IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 22 March 2024, accepted 11 April 2024, date of publication 15 April 2024, date of current version 23 April 2024.

Digital Object Identifier 10.1109/ACCESS.2024.3388841

TOPICAL REVIEW

Raman Spectroscopy and AI Applications in Cancer Grading: An Overview

PIETRO MANGANELLI CONFORTI[1], GIANMARCO LAZZINI[2], PAOLO RUSSO[1], AND MARIO D'ACUNTO[2]

[1]DIAG Department, Sapienza University of Rome, 00185 Rome, Italy
[2]CNR-IBF, Italian National Research Council, Institute of Biophysics, 56124 Pisa, Italy

Corresponding authors: Paolo Russo (paolo.russo@diag.uniroma1.it) and Mario D'Acunto (mario.dacunto@ibf.cnr.it)

ABSTRACT Raman spectroscopy (RS) is a label-free molecular vibrational spectroscopy technique that is able to identify the molecular fingerprint of various samples making use of the inelastic scattering of monochromatic light. Because of its advantages of non-destructive and accurate detection, RS is finding more and more use for benign and malignant tissues, tumor differentiation, tumor subtype classification, and section pathology diagnosis, operating either *in vivo* or *in vitro*. However, the high specificity of RS comes at a cost. The acquisition rate is low, depth information cannot be directly accessed, and the sampling area is limited. Such limitations can be contained if data pre- and post-processing methods are combined with current methods of Artificial Intelligence (AI), essentially, Machine Learning (ML) and Deep Learning (DL). The latter is modifying the approach to cancer diagnosis currently used to automate many cancer data analyses, and it has emerged as a promising option for improving healthcare accuracy and patient outcomes by abiliting prediction diseases tools. In a very broad context, Artificial Intelligence applications in oncology include risk assessment, early diagnosis, patient prognosis estimation, and treatment selection based on deep knowledge. The application of autonomous methods to datasets generated by RS analysis of benign and malignant tissues could make RS a rapid and stand-alone technique to help pathologists diagnose cancer with very high accuracy. This review describes the current milestones achieved by applying AI-based algorithms to RS analysis, grouped according to seven major types of cancers (Pancreatic, Breast, Skin, Brain, Prostate, Ovarian and Oral cavity). Additionally, it provides a theoretical foundation to tackle both present and forthcoming challenges in this domain. By exploring the current achievements and discussing the relative methodologies, this review offers recapitulative insights on recent and ongoing efforts to position RS as a rapid and effective cancer screening tool for pathologists. Accordingly, we aim to encourage future research endeavors and to facilitate the realization of the full potential of RS and AI applications in cancer grading.

INDEX TERMS Raman spectroscopy, cancer diagnosis, artificial intelligence, automated cancer diagnosis, machine learning, deep learning, tumor classification.

I. INTRODUCTION

The development of effective approaches to deal with cancer grading still remains one of the greatest challenges for pathologists [1]. Cancer is the leading cause of death in developed countries, and the number of cases is expected to grow as the population ages [2], [3]. Nowadays, many effective treatments are available (e.g., radiotherapy, chemotherapy, and chemoradiation). The effectiveness of the aforementioned therapeutic approaches is positively correlated to the low invasiveness, high speed and high accuracy of the diagnostic approaches [4]. In this sense, well-established approaches for cancer detection, such as Positron Emission Tomography (PET) or Computed Tomography (CT) show the drawback of employing ionizing radiation, a potential source of risk for the patient's health. On the other hand, Magnetic Resonance Imaging (MRI) cannot be employed on patients with metallic devices permanently implanted within the body,

The associate editor coordinating the review of this manuscript and approving it for publication was Larbi Boubchir.

due to the strong magnetic fields required. Furthermore, the frequent use of contrast liquids in combination with these techniques turns out to be a possible source of side effects, such as allergic reactions. Finally, ultrasound tomography exhibits good performance in detecting some types of tumor masses, but unfortunately produces images at low resolution. In recent years, novel and promising diagnostic approaches based on targeting cancer-related chemical compounds with so-called anti-cancer bioactive peptides (ACPs) have been developed. These molecules found several applications, either in tissue imaging, or as markers of cancer-related substances. Among them, we mention cACP [5], cACP-2LFS [6], iACP-GAEnsC [7], cACP-DeepGram [5], AIPs-SnTCN [8] and pAtbP-EnC [9]. Despite the advantages of ACPs in terms of rapid tissue uptake or rapid body release, future challenges will be represented by the high large-scale production cost and possible interference with the human immune system [10].

Optical diagnostic techniques represent a powerful, non-invasive, and cost-effective alternative to the approaches mentioned above. In particular, Fourier Transform Infrared spectroscopy exploits light absorption in the infrared spectral region to retrieve information about the molecular composition of a material [11]. However, the strong infrared absorption of water with consequent sample heating limits the use of this approach, especially for *in vivo* applications. Fluorescence spectroscopy exploits fluorescence to retrieve molecular information [12], [13]. Despite the wealth of fields of application [14], [15], [16], fluorescence spectroscopy often requires difficult and time-consuming staining procedures, aimed at increasing the fluorescence signal of the sample under investigation. Furthermore, fluorescence is accompanied by non-radiative relaxation processes responsible for the thermal damage of the sample and/or the progressive inactivation of the fluorescent emitters, with a consequent decrease of the fluorescence signal.

RS is a widely studied optical diagnostic technique that exploits the Raman effect [17] to identify molecule-specific vibrational states of the sample of interest [18]. The strong sensitivity of RS to the meolecular groups originating the Raman effect makes this approach capable of retrieving information about the chemical composition of several types of materials in a label-free and nondestructive fashion [19]. As well as for cancer diagnosis [19], [20], [21], [22], [23], this technique is already being employed in several clinical settings, for instance to analyze bacteria [24], [25], viruses [26], fungal infections in blood [27] and blood vessels conditions [28] but also skin and tissue damages as in the case of burns categorization [29]. Additionally, the non-destructive nature of RS makes it particularly suitable for the healthcare sector. For instance, it is experimentally reported in [22] that multiple RS applications with fixed laser and location parameters produced independent spectra without producing differences or damages in the analyzed tissue, thus highlighting the non-invasiveness property of the technique.

Traditionally, biopsy has been the standard methodology for the pathological diagnosis of cancerous tissues. However, this technique typically involves tissue sectioning and staining, while requiring pathologists with specialized training in order to interpret tissue samples [30]. The RS analysis can instead be applied directly to cancerous tissues to obtain the corresponding chemical composition, from which multiple information and patterns can be extracted and then employed to conduct a precise and fast analysis of the disease. In fact, improvements in the categorization method accuracy can lead to reductions in analysis costs, diagnostic delays, test pervasiveness, and the development of new technologies and software to support healthcare practitioners.

A very important limitation of RS imaging has always been its intrinsically low acquisition speed. The collection of RS images with lateral resolution between 10 and 20 μm (i.e. from 2500 to 10,000 spectra per mm^2) usually takes between 40 min to > 20 h/mm^2 (considering an acquisition time of 1 s per spectrum) due to the intrinsically weak signal generated by RS. Since histological samples are typically between 30 mm^2 and 1 cm^2 in size (reaching millions of spectra), diagnosis of entire tissue specimens, with a spatial resolution almost two orders of magnitude worse than that achievable by staining tissues with Hematoxylin and Eosin (H&E), as traditionally made by pathologists, would require many hours or even days. This is surely not compatible with intra-operatory diagnosis (total time from 30 to 120 min depending on the type of tumor) and still far from standard diagnostic procedures considering that the current protocols to obtain H&E-stained slices from the excised sample usually take 12–72 h [31]. Currently, RS cannot be used as a unique diagnostic approach without the intervention of pathologists. Some strategies have been suggested to circumvent the low speed of imaging procedures without necessarily increasing the acquisition speed or changing the acquisition modality [32], [33], [34]. These strategies, which are the focus of the present review, are used to combine RS with AI in order to accelerate diagnosis by data mining of spectral data.

The last decade saw a significant increase in the availability of data resources, as well as new, more effective, and efficient AI-based techniques. With the fast progress of computer science, several conventional study domains have been transformed by novel techniques, but numerous challenges remain in order to successfully integrate AI in clinical oncology [2]. In the remaining of the manuscript, with ML we consider the set of shallow learning techniques of Machine Learning, excluding the deep model ones, while with DL we refer to the subfield of ML that employs deep neural networks as learning techniques. Because ML and DL are applicable to many heterogeneous types of data, ranging from 3D medical images [35] to 1D signals, the implementation of tailored AI-based models can be particularly useful for data pre-processing and data modeling of Raman spectra. Generally, the automated analysis of RS data can be divided into two main steps: data pre-processing and data modeling.

The first procedure allows for the identification of patterns in the data that are not immediately detectable in the original feature space. For example, commonly employed techniques are wavenumber calibration, baseline or spike corrections, smoothing, noise removal, and more [36], [37], [38]. Regarding data modeling, several ML and DL models can be applied for the automated analysis of RS data. For instance, discriminant analysis, support vector machines (SVM), pre-trained neural networks, and other techniques are often used for classification tasks because they provide high accuracy and good generalization capabilities. A thorough discussion of the most promising approaches can be found in the next chapters.

In particular, our goal is to provide a concise yet informative summary of recent techniques utilized in the literature, with an emphasis on the most recent ML and DL techniques applied to RS data from both methodological and application aspects. Concerning the past reviews in the same field, our work stands out due to its exclusive attention to recent literature and the systematic categorization of results according to various types of tumors, as described in Section III. We will focus individually on each research work, and the overall analysis will be accompanied by a theoretical overview of the main RS and AI methodologies encountered. This approach will equip readers with the essential knowledge that is useful to test, improve, and advance implementations in this field.

The rest of the manuscript is organized as follows: In Section II a brief summary of recent reviews is provided. Then, the criteria for collecting and selecting relevant material from the publicly available resources are addressed in Section III, complemented by the methodology employed to compose this review. A comparison table 1 is presented hereafter, wherein the selected articles are systematically compared based on a set of salient parameters.

To understand the primary methodologies that can be encountered in the results, the fundamental steps associated with RS and the principles of ML theory are discussed in Section IV. In Section VI we present the results by summarizing and discussing all the selected publications according to seven major cancer categories. Finally, Section VII-B discusses our conclusive remarks and future perspectives on this field of study.

II. RELATED WORKS

With the advancement of technology, artificial intelligence provides an unprecedented opportunity to extract precise information from complex or large datasets in a variety of academic disciplines. Thanks to the flexibility of the RS technique, several ML and DL approaches have been utilized in the literature for the analysis of Raman spectra obtained from chemical structures, leading to successful classification, detection, and predictive systems. In particular, for the medical field, several reviews have been published in recent years concerning different characteristics and aspects of AI-based methodologies.

In the 2022 review of Luo et al. [36] typical algorithms used for the analysis of Raman spectra are described, with a focus on the DL models implemented in the field. In addition, the latter presents the recent applications for RS, with an emphasis on data pre-processing, classification, regression, and spectral data highlighting. The DL capabilities are outlined, describing the possibility of skipping feature extraction, data modeling, and pre-processing by implementing a single neural network model to obtain the correct analysis. In conclusion, model training and preprocessing are discussed as the major problems in the field.

Another literature review was conducted in 2022 by Blake et al. [37], to determine the most recent ML techniques used to classify malignant neoplasies while exploiting Raman spectral data. They discovered that DL models usually outperform their traditional ML comparisons, although a variety of methodological issues may have contributed to an overestimation of performance. In particular, the main issues demonstrated were small sample sizes as well as poor sampling and validation procedures. Following a theoretical discussion, several suggestions are provided as a guideline for future efforts.

Lussier et al. [39] focus on the analysis techniques for Raman and surface-enhanced Raman scattering (SERS) effect, as well as the ML methods for extracting chemical information from the resulting data. The review, written in 2020, discusses the principal ML techniques and procedures applied to the four most prevalent RS applications: food and beverage, forensics, bacteria and viruses, and medical diagnostics.

Pan et al. [40] published in 2021 a review that describes AI methods paired with RS to determine the composition of substances. With an emphasis on the sectors of chemicals, food, medicine, and medical diagnostics, they provide an overview of the most commonly used RS preprocessing and classification techniques.

The present study focuses on the examination of the most recent ML techniques employed in the analysis of Raman spectra extracted from oncological tissues, as documented in the existing literature. Building upon previous literature reviews, our analysis is centered on a critical evaluation of the multidisciplinary approaches involved in the combination of RS and AI to perform cancer grading and classification, to provide both a theoretical and practical overview. Ongoing research in the field has yielded a steady stream of results, and the most recent findings have been prioritized in the article selection process. The next section will elaborate on the methodology that was adopted to give a comprehensive and detailed overview of the present state of the art.

III. METHODS

As mentioned, this article will provide an extensive review of studies that have used RS in conjunction with ML or DL techniques to investigate cancerous tissues. A minimum impact factor of 2 was chosen as a criterion for article

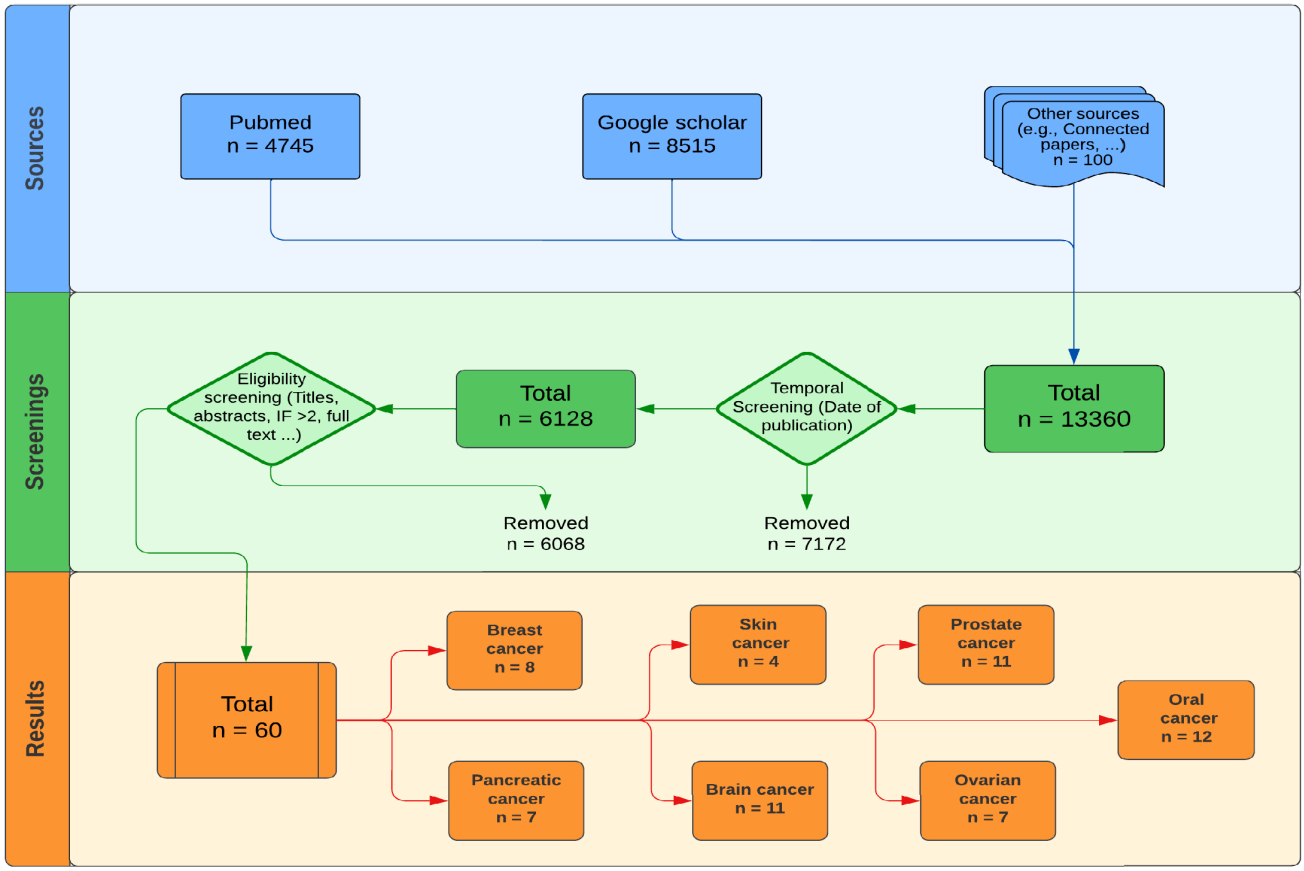


FIGURE 1. The flowchart outlines the systematic process for selecting relevant literature, focusing on the number of studies (n) involved. The main sources utilized for initial article collection are presented in the first row. The second row illustrates the filtration process based on the defined methodology. Finally, the last row displays the total number of selected papers, categorized by the type of tumor under consideration.

selection to eliminate publications of inferior quality. In addition, to maintain a clinical perspective and investigate the potential applications of spectroscopy in a clinical setting, this review considered studies with more than 10 participants to be appropriate, and priority was given to more recent research findings over older ones. This study focuses on several types of malignant tumors, including those originating in the pancreas, the skin, the breast, the brain, the prostate, the ovarian, and the oral cavity, each with its unique properties and characteristics. This is by no means an exhaustive list of all potential medical applications of RS; rather, it is meant to illustrate the breadth of potential diagnostic scenarios and the similarity or diversity of potential automated analyses within them.

The main portals used to find the most suitable articles are the Google Scholar and PubMed websites, and the paper selection is also based on the similarities between the chosen article and other relevant papers cited within it, as determined by the *Connected Papers* [41] website.

The keywords used for the article search are *"deep learning"*, *"machine learning"*, *"Raman spectroscopy"*, *"in vivo"* (preferred when available from the *"ex vivo"* one), together with an identifier of the specific tumor considered (*"pancreatic cancer"*, *"skin cancer"*, *"breast cancer"*, *"brain cancer"*, *"prostate cancer"*, *"ovarian cancer"* and *"oral cancer"*). A selection was made from the thousands of results obtained for each type of tumor based on the title, abstract, and date of publication. A total of 54 articles were selected from the filtered results as the most representative of recent discoveries in the specified categories. Furthermore, a medical viewpoint on each tumor is presented to further emphasize the importance of vibrational spectroscopic methods in healthcare systems. Finally, we present an overview of AI approaches in the context of RS, emphasizing the necessity of applying such technologies in a clinical setting. With these methodologies, this research aims to inform the future development of a cost-effective, efficient, and rapid diagnostic instrument.

IV. RAMAN SPECTROSCOPY

A. BASICS

When a beam of photons impinges on an isolated molecule or a bulk material, several physical processes can occur, which can be classified into two categories: light absorption and light scattering [98]. In the first case, if the energy of the incoming photons matches the difference between two electronic energy levels of the molecule, the electrons occupying the lowest level can absorb the photons, undergoing

TABLE 1. Summary of key studies employed in this review for AI-assisted Raman spectroscopy analysis in oncological applications.

