MALDI-Imaging for classification of epithelial ovarian cancer histotypes from a tissue microarray using machine learning methods.

Oliver Klein[1,4], Frederic Kanter[7], Hagen Kulbe[1, 3,5], Paul Jank[1,2], Carsten Denkert[1,2], Grit Nebrich [1,4], Wolfgang D. Schmitt[1,2], Zhiyang Wu[1,4], Catarina A. Kunze[1,2], Jalid Sehouli[1,3,5] Silvia Darb-Esfahani[1,8], Ioana Braicu [1,3,5], Jan Lellmann[7], Herbert Thiele[6] , Eliane T. Taube[1,2]

[1]Charité – Universitätsmedizin Berlin, corporate member of Freie Universität Berlin Humboldt-Universität zu Berlin, and Berlin Institute of Health

[2]Institute of Pathology, Charité – Universitätsmedizin Berlin, Berlin, Germany

[3]Department of Gynecology, Charité – Universitätsmedizin Berlin, Berlin, Germany

[4]Berlin-Brandenburg Center for Regenerative Therapies, Charité – Universitätsmedizin Berlin, Berlin, Germany

[5]Tumor Bank Ovarian Cancer Network (TOC), Berlin, Germany

[6]Fraunhofer – Institute for Medical Image Computing MEVIS, Lübeck, Germany

[7]Institute of Mathematics and Image Computing, Universität zu Lübeck, Lübeck, Germany

[8]Pathology Department Spandau, Berlin, Germany

Received: 03 01, 2018; Revised: 10 31, 2018; Accepted: 11 22, 2018

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1002/prca.201700181.

Key words: Imaging mass spectrometry, ovarian cancer, histotype classification, machine learning

Correspondence:

Dr. Oliver Klein

Berlin-Brandenburg Center for Regenerative Therapies,

Charité - Universitätsmedizin Berlin

Campus Virchow-Klinikum

Augustenburger Platz 1, 13353 Berlin, Germany

Tel.: 0049 30 450 566143

Fax: 0049 30 450 566 904

Email:oliver.klein@charite.de

Statement of clinical relevance

Epithelial ovarian cancer (EOC) is an inhomogeneous disease, each histological subtype being a separate entity with different morphology, molecular biology and clinical outcome. Reported response to standard therapy already points to different therapeutically approaches needed by EOC subtypes. Unfortunately, clinical studies requiring central histological review showed that misdiagnosis up to 15 % is not that uncommon. In the era of personalized medicine, accurate histological diagnosis is mandatory for taking the right clinical decision.

Therefore, the aim of this study was to examine the potential of MALDI-Imaging in combination with linear and nonlinear machine learning methods to classify EOC histological subtypes from tissue microarray (TMA). In this pilot study, we used linear discriminant analysis, support vector machines with linear and radial basis function kernels, neural network and a convolutional neural network to determine classifiers for EOC- histotypes, based on MALDI-Imaging-derived proteomic signatures. Our results emphasize the high

potential of MALDI-Imaging in combination with machine learning methods to distinguish EOC histological subtypes. Applications might include decision support and image analytics in disease diagnosis and development of new prognostic parameters

Abstract

Purpose:

Precise histological classification of epithelial ovarian cancer (EOC) has immanent diagnostic and therapeutic consequences, but remains challenging in histological routine. The aim of this pilot study was to examine the potential of MALDI-Imaging mass spectrometry in combination with machine learning methods to classify EOC histological subtypes from tissue microarray.

Experimental design:

Formalin-fixed-paraffin-embedded tissue of 20 patients with ovarian clear-cell, 14 low-grade serous, 19 high-grade serous ovarian carcinomas and 14 serous borderline tumors were analysed using MALDI-Imaging. Classifications were computed by linear discriminant analysis (LDA), support vector machines with linear (SVM-lin) and radial basis function kernels (SVM-rbf), a neural network (NN), and a convolutional neural network (CNN).

Results:

MALDI-Imaging and machine learning methods results in classification of EOC histotypes with mean accuracy of 80% for LDA, 80% SVM-lin, 74% SVM-rbf, 83% NN and 85% CNN. Based on sensitivity (69%-100%) and specificity (90-99%), CCN and NN were most suited to EOC classification.

Conclusion and clinical relevance:

The pilot study demonstrates the potential of MALDI-Imaging derived proteomic classifiers in combination with machine learning algorithms to discriminate EOC histotypes. Applications might support the development of new prognostic parameters in the assessment of EOC.

Introduction

Diagnosis of epithelial ovarian cancer (EOC) still means an infaust prognosis. Ovarian cancer is the fifth most common cause of female cancer death in the developed world and the first cause of death among gynaecological malignancies [1].

In the past, EOC was considered as one disease with different morphological manifestations. Last decade findings acknowledged EOC as a heterogeneous group of tumors, characterized by different molecular biology and clinical outcome. The WHO has recognized five main subtypes of EOC [2]. They differ in pathogenesis, molecular changes, and clinical behavior with diverse chemotherapeutic sensitivity and prognosis [3]. The most frequent tumors are serous OC, accounting for approx. 70% of EOC [4]. High-grade serous ovarian cancer (HGSOC) is derived from tubal epithelium and contains frequent p53 and BRCA mutations (deficiency in homologous recombination pathway) resulting in high genetic instability [5]. In contrast, low-grade serous ovarian carcinoma (LGSOC) contains frequent BRAF, KRAS and ERBB2 mutations on the molecular level [6]. Serous borderline tumors (sBOT) are considered the precursors of LGSOC and have similar morphologic and molecular patterns (Figure 1).

Endometrioid and clear cell (OCCC) carcinoma are associated with endometriosis, and published data suggest that some endometriotic foci are able to evolve to cancer. OCCC is a rare subtype of the ovarian cancer being characterized by specific clear cytoplasm and hobnail nuclei in conventional light microscopy and ARID1A mutations and activation of the PIK3/Akt pathway [7].

In case of mucinous and endometrioid carcinomas as tumors with specific features of ovarian cancer, the etiology is unclear, and at least for mucinous carcinoma a thorough clinical search frequently unveils another primary.

The correct histotyping of ovarian carcinomas is more important than ever, because the diverse clinico-pathological and molecular features of OC subtypes require different approaches in clinical management, including chemotherapeutic strategies and targeted therapies (personalized medicine).

Novel classification guidelines [2] and molecular profiling [8, 9] have been developed recently. However, clinical routine typically relies on histology and extensive immunohistochemistry of tumor biopsies. When considering PARP inhibitors, a BRCA

analysis is added. Novel biomarkers are still sparse in the routine management of EOC and therefore the assessment of OC histotyping remains challenging and misdiagnosis ranges around 15% [10]. We included sBOT in the analysis despite it not being an invasive carcinoma, as the differential diagnosis between sBOT and LGSOC, which share many morphologic and molecular characteristics, is often complicated.

Also, frequently serous carcinomas reveal areas with morphologic overlaps in low- or high-grade morphology. Furthermore, HGSOC can contain abundant OCCC changes. All this can hamper a distinct diagnosis [10].

Therefore, there is a clear medical need to define novel molecular classifiers to aid the determination of ovarian cancer histotypes. Commonly-used tissue-based techniques, such as liquid chromatography-based mass spectrometry or gene expression profiling, require large amounts of tissue material. Moreover, these methods do not enable a direct correlation between differentially-expressed molecular profiles and the tissue histology [11].

