLS Lab software (Version 2019c Pro, Bruker Daltonik GmbH).

2.3. Processing of MALDI-MSI Data

MALDI-MSI raw data were converted to SCiLS Lab base data .sbd file and SCiLS Lab extended file .slx format using SCiLS Lab software version 2019c Pro (Bruker Daltonik GmbH). Data were set to total ion count without baseline removal. Patient tissue cores were categorized into ductal adenocarcinoma and non-PDAC (AC+ other pancreatic cancer types) attributes, to split the data into independent data sets (different tumor or patient-characteristics) for analysis. For peak detection and alignment, a standard segmentation pipeline (SCiLS Lab software) was used with the following parameters: width = 0.2 Da, maximal interval processing, total ion-count TIC normalization, medium noise reduction and no smoothing (Sigma: 0.75) [29,30].

2.4. Univariate Statistical Analyses

Supervised receiver-operating-characteristic (ROC) analyses were applied to identify m/z values, which are discriminative between tumor tissue regions of pancreatic ductal adenocarcinoma (PDAC) and ampullary-carcinoma (AC) tumor tissue regions. Area-under-the-curve values (AUC) close to 0 and 1 indicate that m/z values (peptides) are discriminatory. A comparable number of spectra must be used for the ROC analyses, and 10,000 spectra were randomly selected per group. Finally, m/z values with an AUC > 0.7 or <0.3 and a p -value < 0.001 (Wilcoxon rank-sum test) were selected as discriminative markers.

2.5. Protein Identification by Electrospray Ionization Tandem Mass Spectrometry

In order to identify the proteins corresponding to MALDI-MSI-derived m/z values (peptide), bottom-up liquid chromatography-based mass spectrometry (LC MS/MS) was performed on adjacent tissue sections, as previously described [26,31]. Briefly, tissue deparaffinization, antigen retrieval and tryptic digest were performed, as for the MALDI-MSI analyses. Using 40 μ L of 0.1% trifluoroacetic acid, peptides were extracted from the tissue section. The peptide solution was desalted and purified using a ZipTip[®] C18, following the manufacturer instructions. Eluates were vacuum concentrated (Eppendorf[®] Concentrator 5301, Eppendorf AG, Hamburg Germany) and reconstituted separately in 20 μ L 0.1% trifluoroacetic acid. A total of 2 μ L eluate were injected into a NanoHPLC (Dionex UltiMate 3000, Thermo Fisher Scientific, Waltham, MA, USA) coupled with an ESI-QTOF mass spectrometer (Impact II[™], Bruker Daltonic GmbH, Bremen, Germany). All raw spectra from the MS/MS measurement were converted to Mascot generic files (.mgf) using the ProteinScape software. Mascot search engine (version 2.4, MatrixScience; London, UK) and UniProt database were used to analyze mass spectra. The search was performed with the following set of parameters: (i) taxonomy: human; (ii) proteolytic enzyme: trypsin; (iii) peptide tolerance: 10 ppm; (iv) maximum of accepted missed cleavages: 1; (v) peptide charge: 2+, 3+, 4+; (vi) variable modification: oxidation (M); (vii) MS/MS tolerance: 0.8 Da; and (viii) MOWSE score > 25.

The comparison of MALDI-MSI (Supplementary Table S1) and LC-MS/MS m/z values (Supplementary Table S2) required the identification of more than one peptide (mass differences < 0.3 Da). The peptides with highest MOWSE peptide score and smallest mass differences between MALDI-MSI and nanoLC-MS/MS data were accepted as correctly identified.

2.6. Model Architectures for PDAC Classification

In order to classify PDAC, several neural network-based classifiers were employed. Firstly, a 2-layer residual network with skip connections between each layer, where the input of the preceding layer was passed unmodified to the subsequent layer as additional input. In the following, this model is denoted as Residual.

Secondly, an encoder-only variant of the Transformer architecture [32] was implemented. The size of the attention matrices used in this model is $n \times n$, where n denotes the sequence length. As each spectrum consists of several thousand data points, applying a Transformer to the full-scale sequence is unfeasible; therefore, the sequence length was reduced using a pooling layer with a kernel-size of 4, based on each peak consisting of 3 individual data points, before passing the input to the encoder. We refer to this architecture as Transformer-1/2.

All models were implemented using the PyTorch (version 1.3.1) framework, and the trainable weights were initialized by randomly selecting values from a truncated normal distribution. Experiments were conducted using the Adam optimizer and rectified-linear-units (ReLU) activation functions. Different hyperparameters such as learning rate, batch size and kernel size were tested. The configurations resulting in the overall best performance on the test set for each architecture are shown in Table 2.

Table 2. Parameter configurations for network models.

Model	Kernel Size	Channel	Learning Rate	Batch Size	Heads	Pooling Size
Residual	(200, 100)	(16, 8)	1×10^{-5}	100	-	-
Transformer 1	(256, 256)	(16, 8)	1×10^{-4}	100	2	4
Transformer 2	(512, 512)	(16, 8)	1×10^{-4}	500	2	8

2.7. Dataset Design

The data points were converted to NumPy arrays using the NumPy toolbox and a Python-based development environment. We kept the full mass range and stored the spectral data and corresponding label and coordinate information in HDF5 format. A custom tool was used to manually assign spectra to tissue samples. Patient tissue samples with fewer than 20 spectra were excluded, and spectra were normalized to unit median.

The TMA was randomly split into three subsets, which can be seen in Supplementary Figure S1. For both classes (Ductal, non-PDAC), around 70% of the data were used for training the machine learning algorithms, further divided into a training dataset (50% of patient samples) and a validation dataset (20%). The remaining 30% were used as a test dataset to evaluate the classification performance and were not used during the training phase.

We applied 3-fold cross-validation to create distinct data sets for training by repeating the random-splitting process three times (Dataset 1–3). The patient- and spectra-distributions in the training, validation, and test datasets varied slightly among different splits, due to the assignment process. The splitting was performed by selecting full patient core-tissue samples randomly without replacement, and assigning them and all associated spectra to one of the three datasets until the desired size was reached (Table 3, Supplementary Figure S1).

Table 3. Dataset sizes after random split (percentage of total in parentheses).

	Training	Validation	Testing
Dataset 1	267 (51.1%) cores 14,384 (50.3%) spectra	100 (19.2%) cores 5578 (19.5%) spectra	100 (19.2%) cores 8627 (30.2%) spectra
Dataset 2	270 (51.7%) cores 14,323 (50.2%) spectra	97 (18.6%) cores 5611 (19.7%) spectra	155 (29.7%) cores 8619 (30.1%) spectra
Dataset 3	254 (48.7%) cores 14,297 (50.1%) spectra	102 (19.5%) cores 5577 (19.5%) spectra	166 (31.8%) cores 8679 (30.4%) spectra

2.8. Filtering of Noise Spectra

We implemented a filter to remove spectra with little or no relevance, based on an informativeness score. The informativeness of each spectrum was measured in terms of the number of peaks greater than the variance within the spectrum. We considered a spectrum informative whenever the number of such peaks exceeded a predefined threshold. We restricted the range of data points to evaluate the first 60% of data points in a given spectrum, since there were few peaks present in higher Dalton ranges. We provide a visual interpretation of our measure of informativeness in Supplementary Figure S2.