Tumor Type	Authors and Year	Probe wavelength (nm)/Raman apparatus	Type of Sample(s)	Validation Strategy	Dimensionality reduction/Best model	Performance Measure
Pancreatic cancer	Li et al., 2021 [38]	785/QE Pro spectrometer	Human cell cultures transplanted to the dorsa of mice	Five-fold-cross-validation	PCA and none/2D-CNN	Accuracy: close to 100%
	Mandrell et al., 2020 [22]	785/Modular Raman microscope system	Human cell cultures	14-fold cross-validation	T-statistic, MIT correlation, RELIEF, and PCA/SVM (sigmoid kernel)	Accuracy: 98.2%
	Sezer et al., 2022 [20]	532/Confocal Raman microscopy	Human cell cultures	Train/test split (proportion 67/33)	PCA/SVM (RBF Kernel)	Accuracy: 97.6%
	Yan et al., 2020 [21]	532/Confocal Raman microscopy	Human fresh blood	Five-fold-cross-validation	PLS/SVM (Cubic kernel)	Accuracy: 96.4%
	Carmichael et al., 2019 [42]	785/Cofocal Raman microscope	Human blood serum	Not stated	PCA/DFA	Accuracy: 90%
	Aslam et al., 2024 [43]	not stated/conventional RS setup	human cell cultures	15-fold cross-validation	none/Non linear SVM	Accuracy: 98.5 %
	Uthamacumaran et al.,2022 [44]	514/Cofocal Raman microscope	Human blood serum	5-fold cross-validation	DT, SVM	accuracy: 100%
Breast cancer	Zhang et al., 2022 [45]	532/Confocal Raman microscopy	Human cell cultures	Train - Validation - Test split	PCA/SVM, DFA	accuracies: 99%, 93.9%, 70.9%
	Li et al., 2022 [46]	532/confocal Raman microspectrometer	Human fresh blood	simple train/test split, with leave-one-subject-out, with 7: 3 proportion	PCA and none/LDA, SVM (linear kernel), custom NN (with or without PCA)	accuracy: 100%
	Zeng et al., 2023 [47]	532/confocal Raman spectrometer	Human blood serum	6-fold cross-validation, 8:2 split ratio for train:test sets	none/1D-CNN	accuracy: 91.11%, AUC: 0.95
	Fuentes et al., 2023 [48]	785/confocal Raman microscope	Human cell cultures transplanted to the dorsa of mices	random splits or leave-one-(mouse, section or Ramanmap)-out, with train-validation-test splits	none/1D-CNN	Random split, Day 1 (0 vs 15Gy): 85.0, day 3 (0 vs 5Gy): 92.1, day 3 (0 vs 15Gy): 94.6%. leave-one-X-out, Day 3 (0 vs 15Gy) : 92.7% (X=map), 91.4% (X=section), and 83.7% (X=mouse)
	Fuentes et al.,2024 [49]	785/confocal Raman microscopy	human cell cultures	70:20:10 train:test:validation split ratio	none/1D-CNN	mean testing accuracy, sensitivity, specificity, and F-1 score 99.8%
	Ma et al., 2023 [50]	785/confocal Raman microscopy	Human cell cultures	train validation test split, 70:10:20 ratio	none (PCA used but to classify alone, without CNN)/Resnet-based CNN	AUC of 0.999 for MDA-MB-231, and 0.991 for MCF-7, and 0.997 for MCF-10A. sensitivity and specificity: MDAMB-231 and MCF-10A 81.2%, 56.3%; MCF-7 and MCF-10A 89.7%,80.8%; MCF7 and MDA-MB-231 67%, 66.1%
	Ma et al., 2021 [51]	785/A lab-made Raman spectrometer	Human fresh tissues	train/test split, 10-fold-cross-validation	PCA on the ML model used for comparison/1D-CNN	accuracy of 92%, the sensitivity of 98% and the specificity of 86%
	Li et al., 2023 [52]	785/ QE65000 miniature Raman spectrometer	Human fresh tissue	train/test split, with random sub-sampling (10 repetitions) and 10-fold cross-validation	PCA, Relief / MSEA-ALHK	(random sub-sampling) Accuracy: 98.31 Sensitivity: 96.45 Specificity: 100.00, cross-validation (cross-validation average) Accuracy: 94.15 Sensitivity: 89.03 Specificity: 99.12
Skin cancer	Bratchenko et al., 2022 [53]	785/custom portable RS system equipped with a commercially available Raman probe	Human fresh tissues (in vivo)	train/test split with ratio 80:20, and the model is trained on the 80% section with a 10 fold cross-validation with 90:10 train/test split	none/1D-CNN with LSTM	ROC AUCs of 0.96 (0.94 – 0.97 with 95% Confidence Interval(CI)), 0.90 (0.85–0.94; 95% CI), and 0.92 (0.87 – 0.97; 95% CI)
	Araújo et al., 2021 [54]	1064/FT-Raman spectrometer	Human cell cultures	five-fold cross-validation with 80:20 train test ratio	none/LightGBM	Full spectra: AUC 0.98, 95% CI 0.97–0.99, miniaturized spectra: AUC 0.97, 95% CI 0.95–0.98
	Qui et al., 2023 [55]	785/Confocal Raman spectrometer	Human cell cultures	10-fold cross-validation with 70:10:20 train, test, validation splits	none/1D-CNN	Accuracy: 98.5%* *Several other classification tasks can be found in the text
	Wu et al., 2021 [56]	785/Confocal Raman spectrometer + high-powered IR laser	Human fresh tissues	leave-one-out cross-validation	none/1D-CNN	accuracy: around 80-90%* *Several classification setups and results can be found in the text
Brain cancer	Zhang et al., 2023 [57]	532/commercial Raman fiber optic setup, Confocal Raman Microscopy	Human fresh tissues	Leave-One-Out cross-validation, AUROC	selection of characteristic peaks (peaks)/peaks+SVM (Gaussian kernel)	peaks+SVM (normal vs cancer): Accuracy: 95.5%; Sensitivity: 99.1%; Specificity: 50.0%
	Vrazhnov et al., 2023 [58]	532/Confocal Raman Microscopy	Serum from mouses with human glioblastoma	10-fold cross-validation, AUROC	-/XGBoost	XGBoost: AUROC: 0.95
	Quesnel et al., 2023 [59]	785/Confocal Raman Microscopy	Human blood serum, FFPE tissues	5-fold cross-validation	PCA/PCA+Linear SVM, PCA+Bilayed Neural Network	serum, PCA+Bilayed Neural Network, normal vs grade 3: Accuracy: 90%; tissues, PCA+Linear SVM, grade 2 vs grade 3: Accuracy: 90%
	Hollon et al., 2023 [60]	790 nm (pump beam); between 1010 and 1050 nm (Stokes beam)/Stimulated Raman spectroscopy	Human fresh tissues	Leave-One-Out cross-validation	-/DeepGlioma	Accuracy: 93.3 $\pm$ 1.6%

TABLE 1. *(Continued.)* **Summary of key studies employed in this review for AI-assisted Raman spectroscopy analysis in oncological applications.**

	Ma et al., 2022 [61]	532/Confocal Raman Microscopy	Human blood serum	5-fold cross-validation	PLS/PLS+LDA	Accuracy: 97.87%; Sensitivity: 98.10%; Specificity: 98.19%
	Jin et al., 2022 [62]	785/Surface Enhanced Raman spectroscopy	Water put in contact with mouse fresh tissues and placed on a SERS chip	Use of a single training and test set (proportion 80/20)	-/lightweight one-dimensional CNN	Mean Absolute Error (MAE): 45.9%, standard deviation of MAE (SD):53.6%, coefficient of determination (R2): 28.3%, and Sum of the Squared Errors (SSE): 72.8%
	N. Iturrioz-Rodríguez et al., 2022 [63]	532/Confocal Raman Microscopy	Human cell cultures	Use of a single training and test set (proportion 66/33)	PCA/PCA+LDA	Accuracy: 92.5%
	Chen et al., 2023 [64]	532/Confocal Raman Microscopy	Human blood serum	Use of a single training and test set	-/SMOTE- B + CNN-LSTM	Accuracy: 89.33 $\pm$ 4.11; Sensitivity: 85.33 $\pm$ 1.89; Specificity: 93.33 $\pm$ 6.80; Precision: 93.15 $\pm$ 6.73; AUROC: 89.33 $\pm$ 4.11
	Tian et al., 2022 [65]	532/Confocal Raman Microscopy	Human blood serum	5-fold cross-validation	PLS/GoogLeNet	Accuracy: 99.50%; Sensitivity: 98.98%; Specificity: 98.48%
	Bukva et al., 2021 [66]	532/Confocal Raman Microscopy	Extracellular Vesicles for human blood serum	Use of a single training and test set (proportion 90/10), AUROC	PCA/PCA+Linear SVM	Accuracy: 82.9–92.5%; Sensitivity: 80–95%; Specificity 80–90% (in distinguishing the five groups from each other and from control samples); AUROC: 0.8-0.9
	Livermore et al., 2020 [67]	785/Confocal Raman Microscopy	Human fresh tissues	Leave-One-Out cross-validation	PCA/PCA+LDA	Accuracy: 99%; Sensitivity: 96%; Specificity: 99%
Prostate cancer	Ao et al., 2023 [68]	between 680 and 1300 nm (pump beam); 1040 nm (Stokes beam)/Stimulated Raman spectroscopy	Human fresh tissues	Use of 5 randomly extracted subsets, 4 as training sets and 1 as test set	-/Inception-ResNet-V2	Accuracy: 84.4 %
	Chen et al., 2021 [69]	785/Surface Enhanced Raman spectroscopy	Human urine mixed with high viscosity Ag sol	5-fold cross-validation	-/LetNet-5-derived Convolutional Neural Network	Accuracy: 74.95%; Sensitivity: 77.32%; Specificity: 72.46%
	Grajales et al., 2022 [70]	681(-), 785/Custom-built Raman probe, combined with radiomic system	Human fresh tissues	Leave-One-Out cross-validation	Method based on: 1) Exclusion of features with low variance; 2) Exclusion of features not correlated to each other; 3) Lasso regression./SVM	Accuracy: 99 %
	Milligan et al., 2022 [71]	785/Confocal Raman Microscopy	Human fresh tissues	Use of a single training and test set (proportion 75/25)	-/LDA	Accuracy: 66 % (Gleason score); 70% (CAPRA score)
	Milligan et al., 2022 [72]	785/Confocal Raman Microscopy	Human fresh tissues	-	PCA/PCA+Random Forest Classifier	Accuracy: 84.3%; Sensitivity: 78.6%; Specificity: 89.8%
	Picot et al., 2022 [73]	785/miniature Raman fiber-optics system	Human fresh tissues (in vivo)	Leave-One-Out cross-validation	-/SVM	Accuracy: 79%; Sensitivity: 86%; Specificity: 72%
	Zhao et al., 2022 [74]	785/Surface Enhanced Raman spectroscopy	Human blood plasma	Leave-One-Out cross-validation	PLS/PLS+SVM	Accuracy: 71.06%; Sensitivity: 65.40%; Specificity: 75.60%
	Grosset et al., 2020 [75]	785/Confocal Raman Microscopy	Human fresh tissues	5-fold cross-validation	SVM+L1 regularization/SVM+L1 regularization+SVM with a nonlinear Gaussian kernel	Accuracy: 95% $\pm$ 2%; Sensitivity: 96% $\pm$ 4%; Specificity: 94% $\pm$ 2%
	Ma et al., 2020 [76]	633/custom-built Raman spectrometer	Human urine mixed with high viscosity Ag sol	Leave-One-Out cross-validation	Genetic Algorithm+PLS/Genetic Algorithm+PLS+LDA	Accuracy: 86.6%; Sensitivity: 87.0%; Specificity: 86.1%
	Medipally et al., 2020 [77]	785/Confocal Raman Microscopy	Human blood plasma, human Lymphocytes	-	PCA/PCA+PLS-DA	Sensitivity and specificity ranging from 90 to 99%
	Lee et al., 2019 [78]	647/Confocal Raman Microscopy	Extracellular Vesicles derived either from blood or from prostatic cell lines	-	-/custom-built Convolutional Neural Network	Minimum accuracy: 93%
Ovarian cancer	Chen et al., 2022 [79]	785/Confocal Raman Microscopy	Human blood plasma	5-fold cross-validation	PCA/Two-stage model based on distinguishing normal and anomalous classes and to distinguish cyst from cancer within this last sub-group through a Back Propagation Neural Network	Cancer: Sensitivity 81.0%, Specificity: 97.3%; Cyst: Sensitivity: 63.6%, Specificity: 91.5%; Normal: Sensitivity: 100%; Specificity: 90.6%
	David et al., 2022 [80]	785/Confocal Raman Microscopy	Human fresh tissues	Leave-One-Out cross-validation	SVM (regularization Lasso regression)/SVM (Gaussian kernel)	Accuracy: 90%; Sensitivity: 93%; Specificity: 98%
	Giamougiannis et al., 2021 [81]	785/Confocal Raman Microscopy	Human blood plasma, human ascitic fluid	10-fold cross-validation	PCA/PCA+LDA, PCA+SVM	Plasma, PCA+SVM: Accuracy: 68% - Serum, PCA+SVM: Accuracy: 73% - Ascitic fluid, PCA+LDA: Accuracy: 82%
	Hunter et al., 2021 [82]	785/Surface Enhanced Raman spectroscopy	Human exosomes derived from cell cultures, functionalized with gold nanoparticles	Use of a single training set and a single test set (proportion 60/40)	PLS/SVM (radial quadratic kernel)	CASE 1: Accuracy: 75%; CASE 1: Accuracy: 75; CASE 2: Accuracy: 88% (HGS); 70% (EM)
	Moisoiu et al., 2019 [83]	532/Surface Enhanced Raman spectroscopy	Human blood serum functionalized with silver nanoparticles	5-fold cross-validation (proportion between training and test set: 80/20)	PCA/PCA+LDA	Accuracy: 94% (breast); 78% (colorectal); 86% (lung); 95% (ovarian); 93% (oral)

TABLE 1. *(Continued.)* **Summary of key studies employed in this review for AI-assisted Raman spectroscopy analysis in oncological applications.**

	Perumal et al., 2019 [84]	633/Surface Enhanced Raman spectroscopy	Haptoglobin from ovarian cyst fluid functionalized with gold nanoparticles	1000-fold cross-validation (proportion between the training and the test set: 80/20)	-/LDA, Logistic Regression, k-Nearest Neighbor Classifier	Sensitivity: 94%; Specificity: 91%
	Paraskevaidi et al., 2018 [85]	785/Confocal Raman Microscopy	Human blood plasma	5-fold cross-validation	-/SVM (Gaussian kernel)	Sensitivity: 94%; Specificity: 96%
Oral cancer	Chang et al., 2023 [86]	785/Custom-built Raman probe	Human fresh tissues	The dataset was divided in 5 subset approximately of the same size, 4 of which were used as training set, while the other is used as test set	-/ResNet50	Accuracy: 92.81%; precision: 92.93%, sensitivity: 92.86%
	Ghosh et al., 2022 [87]	785/Confocal Raman Microscopy	Exfoliated cells	Use of a single training set made of 176 spectra and a test set made of 24 spectra	-/Deep Reinforced Neural Network	Accuracy: 83.30% (normal); 87.00% (oral leukoplakia); 95.24% (oral squamous cell carcinoma)
	Koster et al., 2022 [88]	785/Confocal Raman Microscopy	Human blood palsma, human saliva	Use of a single training set and a single test (the proportion is not clear)	PLS/LDA/QDA	Accuracy: 91.7%; Sensitivity: 96.3%; Specificity: 85.7%
	Wang et al., 2023 [89]	785/Surface Enhanced Raman spectroscopy	Human saliva and human blood serum functionalized with silver nanoparticles	-	-/SVM (Gaussian kernel)	Serum, 92.9% (adenocarcinoma of the salivary gland, normal); - Saliva, 85.7% (adenocarcinoma of the salivary gland, normal)
	Xia et al., 2020 [90]	785/Custom-built fiber-optic Raman system	Human fresh tissues	5-fold cross-validation	Convolutional Neural Network employed as the feature extractor/SVM (Gaussian kernel)	Accuracy, sensitivity and specificity all above 99%
	Yan et al., 2020 [91]	785/Custom-built fiber-optic Raman system	Human fresh tissues	5-fold cross-validation	-/custom-built Convolutional Neural Network	Accuracy, sensitivity and specificity all above 98%
	Jeng et al., 2019 [92]	532/Confocal Raman Microscopy	Human fresh tissues	Leave-One-Out cross-validation	PCA/QDA	Accuracy: 87.5%; Sensitivity: 90.90%; Specificity: 83.33%
	Yan et al., 2019 [93]	785/Custom-built fiber-optic Raman system	Human fresh tissues	-	-/custom-built Convolutional Neural Network	Accuracy: 97.2%; Sensitivity: 99.7%; Specificity: 95.0%
	Zhang et al., 2019 [94]	between 802 nm (pump beam); 1040 nm (Stokes beam)/Stimulated Raman spectroscopy	Human fresh tissues	5-fold cross-validation (proportion between the training and the test set in a single fold: 80/20)	-/ResNet34	Accuracy: 95.9% $\pm$ 0.4%
	Yu et al., 2019 [95]	785/Custom-built fiber-optic Raman system	Human fresh tissues	5-fold cross-validation (proportion between the training and the test set in a single fold: 80/20)	-/custom-built Convolutional Neural Network	Accuracy: 96.90%; Sensitivity: 99.31%; Specificity: 94.44%
	Sharma et al., 2023 [96]	532/Confocal Raman Microscopy	Human fresh tissues	k-fold cross validation (value of k not specified)	PLS/SVM	Accuracy: 94.74%; Sensitivity: 95.65%; Specificity: 93.33%
	Chaudhuri et al., 2022 [97]	785/Confocal Raman Microscopy	Exfoliated epithelial cells	10-fold cross validation	LDA/SVM	Accuracy: 88.8% (Binary); 95% (Three classes)

an electronic transition. After light absorption, the irradiated target can relax by emitting energy through several radiative or non-radiative mechanisms, within time scales ranging from nanoseconds to milliseconds. In the second case, the light scattering is usually described as the interaction between the incoming photons and short-lived virtual states [99], i. e. not well-defined energy states of the molecule [100]. With respect to light absorption, light scattering processes occur within far smaller time scales, ranging between hundreds and thousands of femtoseconds. In addition, the most part of the scattered photons usually has the same frequency as the incident photons. The corresponding physical mechanism is referred to as Rayleigh scattering. The other scattered photons can have frequencies smaller or larger than the frequencies of the incoming photons. In the first case, the corresponding phenomenon is called Stokes scattering, while in the second case, it is defined as anti-Stokes scattering [101]. Stokes and anti-Stokes scattering can be included within the so-called Raman effect, observed for the first time by Raman et al. in 1928 [17].

The aspect that makes the Raman effect interesting for the assessment of the molecular properties of materials is represented by the fact that the difference between the frequency ν of the scattered and the frequency ν_{ex} of the incident photons is intimately correlated with the molecular vibrational motions of the irradiated material. Since these oscillations depend on the chemical nature of the atoms and of the bonds between them, the Raman effect turns out to be a potential source of information about the molecular properties of the substance under investigation.

The Raman effect is the physical phenomenon on the basis of RS. In a generic RS apparatus, the intensity of the scattered light is recorded as a function of $|\nu - \nu_{ex}|$ or other related quantities, resulting in a Raman spectrum. In particular, the Raman signal is often reported as a function of the so-called wavenumber k, defined as

$$k = \frac{2\pi|\nu - \nu_{ex}|}{c}, \tag{1}$$

where c represents the speed of light in vacuum. In Fig. 3 a typical Raman spectrum of a biological sample in the spectral range of wave numbers between 400 and 3400 cm^{-1} is shown. The Raman signal observed at specific intervals, or bands, in the wave number domain, is proportional to the

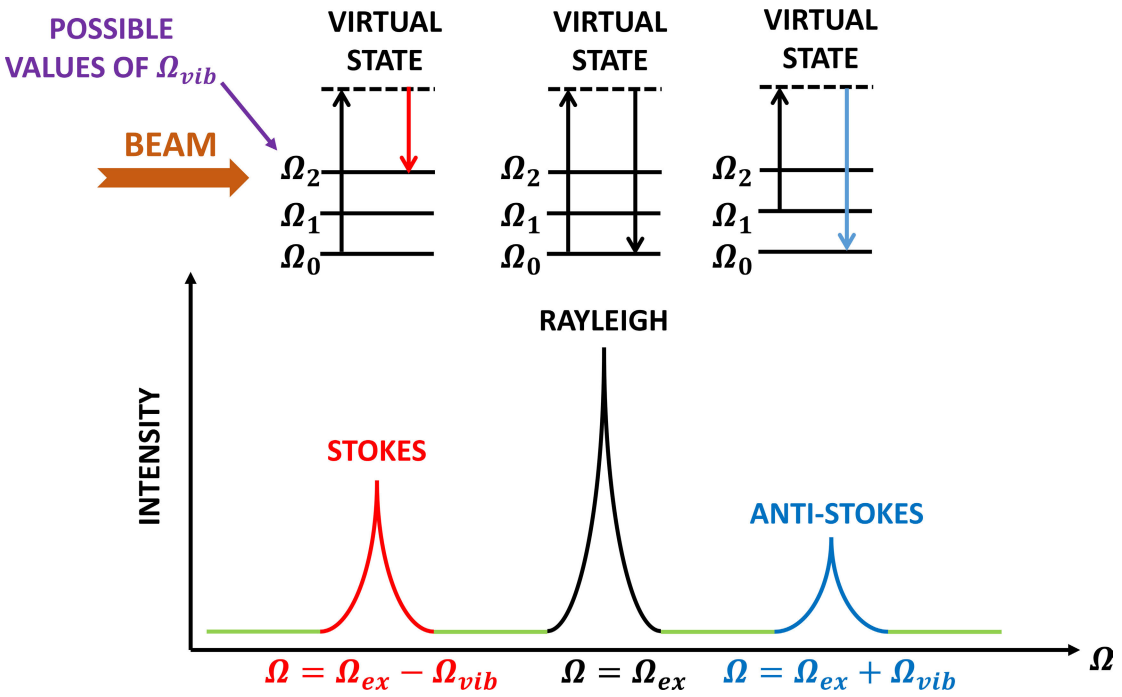


FIGURE 2. Schematic diagram illustrating the electronic levels involving Stokes, Rayleigh and Anti-Stokes scattering and corresponding representative Raman spectra.