Matrix-assisted laser desorption/ionization (MALDI) imaging mass spectrometry (IMS) (in short MALDI-Imaging) has the advantage of combining morphological features with protein expression in tissue. This technique enables the spatially resolved tissue assessment via specific molecular signatures (e.g. proteins, peptides, lipids, and molecules of cell metabolites) and allows their correlation with alterations in tissue histology [12, 13]. Thus, MALDI-Imaging is promising for individualized pathological assessment and in particular for cancer diagnostics [14, 15]. Moreover, MALDI-Imaging enables the large-scale analysis of tumor tissue microarrays (TMA), which are an important tool in clinical studies [16] .

Despite the unquestionable advantages of MALDI-Imaging in visualizing the spatial distribution and relative abundance of biomolecules directly on-tissue, the produced data are complex and high-dimensional, which makes analysis and interpretation mathematically and computationally challenging. Given the large size of typical MALDI-Imaging datasets, robust bioinformatic strategies are necessary to extract meaningful information from the raw measurements. This is typically achieved by separating feature extraction (dimensionality-reduction) and classification stage. A common approach for the extraction of characteristic features is based on finding discriminating m/z- values in the data, often referred to as peak detection. The mass spectrum of a tissue sample has a high molecular information content

with a multitude of signals. Therefore the tissue condition is usually not determined by a single signal but by a number of different peaks, a signal pattern. In order to combine information from several correlated spectral features advanced methods are designed to retrieve molecular signatures [17] or characteristic patterns in the data [18].

Various mathematical tools can be used for objective evaluation and classification of the IMS profiles to determine the histotype of a tissue sample. These are - amongst others - genetic algorithms for cluster analysis, Hierarchical Cluster Analysis (HCA) [19], Principal Component Analysis (PCA), Probabilistic Latent Semantic Analysis (pLSA) [20], Nearest Neighbour Classifiers (kNN), as well as machine learning approaches such as Neural networks (NN) and derivatives (such as Linear Vector Quantization (LVQ), Neural Gas (NG), Self-Organizing Map (SOM)), Support Vector Machine (SVM) [21] [9] and Random Forest (RF) classifiers [22]. Boskamp et al., 2017 [18] used linear discriminant analysis (LDA) to classify features based on spectral patterns into tumor types. Multivariate classification methods, such as principle component analysis (PCA) and probabilistic latent semantic analysis (pLSA), are used to identify local tissue changes within injured skeletal muscles [23].

Besides the classical approach of a separate feature extraction and classification stage, a promising alternative is to learn features and classification in one step, in particular in light of recent developments in the area of large-scale neural-network based “deep learning" [24]. By combining layers of many simple non-linear transformations into a network, such methods can represent very complex decision functions after a suitable training stage. Unlike classical machine learning approaches, they are typically applied on the original data without substantial preprocessing: while some form of feature extraction often occurs in the first few layers, it is implicit and happens as an effect of the training process, rather than being prescribed.

In the area of image processing, deep learning methods typically implement convolutional neural networks (CNNs) in order to exploit spatial relations and retrieve high-level abstractions for the final classification stage of the network [25]. In Behrmann et al. 2018 [26] the authors propose an adapted architecture based on deep convolutional networks to handle the characteristics of mass spectrometry data, as well as a strategy to interpret the learned model in the spectral domain based on a sensitivity analysis. In Inglese

et al. 2017 [27] the authors found linear methods for unsupervised dimensionality reduction to be inadequate for tumor classification based on IMS data sets. For this reason, they adopted a deep unsupervised neural network based technique to map a 3D-DESI-MS dataset onto a 2-dimensional manifold.

In order to aid the histopathological assessment of EOC, the present pilot study investigates the potential of MALDI-Imaging derived proteomic classifiers in combination with machine learning algorithms to discriminate between different EOC histological subtypes (ovarian clear- cell, high-grade serous, low-grade serous carcinomas and serous borderline ovarian tumor) from tissue microarray (TMA).

2. Materials and Methods

Study Population and histopathological examination

Formalin-fixed, paraffin-embedded (FFPE) tumor tissue of 67 EOC patients with OCCC (20 patients) LGSOC (14 patients) HGSOC (19 patients) and sBOT (14 patients) were composited in one tissue microarray. The TMA was prepared as previously described [28] with two tissue cores for each patient. Patients were diagnosed at the Institute of Pathology of the Charité University Hospital. Confirmation of histologic type according to WHO criteria [WHO 2014] was performed by an experienced, board-approved pathologist (SDE). The patients included cases from the tumor bank ovarian cancer project (http://www.toc-network.de). Within TOC, clinical data regarding stage, surgical outcome, age, relapse, chemotherapy and response, as well as survival data were available (table 1). As sBOT is not a malignant disease, characterization was restricted to age.

Chemicals

Alpha-cyano-4-hydroxycinnamic acid (HCCA) was obtained from Bruker Daltonik (Bremen, Germany), ammonium bicarbonate, acetonitril (ACN), 0.1% trifluoroacetic acid from Fluka (St. Louis, USA), and trypsin from Promega (Madison, WI, USA). 100% trifluoroacetic acid

(TFA) was purchased from Merck (spectroscopy purity, Darmstadt, Germany) and ammonium hydrogen phosphate from Sigma (Sigma St. Louis, USA).

Tissue Preparation

Tissue preparation and MALDI-Imaging data acquisition were performed as previously described [13]. Briefly, paraformaldehyde (PFA)-fixed specimens were dehydrated by washing sequentially with increasing concentrations of ethanol, subsequently cleared in xylene, and embedded in paraffin. 6 µm formalin-fixed, paraffin-embedded sections (FFPE) from each subtype were transferred onto indium-tin-oxide slides (Bruker Daltonik, Bremen, Germany). Sections were dewaxed and passed through decreasing concentrations of ethanol according to an adapted protocol from Casadonte R et al. [29]. Antigen retrieval (AR) was performed for 20 min by 90°C in deionized water. After drying slides for 10 min, tryptic digestion was performed. Using an automated spraying device (HTX TM-Sprayer, HTX Technologies LLC, ERC GmbH Riemerling Germany) sixteen layers of tryptic solution (25µg, 20mM ammonium bicarbonate 1% glycerol) were applied onto the section with the following settings: Spray head temperature 30°C, flow rate 15 µl/min, 750 mm/min velocity, 2mm tracking space, CC pattern, 10 psi pressure, 2l/min gas flow, 40 mm nozzle height, without drying time. After tissue incubation (2 h at 50°C; moist chamber with statured potassium sulfate solution), matrix solution (7 g/L a-cyano-4-hydroxycinnamic acid in 70% acetonitrile and 1% trifluoroacetic acid) was applied using the HTX TM Sprayer (spray head temperature 75°C, flow rate 120 µl /min, 1200 mm/min velocity, 3 mm tracking space, HH pattern, 10 psi pressure, 2l/min gas flow, 40 mm nozzle height without drying time).

MALDI-Imaging analysis

MALDI-Imaging data acquisition was performed on a rapifleX MALDI Tissuetyper system (Bruker Daltonik GmbH, Bremen, Germany) operating in reflector mode, m/z 600–3500 detection range, 500 laser shots per spot, sampling rate of 1.25 GS/s, and raster width of 50 µm. The imaging run was coordinated by flexImaging 5.1 and flexControl 3.0 (Bruker Daltonik GmbH). External calibration was performed using a peptide calibration standard (Bruker Daltonik). Spectra processing was performed in flexAnalysis 3.0 (Bruker Daltonik). After the MALDI-Imaging experiments, the matrix was removed with 70% ethanol and the

tissue sections were stained with haematoxylin and eosin (HE) as histological overview staining [29]. Tumor regions (Region of interest ROI, core tumor region) were annotated in SCiLS cloud and transferred into SCiLS Lab software (Version 2019a Pro, Bruker Daltonik GmbH).