2.9. Classification

Following the preprocessing steps above, the spectra and (during training) class-label information were directly passed to the neural networks. All network classifiers were trained on batched single spectra from the training data-set, monitoring the performance on the validation data-set. The model with the highest accuracy score on the validation data-set was selected to evaluate the performance on the unseen test-data.

Accuracy, sensitivity, and specificity metrics were computed on the unseen test-data in order to measure the performance for each classifier. This evaluation was performed on an individual spectrum level, counting the number of correctly classified spectra to determine the metric scores. In addition, a majority voting-strategy was employed to assign one class to each patient, based on the classes of all associated spectra in that patient's sample, and metrics were also evaluated on the patient level.

All metrics were averaged over the three different data splits, described in Section 2.7. All experiments were performed on a 2×6 -core Intel Xeon Gold 6128 CPU @ 3.40 GHz with 24 logical cores and $3 \times$ GeForce RTX 2080 Ti GPUs with 11 GB of memory each.

3. Results

3.1. Acquisition of MALDI-MSI Data

We evaluated the feasibility of MALDI-MSI to classify pancreatic ductal adenocarcinoma. Tissue samples were diagnosed from 446 patient-derived tumor-tissue specimens categorized as ductal adenocarcinoma patient ($n = 260$), non-PDAC ($n = 186$, including 103 ampullary carcinoma). The analyzed tumor-tissue-regions underwent peptide signature extraction, which resulted in 435 aligned m/z values in the mass range m/z 800–3200 Da for tryptic peptides. Representative average spectra from AC (the most common type in the non-PDAC group) and PDAC tissue sections (tumor region) are shown in Figure 2.

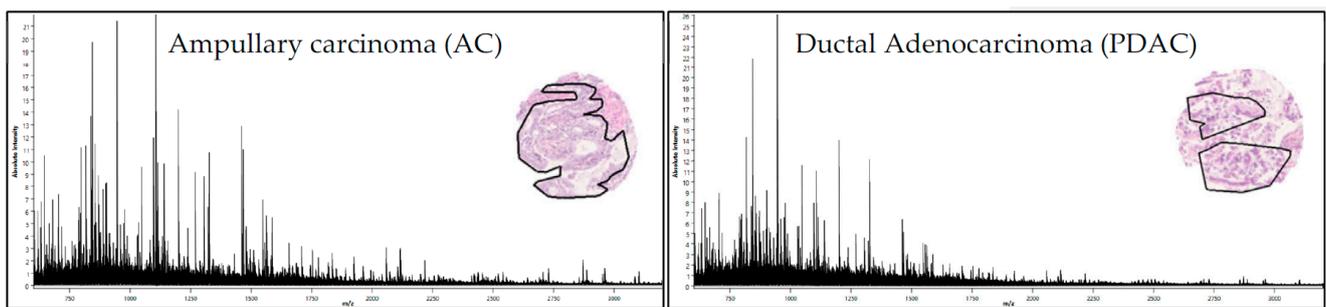


Figure 2. MALDI-MSI derived signature. Average spectra (435 m/z values) from AC (left) and PDAC (right) tumor regions (black line in the H&E staining of representative tissue cores).

3.2. Discriminative m/z Values between Pancreatic Ductal Adenocarcinoma and Ampullary Carcinoma

As pancreatic ductal adenocarcinoma and ampullary carcinoma arise in close proximity to each other, a differential diagnosis of PDAC remains clinically challenging. In order to demonstrate that the MALDI-MSI derived m/z values in PDAC are biologically relevant, we performed a univariate pairwise test (ROC) between PDAC and AC. Univariate analysis of MALDI-MSI data was used to determine single m/z locations that are discriminative between pancreatic ductal adenocarcinoma and ampullary carcinoma tissue (tumor regions). Receiver-operator-characteristic (ROC) analysis was applied to the total 435 aligned m/z peaks from tumor-region areas of tissue sections from PDAC and AC (pairwise comparisons). The intensity distributions of 131 m/z values could be identified as discriminative between PDAC and AC (AUC values of >0.7 or <0.3 ; $p < 0.001$). Representative selections are shown in Figure 3.