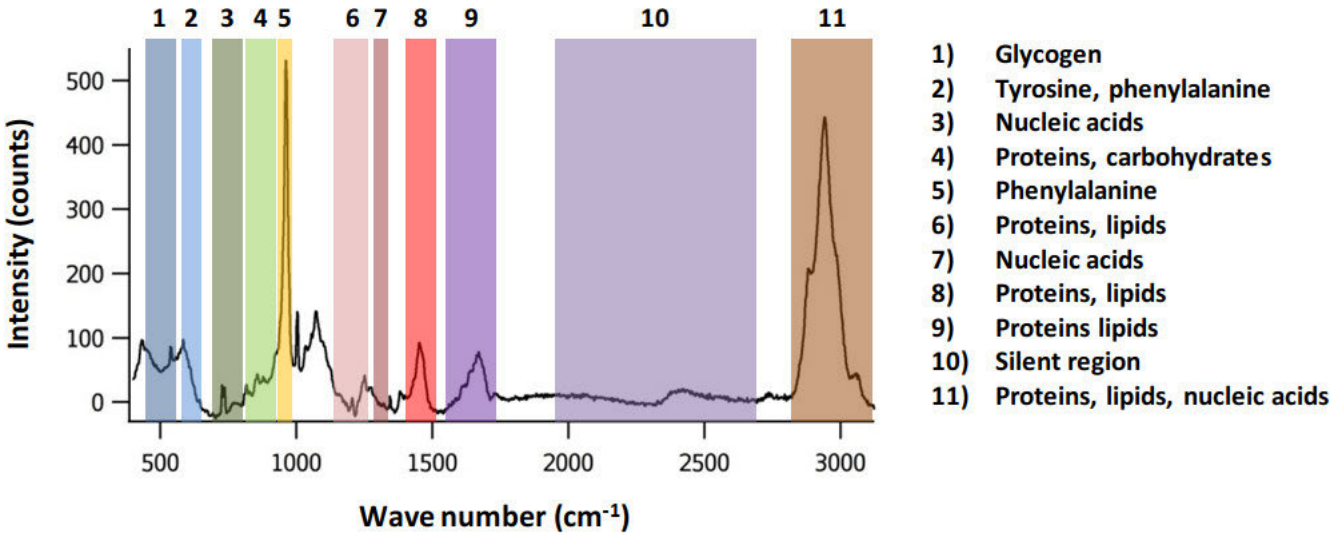


FIGURE 3. Example of a Raman spectrum of a biological sample, highlighting characteristic Raman bands and the related molecular assigment.

concentration of specific molecular groups. In particular, the main Raman bands of interest in biology are highlighted in Fig. 3 as numbered shaded areas.

B. CONFOCAL RAMAN MICROSCOPY (CRM)

Despite the aforementioned strengths, the Raman effect suffers from some drawbacks.

First of all, the Raman signal is usually very weak. In particular, Raman scattering is characterized by cross sections up to five orders of magnitude smaller than for fluorescence emission [102]. Therefore, if the sample under investigation is fluorescent, the contribution of fluorescence to the measured signal can mask the Raman component. Since fluorescence is usually observed at low wavelengths, a possible strategy to minimize fluorescence emission could be represented by the use of infrared light sources. However, the use of infrared light is not always possible, e.g., when the sample of interest is highly transparent within this spectral region. Furthermore, the classical theory of the Raman effect predicts that the intensity of the Raman signal is roughly proportional to the fourth power of the frequency of the incident photons. Therefore, the intensity

of the Raman signal is expected to be low in the infrared spectrum.

In addition, for several applications [103], [104], [105], high spatial resolutions are required in order to detect spatial variations of the Raman signal over length scales compatible with biological structures relevant for several biological processes of interest. Due to light diffraction, conventional optical microscopes do not allow to reach such resolving powers, thus affecting their ability to concentrate the light into small spatial volumes. Confocal Microscopy answers the aforementioned needs.

As shown in Fig. 4 (a), in a conventional optical microscope, light is focused in a relatively large ellipsoidal volume. Therefore, the light collected by the objective contains contributions coming from a relatively thick layer of material around the theoretical focal plane. As a result, the recorded image appears blurry. Unlike the conventional optical microscope, in a confocal setup (see Fig. 4 (b)), the focal volume is far smaller, reaching sizes of 0.25 μm along the direction perpendicular to the optical axis and 0.5 μm along the optical axis [106]. This property results in the suppression of out-of-focus components in the recorded signal and in an increase of contrast and sensitivity. The most peculiar feature responsible for the aforementioned suppression of out-of-focus light is represented by the combination of two spatial filters. As shown in the scheme in Fig. 4 (c), a first spatial filter has the function to suppress the out-of-focus components of the incident light and it is located in the focus of the condenser lens. A second spatial filter is then placed in the focus of the objective, cutting the out-of-focus components of the scattered light. One of the most widespread confocal apparatus is represented by the Minsky's configuration (see Fig. 4 (d)), where a single spatial filter, called pinhole, performs the function of both the aforementioned spatial filters.

As mentioned before, confocal apparatus allows to record the optical information related to a thin layer of material in the neighborhood of the theoretical focal plane. In the Confocal Laser Scanning optical Microscopes (CLSM), the focal volume can be translated along the three directions. This aspect allows to record three-dimensional optical maps, where each pixel corresponds to a single Raman spectrum.

Despite the aforementioned qualities, in several cases CRM is not sufficient to reach adequate Raman signal-to-noise ratios, especially in case of low concentrations of the Raman active molecules or poor Raman scatterers [107]. In this sense, SERS has been introduced to fix this issue. The physico-chemical nature of SERS is still an object of debate. SERS is an effect traditionally attributable to two distinct pictures [108], [109], [110], [111]. First of all, the theory called Electromagnetic Enhancement (EE) attributes SERS to a local enhancement in the electric field, induced by the presence of collective oscillations of the conducting electrons in the metal. In case of matching between the frequency of the radiation and the frequency of the aforementioned electronic oscillations, an excitation phenomenon called Surface Plasmon Resonance (SPR) occurs. In case of Localized Surface Plasmon Resonance (LSPR), i. e. stationary electronic oscillations, the local electric field E_{loc} generated by the corresponding charge distribution is larger in magnitude than the incident electric field. In particular, this increase in magnitude is usually quantified in terms of the so-called enhancement factor $\Gamma_{ex} = \left|\frac{E_{loc}(\nu_{ex})}{E_{ex}(\nu_{ex})}\right|^2$, where E_{ex} and ν_{ex} are the magnitude and the frequency of the incident electric field, respectively. Such an enhanced local electromagnetic field can be further increased by the presence of dipoles or quadrupoles in proximity of the metallic surface. In fact, in presence of an enhanced electromagnetic field, such molecules typically increase their dipole moment. The resulting electromagnetic field generated by the Raman scattering of such molecules further contributes to the LSPR in a mutual excitation process. The corresponding enhancement factor is thus represented by $\Gamma_{Raman} = \left|\frac{E_{loc}(\nu)}{E_{ex}(\nu)}\right|^2$, where ν is the frequency of the Raman electromagnetic field. If $\nu \sim \nu_{ex}$, i.e., when the Raman-active molecules are characterized by vibrational modes of low frequency, the total enhancement factor Γ_{tot} can be written as

$$\Gamma_{tot} = \Gamma_{Raman}\Gamma_{ex} \sim \left|\frac{E_{loc}(\nu)}{E_0(\nu)}\right|^4 \tag{2}$$

Another theory about the origin of SERS is referred to as Chemical Enhancement (CE). This model leads SERS back to the modification of the molecular polarizability induced by chemical mechanisms.

Since the contribution of EE is usually between two and six orders of magnitude larger than CE, this last effect is often considered negligible [102].

In cancer diagnosis, SERS is often realized by exploiting the field enhancement effect in the proximity of metallic nanoparticles. By properly choosing the metal and at appropriate concentrations of such nanoparticles it is possible to increase sensitively the Raman signal-to-noise ratio even at low concentrations of the Raman-active biomarkers. This makes the technique particularly suitable for liquid biopsies, such as saliva, blood, urine, etc., offering the advantage of being collected in a non-invasive manner.

Another common Raman-based approach adopted to increase the signal-to-noise ratio is the so-called Coherent Raman Scattering [112].

In the classical picture of the Raman effect, the illuminated molecular target is represented as an electric dipole, under the effect of a monochromatic electric field of frequency ν_{ex}. In adiabatic conditions, i. e. when electronic and nuclear vibrations can be considered decoupled, the dipole oscillates at the same frequency of the incident electric field. On the other hand, if a coupling between the nuclear and the electronic vibrations occurs, the dipole oscillates at a frequency $\nu_{vib} - \nu_{ex}$, where ν_{vib} is a characteristic vibrational frequency of the molecular target. In Coherence Raman spectroscopy, the Raman effect is induced by irradiating the molecular target with two laser beams corresponding to

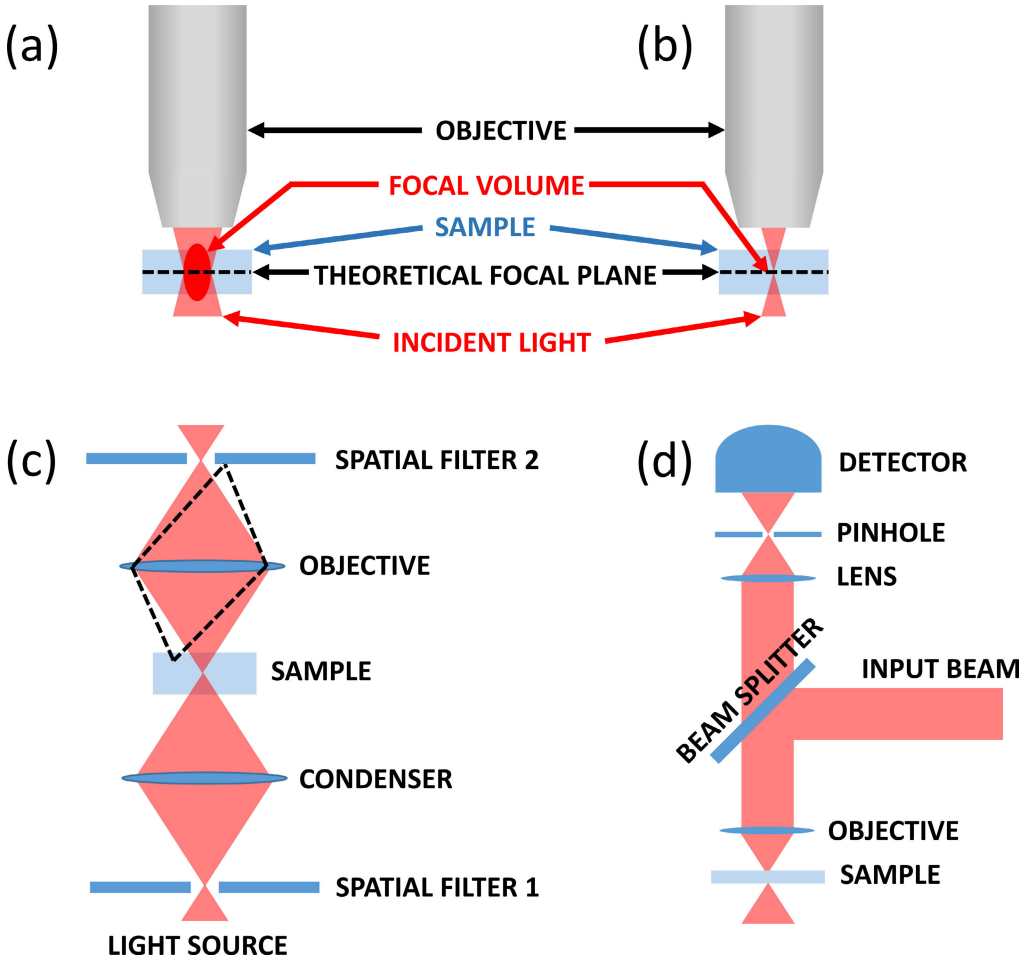


FIGURE 4. (a) Schematic representation of the focal volume in a conventional optical microscope; (b) Schematic representation of the focal volume in a confocal microscope; (c) Scheme of a generic confocal microscope, highlighting the two spatial filters; (d) Schematic representation of a confocal microscope in the Minsky's configuration.

the monochromatic electric fields $E_{ex} = \hat{E}_{ex}e^{-2\pi i\nu_{ex}t}$ and $E_s = \hat{E}_s e^{-2\pi i\nu_{es}t}$. In this case, the classical theory predicts this expression for the induced oscillating dipole $\mu_{Coherent}(t)$ as a function of time t:

$$\mu_{Coherent}(t) = \hat{E}_{ex}\hat{E}_s^*\left(\frac{\delta\alpha}{\delta q}\right)_0^2 f(N)q_0 e^{-2\pi iNt}[E_{ex}(t) + E_s(t)] + c.c., \tag{3}$$

where $N = \nu_{ex} - \nu_s$ and $\left(\frac{\delta\alpha}{\delta q}\right)_0$ is the functional derivative of the electronic polarizability with respect to the nuclear coordinates q, calculated at the equilibrium configuration q_0. First of all, the factor $f(N)$ is maximized when $N \sim \nu_{vib}$. In addition, the presence of the electric fields $E_{ex}(t)$ and $E_s(t)$ produces four contributions at frequencies $\nu_{ex} \pm N$ and $\nu_s \pm N$. In particular, the scattering radiation emitted at frequency $\nu_{ex} + N$ corresponds to the so-called Coherent Anti-Stokes Raman Scattering (CARS).

The use of two laser beams of different frequencies in CARS is mainly justified by the resulting phase correlation of the molecules that constitute the sample. This feature results in a much higher signal in comparison to the Spontaneous Raman effect. In addition, this coherent Raman emission is strongly directional and therefore, easily collected by placing the detector in the forward direction.

C. APPLICATIONS OF COHERENT ANTI-STOKES RAMAN SCATTERING IN CANCER DETECTION

One of the most frequent applications of CARS in the cancer research regards the monitoring of liquid droplets in cells. In fact, the study of the distribution and of the temporal evolution of such cellular structures is considered a relevant marker to assess the progression of tumors. In particular, the high concentration of hydrocarbon functional groups characterizing the lipid droplets makes them particularly sensitive to Raman-based approaches, in particular to CARS.

Only to cite few examples, Weng et al. [113] employed CARS for the DL-assisted recoginition of healthy and

cancerous lung tissues. The authors observed an increased capability of pre-trained neural network in recognizing cellular structures by CARS experimental data, such as lipid droplets. Galli et al. [114] succeded to resolve lipid droplets in a cancerous brain tissue subject to necrosis after chemotherapy treatment.

Another interesting application of CARS in the study of cancer regards the treatment of tissue images, aimed at enhancing the contrast between relevant structures. This field of research could constitute a valid support to pathologists in the qualitative observation of biopsies, complementary to the common practice, based on the staining with Hematoxylin and Eosin. Furthermore, the inherent confocality of CARS allows to put in evidence characteristic phenomena with high spatial resolution. For example McCullagh et al. [115] employed a CARS based apparatus to asses the changes in the concentration of myeline in mice and gerbils brain tissues. The high spatial resolution allowed to distinguish myelin from the Nissl bodies. Petrov et al. [116] coupled CARS with hierarchical clustering ML model for the individuation of tissues microcalcifications associated to breast cancer.

D. MATERIALS

When RS is performed directly on tissues, fixation could represent an useful and practical solution to preserve the state of the tissue and to store it for long periods of time. In this sense, Formalin Fixing and Paraffin Embedding representes a widespread protocol for the tissue fixation. In this case, the tissue is immersed in aqueous formalin solution. This procedure preserves some lipid molecules through chemical reactions with the double bonds of unsaturated hydrocarbons. Furthermore, the primary and secondary amine groups of proteins are cross-linked, thus resulting in protein fixation [117]. The tissue is than immersed in acqueous solutions of ethanol of increasing concentrations, with the aim of replacing water with the ethanol itself. This operation could determine the coagulation of the proteins of cytoplasm and consequent degradation of organelles. This technique has an additional drawback of being aggressive for some lipid molecules, not preserved in the aforementioned formalin treatment. After these steps, ethanol is removed and substitute by an organic solvent, e. g. xylene, by immersing the sample the solvent itself. The sample is then immersed in liquid paraffin and left at room temperature until complete paraffin solidification. The sample is then ready to be cut in micrometric sections for the following analysis. The main drawback of this technique is represented by possible chemical alterations of proteins attributable to formalin and to the Raman activity of paraffin in spectral intervals overlapped to the bands of interest in biology. This last drawback can be limited by employing specific solvents to remove paraffin. However, such de-waxing procedure could decrease the concentration of lipids, often considered target molecules for cancer diagnosis [118]. Alternatively, another route is represented by the removal of the formalin peaks within the post-processing of the Raman spectra, suitable when the paraffin signal doesn't dominate significantly over the contribution of the sample itself. For these reason, the employment of snap-frozen fresh tissues is often preferred, especially for the study of soluble lipids and/or enzymes. This process is carried out in isopentane, previously cooled in liquid nitrogen. The rapid cooling prevents the formation of relatively large crystals of ice, responsible for the damage of the tissues. The result of this procedure is a frozen tissue that can be cutted in sections for the following investigations.

Besides the choice of the type of sample and or the protocol to properly treat it, the choice of the substrate represents another crucial factor. In particular, the glass slides commonly employed in microscopy represent a low-cost solution, with the drawback of a strong contribution to the measured signal due to the fluorescence of the substrate itself. This behavior limits the probing of the Raman effect in the infrared spectrum. Calcium Fluoride (CaF_2) or Magnesium Fluoride (MgF_2) substrates represent a more expensive but effective solution to suppress the background contribution.

Several types of tumors, either solid or not, show evident symptoms only when they are at a late stage. For this reason, the employment of diagnostic methods based on liquid biopsies, e. g. blood, saliva, urine, seminal fluid etc. could potentially represent a non-invasive and effective route towards the early-stage screening of cancer. In this case, the Raman measurement can be performed with a liquid samples as-is through an immersion objective or by keeping the liquid in amicrofluidic device. Another option could be represented by drying a droplet of liquid previously deposited on a substrate. The resulting coffee-ring effect [119] determines an increase of the concentration of the Raman-ective molecules at the margins of the droplet, with a consequent increase of the signal-to-noise ratio.

Another option to enhance the Raman signal is represented by the use of SERS. In this case, the liquid sample is placed incontact with a substrate, mainly represented by gold or silver nanoparticles of diameters typically ranging between 10 and 150 nm.

V. MACHINE AND DEEP LEARNING

In recent years, the escalating need to analyze and comprehend complex patterns within diverse datasets has led to a spike in interest in the fields of ML and DL, two major domains of AI. While their roots trace back to the mid-20th century, it wasn't until the 1980s and 1990s that ML potentialities were identified and started to grow [120]. In particular, DL is a younger subfield of ML that appeared in the early 2000s, and focuses on a specific ML methodology. Despite their widespread acceptance and the fact that they share numerous similarities, such as the extensive use of data, it is important to note that there are several fundamental distinctions between the two.

ML is a branch of AI that mainly involves the development of algorithms capable of learning from input distributions and making predictions based on them. These algorithms are usually designed to acquire implicit knowledge by exploit-

ing large quantities of information. There are two major categories of ML algorithms: supervised and unsupervised. In supervised learning, the objective is to learn a mapping between the input and output variables. The model is trained using labeled data, where the correct output is known. On the other hand, unsupervised learning involves training the model on unlabeled data with the aim of understanding its underlying structure. For instance, it can be used for clustering, the generation of new data, or to reduce the dimensionality of the input.