Data processing

Only the tumor regions from each biopsy ("core tumor regions") were used for the data evaluation, i.e., one patient "core" contains only tumor tissue. The EOC tissue specimens contained histotypes sBOT (19 core tumor regions/14 patients), HGSOC (31 core tumor regions/19 patients), LGSOC (26 core tumor regions/14 patients), and OCCC (35 core tumor regions /20 patients). MALDI-Imaging raw data were imported into the SCiLS Lab software (Version 2019a Pro, Bruker Daltonik GmbH) and converted to the SCiLS Lab file format. Tumor regions (region of interest, ROI) were annotated in SCiLS cloud, transferred into SCiLS Lab software and subjected to the built-in baseline removal. For classification, the data including spectra, m/z values, spectra coordinates, and ground-truth labels were exported from SCiLS Lab and converted to HDF5 format. Spectra were manually assigned to core tumor regions using a custom tool, and very small core tumor regions with less than 20 associated spectra were excluded from the evaluation. All spectra were normalized to unit median. After conversion, the dataset was randomly divided into subsets (Figure 2): for each of the four epithelial ovarian cancer histotype classes (sBOT, HGSOC, LGSOC, OCCC), approximately 70% of the spectra were used for training the classifiers, further randomly split into training dataset (50% of total) and validation dataset (approximately 20% of total). The remaining 30% were set aside until after the training process as a test dataset for evaluation. Exact percentages vary slightly, as splits were performed by selecting full core tumor regions randomly without replacement and assigning all spectra associated to each core tumor region to one of the three datasets until the desired size was reached (see table 2 for details). The splitting process was repeated three times with different randomization (Dataset 1-3).

Classification using linear discriminant analysis

Linear discriminant analysis (LDA) is a classification strategy based on modelling features in each class as a multivariate normal distribution and predicting class membership by maximizing the posterior probabilities. We applied LDA to our datasets using the scikit-learn library (version 0.19.1) [30].

Classification using support vector machines

Support Vector Machine (SVM) are a learning method used especially in the context of classification problems. SVMs use linear decision functions to solve classification tasks and can be augmented to non-linear decision functions using so-called kernel SVMs [31]. We investigated linear SVMs as well as radial basis function (RBF) kernel SVMs, which can solve more complex classification problems. Again, the pipeline was implemented using scikit-learn. For multi-class classification, binary classifiers are trained in a „one-against-one“ strategy, resulting in 6 trained binary classifiers for the 4-class classification problem. The final class label is determined by a voting strategy based on the outputs of the pairwise classifiers [32].

Classification using neural networks and convolutional neural networks (CNN),

Neural networks (NN), in particular in the form of convolutional neural networks (CNN), are widely-used machine learning algorithms employed for a number of tasks such as image- and speech recognition, and image segmentation. Neural network classifiers are based on specifying the structure of the decision function using a network of weighted non-linear “neurons” stacked into layers and tuning the weights during the training phase, which allows to solve highly complex classification problems, see Goodfellow *et al*., 2016 [25] for a more detailed introduction. We based our implementation on the TensorFlow (version 1.7) and Keras (version 2.1.6) frameworks. Two architectures were evaluated: Firstly, a simple linear network (NN) with a single fully-connected layer and softmax activation function based on one-hot encoding of the classes. Weights were randomly initialized from a truncated normal distribution and (crucially) L1-regularized with a weight of 0.001 in order to avoid overfitting,

given the limited amount of data and large feature dimension. Secondly, a two-layer network (CNN) with one convolutional layer (4 feature maps, kernel width 120, stride 1) with ReLU activation, one fully-connected layer, softmax activation and L1-regularization with weight 0.0001 was applied. Among all tested variants, these simple architectures yielded the best results, which we attribute to the relatively small number of qualitatively different tissue samples in this dataset, high feature dimension, and the consequential susceptibility to overfitting. Training was performed using the Adam optimizer at a learning rate of 0.001 with a reduction strategy and batch size of 150.

Evaluation

All classifiers were first trained on and applied to each spectrum individually on the training and validation datasets. Performance was recorded in terms of accuracy, specificity and sensitivity measured by the number of correctly classified spectra on the (unseen during training) test datasets. This serves as an estimate of how accurate the histotype can be predicted from a single MALDI spectrum, i.e., one point measurement. In a second step, each core tumor region was assigned a class (histotype) based on which class received the most votes on the set of all associated spectra. Performance was again measured on the unseen test dataset in terms of accuracy, sensitivity and specificity, in order to investigate how accurate the predominant histotype of a core tumor region – which comprises multiple spectra – can be predicted. All measurements were averaged over three different random splits (Dataset 1-3).

3. Results

Tumor samples (111 core tumor regions from 67 patients, 23,091 spectra at 8,668 points per spectrum) of four EOC histotypes (OCCC, sBOT, LGSOC, HGSOC) were measured using MALDI-Imaging and evaluated using linear discriminant analysis (LDA), support vector machines with linear (SVM-lin) and radial basis function kernels (SVM-rbf), a neural network (NN), and a convolutional neural network (CNN). The full dataset was split randomly into training, validation, and test datasets. The classifiers were trained on the training dataset and evaluated on the test dataset. The process was repeated three times (Dataset 1-3) and

yielded 11,749 to 12,059 spectra for the training set, 3,470 to 3,584 spectra for the validation set and 7,361 to 7,512 spectra for the test set (table 2).

EOC histotype prediction accuracy

The linear SVM (SVM-lin) results in a mean overall accuracy of 75% on an individual spectrum level. After voting, 80% of the core tumor regions (patient tumor samples) are assigned to the correct histotype. SVM with radial basis functions (SVM-rbf) yields a mean overall accuracy of 71% for individual spectra and 74% after voting (core tumor regions). The linear neural network performs similar to the linear SVM (NN, 74% spectra 83% core tumor regions). The convolutional neural network (CNN, 76% spectra 85% core tumor regions) performs slightly better. Linear discriminant analysis (LDA, 59% spectra 80% core tumor regions) yields the least accurate results (table 3). Figure 3 shows the EOC histotype predictions for the full core tumor regions on Dataset 1; see also supplementary figure 1 for the corresponding predictions on the individual spectra level.

EOC histotype specificity and sensitivity of the classification

On core tumor regions, all methods (LDA, SVM-lin, SVM-rbf, NN, CNN) achieved overall specificity above 91% on OCCC, sBOT and LGSOC, and 72%-93% on HGSOC (table 5). Specificity values for individual spectrum classification were above 84% (OCCC, sBOT, LGSOC) and 75%-87% (HGSOC).

Regarding sensitivity, LDA and SVM-rbf achieved the lowest scores down to 56% on core tumor regions 49% on spectra (table 4). The other methods scored in the range of 80%-100% on core tumor regions 75%-89% on spectra for OCCC, sBOT, and HGSOC. For the LGSOC histotype, sensitivity was lower at 64%-75% on core tumor regions 58%-59% on spectra.

Figure 4 shows confusion matrices for dataset 1 for all methods (LDA, SVM-lin, SVM-rbf, NN, CNN) for core tumor region classification; see supplementary figure 2 for the confusion matrices for spectra classification.