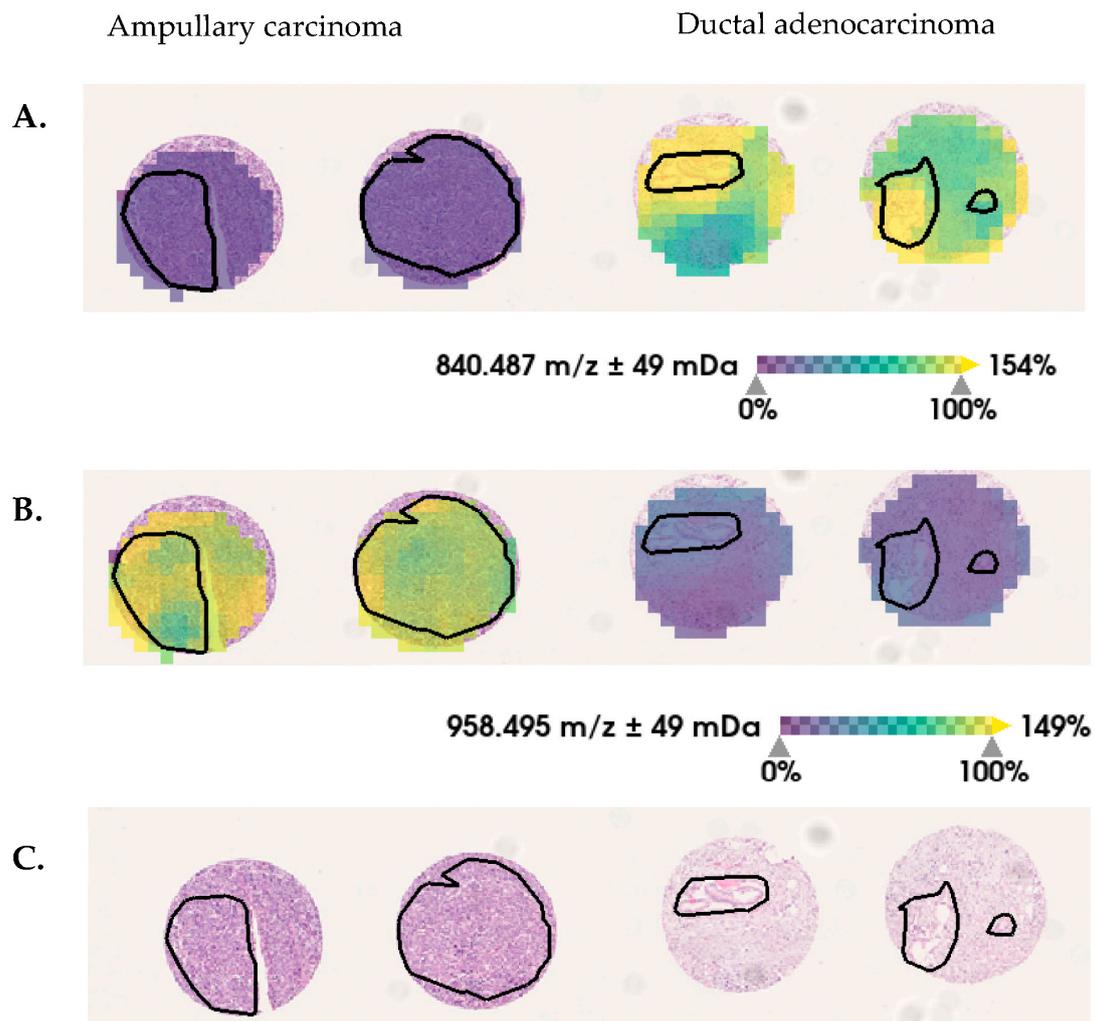


Figure 3. Selected discriminative m/z values between ductal-adenocarcinoma and ampullary-carcinoma pancreatic tumor tissue. In tumor regions (black lines), the relative intensity distribution of (A) m/z 840 Da are decreased and (B) m/z 958 Da are increased in ampullary carcinoma, in comparison with ductal adenocarcinoma. (C) Hematoxylin and eosin (H&E) staining is shown, for orientation.

3.3. Discriminative Proteins Identified from Pancreatic Carcinoma Tissues

MALDI-MSI-derived m/z values were identified using complementary nanoLC-MS/MS (Supplementary Table S2). Three proteins, corresponding to 7 MALDI-MSI derived m/z values (Table 4), fulfilled the criteria to be assumed as identified [26]. The m/z values (peptides; Table 4) from PLEC, AHNAK and COL6A3 proteins show significantly higher intensity-distributions in the tumor region of ductal adenocarcinoma, in comparison with ampullary carcinoma (Figure 4). The PLEC protein drives proliferation, migration and invasion in PDAC [33] and is a potential marker for identifying preinvasive lesions [34]. The AHNAK protein is involved in cell proliferation and migration, and thus may affect PDAC outcomes [35]. The COL6A3 protein is described in tissue as well as serum, and is a potential prognostic factor for pancreatic adenocarcinoma [36,37]. These findings support potentially interesting and plausible biological roles in pancreatic ductal adenocarcinoma.

Table 4. Differential intensity distributions of m/z values (MALDI-MSI) and their corresponding proteins in tissue sections from pancreatic ductal adenocarcinoma and ampullary carcinoma tissue (tumor regions).

MALDI IMS m/z Value	ROC [AUC] AC/PDAC	LC-MS/MS [Mr+H ⁺ cal.]	Deviation [Da]	MOWSE Score	Sequence	Protein	Gen Symbol
1459.7	0.721	1459.8631	−0.16	38.2	K.IGDLHPQIVNLLK.S	Collagen alpha-3(VI) chain	COL6A3
1586.8	0.700	1586.9265	−0.16	47.7	R.LQPVLQPLPSPGVGGK.R		
1267.7	0.715	1267.6529	0.01	53	K.AEGPEVDVNLPK.A	Neuroblast differentiation-associated protein AHNAK	AHNAK
1655.8	0.719	1654.8170	0.96	107.4	K.VDIEAPDVSLEGPEGL.L		
1461.7	0.710	1461.7366	−0.04	78.1	R.SQVMDEATALQLR.E	Plectin	PLEC
1479.8	0.714	1479.7914	−0.03	58.9	R.SLQEEHVAVQLR.E		
2115.1	0.701	2115.0175	0.08	69.6	R.AGTLSITEFADMLSGNAGGFR.S		