DL, a subfield of ML, leverages the use of neural networks to acquire accurate knowledge about data representations by introducing multiple linear transformations. At its core, DL relies on the neural network model, a mathematical structure loosely inspired by the biological neural network. These networks are built to perform a sequence of mathematical approximations on input data, estimating desired results from potentially any kind of distribution. In the medical field, ML and DL are especially helpful because of their capacity to handle massive amounts of complex data with low inference time, building hierarchies of features that capture underlying data relations. Notably, DL has demonstrated exceptional promise in the automated analysis and interpretation of medical images such as X-rays, CT scans, and MRI scans [121], [122], [123].

After the raw data collection procedure, described in the previous section, the fundamental steps involved in learning procedures are the data preprocessing, the model selection, and the model evaluation. The next section will briefly describe these operations.

A. DATA PREPROCESSING

The effective use of ML and DL models with real-world data requires tackling several technical challenges. In the case of RS applications, it is common to deal with potential artifacts, such as background fluorescence noise. These artifacts can adversely affect the learning process, compromising the model's precision. To address these challenges, ML models require the use of data preprocessing techniques. This involves transforming the raw Raman spectra, which may suffer from incompleteness, inconsistencies, a lack of discernible patterns, and human errors, into more understandable and practical formats. The preprocessing of RS typically includes multiple stages, including but not limited to dimensionality reduction, cosmic ray removal, signal smoothing, and baseline subtraction. It is worth noting that, in principle, it is possible to train NNs or CNNs directly on raw data, with the preprocessing step implicitly performed within the network layers. However, this approach may be suboptimal due to constraints related to the quantity and quality of the data. When explicitly performed, the process of preparing data for analysis involves a number of sequential stages, which we summarize into the following steps: data cleaning and transformation, data augmentation, and data reduction.

1) DATA CLEANING AND TRANSFORMATION

Preparing data for analysis often requires an important step of cleaning, which includes procedures such as eliminating outliers, correcting inconsistent data points, and smoothing noisy data. In the RS case, because the Raman scattering phenomenon is significantly less intense (by several orders of magnitude) than fluorescence emission related to biological tissues, the signal may exhibit a strong low-frequency noise pattern. This noise pattern hinders the precise measurement of spectral components, obscuring Raman signals. Therefore, it is common to use correction algorithms before conducting an analysis on the Raman spectrum [124].

Several practical strategies that involve hardware modifications, like shifted excitation and time gating, can be employed to tackle this problem. Nevertheless, computational methods, due to their affordability and simplicity of implementation, have been widely utilized for baseline correction [125]. Such correction techniques, including but not limited to Fourier and Wavelet transforms, polynomial fitting, first- and second-order differentiation, and multiplicative signals [124] can significantly improve the signal clarity and help to enhance the signal-to-noise ratios. In addition, semi-automated preprocessing of Raman spectral data, achieved through algorithms like the Vancouver [125] or the Savitzky-Golay [126] ones, can highlight relevant information and patterns for the training phase while removing unnecessary details. These algorithms could be considered a gold standard in the field when looking at their widespread usage. Furthermore, data transformation into a format compatible with ML or DL algorithms enhances its effectiveness in data science applications. Techniques such as rescaling, normalization, merging, and discretization are employed to facilitate the training process and improve overall accuracy.

2) DATA AUGMENTATION

There are several available strategies to generate synthetic data that closely mirror the original data distribution. They are usually applied to enhance the quantity and quality of information that can be gathered from the dataset, increasing the model's generalization capabilities. For instance, starting from a single sample, a sequence of modifications can be applied to generate additional samples. This methodology is widely used in the AI domain to meet the critical demand for large quantities of data. In the realm of AI-driven RS analysis, simple but effective data augmentation techniques include a left-right shift of the spectra (up to a small shift value, usually tuned accordingly to the relative type of spectra), and a slight additive noise (e.g., sampled from a Gaussian distribution with mean and variance proportional to the data values) to generalize the input and mitigate low-magnitude noise. Additionally, the process of augmentation can be implemented concurrently with the training phase, thus providing the model with samples that exhibit varying levels of detail with each iteration. This dynamic augmentation contributes to the model's ability to generalize and handle diverse inputs effectively.

a: GENERATIVE ADVERSARIAL NETWORKS (GAN)

An interesting possibility is the generation of synthetic data using GAN models [127]. These architectures are built on the concept of competition between two concurrently trained networks, a discriminator and a generator, with the end goal of producing realistic samples. The generator is a network that, given a known distribution as input, can produce realistic samples sampling from a learned distribution. At the same time, the discriminator is taught to differentiate between authentic samples and generated ones using a binary classification. While learning together, the generator seeks to minimize the number of correctly classified samples (thus improving the realism of the produced data), while the discriminator aims to maximize it. It is significant to consider that, due to the peculiarities of adversarial training (which is prone to collapses and instabilities), the discriminator tends to be smaller than the generator with respect to size and layer depth.

3) DATA REDUCTION

This process commonly involves the implementation of techniques for reducing dimensionality, such as PCA, PLS, and LDA. By utilizing denser information during the training process, potential noise patterns within the data can be eliminated, and the more significant features can be emphasized. In this manner, the data, reduced in dimensionality and complexity, can be given as input to the classification model.

a: PRINCIPAL COMPONENT ANALYSIS (PCA)

PCA is a very popular methodology for reducing the dimensionality of inputs in ML-based RS applications (see tab. 1). Its objective is to enhance the interpretability of data without losing its underlying information, as well as to enable the representation of data with multiple dimensions on a lower-dimensional space through eigenvalues and eigenvectors analysis [128]. The process of reducing the dimensionality of a given data set involves the transformation of the data into another coordinate system. The newly established coordinate system displays distinct features in which the first axis, commonly referred to as the first principal component, is capable of capturing the highest degree of variance present in the data, with the following ones capturing a decreasing degree of variance. Another important PCA characteristic is the ability to mitigate the effects of noise and redundancy. PCA can be utilized for a variety of purposes, including but not limited to data reduction, visualization, clustering, and feature extraction.

b: PARTIAL LEAST SQUARES (PLS)

PLS is an algorithm that is commonly used for dimensionality reduction and regression purposes. It iteratively identifies latent variables that capture the maximum covariance between a set of predictors and a set of response variables. It is particularly suitable when predictor variables are highly correlated or higher in number than the response ones, similarly to the PCA. After standardizing the variables, the algorithm extracts components that explain as much of the covariance as possible between response and predictor variables and represent their directions in the data space. In this manner, when generating a new component, the objective differs from PCA, as it doesn't prioritize achieving the highest degree of variance. Instead, its goal is to maximize the response variable predictability from a model. PLS iteratively repeats this process until a specified number of components or a satisfactory level of variance is reached. In the case of a regression task, the final components are used to build a predictive model, and the regression coefficients obtained facilitate making predictions for new observations.

c: LINEAR DISCRIMINANT ANALYSIS (LDA)

LDA is a supervised learning algorithm used for classification, dimension reduction, and data visualization. It is a technique used to find a linear combination of features that best separates the classes in a dataset. The technique operates by reducing the dimensionality of the data and optimizing the distance between the categories. To do this, a group of linear discriminants that refine the proportion of between-class and within-class variances are identified. By maximizing the distance between the means of two classes and minimizing the variance within the individual class, it can be effectively used both as a dimensionality reduction algorithm and as a classifier. Moreover, LDA is also known as Normal Discriminant Analysis (NDA) or Discriminant Function Analysis (DFA) and can be used alternatively (or, if employed as a classifier, sequentially) after the PCA application.

B. MODEL SELECTION

This section provides a catalog of the most frequently employed ML and DL techniques to analyze oncological data obtained through the RS, with the aim of presenting a concise theoretical outline for all the most important methodologies encountered in this review. Each of the mentioned models has been suggested as the best model in at least one of the studies considered in the comparison table 1.

1) LINEAR REGRESSION (LR)

LR is a simple supervised learning technique exploited to establish a statistical relationship between a dependent variable and one or more independent variables. This algorithm requires a linear correlation between these variables and is commonly applied for regression analyses.

2) K-NEAREST NEIGHBORS (K-NN)

The K-NN algorithm [129] is a straightforward non-parametric technique often used for classification and regression tasks. Starting from a specific data point within the dataset, it works by computing which class is the most common among its k-nearest neighbors. The corresponding class category is then assigned to the evaluated point. On the other hand, the k-NN regression case instead evaluates the

average of the values of the k nearest neighbors and returns it as the prediction value.

3) SUPPORT VECTOR MACHINES (SVM)

SVM [130] is a type of supervised learning algorithm that is very frequently employed for both classification and regression tasks. The core idea behind SVM involves the creation of a hyperplane that maximizes the margin, or distance, between itself and the nearest elements of each class, commonly referred to as support vectors. A regularization term denoted as C is typically adjusted to balance the tradeoff between the mentioned margin size and the tolerance for misclassification The algorithm can also be applied to multiclass classification scenarios by employing a one-on-one approach across all classes, sequentially, to determine their respective categories. With the SVM, it is also possible to effectively execute a non-linear classification through an implicit mapping (with the so-called kernel trick) of the inputs into a feature space of higher dimensionality. In this way, the technique can also be applied when dealing with data that is not linearly separable. Several possible kernels are available for this purpose, as shown in the table 1.

4) EXTREME GRADIENT BOOSTING (XGBOOST)

The XGboost algorithm [131] is a decision tree (DT) based ML algorithm commonly used for classification and regression and provided by the XGBoost open-source software library. The DT is a popular non-parametric ML algorithm employed to make inferences about a given set of observations. It employs simple tree-like structures known as classification (or regression) trees, in which leave nodes represent class labels and branches represent feature arrangements that lead to those class labels. An example of a DT-based ensemble learning algorithm is Random Forest, usually applied to simple classification and regression tasks. Random Forest builds multiple decision trees on randomly selected subsets of the training data and combines their predictions to enhance overall accuracy. In contrast, XGBoost follows an ensemble learning approach, using a gradient-boosting framework to sequentially train multiple decision trees while minimizing a specified loss function. This approach often leads to superior accuracy and performance compared to a single decision tree, although it comes at the expense of some interpretability.

5) MULTI-LAYER PERCEPTRON (MLP) OR ARTIFICIAL NEURAL NETWORK (ANN OR NN)

The MLP methodology draws inspiration from the structural organization of the human brain [132] and involves a series of layers composed of interconnected nodes, commonly referred to as neurons. This approach interleaves layers of neurons, which perform weighted sums of inputs, with non-linear activation functions. By introducing non-linearities within the system, the network gains the ability to model and capture complex, non-linear relationships within the data. Complex network structures can be constructed by adding multiple layers with different shapes and operators, increasing the overall number of neurons, and allowing the identification of more intricate patterns within the input distribution.

6) CONVOLUTIONAL NEURAL NETWORK (CNN)

CNN is a widely utilized ANN-based architecture [133] for tasks such as classification, detection, or segmentation of signals, images, and videos. CNNs are based on the application of a convolutional operator within the network layers to improve the extraction of spatial features from the input data. This allows the model to learn hierarchical representations of patterns while decreasing the number of learnable parameters, therefore improving the training time and reducing the likelihood of overfitting with respect to a traditional ANN. Inside the CNN layers, multiple filters are employed to extract relevant information from the input data. A sequence of convolutional layers is capable of achieving optimal performance thanks to the decomposition of the initial information into meaningful and hierarchical features. This enables CNNs to distinguish increasingly complex patterns while processing the information through the network. Additionally, to further generalize the learned features, pooling and dropout layers are commonly employed in this type of network. Through them, input spatial dimensions are reduced by some criteria (e.g., maximum value, mean value, random shutdown), making the network more efficient and robust to small input perturbations.

7) LONG SHORT-TERM MEMORY (LSTM)

the LSTM [134] is a specialized type of recurrent neural network extensively applied in sequence prediction and classification tasks. In particular, a recurrent neural network is a type of ANN designed to process sequential data thanks to recurrent connections built within the network. These connections enable the model to link with previously generated predictions obtained from data already analyzed to capture temporal or sequential information. This methodology faced some issues with longer sequences, where the information was not retained correctly by the network and was lost after some time. This issue was overcome through the LSTM model, in which knowledge of the past is maintained through a special memory and gating mechanisms to selectively forget or remember specific elements in the sequential data. In this way, the architecture can effectively handle long-term dependencies while retaining contextual information over extended periods.

C. MODEL EVALUATION

As Raman-based early diagnosis in oncology is crucial due to its potential impact on cancer treatments, it is fundamental to use the appropriate model evaluation metrics, validate them from various perspectives, and conduct investigations that assess the model's performance in real-world clinical settings. This work will not cover potential human feedback in the evaluation pipeline, but it is important to note that

collaborating with human professionals with expertise in the field is essential for a comprehensive evaluation of the techniques.

The most common and traditional performance measure for a model is the **accuracy** value. This metric evaluates the overall reliability of the output of a model by determining the number of correct predictions divided by the total number of them. However, because it doesn't differentiate or weigh the specific types of data, it can be misleading when dealing with unbalanced datasets. In medical diagnosis scenarios, for example, class imbalances often occur due to data availability and patient conditions, and a simple accuracy assessment may hide significant weaknesses of the model, especially when dealing with anomalies or rare cases. For instance, a significantly higher number of non-disease cases may obscure a rare disease. Moreover, accuracy does not take into account that in a medical setting predicting as cancer a benign tumor sample (false positive) is usually considered a less severe error than classifying as benign tumor a cancer sample (false negative). To deal with these problems, other types of measures are commonly employed and available to support an effective analysis, as can be seen in the performance metric column in Table 1.

The **sensitivity**, also called recall or true positive rate, measures how effectively the model recognizes positive cases. This metric estimates the proportion of accurately predicted positive cases (true positives) in relation to the total number of actual positive instances. A high level of sensitivity demonstrates the model's capacity to accurately detect a majority of positive situations, thereby reducing the occurrence of false negatives. High sensitivity ensures that a medical-oriented model can find as many true positive cases as possible, therefore reducing the chances of missing a potentially serious condition.

Alongside sensitivity, the metric of **specificity** is commonly employed to assess the capability of a model to accurately identify negative occurrences. It is defined as the ratio of correctly predicted negative cases (true negatives) in relation to the total number of actual negative instances. A high level of specificity indicates that the model is successful in minimizing the occurrence of false positives. A high specificity value ensures that the model minimizes the risk of incorrectly identifying a negative occurrence, contributing to more reliable decision-making and for instance avoiding unnecessary diagnostic procedures and treatments.

Among other possible performance measures, the Receiver Operating Characteristic (ROC) curve and the Area Under the Curve (AUC) [135] are two of the most frequent ones. These metrics are used to evaluate the discriminatory capability of a classification model in distinguishing between two distinct categories, as described below:

The **ROC** curve demonstrates the model's ability to classify positive instances while simultaneously minimizing the occurrence of false positives. The curve is a graphical representation of the sensitivity (true positive rate) against the specificity (false positive rate).

The **AUC** is a quantitative metric commonly used in ML applications [136], frequently in conjunction with the ROC curve, representing the region beneath it. A higher AUC value indicates superior performance in classifying instances, while a low value indicates poorer model performance. Fundamentally, AUC metric measures the overall performance of the model across different threshold values. For example, an AUC value of 0.5 implies that the model's performance is equivalent to random chance. On the contrary, an AUC equal to 1.0 means a classifier that is flawless. Hence, a greater AUC indicates optimal classification performance, while the ROC curve visually illustrates the model's effectiveness across various sensitivity and specificity trade-offs.

Although the described metrics offer meaningful insight into the performance of a model, it is important to acknowledge that they may not provide a comprehensive representation of the model's capability to generalize the input distribution properly. Validation approaches are a helpful addition to these measures, assessing how well a model can generalize to previously unseen data and its consistency in real-world scenarios. Selecting an appropriate validation strategy, considering factors such as the data characteristics, model complexity, and validation criteria is fundamental for evaluating the performance and generalization capabilities of both ML and DL models. These techniques can support the model design and development process to avoid issues such as overfitting or underfitting and address challenges related to the bias-variance tradeoff problems.

A basic but fundamental approach to validation involves the use of the **train-test split** technique. This method involves partitioning the dataset into two distinct subsets: a training set, which is employed for model training purposes, and a testing set, which is utilized to assess the model's performance. This methodology is used to understand the model's performance on novel and unobserved data, thereby reducing possible overfitting concerns. However, it is important to acknowledge that the reliability of performance estimation can be compromised due to the specific split of data into training and testing sets considered.

To overcome this issue, the **K-Fold Cross-Validation** is considered a robust approach, especially when dealing with small datasets. Firstly, the dataset is partitioned into K subsets, known as *folds*. During each iteration, one fold serves as the test set, while the rest are employed for training purposes. Then, this procedure is iterated K times, and each fold is utilized as the test set exactly once. The outcomes produced from these iterations are then averaged to get an assessment of the model's overall capabilities in generic settings.

One critical problem is the presence of relevant variability among individual subjects and of different acquisition sessions that must be taken into account in medical disciplines, as in the case of RS. The biological composition of each individual is distinct, and its inherent variability (that can be inter- or intra-subject variability) can result in substantial variations of Raman spectra. In fact, the presence of genetic

diversity, variations in health conditions, and differences in lifestyle factors can lead to significant signatures in the Raman data patterns collected. Investigators must consider and address these distinctions when developing diagnostic or analytical models to guarantee the generalization and dependability of such models across a wide range of patients.

The **leave-One-Patient-Out** (or similar multiple patient variations) validation or cross-validation method is a valuable technique useful to mitigate the inherent bias present in data collected from a small group of patients. This bias is a common challenge in ML and it assumes even greater significance in the medical domain (as highlighted for example in [137] for automated EEG analysis). By employing this technique, the data gathered from specific patients and employed for training purposes is then excluded from the testing phase. Consequently, this eliminates any model bias associated with known patients, preventing it from artificially inflating performance metrics and leading to misleading results that might not adequately generalize to the broader population. When employing leave-one-patient-out k-fold Cross-Validation, the approach is iterated for all the k-folds considered, ensuring that the model is tested each time on unseen patients with respect to the training group considered.

D. DATASETS

One of the most debated issues of ML regards the properties of the considered dataset, in particular the dataset size. In fact, large-sized datasets prevent the risk of overfitting, with important consequences in terms of the model robustness and generalization capabilities to unseen data.

In the special case of RS employed in combination with ML for cancer detection, the dataset size refers to the number of patients involved, the number of samples available for each patient, and the number of Raman spectra collected for each sample. In particular, among the references analyzed in this work, we couldn't find any significant correlation between the size of the collected datasets and the particular Raman technique adopted or the type of samples under interest. However, a possible limiting factor could be represented undoubtedly by the availability of patients corresponding to the cases of interest. This issue is particularly relevant for relatively uncommon cancer among the population, such as the melanoma cancer. Another limiting factor for the dataset size is represented by the availability of samples suitable for Raman spectroscopy. In particular, in the references taken into consideration in this survey paper, the samples treated for RS are mostly fresh tissues, immersed in liquid nitrogen or liquid biopsies, often functionalized with nanoparticles to exploit enhancement effects. These fabrication techniques are supposed to be specifically designed for RS. However, especially in RS of tissues, the samples available are often obtained from biopsies, stored in tissue banks, previously treated with formalin, and embedded in paraffin, not specifically conceived for RS. Due to not well-defined fabrication protocols and/or random causes, the resulting samples are not always suitable for RS.

These limiting factors for the dataset size are particularly crucial for the DL algorithms, which need particularly large datasets to avoid overfitting. In this case, to reach the desired dataset size, a solution could be represented by the obtainment of a large number of spectra from the same samples, with possible drawbacks in terms of data redundancy. Another strategy, already mentioned in Section V, could be represented by data augmentation. The properties related to the size of the dataset employed in the references analyzed in this survey paper are shown in Table 2.

VI. RESULTS

A total of 54 studies from the selected cancer-types have been considered to describe the current state of the art concerning the automated analysis of Raman spectra obtained from cancerous tissue. The following types of cancer have been considered the most representative of the field, on which this review focuses: *pancreatic cancer, breast cancer,, skin cancer, brain cancer, prostate cancer, ovarian cancer, oral cancer*.