Besides indicating the percentage of correctly classified samples per class, the confusion matrices allow to characterize the frequency of misclassifications between all pairs of classes. In our experiments, OCCC, HGSOC, and sBOT histotypes were mostly classified correctly. In contrast, the percentage of correctly classified LGSOC tissue was considerably lower and below 50% for all methods, with LGSOC mainly being confused with OCCC and, to a lower extent, with sBOT.

4. Discussion

Our study demonstrates for the first time that MALDI-Imaging combined with machine learning approaches can classify different histologic subtypes of epithelial ovarian cancer. The different histologic types of EOC have different prognosis [33] and require different clinical management. Gene-expression profiling and proteomics methods commonly applied (e.g. liquid chromatography-based mass spectrometry) use tissue homogenates and require considerable amounts of tissue material. Furthermore, tissue heterogeneity can negatively influence the correctness of diagnostic characterization. This may explain that indeed gene-expression and proteomics analysis improve our understanding of epithelial ovarian cancer, but actionable biomarkers are still sparse in routine management of ovarian cancer. In contrast, MALDI-Imaging mass spectrometry requires a single section, e.g., from endoscopic biopsy, and retains the morphology tissue structure during analysis [34].

So far, most mass spectrometry imaging studies using classification attempts have addressed tumors in different organs or of major different histologic types, e.g., squamous and adenocarcinoma [11, 14, 35]

Aiming at differentiating different histologic subtypes of ovarian cancer, we addressed a more detailed question in pathologic routine. Morphology of sBOT, LGSC and HGSC can be rather similar and only since 2014 it is widely accepted that LGSC and HGSC are in fact different tumors [2], based on precursor lesions and molecular studies. In ovarian cancer MALDI-Imaging studies have been concerned to individual single markers distinguishing stroma and tumor tissue [36]. They indicate tumor characteristics such as proliferation or modulation of immune response [37]. Furthermore, Meding *et al*. demonstrate that tryptic peptide datasets generated by complementary liquid chromatography mass spectrometry facilitate peptide identification in MALDI-Imaging studies on ovarian cancer [38, 39]. Indeed,

it can be very interesting for the understanding of carcinogenesis to study single peaks. However, none of the previous studies incorporated histologic subtyping of epithelial ovarian cancer. Therefore results might confuse with differences between pathologic subtypes.

Supervised classification methods, like leave-one-out cross-validation methods (LMN) [40] or linear discriminant analysis [18] are used to characterize features in tumor subtypes by MALDI- Imaging derived proteomics signatures. Dedicated machine learning algorithms, like Support Vector Machine and Random Forest, are successfully applied for classification of six common cancer types based on proteomic profiling by MALDI-Imaging [11]. Using deep learning approaches [24], discrimination between head and neck cancer and non-cancerous epithelium based on nonlinear microscopic images could be successfully performed [41]. Other studies propose that convolutional neural network (CNN) enables to integrate mass spectrometry data and to determine challenging tumor classification tasks [42].

The previous studies emphasizes the potential of algorithms, to explore MALDI-Imaging derived proteomic classifiers for pathological aid. Therefore, in our pilot study we investigates the potential of different machine learning methods to be used in classification strategies based on MALDI-Imaging proteomic profiles to distinguish EOC. A tissue microarray (TMA) was constructed which includes four different histological types of EOC (OCCC, sBOT, LGSOC, HGSOC). According to studies from Meding et al 2012 [11] a splitting process was performed three times (Dataset 1-3) with different randomization to create a validation, training and test set.

The best overall EOC histotype prediction was achieved by using a convolutional neural network as classifier. For all three Datasets it yielded mean overall accuracies between 79% and 97% for core tumor region classification and between 70% and 82% for spectra classification. Linear SVM and neural network classifiers performed slightly worse with overall classification accuracies in the range from 76%-86% for core tumor region classification and up to 77% spectra classification on the tested Datasets. Furthermore, CNN yielded high specificity for all classes on both core tumor region and spectra classification. All tested classifiers achieved higher specificity (above 72% for tumor core region classification) than sensitivity. Linear discriminant analysis (LDA) and SVM-rbf performed worst in terms of sensitivity (71% and 56%) for LGSOC core tumor region classification. The performance of LDA and SVM-rbf was considerably worse compared to the other three algorithms, only achieving mean accuracies of 80% and 74% core tumor region classification, as well as 59% and 71% for spectra classification. CNN, NN, and SVM-lin

were able to distinguish between the three EOC types OCCC, sBOT, and HGSOC resulting in sensitivity values from 82% to 100% (core tumor region classification).

In contrast, LDA and SVM-rbf instead more often confused sBOT with the other three types. An explanation might be that morphology and molecular composition are a continuous accumulation of malignant criteria in the transition from sBOT in LGSOC [43], which could influence the classification.

Further, the classification of LGSOC tissue appeared to be much more challenging for methods, with sensitivities in the 56%-75% range (core tumor region). The low-grade serous ovarian carcinoma (LGSOC) have only recently been described as a distinct an autologous histotype of EOC [2]. The molecular mechanisms and morphological properties are therefore not well known. Additionally, the number of samples was lower in comparison to the other groups, which can also lead to statistical complications. We expect that studies with a larger number of tissue samples and patients will allow to overcome these limitations.

Due to limitation in the co-registration of analysis and HE staining data even smaller stromal regions may be present in the tumor regions. Hence, classification based on all individual spectra results in decreased values of accuracy, sensitivity and specificity.

The present pilot study used one single TMA to explore MALDI-Imaging and machine learning methods for the discrimination of EOC histotypes. Subsequent studies will have to investigate, based on a larger cohort, which impact technical variables could have on the classification robustness.

In conclusion, we demonstrate with the present pilot study that the proteomic information provided by MALDI-Imaging data combined with machine learning approaches can be a valuable add-on diagnostic tool for the histological assessment of epithelial ovarian cancer subtypes.

Conflict of interest

The authors declare no conflict of interest.

Acknowledgments

We thank Angelika Krajewski and Sylwia Handzik for providing excellent technical assistance. This work was supported by grants from the BCRT through funding by the German Federal Ministry of Education and Research (BMBF) and NVIDIA Corporation. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. BMBF supported this study under the Transcan project TH4Respons, grant no.: JTC 2014-121.

References

[1] Siegel RL, Miller KD, Jemal A. Cancer Statistics, 2017. CA: a cancer journal for clinicians. 2017;67:7-30.

[2] Meinhold-Heerlein I, Fotopoulou C, Harter P, Kurzeder C, Mustea A, Wimberger P, et al. The new WHO classification of ovarian, fallopian tube, and primary peritoneal cancer and its clinical implications. Archives of gynecology and obstetrics. 2016;293:695-700.

[3] Prat J, D'Angelo E, Espinosa I. Ovarian carcinomas: at least five different diseases with distinct histological features and molecular genetics. Human pathology. 2018;80:11-27.

[4] Lu Z, Chen J. [Introduction of WHO classification of tumours of female reproductive organs, fourth edition]. Zhonghua bing li xue za zhi = Chinese journal of pathology. 2014;43:649-50.

[5] Kurman RJ, Shih Ie M. Molecular pathogenesis and extraovarian origin of epithelial ovarian cancer--shifting the paradigm. Human pathology. 2011;42:918-31.

[6] Della Pepa C, Tonini G, Santini D, Losito S, Pisano C, Di Napoli M, et al. Low Grade Serous Ovarian Carcinoma: from the molecular characterization to the best therapeutic strategy. Cancer treatment reviews. 2015;41:136-43.