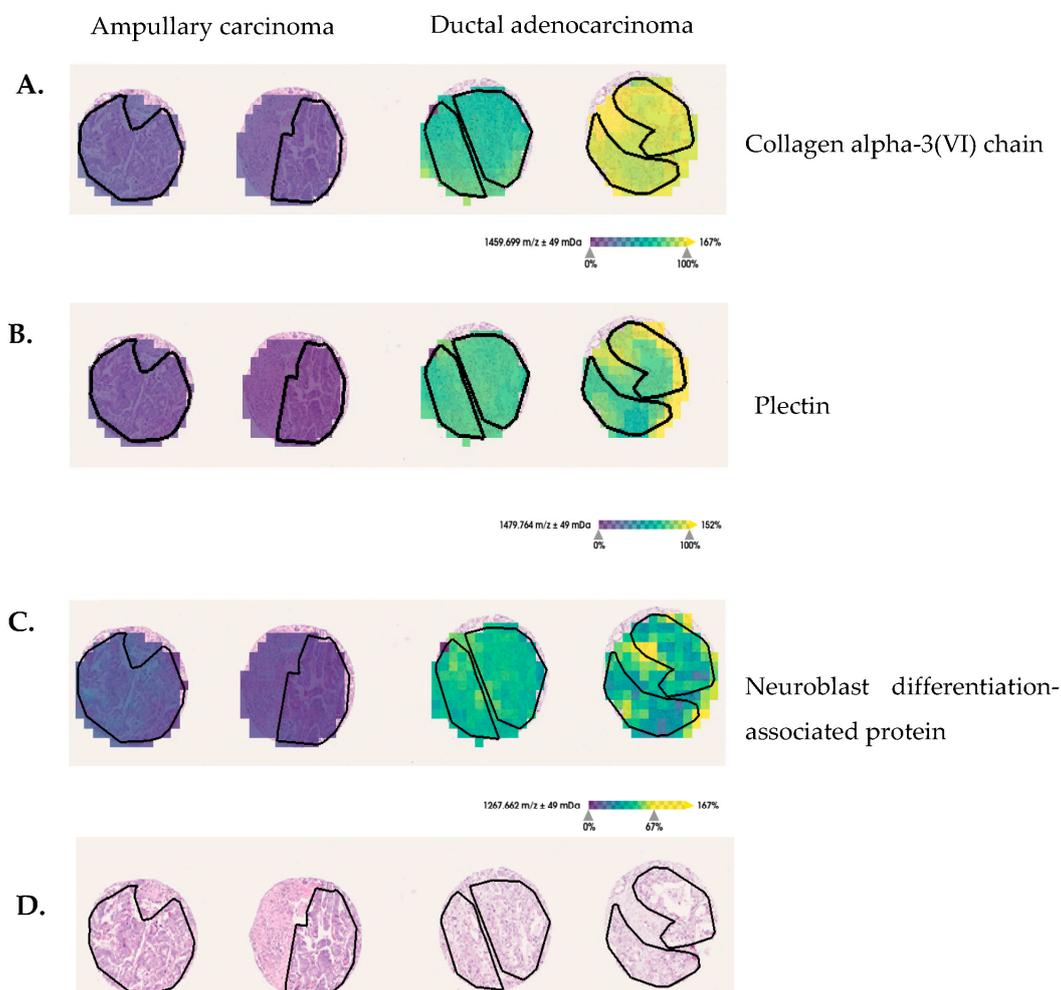


Figure 4. Discriminative protein markers for pancreatic-ductal-adenocarcinoma and ampullary-carcinoma tissue sections. (A) The m/z 1459 Da assigned to COL6A3, (B) m/z 1479 Da assigned to PLEC and (C) m/z 1267 Da assigned to neuroblast differentiation-associated protein show increased intensity-distributions in ductal adenocarcinoma, in comparison with ampullary-carcinoma tumor regions (black lines delineate tumor-border area). (D) For orientation, hematoxylin and eosin (H&E) staining are provided.

However, the highest AUC is 0.7, which needs to be improved to support PDAC assessment in a clinical setting. Mc Combie et al. [38] demonstrated the multivariate statistical treatment of MALDI imaging-data initially. Multivariate approaches are advantageous in

MSI, since they involve the complete mass-spectral information, reduce data dimensions, and obtain differences that are impossible to detect using univariate tests [39–41]. More sophisticated machine-learning algorithms aim at revealing more of the intrinsic information hidden in the data by exploring statistical correlations, retrieving data subsets with spectra similarities as well as applying supervised classification methods.

3.4. Pancreatic-Ductal-Adenocarcinoma Classifier Identification by Using Neuronal-Network-Evaluation Strategies

In order to increase prediction accuracy of PDAC, we explore the feasibility of MALDI-MSI combined with neuronal-network-evaluation strategies. To reduce the impact of imbalanced training data for convolutional neural networks [42], the classification was performed for PDAC ($n = 260$) against all non-PDAC tissues ($n = 186$, which includes 103 AC and 86 other pancreatic cancer types). The parameter search for all models was carried out using a random-search routine [43]. For the Residual and Transformer models, a high sensitivity to the choice of learning was observed on this data set; other parameters such as batch size and number of epochs did not show a similarly pronounced effect. Lower learning-rates result in the reported accuracies, whereas higher learning-rates hinder convergence during training. Different network depths were tested, but experiments showed that increasing the layer count ultimately harms the overall classification accuracy. Filtering out noise spectra as described in Section 2.8, resulted in the exclusion of approximately 13% of spectra (2183 measurements), resulting in a 3% increase in prediction accuracy for all tested models. The filter was applied to the whole data set before the creation of subsets for training and testing.

Interestingly, some of the samples were consistently labeled incorrectly, across the different model architectures. These samples may express low tumor-cell content and be dominated by the morphological structures of the tissue. Furthermore, the variance within the second class Non-PDAC is high, with multiple subtypes, which may express different spectral patterns. Neural networks tend to work better if the number of features (spectrum size) is relatively small, compared to the number of samples in the data set. However, further investigations with larger cohorts are needed to confirm this trend. Both the Residual model and the Transformer model achieve an average accuracy of 80% on the spectral and patient level over all splits. The highest accuracy on a single split is achieved by the Residual model, with a prediction accuracy of up to 86% on the spectral level and 86% on the patient level. The performance of the Transformer model on the same split is less accurate, with 85% on the spectral level and 86% on the patient level (see Table 5). The attention mechanism in the Transformer architecture does not improve prediction accuracy, compared to the skip connections in the Residual network.

The sensitivity of the Residual network for the ductal histotype is 0.79 on the spectral level and 0.82 for patient prediction. Similarly, the prediction sensitivity of the Transformer model is 0.81 (spectrum) and 0.83 (patient) for the ductal class. The Transformer model achieves a 7% gain on sensitivity compared to the Residual model. Due to the nature of the given two-class problem, the specificity of the Transformer model is 2% lower than the Residual specificity for the ductal class on the patient level (see Table 6).

Table 5. Results for neural-network models. Given are the model classification-accuracies for single spectra (spec) and full patient (sample) predictions.

Model	Split	Accuracy (Spec)	Accuracy (Sample)
Residual	I	0.86	0.86
	II	0.76	0.77
	III	0.77	0.76
		0.80	0.80

Table 5. *Cont.*

Model	Split	Accuracy (Spec)	Accuracy (Sample)
Transformer 1	I	0.85	0.86
	II	0.78	0.77
	III	0.77	0.77
		0.80	0.80
Transformer 2	I	0.83	0.84
	II	0.77	0.76
	III	0.78	0.79
		0.80	0.80

Table 6. Sensitivity and specificity metrics distinguished by histotype (PDAC and non-PDAC). Non-PDAC denotes all tumor and non-PDAC types labeled in the given data set. Reported are the results for all models.

Model	Class	Spot/Patient	Sensitivity	Specificity
Residual	PDAC	Spot	0.79	0.90
	Non-PDAC	Spot	0.90	0.79
	PDAC	Patient	0.82	0.90
	Non-PDAC	Patient	0.90	0.82
Transformer 1	PDAC	Spot	0.81	0.89
	Non-PDAC	Spot	0.89	0.81
	PDAC	Patient	0.83	0.88
	Non-PDAC	Patient	0.88	0.83
Transformer 2	PDAC	Spot	0.76	0.87
	Non-PDAC	Spot	0.87	0.76
	PDAC	Patient	0.78	0.89
	Non-PDAC	Patient	0.89	0.78

4. Discussion

MALDI-MSI combines spatial molecular (mass-spectrometric) analysis with conventional histological tissue-assessment. This technology enables the simultaneous analyzing of the spatial distribution of hundreds of m/z values without prior knowledge (label-free). MALDI-MSI is performed in high-throughput format (less than 5 min/mm² analysis time) with relatively low consumable cost (less than EUR one hundred per glass-slide). Tissue microarrays can be used to transfer up to 100 samples onto a single slide, again reducing the cost per patient-sample. These advantages make MALDI-MSI promising for identifying biomarker signatures and exploring tumor complexity in a clinically relevant format. In the presented study, we were able to demonstrate that the MALDI-MSI analysis results in biologically relevant m/z values to discriminate AC and PDAC by using univariate statistical analysis in combination with complementary nanoLC-MS/MS. The statistical analysis (receiver-operating-characteristic; ROC) results in 131 discriminative m/z values between pancreatic ductal adenocarcinoma and ampullary adenocarcinoma. Using complementary nanoLC-MS/MS, MALDI-MSI-derived m/z values could be assigned to three discriminative proteins: PLEC, Collagen CO6A3, and AHNK. However, direct identification of proteins, from which the m/z values (peptides, acquired by MALDI-MSI) stem, remain limited to only a few abundant proteins.