A. PANCREATIC CANCER

Pancreatic cancer is one of the leading causes of cancer death in the United States for both men and women [38], [138], with a very low 5-year survival rate (12%) for all its stages combined. It is estimated that by 2030, it will become the second leading cause of death in malignant tumors [21]. The symptoms usually do not appear until the disease has progressed, among which we can find weight loss, abdominal discomfort, and sometimes diabetes. Tumors that develop near the common bile duct can cause the yellowing of the skin and eyes (jaundice), which sometimes helps with the diagnosis. In the advanced stages, the disease may cause severe abdominal pain, nausea, and vomiting [138]. Moreover, the use of alcohol and tobacco, diabetes, obesity, and genetics are some of the most prevalent factors that increase the risk of developing pancreatic cancer. Surgical interventions have the potential to enhance survival and symptomatology [42], as well as radiation therapy and chemotherapy. Intraoperative frozen section analysis and postoperative histopathology examination are two common technical procedures used to diagnose pancreatic cancer [38]. These conventional methods have significant drawbacks because they can be time-consuming, subjective to expert interpretations, and dependent on tissue preparation; at the same time, those methods can suffer from sampling bias and be limited by the number of sample points available for the analysis. In contrast, RS-based spectral histopathology offers a new perspective on cancer diagnosis that has the potential to facilitate effective intraoperative tissue diagnosis [139].

Li et al. in [38] reported the first effort to diagnose pancreatic cancer through Raman scattering with a lab-developed system designed for intraoperative applications. The considered dataset is composed of 1305 Raman spectra taken from cancerous pancreatic tissues, and 1224 samples taken from healthy pancreatic tissues. 2D samples were generated by the

TABLE 2. Summary respectively reporting the number of patients, samples, and resulting Raman spectra analyzed in this survey.

Tumor type	Authors and Year	Number of Patients /Samples /Spectra
Pancreatic cancer	Li et al., 2021 [38]	20 mice / 1305 cancerous, 1224 healthy
	Mandrell et al., 2020 [22]	- /13 / 51
	Sezer et al., 2022 [20]	- / - / 254
	Yan et al., 2020 [21]	18 / 18 / -
	Carmichael et al., 2019 [42]	- / 121 / 121
	Aslam et al., 2024 [43]	20 mice / - / 2529 spectra
	Uthamacumaran et al. [44]	9 / - / 18
Breast cancer	Zhang et al., 2022 [45]	6 cell lines / 900 cells / 4500 spectra
	Li et al., 2022 [46]	271 / - / 813
	Zeng et al., 2023 [47]	- / 75 / 225
	Fuentes et al., 2023 [48]	- / - / 3054 (day 1), 4203 (day3, 5Gy), 4705 (day 3, 15Gy)
	Fuentes et al., 2024 [49]	- / - / 3480
	Ma et al., 2023 [50]	- / - / 1160 (400 from MCF-10 Aexosomes, 380 from MDA-MB-231 exosomes and 380 from MCF-7 exosomes)
	Ma et al., 2021 [51]	20 / - /600
	Li et al. [52]	16 / 16 / 125
Skin cancer	Bratchenko et al. [53]	54/617/617
	Araújo et al., 2021 [54]	-/84/436
	Qui et al., 2023 [55]	-/10/20000
	Wu et al., 2021 [56]	-/149/149 + a variable number of GAN generated samples
Brain cancer	Zhang et al., 2023 [57]	52/63/2220
	Vrazhnov et al., 2023 [58]	60/60/-
	Quesnel et al., 2023 [59]	FFPE tissues: 30/30/750; serum: 30/30/150
	Hollon et al., 2023 [60]	153/-/-
	Ma et al., 2022 [61]	83/83/415
	Jin et al., 2022 [62]	-/-/10361
	N. Iturrioz-Rodríguez et al., 2022 [63]	4 / 4/ 60(-)
	Chen et al., 2023 [64]	-/-/83
	Tian et al., 2022 [65]	83/83/415
	Bukva et al., 2021 [66]	138/138/ 690
	Livermore et al., 2020 [67]	73/73/1825 (normal brain), 9990 (fresh tissues)
Prostate cancer	Ao et al., 2023 [68]	83/83/-
	Chen et al., 2021 [69]	84/84/501
	Grajales et al., 2022 [70]	18/18/47
	Milligan et al., 2022 [71]	09/24/05
	Milligan et al., 2022 [72]	9/36/5950
	Picot et al., 2022 [73]	32/32/228
	Zhao et al., 2022 [74]	106/106/-
	Grosset et al., 2020 [75]	483/483/ 3381 (on average)
	Ma et al., 2020 [76]	75/75/-
	Medipally et al., 2020 [77]	76/76/-
	Lee et al., 2019 [78]	4(subtypes) /4/1200 (augmented set)
Ovarian cancer	Chen et al., 2022 [79]	174/174/870
	David et al., 2022 [80]	9/9 /76
	Giamougiannis et al., 2021 [81]	38/38/345
	Hunter et al., 2021 [82]	-/-/-
	Moisoiu et al., 2019 [83]	253/253/-
	Perumal et al., 2019 [84]	111/111/-
	Paraskevaidi et al., 2018 [85]	55/55/-
Oral cancer	Chang et al., 2023 [86]	90/180/ 16200
	Ghosh et al., 2022 [87]	42/42/176
	Koster et al., 2022 [88]	53/53/-
	Wang et al., 2023 [89]	183/68/-
	Xia et al., 2020 [90]	24/24/-
	Yan et al., 2020 [91]	22/44/1604
	Jeng et al., 2019 [92]	80/80/400
	Yan et al., 2019 [93]	24/24/216
	Zhang et al., 2019 [94]	78/78/ 18750
	Yu et al., 2019 [95]	24/24/288
	Sharma et al., 2023 [96]	64/128/640
	Chaudhuri et al., 2022 [97]	-/-/65

dot product of the Raman spectra with their transpose, creating a two-dimensional representation by changing the input shape while maintaining the same quantity of information. The latter were used as input of a DL pipeline, together with one-dimensional RS samples and the principal component of 2D samples projected into one dimension. A 1D-CNN and a 2D-CNN were used to classify the data, and a five-fold cross-validation strategy was applied to test them. The final

accuracy is 98–99% for almost all the models presented, with a very fast convergence rate. Additionally, an explainable AI effort was made to identify the key Raman characteristics that aid in this differentiation. After training the networks, the mean Raman spectra for both the one-dimensional and two-dimensional cases of cancerous and normal pancreatic tissues were loaded into their corresponding CNNs. The strongest activation channels were then extracted from the max-pooling layer outputs, and the learned CNN features were plotted on the Raman shifts. Therefore, some critical CNN Raman features were visualized in the max-pooling layers. The results of this study were compared with GoogleNet [140] and random forest models, with the proposed model obtaining better performances than those models.

A ML approach taking as input Raman spectra of a human pancreatic cell line is reported in the work of Mandrell et al. [22]. The article tackles the task of distinguishing between tumor-repopulating cells (TRCs) and parental control cells while determining the best combination of data type, dimension size, and classification technique to differentiate the cell types. 51 Raman spectra were collected from 13 samples: eight parental controls for the 37 spectra and five TRCs for the 14 spectra. An accuracy of 98% is obtained from SVM and kNN classifiers. Moreover, two refinement steps are applied to the input data: the dimension elimination (discarding unwanted data) and the dimension selection (selecting only certain wavenumbers) processes. In the first case, very noisy regions of a spectrum are chosen for elimination. Because the Raman fingerprint region of the skeletal vibrations is reported to lie between 600 and 1800 cm^{-1}, with the high energy signals of biologicals appearing in the region 2500–3500 cm^{-1}, rather than scanning from 250 to 3500 cm^{-1}, a targeted scanning from 600 to 1800 cm^{-1} and 2500–3400 cm^{-1} produces the required data while limiting the dimension size to 1900. This approach was motivated by prior knowledge of the salient Raman fingerprint regions of the skeletal vibrations and the high-energy signal positions. Moreover, in order to further decrease the dimensions, the spectral values that correspond to the local maxima were taken into account. Finally, the authors compared traditional continuous spectrum data with peak data. The peak data considered only the wavenumbers that corresponded to a local maximum. If any spectrum had a peak at a particular dimension, then that wavenumber was used in the analysis; otherwise, a value of zero was assigned to that dimension. For the dimensions selection, a series of techniques are employed to accomplish this task: T-statistic, MIT correlation [141], Relief algorithm [142], and PCA. The first two provide a statistical measure of the correlation of each wavenumber to the class where the sample belongs. The third is a ML-based algorithm to find a weight for each dimension based on the distance to its nearest neighbor from each class. The last method, PCA, converts a sample space with dependent variables into an independent space while retaining nearly all of the original space variances. A 14-fold cross-validation strategy was applied, in which the dataset was divided into 14 subsets (one for each TRC cell), and then the classifier performed 14 training and testing phases. After the cells were extracted, they were evaluated using modular Raman microscopy equipment to prevent excessive cell degeneration. In addition, a number of optimization procedures were implemented to improve the signal-to-noise ratio and repeatability of the Raman signal. For example, a baseline correction was implemented by matching the baseline segment to a polynomial function using an in-house-developed algorithm.

Sezer et al. [20], the pancreatic ductal adenocarcinoma (PDAC) is analyzed using RS, and the data is classified with 97.6 percent accuracy using a PCA-assisted vector machine algorithm. PDAC is the most frequent form of pancreatic cancer. It has a very low survival rate, and an early diagnosis can not only prevent metastasis development but also increase the efficacy of chemotherapy treatment. The phenomenon analyzed in the article is called epithelial-to-mesenchymal transition (EMT) and occurs in a variety of processes, such as embryonic development, wound healing, fibrosis, and early-stage tumor metastasis. A particularly useful tool for the monitoring of cellular processes such as cell division, apoptosis, and EMT is the RS. By varying peak intensities and shifting peak positions in the Raman spectra, the researchers wanted to see if indomethacin (a nonsteroidal anti-inflammatory drug) had any effect on an induced EMT (with the growth factor TGF-β) in a pancreatic cancer cell line. Additional details are available in section 3.3 of [20], where can be found a detailed description of the spectral changes reflecting the chemical composition before and after TGF-β treatment. Raman signals collected from untreated (control) and TGF-β1 treated cells in the presence or absence of indomethacin (of 100 μM and 300 μM) revealed EMT-related chemical changes with differences in both signal intensities and band positions. The spectral data were preprocessed employing background subtraction and average smoothing. The PCA algorithm was applied on 254 total Raman spectra (101 TGF-β, 48 TGF-β + indomethacin 100 μM, 50 TGF-β + indomethacin 300 μM, 55 control) to reduce the input dimensionality. Four different ML classifiers were run on Raman data: SVM, AdaBoost, kNN, and decision tree. A combined approach of a PCA-SVM model obtained the best result of 97.6% of accuracy. Other molecular techniques (i.e., immunostaining and qRT-PCR) were used to confirm the treatment results in order to evaluate the outcome with a more time-consuming procedure that requires specific reagents. Unlike the latter, the study demonstrates that Raman imaging is an excellent method for obtaining label-free, nondestructive spectral classification of the samples with minimal sample preparation and time constraints.

In Yan et al. [21] the blood serum is studied by using a confocal Raman micro-spectrometer able to distinguish between benign and malignant pancreatic human tumor samples. Fresh blood samples from 18 patients with pancreatic cancer and 10 patients with a benign tumor stage were utilized. The samples were analyzed in the 400 to 1800 cm-1 wavelength

region, and the major molecular differences in the serum of patients with pancreatic tumors were identified by changes in alcohols, lipids, amides, and nucleic acids. Fluorescence interference, a typical phenomenon that affects Raman scattering, was avoided by performing a preprocessing step. The authors adopted partial least squares (PLS) to extract the Raman spectrum information of patients with pancreatic tumors to minimize complications caused by models with a large number of parameters. 815 features available in the high-dimensional input spectrum were projected into the low-dimensional data space via PLS, and then the samples were classified employing LDA, SVM, and KNN classification algorithms. Both the SVM model (with a linear, quadratic, and cubic kernel) and the KNN technique (with a neighbor count of 1) obtained a promising result of 96.4% of accuracy. Moreover, the results were verified using a five-fold cross-validation strategy.

Conforming to the previously reported articles, a dimensionality reduction technique and a ML classifier are frequently described as standard procedures for approaching the discussed task. For instance, to classify spectral pancreatic samples, Carmicheal et al. [42] use a principal component-discriminant function analysis model (PC-DFA). Within the proposed system, exosomes are employed as biomarkers in order to classify the input samples. The exosomes are extracellular vesicles that form inside multivesicular compartments of eukaryotic cells and are secreted when these compartments fuse with the plasma membrane [143]. Their identification can be fundamental to understand their role in cancer progression. Soung et al. [144] employed exosomes to distinguish pancreatic cancer cells from healthy pancreatic epithelial cell lines. Moreover, using the Vancouver Raman method, the raw spectra collected from SERS measurements were normalized and baseline-corrected. Then, applying the PC-DFA 121, Raman spectra were first reduced from 1004 variables (from 719 cm^{-1} to 1800 cm^{-1}) to 20 principal components (PCs), then classified with 90% of accuracy by the aid of PC-DFA.

In their study, Aslam et al. [43] tackled the task of detecting PDAC using a variety of analytical techniques applied to distinctive explainable features acquired from RS data. Through these features the authors successfully identified mutations related to Kirsten-rat-sarcoma-viral-oncogene-homolog and tumor-suppressor-protein-53 in the fingerprint region of PDAC, providing valuable insights into the molecular characteristics of the disease. Their study was based on RS data collected from 20 mice, each injected with human cells subcutaneously. The dataset comprised over 2500 RS signals, including readings from both tumor and normal pancreas tissues. To ensure the reliability of their results, the analysis was conducted with a 15-fold cross-validation process. In addition, to uncover crucial information regarding the spectra, several feature extraction techniques (statistical analysis, empirical mode decomposition, and peak feature extraction) have been used, obtaining valuable insights exploited in the analysis process. The analysis was performed employing a SVM with a recursive feature elimination [145] and the correlation bias reduction methods. This approach outperformed other traditional techniques with a final classification accuracy of 98.5%, while reducing testing time and memory usage.

Uthamacumaran et al. [44] conducted a study using a small dataset comprising nine patients across four different cancer subtypes (colorectal cancer, hepatocellular carcinoma, breast cancer, and pancreatic cancer), along with five healthy control patients. Spectra were obtained from RS analysis of blood serum samples, paired with the Fourier Transform Infrared (FTIR) spectroscopy, obtaining from these techniques 19 and 15 spectra respectively. Their investigation focused on extracellular vesicles (EVs) found in the sera of cancer patients, which contain a diverse range of biomarkers reflective of their cell of origin. These EVs are of particular interest in liquid biopsy and cancer screening research. To analyze them, the the authors employed several machine learning algorithms including AdaBoost Random Forest Classifier, Decision Trees, and Support Vector Machines, which successfully distinguished between the baseline-corrected Raman spectra of cancer EVs and those of healthy controls. This discrimination achieved a classification accuracy of 100% when the spectral frequency range was narrowed down to 1800-1940 cm^{-1} and subjected to a 50:50 training-testing split. In addition, the result also incorporated a five-fold cross-validation approach, although there were noted uncertainties in the outcomes. Moreover, given the very small dataset size and the perfect result scores, the outcome raises possible concerns about potential overfitting.

It is important to remark that the state-of-the-art research concerning pancreatic cancer led to very few available articles of recent years that analyzed the disease with the aid of RS and artificial intelligence techniques, thus highlighting the novelty in the field. For instance, in the thorough summary of the AI applied to pancreatic cancer tissue in the literature done by [146], only one study applied the RS technique to analyze the pathology differently from the most traditional approaches (PET, MRI, CT, EUS, etc.), which have currently a much broader use.

B. BREAST CANCER

Breast cancer is the second-leading cause of cancer death as well as the most frequently diagnosed cancer in women [138]. The death rate has declined in the last 20 years, thanks to improvements in early detection and treatment. The most common symptom of breast cancer is a lump or mass in the breast, which is the main indicator for a diagnosis together with the analysis of external factors such as weight gain, physical inactivity, heavy smoking, and alcohol consumption [138]. X-ray mammography is the current gold standard technique for breast mass screening, and it is commonly used to diagnose breast cancer. However, mammography performs poorly in thick breasts and cannot discern easily between malignant and benign lesions. Depending on the

specific cases, the same limitations can apply to ultrasound and magnetic resonance imaging (MRI) processes [147]. Consequently, a needle or surgical excision biopsy is still necessary to develop a precise diagnosis. The latter may add weeks to the diagnostic procedure, increasing patient anxiety and medical treatment expenses. Depending on the severity of the disease, treatments typically involve mastectomy or breast-conserving surgery. Moreover, the biological behavior can significantly differ between patients, despite common histopathological features at diagnosis, with the current 5-year relative survival rate equal to 90% [148]. Complex gene expression patterns, which may be exploited to create molecular signatures of breast cancer, have practical significance for long-term patient outcomes but are not yet utilized to guide therapies. A more thorough investigation and evaluation of the molecular patterns of the illness in each individual patient could be fundamental to achieving an effective ad-hoc breast cancer therapy. Towards this objective, the development of a novel approach for the detection of tumor molecular phenotype is one of the possible next fundamental steps in the field of tumor classification and diagnostics [148], [149].

Zhang et al. [45] tackle this problem through a series of classification tasks to distinguish cancerous breast cells from different levels of analysis. In the study, roughly 150 cells were recorded for each of the 6 cell lines, resulting in approximately 4500 Raman spectra with a Raman shift ranging from 300 cm^{-1} to 1800 cm^{-1}. A series of transformations were applied to clean the input data and remove noise, such as cosmic ray removal, background subtraction by polynomial curve fitting, and normalization. The dataset was divided into three subsets of equal size, two of which were designated as training sets and one as a test set. It is important to note that this splitting approach is more prone to overfitting than other validation techniques, such as cross-validation, but can be considered adequate for preliminary analysis with enough sample sizes and is thus considered informative for a comparison with similar studies. The initial phase of the analysis involves a binary classification task aimed at distinguishing between normal and cancerous breast cells. The task is carried out employing PCA-SVM and PCA-DFA techniques, with very promising accuracy results (90%). The study continues with a multi-classification procedure that categorizes breast cancer into its four possible subtypes: luminal A type, luminal B type, human epidermal growth factor receptor (HER2)-positive type, and triple-negative breast cancer (TNBC). The analysis was performed using two different approaches commonly employed in the field: PCA-DFA and PCA-SVM. The PCA was used to reduce the number of dimensions in the data from 1024 to 351, keeping 95.0% of the variance. The accuracy of both algorithms in distinguishing cell subtypes was greater than 92%. Lastly, using two cell lines (SUM149 and MDA-MB-231) a supplementary classification of TNBC into two subtypes is performed. The overall performance of PCA-DFA and PCA-SVM models is 70.9% and 70.7%, respectively.

Li et al. [46] analyzed the serum Raman spectra of 171 patients with invasive ductal carcinoma and 100 healthy participants. Using a leave-one-subject-out validation paradigm, the samples were randomly divided according to a train-test ratio equal to 7:3. After normalization, the data were filtered using the adaptive iteratively reweighted penalized least squares algorithm and the Savitzky-Golay algorithm [150] for baseline correction, smoothing, and denoising. The PCA algorithm was used to extract 102 features, and then multiple ML algorithms (SVM, DT, LDA) and a custom fully connected network were used to classify samples based on both raw and reduced data. The neural model achieved the highest level of accuracy, scoring a perfect 100%. It is worth noticing that the models trained on the raw data performed better than the ones working on PCA-reduced data. This finding goes in the opposite direction with respect to the majority of other research works and indicates that the PCA noise reduction and data cleaning capabilities are usually crucial to obtaining the best performances. In fact, several other studies performed the classification through a dimensionality reduction technique and a ML classifier, as in the case of [151], [152], and [153].

Together with traditional ML models, the current lines of research focus also on deeper and more complex AI models, which exploit a greater number of trainable parameters as well as more expressive power, such as deep neural networks. Zeng et al. [47] proposed a diagnostic system based on serum RS and DL algorithms, where a fast and low-cost diagnosis method is used for screening 75 serum samples for breast cancer and healthy controls. In particular, 69 spectral samples are obtained from 23 samples of TNBC, 60 spectra from 22 samples of HER2-positive breast cancer, and 90 spectra from 30 healthy control samples. In total, 225 spectral measurements were obtained and exploited for this study.