[7] Mabuchi S, Sugiyama T, Kimura T. Clear cell carcinoma of the ovary: molecular insights and future therapeutic perspectives. Journal of gynecologic oncology. 2016;27:e31.

[8] Feng Z, Wen H, Bi R, Ju X, Chen X, Yang W, et al. A clinically applicable molecular classification for high-grade serous ovarian cancer based on hormone receptor expression. Scientific reports. 2016;6:25408.

[9] Rojas V, Hirshfield KM, Ganesan S, Rodriguez-Rodriguez L. Molecular Characterization of Epithelial Ovarian Cancer: Implications for Diagnosis and Treatment. International journal of molecular sciences. 2016;17.

[10] Kobel M, Bak J, Bertelsen BI, Carpen O, Grove A, Hansen ES, et al. Ovarian carcinoma histotype determination is highly reproducible, and is improved through the use of immunohistochemistry. Histopathology. 2014;64:1004-13.

[11] Meding S, Nitsche U, Balluff B, Elsner M, Rauser S, Schone C, et al. Tumor classification of six common cancer types based on proteomic profiling by MALDI imaging. Journal of proteome research. 2012;11:1996-2003.

[12] Walch A, Rauser S, Deininger SO, Hofler H. MALDI imaging mass spectrometry for direct tissue analysis: a new frontier for molecular histology. Histochemistry and cell biology. 2008;130:421-34.

[13] Klein O. S, K., Nebrich, G., Oetjen, J., Trede, D., Thiele, H., Alexandrov, T., Giavalisco, P., Duda, G. N., von Roth, P., Geissler, S., Klose, J., Winkler, T. MALDI imaging mass spectrometry: discrimination of pathophysiological regions in traumatized skeletal muscle by characteristic peptide signatures. Proteomics. 2014;14:2249-60.

[14] Aichler M, Walch A. MALDI Imaging mass spectrometry: current frontiers and perspectives in pathology research and practice. Laboratory investigation; a journal of technical methods and pathology. 2015;95:422-31.

[15] Kriegsmann J, Kriegsmann M, Casadonte R. MALDI TOF imaging mass spectrometry in clinical pathology: a valuable tool for cancer diagnostics (review). International journal of oncology. 2015;46:893-906.

[16] Casadonte R, Longuespee R, Kriegsmann J, Kriegsmann M. MALDI IMS and Cancer Tissue Microarrays. Advances in cancer research. 2017;134:173-200.

[17] Harn YC, Powers MJ, Shank EA, Jojic V. Deconvolving molecular signatures of interactions between microbial colonies. Bioinformatics. 2015;31:i142-50.

[18] Boskamp T, Lachmund D, Oetjen J, Cordero Hernandez Y, Trede D, Maass P, et al. A new classification method for MALDI imaging mass spectrometry data acquired on formalin-fixed paraffin-embedded tissue samples. Biochimica et biophysica acta. 2017;1865:916-26.

[19] Deininger SO, Becker M, Suckau D. Tutorial: multivariate statistical treatment of imaging data for clinical biomarker discovery. Methods in molecular biology. 2010;656:385-403.

[20] Hanselmann M, Kirchner M, Renard BY, Amstalden ER, Glunde K, Heeren RM, et al. Concise representation of mass spectrometry images by probabilistic latent semantic analysis. Analytical chemistry. 2008;80:9649-58.

[21] Lagarrigue M, Alexandrov T, Dieuset G, Perrin A, Lavigne R, Baulac S, et al. New analysis workflow for MALDI imaging mass spectrometry: application to the discovery and identification of potential markers of childhood absence epilepsy. Journal of proteome research. 2012;11:5453-63.

[22] Hanselmann M, Kothe U, Kirchner M, Renard BY, Amstalden ER, Glunde K, et al. Toward digital staining using imaging mass spectrometry and random forests. Journal of proteome research. 2009;8:3558-67.

[23] Klein O, Strohschein K, Nebrich G, Fuchs M, Thiele H, Giavalisco P, et al. Unraveling local tissue changes within severely injured skeletal muscles in response to MSC-based intervention using MALDI Imaging mass spectrometry. Scientific reports. 2018;8:12677.

[24] LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521:436-44.

[25] Ian G, Yoshua B, Aaron C. Deep Learning: The MIT Press; 2016.

[26] Behrmann J, Etmann C, Boskamp T, Casadonte R, Kriegsmann J, Maass P. Deep learning for tumor classification in imaging mass spectrometry. Bioinformatics. 2018;34:1215-23.

[27] Inglese P, McKenzie JS, Mroz A, Kinross J, Veselkov K, Holmes E, et al. Deep learning and 3D-DESI imaging reveal the hidden metabolic heterogeneity of cancer. Chemical science. 2017;8:3500-11.

[28] Darb-Esfahani S, Wirtz RM, Sinn BV, Budczies J, Noske A, Weichert W, et al. Estrogen receptor 1 mRNA is a prognostic factor in ovarian carcinoma: determination by kinetic PCR in formalin-fixed paraffin-embedded tissue. Endocrine-related cancer. 2009;16:1229-39.

[29] Casadonte R, Caprioli RM. Proteomic analysis of formalin-fixed paraffin-embedded tissue by MALDI imaging mass spectrometry. Nature protocols. 2011;6:1695-709.

[30] Pedregosa F, Ga, #235, Varoquaux l, Gramfort A, Michel V, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res. 2011;12:2825-30.

[31] Ingo S, Andreas C. Support Vector Machines: Springer Publishing Company, Incorporated; 2008.

[32] Chih-Chung C, Chih-Jen L. LIBSVM: A library for support vector machines. ACM Trans Intell Syst Technol %@ 2157-6904. 2011;2:1-27.

[33] Kommoss S, Gilks CB, du Bois A, Kommoss F. Ovarian carcinoma diagnosis: the clinical impact of 15 years of change. British journal of cancer. 2016;115:993-9.

[34] Kim HK, Reyzer ML, Choi IJ, Kim CG, Kim HS, Oshima A, et al. Gastric cancer-specific protein profile identified using endoscopic biopsy samples via MALDI mass spectrometry. Journal of proteome research. 2010;9:4123-30.

[35] Kriegsmann M, Casadonte R, Kriegsmann J, Dienemann H, Schirmacher P, Hendrik Kobarg J, et al. Reliable Entity Subtyping in Non-small Cell Lung Cancer by Matrix-assisted Laser Desorption/Ionization Imaging Mass Spectrometry on Formalin-fixed Paraffin-embedded Tissue Specimens. Molecular & cellular proteomics : MCP. 2016;15:3081-9.

[36] Everest-Dass AV, Briggs MT, Kaur G, Oehler MK, Hoffmann P, Packer NH. N-glycan MALDI Imaging Mass Spectrometry on Formalin-Fixed Paraffin-Embedded Tissue Enables the Delineation of Ovarian Cancer Tissues. Molecular & cellular proteomics : MCP. 2016;15:3003-16.

[37] El Ayed M, Bonnel D, Longuespee R, Castelier C, Franck J, Vergara D, et al. MALDI imaging mass spectrometry in ovarian cancer for tracking, identifying, and validating biomarkers. Medical science monitor : international medical journal of experimental and clinical research. 2010;16:BR233-45.

[38] Meding S, Martin K, Gustafsson OJ, Eddes JS, Hack S, Oehler MK, et al. Tryptic peptide reference data sets for MALDI imaging mass spectrometry on formalin-fixed ovarian cancer tissues. Journal of proteome research. 2013;12:308-15.