Recent studies have shown that high-resolution MSI data combined with microproteomics (high-resolution mass spectrometry) can be a valuable tool for protein assignment with high mass-accuracy and spatial specificity [44–46]. As a result, this strategy is a promising candidate for exploring potentially disease-causing protein changes in small patient collectives, but inadequate for large-scale studies because the processing time for both microdissection and mass spectrometry is longer, and the cost is higher.

Moreover, due to tissue heterogeneity, single-protein-marker classification usually does not result in sufficient accuracy for clinical routine. Several previous studies described the fact that MALDI-MSI-derived m/z value (signatures) in combination with supervised machine learning models enable a robust cancer-tissue classification [28,47–49]. In previous studies, we used MALDI-MSI to distinguish among four different epithelial ovarian-cancer histotypes [50] to predict a proteomic signature in early-stage ovarian cancer disease, which is a prognostic marker for recurrence [51], and to classify molecular subtypes of high-grade serous ovarian cancer [28].

In the present study, we apply this technique in combination with neural-network strategies to expose spatially resolved proteomic-signatures for pancreatic-ductal-adenocarcinoma classification. As data size and the high dimensionality of MALDI-MSI analyses still pose complex computational and memory requirements that hinder highly accurate identification of relevant molecular patterns, we explore the feasibility of MALDI-MSI in combination with neuronal-network-evaluation strategies [52]. We demonstrate that machine learning tools, in particular neural networks, given high-dimensional MALDI-MSI data are able to identify ductal carcinoma, giving high-dimensional MALDI-MSI data with an accuracy of up to 86%. MALDI-MSI combined with machine learning enables an accurate and quick PDAC prediction of large data sets, with a minimum of data preprocessing.

Fast and robust prediction is needed to enable the integration of MALDI-MSI in the clinical workflow. However, most studies apply a two-step pipeline for data processing [53], consisting of feature selection and subsequent classification steps. Firstly, features are selected either by hand, using the well-known practice of peak selection [54], or using a dimension-reduction algorithm such as principal-component analysis or non-negative matrix factorization [27], before applying an—often linear—classifier or a thresholding approach. This data-processing pipeline is time consuming, and selecting peaks (features) to increase classification performance can result in potentially valuable data being discarded, thus resulting in negative robustness. Therefore, in this work, we make use of the non-processed spectra, and do not apply any type of explicit feature selection. Non-linear spectral diversity has the potential to determine biologically relevant clusters for tissue assessment and clinical phenotypes prediction.

Neural networks can use non-linear mapping to reveal correlations in the spectral data which are not accessible with the established linear methods. In addition, these methods enable the combination of the feature-selection and classification stage in an end-to-end fashion [49]. Our model architectures are well suited to deal with sequential data such as MALDI-MSI data.

Consequently, we directly apply neural networks to the raw spectrum-data, and utilize the inherent-feature-extraction capabilities, similar to in the work of Behrmann et al. in [49]. In our previous work, we demonstrated that a convolutional network with skip-connections can differentiate four different subtypes of ovarian cancer when applied to MALDI-MSI data [50]. Our implementation is based on the original ResNet architecture [55], which famously increased the performance on the famous ImageNet classification challenge [56] by a large margin. We also considered the Transformer architecture, which has gained increasing popularity in fields such as language translation or caption generation, and is one of the most effective tools for processing sequential data: MALDI-MSI-derived spectra can be seen as sequential data, as each mass peak is linked to a specific detection-time. As we are interested in classification rather than sequence-to-sequence transformation, we only employ the encoder part of the original Transformer’s encoder-decoder design, with an added classification layer.

The capability of the proposed methods to extract robust features from the given spectral data is limited by the extent of noise. These features are crucial for the subsequent classification steps. The MALDI-MSI acquisition technique, due to its high resolution, results in noisy data. This sensitivity to noise hinders the learning process. The problem of acquisition-related noise is described in [29], and can be compensated for, but the problem of noise due to structurally non-informative spectral data points remains. In our work, we

implemented a filter to manage these data points. The proposed filter aims to reject noisy spectra before the classification step. Applying the filter to each spectrum results in a total of 2183 rejected spectra, due to their low informativeness, and improves classification by 3% across the models.

In total, our proposed method allows us to correctly classify ductal carcinoma with an accuracy of 86% and a sensitivity of 82%. The entire spectral data (full m/z range) can be used without time-consuming feature (m/z) selection. In upcoming studies, larger cohorts will be tested by a trained network to verify these findings. This will allow us to elevate our method from a classification algorithm to a more broadly applicable tool for diagnostic research.

5. Conclusions

Accurate and reliable diagnoses of pancreatic ductal adenocarcinoma (PDAC) are currently not adequately available. This pilot study demonstrates that (1) MALDI-MSI, combined with nanoLC-MS/MS, is a feasible pathway for identifying discriminative m/z values from corresponding proteins in pancreatic ductal adenocarcinoma versus ampullary carcinoma, which might have an important role in tumor progression, and (2) MALDI-MSI with neural-network strategies provides an accurate classification of PDAC without time-consuming feature-extraction methods such as peak picking. Moreover, we address the caveat of noisy data which is inherent to MALDI-imaging, utilizing only properties of the spectrum signal itself, without the need for user interaction. This study provides a proof-of-concept for the usefulness of the technology in assisting pancreatic-ductal-adenocarcinoma classification in a fast and cost-effective manner.

Supplementary Materials: The following supporting information can be downloaded at <https://www.mdpi.com/article/10.3390/cancers15030686/s1>, Supplementary Figure S1: TMA dataset and split. Supplementary Figure S2: Example for spectra with low and high topology. Supplementary Table S1: Aligned m/z values from cell-rich tumor region in tumor microarray cores from AC and PDAC. Supplementary Table S2: Identified m/z values by using nanoLC-MS/MS.

Author Contributions: O.K. and A.W. designed the study. S.K. provided patient samples and patient information. A.W. determined tumor content and histological sections. O.K. performed all MS experiments. F.K. programmed the evaluation tools and trained all machine learning algorithms. F.K. and J.L. performed data evaluation and interpretation on the machine-learning-model performances. J.L. served as student supervisor. F.K., O.K. and H.T. performed interpretation of the results in the context of MS and tumor typing. J.L., H.T., A.W., S.K. and D.F.S. provided intellectual input and advice on the experiments, as well as on the manuscript. F.K., O.K., H.T. and J.L. wrote the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the German Ministry for Education and Research in MSTAR (#031L0220A).

Institutional Review Board Statement: Ethics Statement: the use of these samples is covered by ethical approval from the University of British Columbia Research Ethics Board H22-00073.