To conduct the statistical analysis, the authors filtered the data through the Savitzky-Golay algorithm, used for smoothing and denoising, after the application of a linear normalization step. Following that, a 6-fold cross-validation paradigm is used to ensure the model reliability, with an 8:2 split ratio between the train and test sets. The custom models employed in this work are based on the NN, CNN, and BiLSTM architectures, with resulting accuracy rates of 87.78%, 90.37%, and 91.11%, respectively. For the healthy control group, all models achieve perfect classification accuracy (100%), which could be attributed to the large difference in bio-molecule concentrations in serum between the control group and breast cancer patients.

Another interesting possible RS application is to monitor radiation-based treatment response up to a molecular level. In particular, the application of those treatments can result in structural changes in tissues, which can be measured by the RS. Fuentes et al. [48] propose the use of a Raman microscope to identify biochemical in-vivo radiation

response in xenografted human breast tumors. Radiotherapy employs high-energy ionizing radiation to eliminate tumor tissues while limiting collateral harm to healthy tissues. However, a fraction of breast cancer patients did not react well to the radiation treatment, resulting in high cancer recurrence rates (42%) [154]. The final aim of this work has been to develop customized treatment plans thanks to the ability of RS to detect and monitor changes in radioresistance in metabolic biomarkers. In order to pursue this goal, the authors implemented a 1D-CNN to classify irradiated versus non-irradiated tissue with a leave-one-subject-out validation strategy. A total of 3054 spectra were sampled on day 1 after irradiation, and 6708 spectra were collected on day 3. Moreover, it is important to remark that the Raman spectra used in this study were previously collected in another study that employed female mice [155], thus limiting the generalization capabilities of the model to human subjects. Each spectrum was pre-processed to remove cosmic rays, subtract background noise, adapt to wavenumber calibration drifts, and normalize to the total area under the curve. The system obtained a final accuracy of 92.1% for data collected 3 days post-irradiation, and 85.0% on day 1 post-irradiation. Moreover, the proposed approach shows improved discrimination capabilities compared to a random forest model and to a previously implemented algorithm [156].

In their latest research, Fuentes et al. [49] explored the use of RS and CNN to characterize tumor response to radiotherapy, with a focus on identifying its degree of radioresistance with an explainable AI technique. Their dataset comprised Raman spectra gathered from three distinct human tumor cell lines (breast, lung, and prostate cancer) cultivated in vitro, classified as either radiosensitive or radioresistant across various treatment doses and time points. The CNN architecture employed in this work was refined based on the one described previously in [48], simplifying the model by reducing the network parameters to mitigate over-parameterization and overfitting risks. The model achieved an exceptional mean classification accuracy, sensitivity, specificity, and F1 score with a reported mean matrix value of 99.8%. The evaluation metrics demonstrated again the ability of the model to accurately classify tumor responses to radiotherapy. The training procedure involved a simple 70-20-10 split with early stopping technique, that typically converged within only 35 epochs. Additionally, an interesting explainability approach based on the Gradient-Weighted Class Activation Mapping technique was employed to visualize critical discriminative features captured by the 1D model, analyzing each feature map contribution of the last convolutional layer to the final score. Heatmaps generated with this procedure revealed contributions from Raman bands associated with glycogen, amino acids, and nucleic acids, while radiosensitive cell lines showed activations at lipid and phospholipid bands.

An exosome-based analysis is performed by Ma et al. [50], where the classification is performed through SERS data, acquired with a Raman microscopy system. Exosomes were employed as biomarkers thanks to their ability to reflect the genetic and phenotypic status of the cells from which were obtained. The authors propose an enhanced methodology for their separation, specifically regarding the extraction process through ultracentrifugation. Subsequently, after the sample preparation, the SERS signals are obtained from 130 randomly selected locations and then averaged in order to obtain a data sample. A baseline correction and fluorescence background subtraction of the spectra, as well as a min-max normalization, are performed as data preprocessing. The authors do not explicitly declare the use of a specific validation technique, so it is assumed that no subject-wise split validation paradigm was used, where the dataset division is reported to be 70% training, 10% validation, and 20% testing. In the end, a total of 1160 SERS spectra are classified between normal breast cells and cancer cells with a Resnet-based CNN structure, obtaining a final accuracy of 95%. The findings suggest that the integration of SERS detection and CNN screening holds promise for the early detection of breast cancer. This is supported by the extremely sensitive nature of Raman detection and the consistent presence of exosomes during the incubation period of cancer cells. Consequently, this exosome-based approach may have potential for future applications in the early diagnosis of cancer cells.

Other works leverage DL models to overcome the traditional ML approaches, as in the case of the work of Ma et al. [51]. In the latter, through a DL technique, the authors analyzed a dataset containing 600 spectra taken from 20 patients, where from the same patient both a healthy and a cancerous breast sample were retrieved. As preprocessing, a traditional baseline correction was performed with the automatic background subtraction algorithm. To deal with the issue of limited data, the authors implemented the following data augmentation techniques: a 2 cm^{-1} shift of the original spectra (left or right, considering also the spectral range is 800–1800 cm^{-1} with a resolution of 3 cm^{-1}), an addition of random Gaussian noise, and a scaling process with a random coefficient. In turn, thanks to this data augmentation schema, the total number of spectra increased by more than eight times, from 600 to 5000. A 1D convolutional neural network is selected to perform the binary classification, which obtained the best classification performance of 92% accuracy. Its results are then compared to the PCA-SVM and PCA-LDA algorithms in order to assess the model capabilities. However, although ten-fold cross-validation is used to validate the results on the ML models, it is not clear to the reader if the convolutional neural network has been tested on the same validation paradigm.

Li et al. [52] developed a novel methodology centered on a feature fusion strategy for the analysis of Raman spectra, aimed at mitigating the signal-to-noise ratio inherent in RS. This approach involved a multiparameter serial encoding evolutionary algorithm combined with an adaptive Local Hyperplane K-nearest Neighbor classification algorithm for adaptive hyperparameter optimization. The latter was

employed in the classification task of 125 spectra from sixteen breast tissue samples, gathered through RS from 16 patients. In particular, the experiment utilized a QE65000 miniature Raman spectrometer equipped with a 785 nm Raman laser, covering a scanning range from 0 to 2723 cm^{-1}.

To address challenges such as high-frequency noise, baseline drift, and fluorescence background interference, a combination of techniques including Savitzky-Golay smoothing, Standard Normal Variate, and adaptive iteratively reweighted penalized least squares algorithm were applied to the spectra. Moreover, given the limited spectral data and the imbalance in the number of different types of data, the Synthetic Minority Oversampling Technique algorithm was employed to increase the number of spectra and regulate the proportion between normal and tumor samples. After the preprocessing, the dataset was split into train and test sets with a ratio of 2:1 (train:test), with both classes (normal, tumor) containing an equal number of spectra. Feature selection was initially conducted using the Relief Algorithm to identify a representative train and test a subset of the original data based on a predetermined threshold. Additionally, another train and test subset was derived from PCA processing on the original data. Both the methodologies derived two sets, later employed for the classification. The Classification was performed using the Adaptive Local Hyperplane K-nearest neighbor algorithm, fine-tuned by the Multi-parameter serial encoding evolutionary algorithm to optimize hyperparameters. Notably, the model yielded optimal hyperparameters and accuracy scores, finding very good classification performances and exhibiting enhanced performance compared to manual configuration.

To ensure robustness, the methodology incorporated 10 sets of parallel experiments through random sub-sampling. A mean accuracy score of 98.31%, sensitivity of 96.45%, and specificity of 100% were achieved. Consistent results were obtained with an 80% vs. 20% train/test split ratio. On the other hand, to enhance generalizability, an additional validation approach was employed, with a 10-fold cross-validation approach, resulting in a mean accuracy of 94.15%, specificity of 89.03%, and sensitivity of 99.12%.

In conclusion, the literature demonstrates that the RS technique is effective as a noninvasive method for analyzing various cancer types, including breast cancer. In addition, RS-based in vivo analysis of breast cancer could lead to personalized and fast prediction results, assisting medical personnel and reducing diagnostic delays. As the correct identification of the type of cancer is essential for selecting the appropriate treatment, further investigations may yield novel and more effective methodologies in this domain.

C. SKIN CANCER

Skin cancer is the most prevalent malignancy among white people, accounting for about one-third of all malignant neoplasies diagnosed annually [53] and resulting in more than 40% of global total cancer cases [56]. There are three major types of skin cancer, which are: basal cell carcinoma (BCC), squamous cell carcinoma (SCC), and Melanoma (ME), and in particular BCC and SCC are the most common forms and together are referred to as non-ME skin cancers. Although less common than the others, ME is more lethal and accounts for the majority of skin cancer deaths [157]. Among the possible causes, the majority of skin cancer cases are caused by exposure to UV radiation, and thus are potentially preventable [138]. Visual inspection together with a preliminary clinical screening is generally sufficient to perform an adequate diagnosis, with the addition of procedures like dermoscopy, biopsy, and histological studies to further improve the accuracy [158]. Moreover, to diagnose and treat skin cancer at an early stage, new or changing skin spots or growths must be recognized in time and monitored. Any change in the lesion appearance should be evaluated by a specialist. For instance, by visual inspection, the properties of asymmetry, border irregularities, color, diameter size, and appearance changes are the warning elements of melanoma [138]. Once identified, most cases are cured by removing the lesion through minor surgery, while immunotherapy or targeted drugs might be options for the more advanced stages. The five-year relative survival rate is over 99% if the cancer is entirely confined in the place of origin, while decreases to 32% if the cancer has spread from the primary tumor to remote parts of the body [138]. Several investigations have been conducted to analyze skin cancer with AI-based techniques, leveraging deep neural networks and clinical images to classify with a level of accuracy comparable to dermatologists [158], [159]. On the other hand, a diagnosis based only on visual inspections may be not precise enough in the case of more complex cases, which could benefit instead from a thorough analysis taken on a molecular level. Optical biopsy can be an essential element to detect tumors based on their spectral features resulting from the comparative presence of different chemical components [53]. Among the several possible fields of application, Raman spectra are effectively employed to conduct different types of studies on the skin. Recently, RS combined with AI procedures has been adopted as a standard to distinguish different kinds of skin patterns. For instance, Ye et al. [160] used a ML-based classification system to distinguish burn severities from ex vivo porcine skin tissue, while Kanemura et al. [161] employed a near-infrared RS and ML techniques to classify inflammatory skin diseases. In the case of skin cancer detection, several investigations used RS as the main tool to analyze malignant skin samples, obtaining analysis up to a molecular level.

Bratchenko et al. performed multiple studies [53], [162], [163] on a portable spectroscopic system for in vivo skin neoplasms diagnostics by Raman and autofluorescence analysis. Due to the proposed portable system affordability and simplicity, the device is considered suitable to be employed in small clinics by first-hand professionals as well as expert oncological specialists. In the most recent advancement found in [53], Bratchenko et al. analyzed 617 skin tumor samples, considering 204 malignant and 413 benign

neoplasms. The examined data were gathered in previous ML-based research including 54 white European American patients [163]. They used a train-test split and 10-fold cross-validation, with 90% of the dataset employed as a training set and the remaining 10% as a test set. Finally, the previous splitting procedure was repeated 10 times in order to obtain the final mean receiver operating characteristic area under the curve (ROC AUCs) results. In this investigation, the authors demonstrated that a CNN-based approach produced superior results than the PLS-DA classification of Raman signals of skin tumors. The CNN custom architecture considers a residual block in parallel with two convolutional and two pooling layers. The sum of the outputs is then passed to a recurrent layer and to two fully connected layers. When this model was used to differentiate between malignant vs benign tumors as well as melanomas vs pigmented tumors and melanomas vs seborrheic keratosis, ROC AUCs of 0.96, 0.90, and 0.92 were respectively obtained. The developed model outperformed the PLS-DA approach, which yielded inferior ROC AUCs. Finally, to determine the informativeness of individual predictors, the variable importance in projection (VIP) score for the PLS loadings was evaluated.

Araújo et al. [54] proposed an interesting methodology to select the optimal model and minimum necessary segment of the input spectra to classify with high performances ME and benign melanocytic nevus (MN) samples. In particular, the MN group consisted of superficial and deep melanocytic nevus, while the ME group was composed of primary and metastatic melanoma spectra. Data were collected in vitro using an FT-Raman spectrometer (with 2 cm^{-1} resolution) on 33 MN samples and 51 ME samples, which in turn have been obtained through surgical resection. Despite the traditional implications concerning the autofluorescence phenomenon, they decided to use the unprocessed raw data as input signals. Five to ten spectra were collected from each sample, with a total of 436 Raman spectra, of which 168 were taken from the MN group and 268 from the ME group. After that, the local properties of the spectra were investigated by segmenting them into subsequences, each with a predetermined length. A variety of statistical methods (the arithmetic mean, standard deviation, Kurtosis skewness, derivative, minimum, and maximum) were used to extract different features on these vectors. The evaluated features were fed as input to a fast LightGBM [164] model, which obtained a slightly better AUC ROC measurement (0.98) with respect to other ML-based techniques (Random Forest, KNN, XGBOOST). The result was validated using the traditional five-fold cross-validation technique, where the final result value is the mean of the five test run results. The analysis discovered that a resolution size of 5 cm^{-1} for the subsequences provides the best results. After applying the LightGBM model to the input spectral subsequences, the derivative output was found to be the most informative feature. The latter identification was performed using the SHAP [165] technique, which measures the influence of input features' absence or presence on the model decisions in order to evaluate the relevance of those characteristics. In addition, to discover the most informative interval of the Raman samples, the authors generated multiple datasets composed of cropped spectra, with sample lengths ranging from 20 cm^{-1} up to 300 cm^{-1}. Afterward, thousands of mini-models were trained on these subsets for the ME vs. MN classification task. At the end of the procedure, 158 models with an AUC of 0.96 or higher were found. Among them, the model that obtained the best AUC value (0.973) with the smaller spectral regions (from 896 to 1039 cm^{-1}) was selected as the optimal one, highlighting the fact that even a minimized spectrum of only 143 cm^{-1} can be effectively exploited for the classification task. It is important to note that all the model performances were evaluated using a five-cross-validation paradigm.

Several other applications follow the traditional approach of a dimensionality reduction technique followed by a traditional ML algorithm to classify skin cancers with the use of RS (e.g. the use of in PCA-SVM of Qiu et al. in [166] or of Liu et al. in [167]) on small datasets with very promising results. On the other hand, the literature shows also an intensive use of DL techniques that, combined with the Raman technology, are able to produce autonomous skin cancer diagnosis with high accuracy.

Qui et al. [55] have devised a CNN network model to differentiate between healthy and various types of malignant cells. The Raman spectra samples were collected using a SERS substrate and a confocal Raman spectrometer with a resolution of roughly 1 cm^{-1}. They employed Raman spectra of ten different types of cell line samples, for a total of 20000 spectra obtained by collecting 2000 spectra from each cell line sample. All the spectra were normalized within the measurement range, and pre-processed by cosmic ray removal, smoothing, and baseline correction. To classify the samples, a DL architecture based on a small custom CNN with 4 convolutional layers, 4 max-pooling layers, and two fully connected layers was employed. The proposed model achieved a perfect accuracy of 100%, which converged faster if trained on the preprocessed data with respect to the raw data. In addition, in order to analyze the performances of the proposed approach, their model has been compared with a baseline PCA procedure followed by K-nearest neighbor. This comparison could be unfair given the differences in the number of model parameters and their complexity, and may be inadequate to assess the methodology capabilities. In addition, to validate the proposed CNN model, the authors employed a 10-fold cross-validation approach with a 7-2-1 split for the training, testing, and validation sets. In addition, to categorize multiple subtypes of the tumor under exam, the authors applied the classification procedure at different possible biological levels. They proceeded with a progressive approach: firstly, a binary model was employed to distinguish melanocytes from malignant ME, with 98.5% accuracy. Then, using only samples in the ME category,

an additional classifier has been applied to distinguish between wild-type and mutant ME, with 99% accuracy. After these steps, another CNN model was used to classify two wild-type ME and three mutant ME cell types, with 100% and 98% accuracy, respectively. Afterward, the authors produced a dataset to assess the efficacy of the proposed CNN model in discriminating between several forms of drug-resistant ME. This dataset is composed of ME cell line samples (A375 and M14) that were subjected to varying concentrations (0.5 and 2.0 μM) of vemurafenib, a kinase inhibitor (also known as a cancer growth blocker). The investigation obtained satisfactory results, with A375 and M14 cell lines samples classified with almost 100% and 93% of accuracy, respectively. The described results demonstrate how RS spectra classification through AI can potentially provide a novel direction for the characterization of drug-resistant (or drug-treated) melanoma cells.

Wu et al. [56] investigated the use of generative strategies to address the data scarcity problem, which is a common issue in the field. The authors introduced a DL system capable of identifying skin cancer spectral samples through the exploitation of synthetic data. More in detail, they applied a generative adversarial network to estimate the implicit distribution underneath the input data. This enabled the production of additional input data samples that strongly resembled the original ones. These newly generated samples were consequently fed as input to train the employed classification models. In turn, this study focuses on the multi-category task of discriminating three distinct tissue types, namely BCC, SCC, and healthy samples. In particular, this work employed a previously analyzed dataset [168], used to explore the correlation of Raman spectral features of normal and malignant tissues subjected to a high-powered CO2 laser treatment. Since the lack of correlation in the tissue was proven in previous research, both laser-treated and untreated samples were considered suitable for this study. The dataset consists of 36, 63, and 50 RS samples for BCC, SCC, and healthy categories, respectively. To augment the volume of data, two strategies were considered: the first aimed to reduce the class unbalance in the input training distribution by producing a fixed amount of GAN-generated samples; the second approach considered the relative proportion between the input categories, thus generating samples to keep the dataset balanced. Several models, trained with the same number of epochs (500) and with a leave-one-out cross-validation scheme, have been used to find the best architecture and to demonstrate the methodology effectiveness. The models analyzed included logistic regression and SVM both with or without PCA, and MLP, LSTM, and CNNs. Moreover, almost all the methods have been tested with and without data augmentation. The results of the study indicate that DL models have superior performances compared to ML techniques. Additionally, classifier models that were trained on augmented data demonstrated better accuracies in comparison to those trained only with original data. A customized 1D-CNN architecture, composed by a 1D convolutional and two dense layers, resulted in the best-proposed model.

The health and economic burden of skin cancer treatment is substantial and increasing, as shown by Guy et al. [169], where the authors emphasized the importance of skin cancer prevention and early detection efforts, which may lead to timely treatments as well as healthcare system cost savings. Moreover, skin examinations are an effective method for detecting skin cancer in its early, more manageable phases. In particular, in its first stages, skin cancer has the highest likelihood of being successfully treated. Moreover, the accuracy of RS-based, skin cancer classification models can be greater than or comparable to the accuracy of trained dermatologists [53]. On the other hand, due to the high background noise that heavily affects RS measurements, a precise tumor classification is not trivial to achieve. Hence, it can be challenging to categorize skin cancers effectively with a single type of data. Further research on this subject could contribute to the creation of an autonomous, reliable, and portable analysis device for skin cancer diagnosis, therefore reducing analysis delays and healthcare costs. For instance, by applying the discussed techniques in a multimodal framework, which could be based on dermoscopy- and spectroscopy-based diagnostic systems, the classification accuracy of skin neoplasms could be further improved.

D. BRAIN CANCER

Brain tumors can be classified within the category of primary and metastatic neoplasms of the central nervous system (CNS). With respect to other forms of cancer, brain tumors are considered rare, with around 7 cases per 100000 people registered in Europe in 2018 [170]. However, the brain cancer-related prognosis is commonly considered poor, with a mortality of more than 5 cases per 100000 people every year [170]. Among the various cases of brain cancer, almost 70% are high-grade and invasive. In particular, glioblastoma is characterized by an average survival that doesn't exceed 15 months [171].