[39] Elsner M, Rauser S, Maier S, Schone C, Balluff B, Meding S, et al. MALDI imaging mass spectrometry reveals COX7A2, TAGLN2 and S100-A10 as novel prognostic markers in Barrett's adenocarcinoma. Journal of proteomics. 2012;75:4693-704.

[40] Winderbaum L, Koch I, Mittal P, Hoffmann P. Classification of MALDI-MS imaging data of tissue microarrays using canonical correlation analysis-based variable selection. Proteomics. 2016;16:1731-5.

[41] Litjens G, Kooi T, Bejnordi BE, Setio AAA, Ciompi F, Ghafoorian M, et al. A survey on deep learning in medical image analysis. Medical image analysis. 2017;42:60-88.

[42] Behrmann J, Etmann C, Boskamp T, Casadonte R, Kriegsmann J, Maass P. Deep Learning for Tumor Classification in Imaging Mass Spectrometry. Bioinformatics. 2017.

[43] Hunter SM, Anglesio MS, Ryland GL, Sharma R, Chiew YE, Rowley SM, et al. Molecular profiling of low grade serous ovarian tumours identifies novel candidate driver genes. Oncotarget. 2015;6:37663-77.

Figure Legend

Figure 1. MALDI-Imaging spectra for ovarian cancer histotypes and corresponding haematoxylin and eosin **staining:** (A) serous borderline tumor (sBOT), (B) high-grade serous ovarian carcinoma (HGSOC), (C) low-grade serous ovarian carcinoma (LGSOC), (D) ovarian clear cell ovarian cancer carcinoma (OCCC). The average MALDI-Imaging spectrum of each histotype is shown on the right.

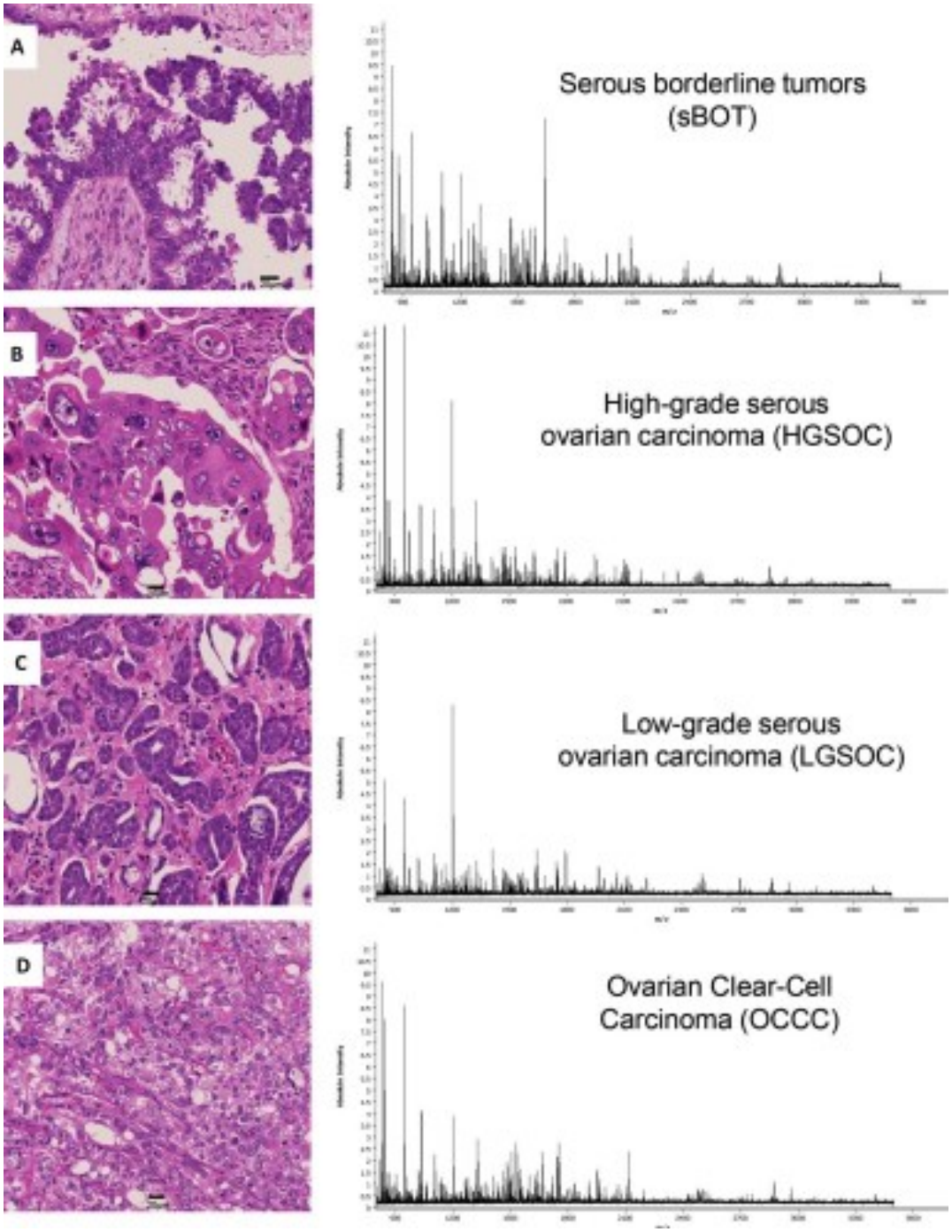

Figure 2. Epithelial ovarian cancer TMA dataset including ground truth class labels (a); exemplary randomized split ("Dataset 1") into 50% training data (b), 20% validation data (c) and 30% test data (d). All classifiers were trained on the training and validation sets (b), (c) and evaluated on the previously unseen, disjoint test dataset (d). Best viewed in color.