Informed Consent Statement: Written consent was given by parents or legal guardians of the patients treated, according to the Ethics Board H22-00073 for use of residual tissue samples in research. No individual patients can be identified from data presented in this paper.

Data Availability Statement: Data is contained within the article or supplementary material. The MALDI-MSI data presented in this study are available on request from the corresponding author.

Acknowledgments: The authors thank the patients for participation in this study. We thank Grit Nebrich for device and Sylwia Handzik for histology support.

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

References

1. Kleeff, J.; Korc, M.; Apte, M.; La Vecchia, C.; Johnson, C.D.; Biankin, A.V.; Neale, R.E.; Tempero, M.; Tuveson, D.A.; Hruban, R.H.; et al. Pancreatic Cancer. *Nat. Rev. Dis. Prim.* **2016**, *2*, 16022. [[CrossRef](#)] [[PubMed](#)]
2. 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 Cancers in 185 Countries. *CA A Cancer J. Clin.* **2018**, *68*, 394–424. [[CrossRef](#)] [[PubMed](#)]
3. Rahib, L.; Smith, B.D.; Aizenberg, R.; Rosenzweig, A.B.; Fleshman, J.M.; Matrisian, L.M. Projecting Cancer Incidence and Deaths to 2030: The Unexpected Burden of Thyroid, Liver, and Pancreas Cancers in the United States. *Cancer Res.* **2014**, *74*, 2913–2921. [[CrossRef](#)]
4. Quante, A.S.; Ming, C.; Rottmann, M.; Engel, J.; Boeck, S.; Heinemann, V.; Westphalen, C.B.; Strauch, K. Projections of Cancer Incidence and Cancer-Related Deaths in Germany by 2020 and 2030. *Cancer Med.* **2016**, *5*, 2649–2656. [[CrossRef](#)]
5. Gapstur, S.M.; Jacobs, E.J.; Deka, A.; McCullough, M.L.; Patel, A.V.; Thun, M.J. Association of Alcohol Intake with Pancreatic Cancer Mortality in Never Smokers. *Arch. Intern. Med.* **2011**, *171*, 444–451. [[CrossRef](#)] [[PubMed](#)]
6. Pelucchi, C.; Galeone, C.; Polesel, J.; Manzari, M.; Zucchetto, A.; Talamini, R.; Franceschi, S.; Negri, E.; La Vecchia, C. Smoking and Body Mass Index and Survival in Pancreatic Cancer Patients. *Pancreas* **2014**, *43*, 47–52. [[CrossRef](#)] [[PubMed](#)]
7. Olson, S.H.; Chou, J.F.; Ludwig, E.; O'Reilly, E.; Allen, P.J.; Jarnagin, W.R.; Bayuga, S.; Simon, J.; Gonen, M.; Reisacher, W.R.; et al. Allergies, Obesity, Other Risk Factors and Survival from Pancreatic Cancer. *Int. J. Cancer* **2010**, *127*, 2412–2419. [[CrossRef](#)] [[PubMed](#)]
8. Delitto, D.; Zhang, D.; Han, S.; Black, B.S.; Knowlton, A.E.; Vlada, A.C.; Sarosi, G.A.; Behrns, K.E.; Thomas, R.M.; Lu, X.; et al. Nicotine Reduces Survival via Augmentation of Paracrine HGF–MET Signaling in the Pancreatic Cancer Microenvironment The Effect of Nicotine on c-Met in Pancreatic Cancer. *Clin. Cancer Res.* **2016**, *22*, 1787–1799. [[CrossRef](#)]
9. Uomo, G. Periampullary Carcinoma: Some Important News in Histopathology. *JOP. J. Pancreas* **2014**, *15*, 213–215.
10. Baghmar, S.; Agrawal, N.; Kumar, G.; Bihari, C.; Patidar, Y.; Kumar, S.; Chattopadhyay, T.K.; Panda, D.; Arora, A.; Pamecha, V. Prognostic Factors and the Role of Adjuvant Treatment in Periampullary Carcinoma: A Single-Centre Experience of 95 Patients. *J. Gastrointest. Cancer* **2019**, *50*, 361–369. [[CrossRef](#)]
11. Ferchichi, M.; Jouini, R.; Koubaa, W.; Khanchel, F.; Helal, I.; Hadad, D.; Bibani, N.; Chadli-Debbiche, A.; BenBrahim, E. Ampullary and Pancreatic Adenocarcinoma—A Comparative Study. *J. Gastrointest. Oncol.* **2019**, *10*, 270. [[CrossRef](#)] [[PubMed](#)]
12. Gonzalez, R.S.; Bagci, P.; Basturk, O.; Reid, M.D.; Balci, S.; Knight, J.H.; Kong, S.Y.; Memis, B.; Jang, K.-T.; Ohike, N.; et al. Intrapancreatic Distal Common Bile Duct Carcinoma: Analysis, Staging Considerations, and Comparison with Pancreatic Ductal and Ampullary Adenocarcinomas. *Mod. Pathol.* **2016**, *29*, 1358–1369. [[CrossRef](#)] [[PubMed](#)]
13. Heijs, B.; Holst-Bernal, S.; de Graaff, M.A.; Briaire-de Bruijn, I.H.; Rodriguez-Gironde, M.; van de Sande, M.A.; Wuhrer, M.; McDonnell, L.A.; Bovée, J.V. Molecular Signatures of Tumor Progression in Myxoid Liposarcoma Identified by N-Glycan Mass Spectrometry Imaging. *Lab. Invest.* **2020**, *100*, 1252–1261. [[CrossRef](#)]
14. Harris, A.; Roseborough, A.; Mor, R.; Yeung, K.K.-C.; Whitehead, S.N. Ganglioside Detection from Formalin-Fixed Human Brain Tissue Utilizing MALDI Imaging Mass Spectrometry. *J. Am. Soc. Mass Spectrom.* **2020**, *31*, 479–487. [[CrossRef](#)]
15. Boyle, S.T.; Mittal, P.; Kaur, G.; Hoffmann, P.; Samuel, M.S.; Klingler-Hoffmann, M. Uncovering Tumor–Stroma Inter-Relationships Using MALDI Mass Spectrometry Imaging. *J. Proteome Res.* **2020**, *19*, 4093–4103. [[CrossRef](#)] [[PubMed](#)]
16. Aichler, M.; Walch, A. MALDI Imaging Mass Spectrometry: Current Frontiers and Perspectives in Pathology Research and Practice. *Lab. Invest.* **2015**, *95*, 422–431. [[CrossRef](#)]
17. Grüner, B.M.; Hahne, H.; Mazur, P.K.; Trajkovic-Arsic, M.; Maier, S.; Esposito, I.; Kalideris, E.; Michalski, C.W.; Kleeff, J.; Rauser, S.; et al. MALDI Imaging Mass Spectrometry for in Situ Proteomic Analysis of Preneoplastic Lesions in Pancreatic Cancer. *PLoS ONE* **2012**, *7*, e39424. [[CrossRef](#)]
18. Prentice, B.M.; Hart, N.J.; Phillips, N.; Haliyur, R.; Judd, A.; Armandala, R.; Spraggins, J.M.; Lowe, C.L.; Boyd, K.L.; Stein, R.W.; et al. Imaging Mass Spectrometry Enables Molecular Profiling of Mouse and Human Pancreatic Tissue. *Diabetologia* **2019**, *62*, 1036–1047. [[CrossRef](#)]
19. Bollwein, C.; Gonçalves, J.P.L.; Utpatel, K.; Weichert, W.; Schwamborn, K. MALDI Mass Spectrometry Imaging for the Distinction of Adenocarcinomas of the Pancreas and Biliary Tree. *Molecules* **2022**, *27*, 3464. [[CrossRef](#)]
20. Casadonte, R.; Kriegsmann, M.; Perren, A.; Baretton, G.; Deininger, S.-O.; Kriegsmann, K.; Welsch, T.; Pilarsky, C.; Kriegsmann, J. Development of a Class Prediction Model to Discriminate Pancreatic Ductal Adenocarcinoma from Pancreatic Neuroendocrine Tumor by MALDI Mass Spectrometry Imaging. *Proteomics Clin. Appl.* **2019**, *13*, 1800046. [[CrossRef](#)]
21. Chorowski, J.K.; Bahdanau, D.; Serdyuk, D.; Cho, K.; Bengio, Y. Attention-Based Models for Speech Recognition. *Adv. Neural Inf. Process. Syst.* **2015**, *28*, 577–585.
22. Le'Clerc Arrastia, J.; Heilenkötter, N.; Otero Baguer, D.; Hauberg-Lotte, L.; Boskamp, T.; Hetzer, S.; Duschner, N.; Schaller, J.; Maass, P. Deeply Supervised UNet for Semantic Segmentation to Assist Dermatopathological Assessment of Basal Cell Carcinoma. *J. Imaging* **2021**, *7*, 71. [[CrossRef](#)] [[PubMed](#)]
23. LeCun, Y.; Bengio, Y.; Hinton, G. Deep Learning. *Nature* **2015**, *521*, 436–444. [[CrossRef](#)] [[PubMed](#)]
24. Goodfellow, I.; Bengio, Y.; Courville, A. *Deep Learning*; MIT Press: Cambridge, MA, USA, 2016.
25. Verbeeck, N.; Caprioli, R.M.; Van de Plas, R. Unsupervised Machine Learning for Exploratory Data Analysis in Imaging Mass Spectrometry. *Mass Spectrom. Rev.* **2020**, *39*, 245–291. [[CrossRef](#)]