Early-stage diagnosis of brain cancer is universally considered a crucial step, decisively affecting the effectiveness of the following therapies and, thus, the patient's quality of life and survival. Today, the most common approaches for brain mass detection are based on PET, MRI, and CT. Besides the poor contrast, long acquisition times, the use of intense magnetic fields or ionizing radiation, etc., such approaches alone do not lead to a final diagnosis. On the other hand, invasive and difficult tissue excisions and consequent histopathological analysis are required to provide a final response. All these difficulties pushed the researchers to find non-invasive and accurate alternatives to complement the commonly used diagnostic approaches. In this sense, the investigations about the possibility of detecting brain cancer from blood-derived samples, such as serum or extracellular vesicles, aroused particular interest [172]. Such approaches

are hampered by the presence of the so-called Blood-Brain-Barrier, i.e., a semi-permeable membrane that selectively regulates the molecular crossing from the brain to peripheral circulation. The presence of Blood-Brain-Barrier determines low concentrations of cancer-related biomarkers in the peripheral blood. This feature makes the use of RS for the detection of such biomarkers particularly challenging, due to the weakness of the Raman signal. However, several works [58], [59], [61], [64], [65] have proved that ML protocols allow to detect brain cancer from Raman spectra obtained from blood-derived liquid biopsies. For example, Tian et al. [65] experimented with various DL techniques for the detection of glioma from Raman spectra of blood-derived serum samples. For this study, they employed a statistical sample made of 83 patients, 38 with glioma and 45 control cases. The corresponding Raman spectra were retrieved with a Confocal Raman Microscope, working with a laser source at 532 nm. the final dataset included 415 Raman spectra, 190 for the malignant class and 225 for the control cases. The DL pipeline consisted in PLS feature extraction, data augmentation, and the application of several Convolutional Neural Networks. Thanks to 5-fold cross-validation, the study revealed that PLS+GoogLeNet corresponded to the maximum performance with sensitivity, specificity, and accuracy exceeding 98%. The excellent prediction capability shown in this work suggests the applicability of RS and ML for brain cancer detection from non-invasive liquid biopsies.

As regards the treatment of brain cancer, the gold standard is made by the combination of surgical excision and Chemo- or Radiotherapy. In this sense, one of the crucial issues is represented by the need to maximize the amount of the resected malignant mass by preserving the surrounding healthy tissues. Despite several approaches, such as intraoperative MRI [173], fluorescence-guided surgery [174], and electrophysiological mapping [175], have been introduced to help the surgeon during the mass resection, today one of the most relevant causes of morbidity and mortality related to brain tumors is researched in the recurrence of the tumor mass in the proximity of the boundaries of the resection cavity [63]. This finding suggests the failure of the surgical treatment, with portions of malignant tissues left in the resection cavity [176]. Such dramatic statistics pushed the researchers to find fast and accurate approaches for the estimation of the tumor boundaries. Due to the fast response and label-free character of RS, this technique has been widely studied in combination with ML for the accurate estimation of brain tumor boundaries [57], [60], [62], [67]. In particular, Jin et al. [62] proposed a protocol to detect brain tumor boundaries that exploit the reduction of the pH induced by the tumor mass on its boundaries. The concept of this assessment is based on putting a water droplet in contact with brain tissues. Then, the droplet is placed on a chip coated with metal nanoparticles for SERS investigation. The resulting Raman spectra of these samples were employed to feed a custom Convolutional Neural Network for an ultrasensitive prediction of the pH of the portion of tissue that has come into contact with the water droplet. Based on this principle, it is possible to probe the tissue pH in several points of the intracranial area and to reconstruct pH maps, highlighting the pH of the tissues point-by-point. To verify the effectiveness in detecting the tumor boundaries, the aforementioned principle was tested in SERS-guided surgical excisions of glioma in mice. Concerning the conventional strategies, which are usually based on MRI-assisted surgery, SERS resulted in a significantly increased survival rate of the mice, revealing itself promising for future implementations in common practice.

Alongside SERS, Stimulated RS constitutes an alternative solution to the problem of increasing the Raman signal-to-noise ratio in the assessment of brain tumor boundaries. In this sense, Stimulated Raman Histology represents a promising approach, providing spatially-resolved molecular and morphological properties of brain tissues. In particular, Hollon et al. [60] performed a study aimed at comparing Stimulated Raman Histologic images of brain tissues to traditional Histologic images. This investigation was based on the use of a sophisticated DL protocol, called DeepGlioma, in which a Convolutional Neural Network was fed with 2.5 millions of Stimulated Raman Histologic Images, obtained from an initial cohort of 278 patients. The trained Convolutional neural network reached accuracies exceeding 93%, in line with the performances of the traditional histopathological exam. Furthermore, the DL model was capable of detecting peculiar features of histopathological images, i. e. tumor infiltrates within healthy tissues.

The promising results regarding brain cancer detection on brain tissues through RS coupled with ML pushed the researchers to realize engineering tools aimed at helping the surgeon distinguish the tumor boundaries during the excision in a real-time fashion. For example, Zhang et al. [57] adopted a portable VRR-LRR^{TM} Raman analyzer under quasi-clinical conditions to detect glioma from fresh human brain tissues. The apparatus worked with light at wavelength of 532 nm. The Raman technique was coupled with ML pipeline based on the application of PCA and SVM. The model was trained with 2220 spectra, retrieved from 53 patients. The technique allowed to distinguish normal and malignant tissues with an accuracy exceeding 80%, with relevant implications in terms of the future implementations of the aforementioned technologies in the common practice.

E. PROSTATE CANCER

Prostate cancer turns out to be the second most frequent form of tumor in men, with more than one every ten cases of cancer, with almost 7% of cancer-related deaths in male in 2018 worldwide [177]. In addition, among all the prostate cancer cases registered every year, around 97% involve men after 50 years of age [178]. Today, despite the presence of well-established screening protocols, there isn't an universally recognized diagnostic recipe for prostate cancer. Unlike the Digital Rectal Examination, the diagnostic

gold standard is represented by the measurement of the level of Prostate-Specific Antigen (PSA). This antigen is secreted by the epithelial cells of prostate and its level is usually assessed in blood. In particular, high levels of PSA in blood correspond to suspect cases of prostate cancer. It is commonly recognized that this protocol suffers from lack of sensitivity and accuracy, especially in distinguishing cancer from the so-called Benign Prostatic Hyperplasia. For this reason, PSA alone leads to high percentage of false positive cases and invasive and painful biopsies are generally required to obtain a final diagnosis through histopatological analysis [179], [180], [181]. The result is the so-called Gelason Score (GS), a parameter that, according to the original definition, assumes integer values ranging between 1 and 5, representative of the appearance and the differentiation of tissues in the histological pattern. The definition of GS has been subjected through the years to several revisions and corrections [182]. Therefore, novel and non-invasive screening protocols are needed for a more effective detection of prostate cancer and for sensitively reducing painful and unnecessary biopsies. In this sense, a possible strategy is represented by the analysis of liquid samples, such as blood, serum, seminal fluid or urine, by measuring the presence of specific biomarkers, such as micro-RNA, circulating tumor cells or exosomes [178]. For example, Medipally et al. [77] conducted a study whose goal was to search prostate cancer from blood plasma. To this aim, 76 patients were selected, 43 of whom were affected by prostate cancer, while the others were healthy. The technique Raman adopted was based on the use of Confocal Raman Microscope equipped with a laser emitting light of wavelength 785 nm. PCA applied to the resulting spectral components put in evidence differences between the benign and the malignant sub-groups. Furthermore the algorithm revealed also differences related to the Gleason Score, i. e. to the degree of malignancy. The principal components were than used to train a ML model based on Partial Least Squares-Discriminant Analysis, leading to sensitivities and specificities ranging from 90% to 99%. Finally, a deconvolution of the characteristic peak convolution was carried out through a Classical Least Squares fitting procedure, highlighting an increase in the content of interleukin-6, phosphatidylethanolamine, creatinine and RNA and a decrease in the concentration of albumin and β-carotene in the malignant patients with respect to the healthy ones.

RS and ML were also employed in the monitoring of the effect of therapies on the progression of prostate cancer. For example, Picot et al. [73] developed a Raman fiber-optics system for the trans- rectal examination of prostatic tissues, aimed at monitoring the effects of brachytherapy on the disease recurrence. The apparatus was fed with a laser emitting light at 785 nm and it was tested on a sample made of 18 patients with prostate cancer and 14 tissues retrieved from patients subjected to prostatectomy. The resulting dataset was composed of 138 Raman spectra. A ML model based on SVM was applied with the purpose of distinguishing normal from cancerous tissues. The model was tested *in situ* in 28 benign and 21 malignant cases, resulting in 79% accuracy, 86% sensitivity and 72% specificity.

F. OVARIAN CANCER

Ovarian cancer turns out to be the second leading cause of cancer-related deaths among women worldwide, with a mortality of around 60% registered in 2018 [183]. The reasons for such dramatic statistical data can be found in the absence of symptoms representative of the early stage of the disease. Furthermore, despite the efforts devoted by the researchers, today there are no biomarkers conceived for an early-stage diagnosis of the tumor. Recent screening campaigns [184], performed on women with suspect symptoms, led to a reduction in tumor mortality, highlighting the need for novel, easy in use and powerful techniques to perform an early-stage diagnosis and thus increase the 5-year survival rate of the patients.

Today the non-invasive screening protocols for the detection of ovarian cancer are based on the research in blood samples of specific biomarkers, such as Cancer Antigen 125 (CA-125) or HE4 [185]. The main drawback of these approaches can be found in the lack of sensitivity, especially in the detection of early-stage cases. For this reason, the employment of multiple biomarkers could represent a possible solution to improve the detection performances. In this case, future challenges are represented by the choice of combinations of biomarkers representing a good compromise between the need to maximize the detection performances and the need to avoid the employment of useless and/or redundant biomarkers [186]. As regards the imaging of ovarian tissues, today the most widespread approach is based on the use of ultrasounds, either in a transvaginal or in a transabdominal fashion, with the weakness of not distinguishing benign from malignant masses.

The common therapy for ovarian cancer consists of surgical removal and chemotherapy, accompanied by the monitoring of the disease recurrence through specific biomarkers, such as CA-125, and imaging techniques. The pitfall of this stage of the therapy is represented by the possible presence of micro-metastasis, possible causes of cancer recurrence, and often hard to detect [80]. In particular, it has been demonstrated a positive correlation between the presence of micro-metastasis after surgical therapy and an increase in tumor morbidity [187]. For example, David et al. [80] developed an RS-ML protocol for the detection of ovarian cancer by directly measuring ovarian tissues, with the aim of offering an additional tool to detect the presence of micro-metastasis. Their sample was constituted of 9 patients, 4 of whom with ovarian cancer, 3 with endometrial cancer, 1 with benign mucinous cystadenofibroma, and 1 with unknown diagnosis. The Raman analysis was performed on fresh ovarian tissues and with a Confocal Raman Microscope, equipped with a laser emitting light at 785 nm. The resulting dataset was subjected to an SVM algorithm with a LASSO regression. The resulting dataset was subjected to

k-NN analysis. The protocol allowed the detection of eight characteristic spectral bands, used for the prediction. The final performances in distinguishing cancer were assessed through a Leave-One-Out cross-validation, and consisted of accuracy, sensitivity, and specificity of 90%, 93%, and 88%, respectively.

However, as regards the employment of the Raman technique in association with ML for diagnostic purposes, the research about ovarian cancer has progressively devoted its efforts to developing non-invasive protocols to detect cancer from body fluids, such as blood plasma, serum, cystic fluid, ascitic fluid etc [79], [81], [83], [84], [85]. In this sense, the SERS concept has been widely adopted to increase the Raman signal-to-noise ratio of the characteristic biomarkers for ovarian cancer. Moisoiu et al. [83] experimented with the use of SERS in the detection of ovarian cancer in human blood serum samples functionalized with silver nanoparticles to perform signal enhancement. For this study, they gathered a cohort of 253 patients, 39 healthy volunteers, 42 with breast cancer, 109 with colorectal cancer, 33 with lung cancer, 17 with oral cancer, and 13 with ovarian cancer. The Raman spectra were retrieved with a Confocal Raman Microscope by probing the Raman effect with a laser source at 532 nm. The resulting Raman spectra were used as input variables for a standard ML model based on the consecutive application of PCA and LDA. The performances of the model were assessed through the 5-fold Cross-Validation technique. The model showed excellent performances in distinguishing benign from malignant cases with an average sensitivity and specificity of 98% and 91%, respectively. In addition, the model has proved to be effective in distinguishing the types of tumor, with accuracies of 88% for oral cancer, 86% for colorectal cancer, 80% for oral ovarian cancer, 76% for breast cancer, and 59% for lung cancer. Giamougiannis et al. [81] described a technique for the identification of ovarian cancer from human blood plasma and ascitic fluid, based on the use of a standard Raman Confocal Microscope, equipped with a laser emitting light at 785 nm. The study involved 20 healthy patients and 18 with ovarian cancer, with the aim of performing a binary classification. The resulting initial dataset was composed of 385 Raman spectra. After a dimensionality reduction, employing PCA, the resulting data were subjected to two different classifiers, i.e. LDA and SVM. The detection performances were assessed through a 10-fold cross-validation revealing a maximum accuracy of 82% in the detection of cancer from ascitic fluid. Paraskevaidi et al. [85] carried out a study aimed at distinguishing ovarian cancer from the human blood plasma retrieved from 55 patients, by employing a standard Confocal Raman Microscope operating at 785 nm. The Raman spectra were subjected to a Support Vector Machine classifier, adopting a Gaussian kernel, resulting in excellent detection performances, i. e. sensitivity of 94% and specificity of 96%. This last work testifies to the effectiveness of blood samples for the Raman-assisted detection of ovarian cancer.

G. ORAL CANCER

Oral cancer is defined as a form of tumor affecting the floor of the mouth, palate, tongue, alveolus, or buccal mucosa. This form of cancer turns out to be the sixth most frequent form of tumor worldwide, with mortality overcoming 60% within the first five years from the initial diagnosis [188]. Today, the commonly accepted diagnostic protocol for oral cancer is based on a visual examination and palpation of the tumor mass, aimed at distinguishing the tumor boundaries. This phase is followed by tumor resection, whose effectiveness is strictly correlated to the correct estimate of tumor margins. In particular, it was estimated that in the most widespread form of oral cancer, i. e. the oral cavity squamous cell carcinoma, in almost one out of every two cases the marginal resection is incomplete [91]. Furthermore, this operation is often made complex by the presence of tumor margins in deep tissue layers [189]. All these difficulties lead to the need for multiple and invasive resections and to a high probability of cancer recurrence with negative consequences in terms of the 5-year survival rate. In this sense, the combination between RS and ML has taken two paths. A first research line is focused on the estimation of tumor boundary from the analysis of fresh tissues. Jeng et al. [92] conducted a study on human fresh tissues aimed at detecting oral squamous cell carcinoma through a standard CRM operating in the visible range (532 nm). The patients' cohort was composed of 80 individuals, resulting in a dataset of 400 Raman spectra. The spectra were pre-processed through a Savitzky-Golay filter of order 3, followed by the removal of the basline component. Finally, the spectra were normalized by standardizing the Area Under Curve to the group median. The classification was performed by sequentially applying PCA and Quadratic Discriminant Analysis, resulting in an accuracy of 87.5%, a sensitivity of 90.9%, and a specificity of 83.3%. Sharma et al. [96] successfully employed a common Confocal Raman Microscope with laser light at 532 nm for the detection of oral squamous cell carcinoma from fresh biopsies. In particular, an ensemble of 64 patients with oral squamous cell carcinoma was considered. From each patient, two biopsies, of healty and malignant oral tissue, were excised. Five spectra were acquired for each sample of tissue, resulting in a total number of 640 spectra. Prior to the classification, the spectra were subjected to pre-processing operations, aimed at reducing the interference, removing the baseline component and eliminating the data redundancy. The classification was carried out by applying Partial Least Squared and Support Vector Machine sequentially. The resulting ML model discriminated malignant from healthy tissues with an accuracy of 94.74%, a sensitivity of 95.65% and a specificity of 93.33%.

These and other similar studies conducted *ex vivo* pushed the researchers to study measurements on fresh tissues with engineered Raman probes, in view of future applications *in situ*. For example, Chang et al. [86] compared various DL models (AlexNet, VGGNet, ResNet50, MobileNetV2,

and Transformer) for the interpretation of experimental data obtained from human fresh tissues excised from 90 patients. In particular, the employed tissues corresponded to control samples and forms of squamous cell carcinoma located in various positions within the oral cavity. The Raman setup was a custom-built fiber-based Raman apparatus working with a laser emitting light at 785 nm. This led to a final dataset including 16400 Raman spectra, employed to feed the aforementioned DL models. The collected Raman spectra were pre-processed by apllying, in the order, Savitzky-Golay filtering, baseline correction through the asymmetric weighted penalty least squares and MinMax normalization. The ML models were tested through a k-fold cross-validation technique, with $k = 1, \ldots, 5$. In those conditions, the resulting performances highlighted how ResNet50 was the best model in distinguishing normal and cancerous tissues, with a final accuracy of 92.81%, a precision of 92.93%, and a sensitivity of 92.86%. Xia et al. [90] employed DL techniques to detect cancerous oral tissues from para-cancerous tissues excised by patients affected by tongue squamous cell carcinoma. The initial cohort was composed of 12 patients. The Raman spectra were retrieved through a fiber-based setup equipped with laser emitting at a wavelength of 785 nm. The resulting Raman spectra, pre-processed with Savitzky-Golay filtering, asymmetric weighted penalty least squares basline correction and MinMax normalization, were employed to train a custom-built CNN adopted as a feature extractor. In other words, the CNN had the function of detectiong the relevant spectral components for cancer detection. The extracted features were employed to train an SVM with a Gaussian kernel. A 5-fold cross-validation procedure led to an overall accuracy larger than 99%.

As mentioned before, one of the most relevant difficulties addressed during oral cancer removal is related to the presence of tumor margins infiltrated into deep layers of tissues, especially when the tumor is at a late stage. Since the tumor mass develops from the external surface of the oral cavity inwards [190], novel and effective strategies to detect the tumor could represent a decisive step to detect the tumor mass when it is limited to surface layers of tissue, and therefore, easier to remove. A natural and minimally-invasive evolution of the aforementioned techniques based on the analysis of *ex vivo* tissue is based on the efoliation of epithelial cells. Chaudhuri et al. [97] employed this technique to whitdrawal oral cells from patients affected by oral squamous cell carcinoma. In particular, 65 Raman spectra were collected from healthy, pre-cancerous and cancerous patients. The Raman apparatus was a Confocal Raman Microscope employing laser light at 785 nm. The authors focused their attention in two classification tasks: a binary classification, aimed at distinguishing normal tissues from all the other types; a three-class classification, aimed at classifying normal, pre-cancerous and cancerous tissues. After a pre-processing phase, i. e. baseline removal and MinMax normalization, Linear Discriminant analysis was applied to reduce the system dimensionality. Then, the spectra were classified with Support Vector Machine, resulting in an accuracy of 95% for the binary classification problem and of 88.8% for the three-class based classification problem.

Another strategy for an non invasive detection of oral cancer is based on the employment of liquid biopsies such as blood serum [89], blood plasma [89], saliva [88], [89] or saliva-derived cells [87]. For example, Wang et al. [89] experimented with a SERS protocol for the early and non-invasive diagnosis of oral squamous cell carcinoma from blood serum and saliva liquid biopsies. The blood samples were collected from a cohort of 183 patients, while the saliva samples were retrieved from 68 patients. The study was focused on the individuation of three classes, corresponding to oral squamous cell carcinoma, neoplasm of the salivary glands, and healthy patients. The field enhancement was obtained through the use of silver nanoparticles meshed with the liquid samples, while the Raman signal was measured with a conventional Confocal Microscope equipped with a laser emitting light at 785 nm. The resulting Raman spectra were subjected to pre-processing operation, i. e. removal of fluorescence background and area normalization. The Raman spectra of blood serum put in evidence an association of cancer to the Raman peaks attributable to glycogen, lactose, L-tyrosine, and Phenylalanine. An SVM protocol allowed to demonstrate that the resulting Raman spectra allowed the discrimination of the three classes of interest considered pairwise, with an accuracy larger than ~83% and reaching ~92% in distinguishing adenocarcinoma and normal samples. As regards the saliva samples, the resulting performances were lower, with accuracies ranging between ~73% and ~85%. In conclusion, this work demonstrated the effectiveness of this SERS protocol combined with conventional ML for the detection of oral cancer from liquid biopsies.

VII. SUMMARY AND CONCLUSION

In this survey, we provided an overview of the most promising studies employing RS and DL/ML techniques for cancer grading. To this aim, we collected 60 previous works regarding the employment of this technique on seven types of tumors, by taking into consideration the most recent results available. In the papers analyzed, several experimental approaches were adopted to retrieve the Raman signal of interest, ranging from the widespread Confocal Raman Microscopy to the employment of engineered Raman probes. Besides the assessment of cancer through solid biopsies, the use of samples derived from body fluids has gained attention due to the intrinsic non-invasive nature of this kind of biopsies. In literature, the analysis of the Raman-derived experimental data involved several approaches belonging to ML or to its sub-field, i.e., DL, which is gaining more and more interest due to its ability to produce better classification performances.