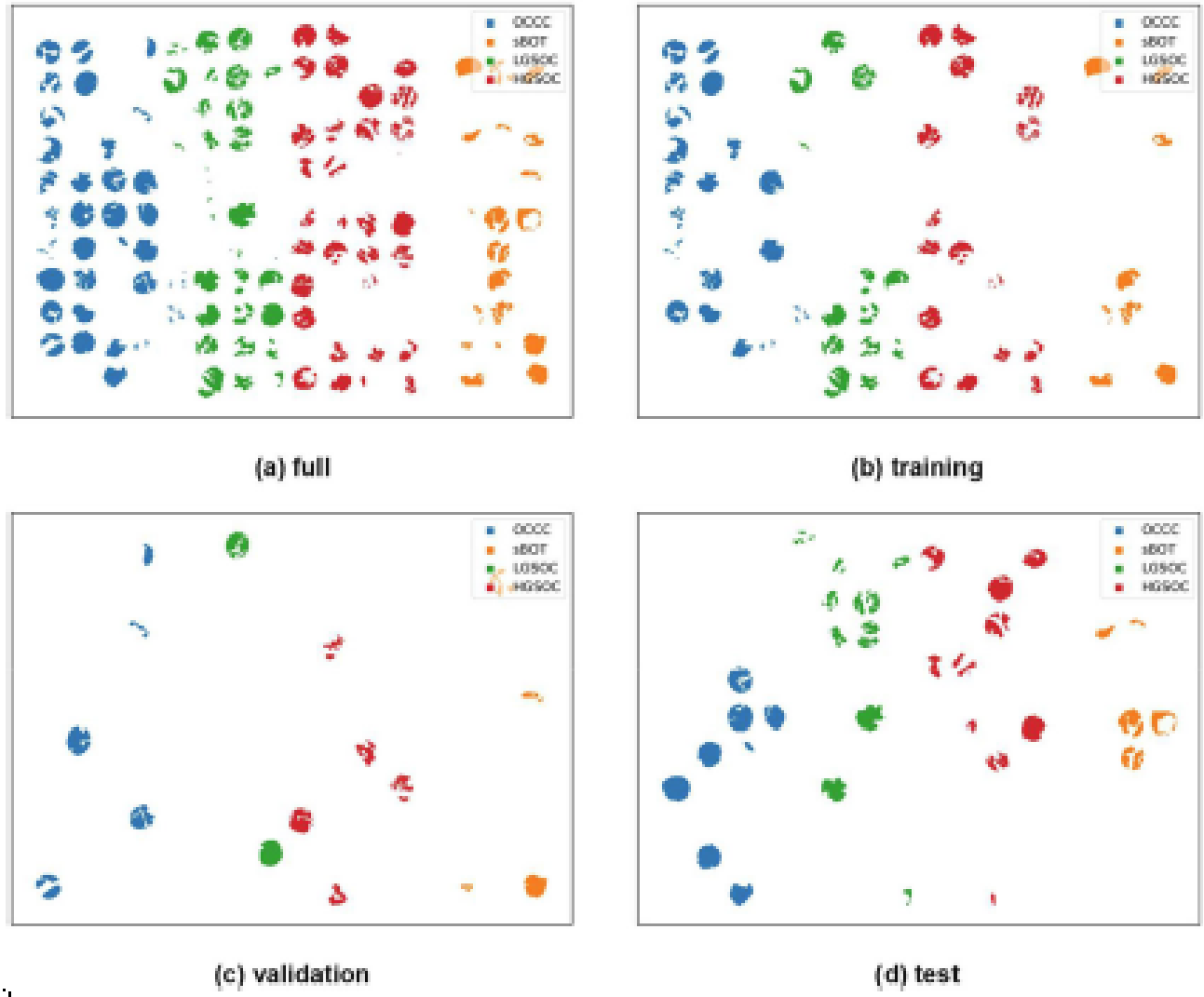

Figure 3. EOC histotype prediction from core tumor regions using MALDI-Imaging (Dataset 1). Classifiers were trained on the training dataset and applied to the unseen test dataset (a), followed by a voting scheme to assign a single histotype to each core tumor region. (b) LDA results in sensitivity of 60% to 93% and specificity of 88%-99%, (c) SVM-lin results in sensitivity of 64% to 93% and specificity 87%-99%, (d) SVM-rbf results in sensitivity of 56% to 93% and specificity of 72%-99%, (e) NN results in sensitivity of 75%-93% and specificity of 91%-99%, and (f) CNN results in sensitivity of 69%-100% and specificity of 90%-97%.

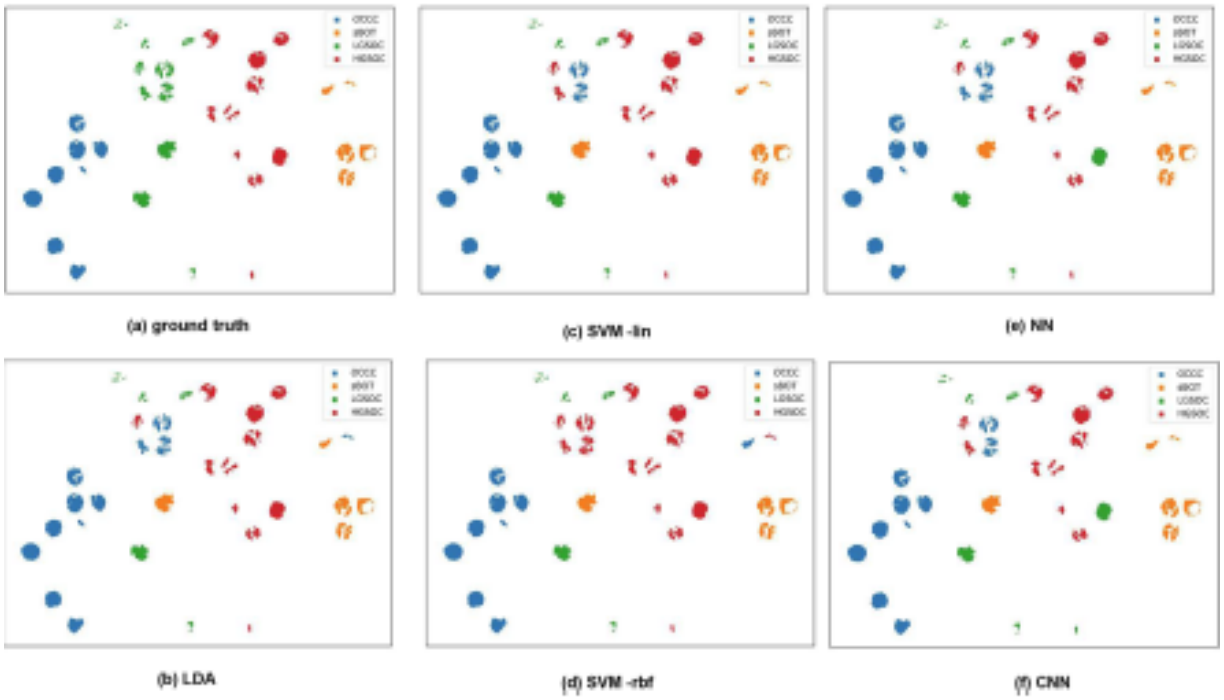

Figure 4. Confusion matrices for core tumor region classification on Dataset 1. Confusion matrices are shown for (a) linear discriminant analysis (LDA), (b) support vector machines with linear (SVM-lin) and (c) radial basis function kernels (SVM-rbf), (d) neural network (NN), and (e) convolutional neural network (CNN) classifiers. For each histotype, the numbers in the corresponding row indicate which fraction of all spectra/core tumor regions with the particular histotype were correctly classified (on the matrix diagonal) and incorrectly classified as each of the other histotypes (off-diagonal). While OCCC, sBOT and HGSOC were classified mostly correctly, accuracy for LGSOC spectra was lower at 50%.