26. Wu, Z.; Hundsdorfer, P.; Schulte, J.H.; Astrahantseff, K.; Boral, S.; Schmelz, K.; Eggert, A.; Klein, O. Discovery of Spatial Peptide Signatures for Neuroblastoma Risk Assessment by MALDI Mass Spectrometry Imaging. *Cancers* **2021**, *13*, 3184. [[CrossRef](#)] [[PubMed](#)]
27. Leuschner, J.; Schmidt, M.; Fernsel, P.; Lachmund, D.; Boskamp, T.; Maass, P. Supervised Non-Negative Matrix Factorization Methods for MALDI Imaging Applications. *Bioinformatics* **2019**, *35*, 1940–1947. [[CrossRef](#)] [[PubMed](#)]
28. Kassuhn, W.; Klein, O.; Darb-Esfahani, S.; Lammert, H.; Handzik, S.; Taube, E.T.; Schmitt, W.D.; Keunecke, C.; Horst, D.; Dreher, F.; et al. Classification of Molecular Subtypes of High-Grade Serous Ovarian Cancer by MALDI-Imaging. *Cancers* **2021**, *13*, 1512. [[CrossRef](#)] [[PubMed](#)]
29. Alexandrov, T.; Becker, M.; Deininger, S.; Ernst, G.; Wehder, L.; Grasmair, M.; Von Eggeling, F.; Thiele, H.; Maass, P. Spatial Segmentation of Imaging Mass Spectrometry Data with Edge-Preserving Image Denoising and Clustering. *J. Proteome Res.* **2010**, *9*, 6535–6546. [[CrossRef](#)]
30. Alexandrov, T.; Becker, M.; Guntinas-Lichius, O.; Ernst, G.; von Eggeling, F. MALDI-Imaging Segmentation Is a Powerful Tool for Spatial Functional Proteomic Analysis of Human Larynx Carcinoma. *J. Cancer Res. Clin. Oncol.* **2013**, *139*, 85–95. [[CrossRef](#)]
31. Klein, O.; Strohschein, K.; Nebrich, G.; Oetjen, J.; Trede, D.; Thiele, H.; Alexandrov, T.; Giavalisco, P.; Duda, G.N.; von Roth, P.; et al. MALDI Imaging Mass Spectrometry: Discrimination of Pathophysiological Regions in Traumatized Skeletal Muscle by Characteristic Peptide Signatures. *Proteomics* **2014**, *14*, 2249–2260. [[CrossRef](#)]
32. Vaswani, A.; Shazeer, N.; Parmar, N.; Uszkoreit, J.; Jones, L.; Gomez, A.N.; Kaiser, L.; Polosukhin, I. Attention Is All You Need. *Adv. Neural Inf. Process. Syst.* **2017**, *30*, 5998–6008.
33. Shin, S.J.; Smith, J.A.; Rezniczek, G.A.; Pan, S.; Chen, R.; Brentnall, T.A.; Wiche, G.; Kelly, K.A. Unexpected Gain of Function for the Scaffolding Protein Plectin Due to Mislocalization in Pancreatic Cancer. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, 19414–19419. [[CrossRef](#)] [[PubMed](#)]
34. Bausch, D.; Thomas, S.; Mino-Kenudson, M.; Fernández-del, C.C.; Bauer, T.W.; Williams, M.; Warshaw, A.L.; Thayer, S.P.; Kelly, K.A. Plectin-1 as a Novel Biomarker for Pancreatic Cancer. *Clin. Cancer Res.* **2011**, *17*, 302–309. [[CrossRef](#)] [[PubMed](#)]
35. Zhang, Z.; Liu, X.; Huang, R.; Liu, X.; Liang, Z.; Liu, T. Upregulation of Nucleoprotein AHNAK Is Associated with Poor Outcome of Pancreatic Ductal Adenocarcinoma Prognosis via Mediating Epithelial-Mesenchymal Transition. *J. Cancer* **2019**, *10*, 3860. [[CrossRef](#)]
36. Svoronos, C.; Tsoulfas, G.; Souvatzi, M.; Chatzitheoklitos, E. Prognostic Value of COL6A3 in Pancreatic Adenocarcinoma. *Ann. Hepato-Biliary-Pancreat. Surg.* **2020**, *24*, 52. [[CrossRef](#)]
37. Kang, C.Y.; Wang, J.; Axell-House, D.; Soni, P.; Chu, M.-L.; Chipitsyna, G.; Sarosiek, K.; Sendeci, J.; Hyslop, T.; Al-Zoubi, M.; et al. Clinical Significance of Serum COL6A3 in Pancreatic Ductal Adenocarcinoma. *J. Gastrointest. Surg.* **2014**, *18*, 7–15. [[CrossRef](#)]
38. McCombie, G.; Staab, D.; Stoekli, M.; Knochenmuss, R. Spatial and Spectral Correlations in MALDI Mass Spectrometry Images by Clustering and Multivariate Analysis. *Anal. Chem.* **2005**, *77*, 6118–6124. [[CrossRef](#)] [[PubMed](#)]
39. Deininger, S.-O.; Becker, M.; Suckau, D. Tutorial: Multivariate Statistical Treatment of Imaging Data for Clinical Biomarker Discovery. *Mass Spectrom. Imaging* **2010**, *656*, 385–403.
40. Jones, E.A.; van Remoortere, A.; van Zeijl, R.J.; Hogendoorn, P.C.; Bovée, J.V.; Deelder, A.M.; McDonnell, L.A. Multiple Statistical Analysis Techniques Corroborate Intratumor Heterogeneity in Imaging Mass Spectrometry Datasets of Myxofibrosarcoma. *PLoS ONE* **2011**, *6*, e24913. [[CrossRef](#)]
41. Veselkov, K.A.; Mirnezami, R.; Strittmatter, N.; Goldin, R.D.; Kinross, J.; Speller, A.V.; Abramov, T.; Jones, E.A.; Darzi, A.; Holmes, E.; et al. Chemo-Informatic Strategy for Imaging Mass Spectrometry-Based Hyperspectral Profiling of Lipid Signatures in Colorectal Cancer. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 1216–1221. [[CrossRef](#)]
42. He, H.; Garcia, E.A. Learning from Imbalanced Data. *IEEE Trans. Knowl. Data Eng.* **2009**, *21*, 1263–1284.
43. Bergstra, J.; Bengio, Y. Random Search for Hyper-Parameter Optimization. *J. Mach. Learn. Res.* **2012**, *13*, 281–305.
44. Dilillo, M.; Ait-Belkacem, R.; Esteve, C.; Pellegrini, D.; Nicolardi, S.; Costa, M.; Vannini, E.; De Graaf, E.; Caleo, M.; McDonnell, L. Ultra-High Mass Resolution MALDI Imaging Mass Spectrometry of Proteins and Metabolites in a Mouse Model of Glioblastoma. *Sci. Rep.* **2017**, *7*, 603. [[CrossRef](#)]
45. Mezger, S.T.; Mingels, A.M.; Bekers, O.; Heeren, R.M.; Cillero-Pastor, B. Mass Spectrometry Spatial-Omics on a Single Conductive Slide. *Anal. Chem.* **2021**, *93*, 2527–2533. [[CrossRef](#)]
46. Spraggins, J.M.; Rizzo, D.G.; Moore, J.L.; Noto, M.J.; Skaar, E.P.; Caprioli, R.M. Next-Generation Technologies for Spatial Proteomics: Integrating Ultra-High Speed MALDI-TOF and High Mass Resolution MALDI FTICR Imaging Mass Spectrometry for Protein Analysis. *Proteomics* **2016**, *16*, 1678–1689. [[CrossRef](#)]
47. Mittal, P.; Condina, M.R.; Klingler-Hoffmann, M.; Kaur, G.; Oehler, M.K.; Sieber, O.M.; Palmieri, M.; Kommoss, S.; Brucker, S.; McDonnell, M.D.; et al. Cancer Tissue Classification Using Supervised Machine Learning Applied to MALDI Mass Spectrometry Imaging. *Cancers* **2021**, *13*, 5388. [[CrossRef](#)]
48. Mascini, N.E.; Teunissen, J.; Noorlag, R.; Willems, S.M.; Heeren, R.M. Tumor Classification with MALDI-MSI Data of Tissue Microarrays: A Case Study. *Methods* **2018**, *151*, 21–27. [[CrossRef](#)]
49. Behrmann, J.; Etmann, C.; Boskamp, T.; Casadonte, R.; Kriegsmann, J.; Maaß, P. Deep Learning for Tumor Classification in Imaging Mass Spectrometry. *Bioinformatics* **2018**, *34*, 1215–1223. [[CrossRef](#)]

50. Klein, O.; Kanter, F.; Kulbe, H.; Jank, P.; Denkert, C.; Nebrich, G.; Schmitt, W.D.; Wu, Z.; Kunze, C.A.; Sehouli, J.; et al. MALDI-Imaging for Classification of Epithelial Ovarian Cancer Histotypes from a Tissue Microarray Using Machine Learning Methods. *Proteomics Clin. Appl.* **2019**, *13*, 1700181. [[CrossRef](#)]
51. Kulbe, H.; Klein, O.; Wu, Z.; Taube, E.T.; Kassuhn, W.; Horst, D.; Darb-Esfahani, S.; Jank, P.; Abobaker, S.; Ringel, F.; et al. Discovery of Prognostic Markers for Early-Stage High-Grade Serous Ovarian Cancer by MALDI-Imaging. *Cancers* **2020**, *12*, 2000. [[CrossRef](#)]
52. Abdelmoula, W.M.; Lopez, B.G.-C.; Randall, E.C.; Kapur, T.; Sarkaria, J.N.; White, F.M.; Agar, J.N.; Wells, W.M.; Agar, N.Y. Peak Learning of Mass Spectrometry Imaging Data Using Artificial Neural Networks. *Nat. Commun.* **2021**, *12*, 5544. [[CrossRef](#)] [[PubMed](#)]
53. Klein, O.; Fogt, F.; Hollerbach, S.; Nebrich, G.; Boskamp, T.; Wellmann, A. Classification of Inflammatory Bowel Disease from Formalin-Fixed, Paraffin-Embedded Tissue Biopsies via Imaging Mass Spectrometry. *Proteomics Clin. Appl.* **2020**, *14*, 1900131. [[CrossRef](#)] [[PubMed](#)]
54. Boskamp, T.; Lachmund, D.; Casadonte, R.; Hauberg-Lotte, L.; Kobarg, J.H.; Kriegsmann, J.; Maass, P. Using the Chemical Noise Background in MALDI Mass Spectrometry Imaging for Mass Alignment and Calibration. *Anal. Chem.* **2019**, *92*, 1301–1308. [[CrossRef](#)] [[PubMed](#)]
55. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 27–30 June 2016; pp. 770–778.
56. Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. Imagenet Large Scale Visual Recognition Challenge. *Int. J. Comput. Vis.* **2015**, *115*, 211–252. [[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.