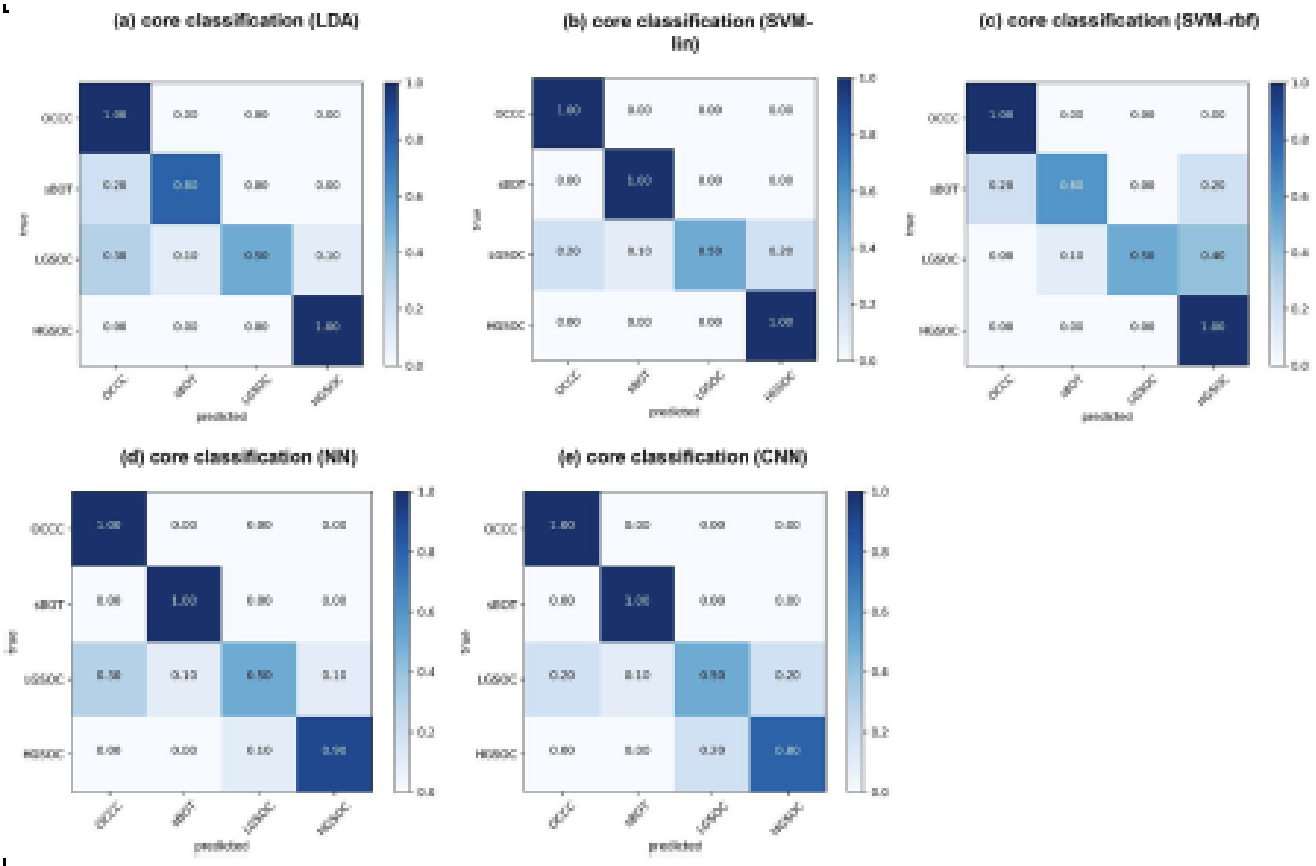

Table 1. Characteristics of the study group

Histotype	Study group n (%)	HGSC	LGSC	OCC	sBOT
Total	67	19	14	20	14
age					
<=60 years	30 (44.8%)	6 (31.6%)	6 (42.9%)	10 (50%)	8 (57.1%)
>60 years	37 (55.2%)	13 (68.4%)	8 (57.1%)	10 (50%)	6 (42.9%)
pT					
pT1	12 (22.6%)		1 (7.1%)	11 (55%)	
pT2	6 (11.3%)	3 (15.8%)	1 (7.1%)	2 (10%)	
pT3	35 (66%)	16 (84.2%)	12 (85.7%)	7 (35%)	
pN					
pN0	20 (37.7%)	5 (26.3%)	3 (21.4%)	12 (60%)	
pN1	22 (41.5%)	9 (47.4%)	8 (57.1%)	5 (25%)	
pNx	11 (20.8%)	5 (26.3%)	3 (21.4%)	3 (15%)	
FIGO stage					
FIGO I	11 (20.8%)		1 (7.1%)	10 (50%)	
FIGO II	3 (5.7%)		1 (7.1%)	2 (10%)	
FIGO III	31 (58.5%)	12 (63.2%)	12 (85.7%)	7 (35%)	
FIGO IV	8 (15.1%)	7 (36.8%)		1 (5%)	
residual tumor					

0 cm	24 (45.3%)	12 (63.2%)	7 (50%)	5 (25%)	
>0 cm	13 (24.5%)	6 (31.6%)	3 (21.4%)	4 (20%)	
Not done (FIGO I)	11 (20.8%)		1 (7.1%)	10 (50%)	
not determined	5 (9.4%)	1 (5.3%)	3 (21.4%)	1 (5%)	
CTX					
Platinum-based CTX	43 (81.1%)	17 (89.5%)	13 (92.9%)	13 (65%)	
No CTX	2 (3.8%)		1 (7.1%)	1 (5%)	
Not defined	8 (15.1%)	2 (10.5%)		6 (30%)	

Table 2. Dataset sizes after random split (percentage of total in parentheses). See Figure 1 for details.

	Training	Validation	Testing
Dataset 1	56 (53.3%) cores, 12051 (52.4%) spectra	16 (15.2%) cores, 3470 (15.1%) spectra	33 (31.4%) cores, 7483 (32.5%) spectra
Dataset 2	56 (53.3%) cores, 11749 (51.1%) spectra	20 (19.0%) cores, 3743 (16.3%) spectra	29 (27.6%) cores, 7512 (32.7%) spectra
Dataset 3	54 (51.4%) cores, 12059 (52.4%) spectra	18 (17.1%) cores, 3584 (15.6%) spectra	33 (31.4%) cores, 7361 (32.0%) spectra

Table 3. Prediction accuracy on the test datasets for spectra (values for cores in parentheses) using different classification algorithms.

	LDA	SVM-lin	SVM-rbf	NN	CNN
Dataset 1	0.61 (0.82)	0.77 (0.85)	0.76 (0.79)	0.76 (0.82)	0.76 (0.79)
Dataset 2	0.58 (0.79)	0.77 (0.79)	0.68 (0.72)	0.77 (0.86)	0.82 (0.97)
Dataset 3	0.57 (0.79)	0.72 (0.76)	0.70 (0.70)	0.69 (0.82)	0.70 (0.79)
Average	0.59 (0.80)	0.75 (0.80)	0.71 (0.74)	0.74 (0.83)	0.76 (0.85)

Table 4. Sensitivity on the test dataset for spectra (values for cores in parentheses) using different classification algorithms. Values are provided per EOC histotype and averaged over three different splits.

Histotype	LDA	SVM-lin	SVM-rbf	NN	CNN
OCCC	0.67 (0.93)	0.80 (0.80)	0.77 (0.79)	0.79 (0.85)	0.85 (0.97)
sBOT	0.49 (0.60)	0.75 (0.93)	0.56 (0.60)	0.78 (0.93)	0.80 (1.00)
LGSOC	0.53 (0.71)	0.59 (0.64)	0.49 (0.56)	0.58 (0.75)	0.59 (0.69)
HGSOC	0.59 (0.89)	0.83 (0.89)	0.89 (0.93)	0.78 (0.86)	0.76 (0.82)

Table 5. Specificity on the test dataset for spectra (values for cores in parentheses) using different classification algorithms. Values are provided per EOC histotype and averaged over three different splits.

Histotyp	LDA	SVM-lin	SVM-rbf	NN	CNN
OCCC	0.85 (0.92)	0.91 (0.94)	0.92 (0.96)	0.91 (0.94)	0.90 (0.97)
sBOT	0.95 (0.99)	0.98 (0.99)	0.98 (0.99)	0.97 (0.99)	0.98 (0.99)
LGSOC	0.84 (0.94)	0.92 (0.93)	0.95 (0.97)	0.91 (0.91)	0.92 (0.93)
HGSOC	0.79 (0.88)	0.86 (0.87)	0.75 (0.72)	0.84 (0.93)	0.87 (0.90)