A. AI ANALYSIS AND INSIGHTS

Based on the reported works findings, several possibilities are available when selecting an ML or DL-based model to

autonomously analyze Raman spectra. Our analysis suggests that each technique excels in specific problem scenarios, with no single method universally standing out as the superior choice. From the range of potential options, prevalent but effective approaches use a dimensionality reduction technique (mainly PCA) with an SVM or a specifically tuned deep neural network, yielding notable results on multiple types of cancer. However, the lack of standardized practices and survivorship bias in studies pose challenges to the field that could limit the ability to draw firm conclusions. Usually, within the same work, DL methodologies show improved accuracy in comparison to the ML approaches, with or without multiple preprocessing steps, often serving as benchmarks for custom implementations. Yet, this advantage might be attributed to a procedural bias in ML-based studies that is not consistently followed by other DL-based investigations. In addition, no thorough comparison with multiple dataset sources combined and/or on different Raman apparatus was found, highlighting the necessity of standardization of the Raman procedure and a joint effort to produce a common, larger dataset to perform further analysis. With more extensive, and widely recognized datasets, better-performing and more complex algorithms that are usually avoided due to data scarcity could be introduced in this field. Many studies (e.g., [47], [55], [56]) implemented custom NN-based architectures, which obtained very high classification accuracies (close to 100%) on small datasets. However, the real-world effectiveness of statistical models relies on their capacity to discern underlying patterns behind all possible types of data. However, a sufficiently large dataset that reflects the population across many scenarios is required, as well as a data variability which should be representative enough of the specific classes of data under consideration.

In other scientific domains, such as computer vision or natural language processing, larger and proven architectures are frequently utilized across various data types. Expanding the size and diversity of Raman datasets could not only mitigate overfitting but also introduce new baseline comparisons, leading to a significant enhancement in the quality and thoroughness of Raman spectroscopy analysis.

Additionally, even if synthetic data cannot fully replicate real-world scenarios, several AI techniques showcase the potential to generate synthetic yet reliable data, potentially bridging the gap between data-intensive, high-performing analyses and the more practical but less generalized current state of Raman data analysis. To tackle these challenges, a variety of deep generative methods can be utilized, such as GANs, variational autoencoders, diffusion models, flow-based models, and energy-based models, leading to new possibilities within the existing resources [191], [192], [193].

The integration of AI approaches with RS plays a crucial role in advancing medical diagnostics, with a primary focus on directly distinguishing between normal and cancerous tissues based on the distinct patterns within their Raman spectra. This application is pivotal in the field, as it enables precise and non-invasive identification of diseased tissues, offering significant potential for improving patient outcomes. However, despite its primary emphasis on tumor classification, the versatility of RS can advance beyond this scope. For instance, researchers like Fuentes et al. have demonstrated its utility in identifying biochemical radiation responses in xenografted human breast tumors, facilitating the monitoring of changes in the radioresistance of metabolic biomarkers [48]. Additionally, studies by Ma et al. explore the potential of Raman spectroscopy in analyzing exosomes, providing valuable insights into their diagnostic relevance [50]. This proves that different or even multiple tasks could be performed at once, enabling parallel and multi-level analysis.

Furthermore, several studies (e.g., [38], [43], [49], [54]) made efforts to provide explanations for their findings, illustrating when the outputs were useful for the implemented algorithms. This aspect holds significant importance as it can improve outcomes across various domains, particularly within healthcare. It is expected that XAI will gain increasing significance among all stakeholders, including users, developers, and those affected by AI systems [194]. For instance, in the healthcare sector, through XAI techniques physicians can understand which pathological features in input data guided the algorithm before exploiting the insight obtained from automatically generated diagnosis reports [194], [195]. This enhances transparency in diagnosis, fostering trust among both medical staff and patients. In addition, implementing these techniques can help to detect bias in the input data distribution or faults in the algorithm decision process, increasing reliability and the ability to face unexpected situations [194]. Thus, it is highly advisable to implement this procedure in the final version of an automated RS analysis system. However, the interpretability comes with a cost, due to the tradeoff between accuracy and intelligibility [196] that must be considered. Thus, simultaneously prioritizing both these aspects may represent the most prudent path forward in this field where responsibility and confidence are fundamental for effective decision-making.

B. FUTURE PERSPECTIVES

As stated in the previous sections, the richness of publications about cancer detection performed by coupling RS with ML/DL makes this combined approach a promising route, either for cancer screening or to help the surgeon during the tumor mass excision to assess the tumor boundaries, as prescribed by precision medicine. In some cases, the ML analysis of Raman-based experimental data allowed us to reach impressive detection accuracy, not only in distinguishing malignant from benign tumors but also in the detection of cancer sub-types, metastasis, tumor grade and location, etc. Specifically, several possibilities are available when selecting an ML or DL-based model to autonomously analyze Raman spectra, as shown in the table 1. These results constitute an encouraging landmark for future applications in engineered devices. Still, several issues remain open, representing potential insights for further investigations:

1) the time needed to obtain the final response inevitably places limitations on the applicability of such techniques. The assessment of the performance in terms of diagnostic time requires comparison with commonly accepted protocols. For example, cancer detection through the qualitative exam of stained solid biopsies requires a few minutes, while RS analysis takes from a few hours to several days [197]. Possible causes of long diagnostic times can be undoubtedly attributed to the width of the spectral range, to the number of spectra needed for a single diagnosis, and, consequently, to the computational cost of the classification.
   The commonly used Raman-based setups turn out to be cost-effective compared with other techniques, such as MRI or CT, with a difference of about one order of magnitude [198]. However, in view of a possible implementation of commonly used screening protocols, the aforementioned strength could not be sufficient.
   In this sense, identifying specific spectral bands relevant to the classification process could represent a crucial step, either for reducing spectral acquisition and classification times or for realizing Raman setups featuring relatively simple and affordable optics and/or detectors suitable for everyday use. Furthermore, the introduction of pre-processing and/or classification algorithms, designed to cope with signal noise resulting from the employment of less technologically advanced instruments, could help in reaching the desired levels of diagnostic accuracy. Finally, High-Performance Computing (HPC) could represent another solution to speed up the classifier algorithms.
2) Novel approaches can be explored by integrating information obtained from multiple sources, leading to comprehensive analyses and potentially yielding increased performance. For example, the integration of multimodal oncological data, such as the combination of Raman imaging and spectra with optical images for nevi monitoring or with an EEG scan for brain tumors could result in improved molecular analysis and more robust predictions, exploiting different points of view of the specific disease considered. This method exploits different viewpoints of the specific disease under consideration and, thanks to the increased quantity of data, opens up the possibility of using more recent architectures, such as Graph Neural Networks [199] and Transformers [200], that showed promising results in oncological scenarios when a multimodal learning approach was available [201].
3) The application of AI techniques introduces several legal and ethical issues beyond their practical effectiveness. It is crucially important to employ transparent methods that clarify the operations and reasoning of the algorithm in order to enhance confidence in the outputs and explain to the medical user the features recognized by the model in the inputs. Achieving knowledge of the model's inner process allows further investigation of signal patterns that are challenging to distinguish through visual inspections and human analysis and can potentially lead to more grounded and justified diagnoses.

VIII. GLOSSARY

AUC	Area Under the Curve.
AdaBoost	Adaptive Boosting.
AI	Artificial Intelligence.
ANN	Artificial Neural Network.
BCC	Basal Cell Carcinoma.
Bi-LSTM	Bilinear Long Short-Term Memory.
CE	Chemical Enhancement.
CLSM	Confocal Laser Scanning Optical Microscopes.
CNN	Convolutional Neural Network.
CNS	Central Nervous System.
CRM	Confocal Raman Microscopy.
CT	Computed Tomography.
DL	Deep Learning.
DT	Decision Tree.
EE	Electromagnetic Enhancement.
EEG	ElectroEncephaloGram.
EUS	Endoscopic Ultrasound.
GAN	Generative Adversarial Network.
GS	Gleason Score.
HER2	Human Epidermal Growth Factor Receptor.
KNN	K-nearest Neighbors.
LDA	Linear Discriminant Analysis.
LSPR	Localized Surface Plasmon Resonance.
LR	Linear Regression.
LSTM	Long Short-Term Memory.
LightGBM	Light Gradient-Boosting Machine.
ME	Melanoma.
ML	Machine Learning.
MLP	Multilayer Perceptron.
MN	Benign Melanocytic Nevus.
MRI	Magnetic Resonance Imaging.
PCA	Principal Component Analysis.
PDAC	Pancreatic Ductal AdenoCarcinoma.
PET	Positron Emission Tomography.
PLS	Partial Least Squares.
PSA	Prostate-Specific Antigen.
ROC	Receiver Operating Characteristic.
RS	Raman spectroscopy.
SERS	Surface-Enhanced Raman Scattering.
SCC	Squamous Cell Carcinoma.
SPR	Surface Plasmon Resonance.
TNBC	Triple-Negative Breast Cancer.
VIP	Variable Importance in Projection.
XGBOOST	Extreme Gradient Boosting.

ACKNOWLEDGMENT

Gianmarco Lazzini and Mario D'Acunto wish to thank TELEMO Project, under 146 Grant Bando Salute 2018, Regione Toscana.

REFERENCES

[1] *International Statistical Classification of Diseases and Related Health Problems*, 10th ed., World Health Org., Geneva, Switzerland, 2015.
[2] H. Shimizu and K. I. Nakayama, "Artificial intelligence in oncology," *Cancer Sci.*, vol. 111, no. 5, pp. 1452–1460, May 2020, doi: 10.1111/cas.14377.
[3] C. Pucci, C. Martinelli, and G. Ciofani, "Innovative approaches for cancer treatment: Current perspectives and new challenges," *Ecancer Med. Sci.*, vol. 13, p. 961, Sep. 2019.
[4] M. Blanco-Formoso and R. A. Alvarez-Puebla, "Cancer diagnosis through SERS and other related techniques," *Int. J. Mol. Sci.*, vol. 21, no. 6, p. 2253, Mar. 2020.
[5] S. Akbar, M. Hayat, M. Tahir, S. Khan, and F. K. Alarfaj, "CACP-DeepGram: Classification of anticancer peptides via deep neural network and skip-gram-based word embedding model," *Artif. Intell. Med.*, vol. 131, Sep. 2022, Art. no. 102349.
[6] S. Akbar, M. Hayat, M. Tahir, and K. T. Chong, "CACP-2LFS: Classification of anticancer peptides using sequential discriminative model of KSAAP and two-level feature selection approach," *IEEE Access*, vol. 8, pp. 131939–131948, 2020.
[7] S. Akbar, M. Hayat, M. Iqbal, and M. A. Jan, "IACP-GAEnsC: Evolutionary genetic algorithm based ensemble classification of anticancer peptides by utilizing hybrid feature space," *Artif. Intell. Med.*, vol. 79, pp. 62–70, Jun. 2017.
[8] A. Raza, J. Uddin, A. Almuhaimeed, S. Akbar, Q. Zou, and A. Ahmad, "AIPs-SnTCN: Predicting anti-inflammatory peptides using fastText and transformer encoder-based hybrid word embedding with self-normalized temporal convolutional networks," *J. Chem. Inf. Model.*, vol. 63, no. 21, pp. 6537–6554, Nov. 2023.
[9] S. Akbar, A. Raza, T. A. Shloul, A. Ahmad, A. Saeed, Y. Y. Ghadi, O. Mamyrbayev, and E. Tag-Eldin, "PAtbP-EnC: Identifying anti-tubercular peptides using multi-feature representation and genetic algorithm-based deep ensemble model," *IEEE Access*, vol. 11, pp. 137099–137114, 2023.
[10] G. Ghaly, H. Tallima, E. Dabbish, N. B. ElDin, M. K. A. El-Rahman, M. A. A. Ibrahim, and T. Shoeib, "Anti-cancer peptides: Status and future prospects," *Molecules*, vol. 28, no. 3, p. 1148, Jan. 2023.
[11] C. Berthomieu and R. Hienerwadel, "Fourier transform infrared (FTIR) spectroscopy," *Photosynthesis Res.*, vol. 101, pp. 157–170, Sep. 2009.
[12] L. Bachmann, D. M. Zezell, A. C. Ribeiro, L. Gomes, and A. S. Ito, "Fluorescence spectroscopy of biological tissues—A review," *Appl. Spectrosc. Rev.*, vol. 41, no. 6, pp. 575–590, 2006.
[13] J. S. D. Mieog, F. B. Achterberg, A. Zlitni, M. Hutteman, J. Burggraaf, R.-J. Swijnenburg, S. Gioux, and A. L. Vahrmeijer, "Fundamentals and developments in fluorescence-guided cancer surgery," *Nature Rev. Clin. Oncol.*, vol. 19, no. 1, pp. 9–22, Jan. 2022.
[14] C. Tang, J. Qiao, Y. Wen, Z. Zeng, S. Shao, and S. Dong, "Quality control of woody edible oil: The application of fluorescence spectroscopy and the influencing factors of fluorescence," *Food Control*, vol. 142, Dec. 2022, Art. no. 109275.
[15] J. Mirwald, B. Hofko, and H. Grothe, "Utilising fluorescence spectroscopy and optical microscopy to investigate bitumen long-term ageing," *Road Mater. Pavement Des.*, vol. 22, no. 1, pp. S23–S36, Jun. 2021.
[16] M. E. Khosroshahi and V. Woll-Morison, "Visualization and fluorescence spectroscopy of fingerprints on glass slide using combined 405 nm laser and phase contrast microscope," *J. Visualizat.*, vol. 24, no. 4, pp. 665–670, Aug. 2021.
[17] C. V. Raman and K. S. Krishnan, "A new type of secondary radiation," *Nature*, vol. 121, no. 3048, pp. 501–502, Mar. 1928.
[18] T. W. Bocklitz, S. Guo, O. Ryabchykov, N. Vogler, and J. Popp, "Raman based molecular imaging and analytics: A magic bullet for biomedical applications!?" *Anal. Chem.*, vol. 88, no. 1, pp. 133–151, Jan. 2016, doi: 10.1021/acs.analchem.5b04665.
[19] M. D'Acunto, R. Gaeta, R. Capanna, and A. Franchi, "Contribution of Raman spectroscopy to diagnosis and grading of chondrogenic tumors," *Sci. Rep.*, vol. 10, no. 1, p. 2155, Feb. 2020.
[20] G. Sezer, M. S. Onses, M. Sakir, F. Sahin, A. Çamdal, Z. Sezer, A. Inal, and Z. Ciftci, "Indomethacin prevents TGF-β-induced epithelial-to-mesenchymal transition in pancreatic cancer cells; evidence by Raman spectroscopy," *Spectrochimica Acta A, Mol. Biomolecular Spectrosc.*, vol. 280, Nov. 2022, Art. no. 121493. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1386142522006424
[21] Z. Yan, C. Ma, J. Mo, W. Han, X. Lv, C. Chen, C. Chen, and X. Nie, "Rapid identification of benign and malignant pancreatic tumors using serum Raman spectroscopy combined with classification algorithms," *Optik*, vol. 208, Apr. 2020, Art. no. 164473. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0030402620303077
[22] C. T. Mandrell, T. E. Holland, J. F. Wheeler, S. M. A. Esmaeili, K. Amar, F. Chowdhury, and P. Sivakumar, "Machine learning approach to Raman spectrum analysis of MIA PaCa-2 pancreatic cancer tumor repopulating cells for classification and feature analysis," *Life*, vol. 10, no. 9, p. 181, Sep. 2020. [Online]. Available: https://www.mdpi.com/2075-1729/10/9/181
[23] P. Manganelli Conforti, M. D'Acunto, and P. Russo, "Deep learning for chondrogenic tumor classification through wavelet transform of Raman spectra," *Sensors*, vol. 22, no. 19, p. 7492, Oct. 2022. [Online]. Available: https://www.mdpi.com/1424-8220/22/19/7492
[24] S. Pahlow, S. Meisel, D. Cialla-May, K. Weber, P. Rösch, and J. Popp, "Isolation and identification of bacteria by means of Raman spectroscopy," *Adv. Drug Del. Rev.*, vol. 89, pp. 105–120, Jul. 2015. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169409X15000629
[25] M. Eady, G. Setia, B. Park, B. Wang, and J. Sundaram, "Biopolymer encapsulated silver nitrate nanoparticle substrates with surface-enhanced Raman spectroscopy (SERS) for Salmonella detection from chicken rinse," *Int. J. Food Microbiol.*, vols. 391–393, Apr. 2023, Art. no. 110158.
[26] M. Saleem, S. Ali, M. B. Khan, A. Amin, M. Bilal, H. Nawaz, and M. Hassan, "Optical diagnosis of hepatitis B virus infection in blood plasma using Raman spectroscopy and chemometric techniques," *J. Raman Spectrosc.*, vol. 51, no. 7, pp. 1067–1077, Jul. 2020.
[27] N. Arend, A. Pittner, A. Ramoji, A. S. Mondol, M. Dahms, J. Rüger, O. Kurzai, I. W. Schie, M. Bauer, J. Popp, and U. Neugebauer, "Detection and differentiation of bacterial and fungal infection of neutrophils from peripheral blood using Raman spectroscopy," *Anal. Chem.*, vol. 92, no. 15, pp. 10560–10568, Aug. 2020.
[28] G. V. Nogueira, L. Silveira, A. A. Martin, R. A. Zangaro, M. T. T. Pacheco, M. C. Chavantes, and C. A. Pasqualucci, "Raman spectroscopy study of atherosclerosis in human carotid artery," *J. Biomed. Opt.*, vol. 10, no. 3, 2005, Art. no. 031117.
[29] L. P. Rangaraju, G. Kunapuli, D. Every, O. D. Ayala, P. Ganapathy, and A. Mahadevan-Jansen, "Classification of burn injury using Raman spectroscopy and optical coherence tomography: An ex-vivo study on porcine skin," *Burns*, vol. 45, no. 3, pp. 659–670, May 2019.
[30] L. Huang, H. Sun, L. Sun, K. Shi, Y. Chen, X. Ren, Y. Ge, D. Jiang, X. Liu, W. Knoll, Q. Zhang, and Y. Wang, "Rapid, label-free histopathological diagnosis of liver cancer based on Raman spectroscopy and deep learning," *Nature Commun.*, vol. 14, no. 1, p. 48, Jan. 2023, doi: 10.1038/s41467-022-35696-2.
[31] R. Vanna, A. De la Cadena, B. Talone, C. Manzoni, M. Marangoni, D. Polli, and G. Cerullo, "Vibrational imaging for label-free cancer diagnosis and classification," *La Rivista Nuovo Cimento*, vol. 45, no. 2, pp. 107–187, Nov. 2021.
[32] K. Kong, C. J. Rowlands, S. Varma, W. Perkins, I. H. Leach, A. A. Koloydenko, H. C. Williams, and I. Notingher, "Diagnosis of tumors during tissue-conserving surgery with integrated autofluorescence and Raman scattering microscopy," *Proc. Nat. Acad. Sci. USA*, vol. 110, no. 38, pp. 15189–15194, Sep. 2013.
[33] D. W. Shipp, E. A. Rakha, A. A. Koloydenko, R. D. Macmillan, I. O. Ellis, and I. Notingher, "Intra-operative spectroscopic assessment of surgical margins during breast conserving surgery," *Breast Cancer Res.*, vol. 20, no. 1, p. 69, Dec. 2018.
[34] Z. Liao, M. G. Lizio, C. Corden, H. Khout, E. Rakha, and I. Notingher, "Feasibility of integrated high-wavenumber Raman imaging and fingerprint Raman spectroscopy for fast margin assessment in breast cancer surgery," *J. Raman Spectrosc.*, vol. 51, no. 10, pp. 1986–1995, Oct. 2020.
[35] S. P. Singh, L. Wang, S. Gupta, H. Goli, P. Padmanabhan, and B. Gulyás, "3D deep learning on medical images: A review," *Sensors*, vol. 20, no. 18, p. 5097, Sep. 2020. [Online]. Available: https://www.mdpi.com/1424-8220/20/18/5097
[36] R. Luo, J. Popp, and T. Bocklitz, "Deep learning for Raman spectroscopy: A review," *Analytica*, vol. 3, no. 3, pp. 287–301, Jul. 2022. [Online]. Available: https://www.mdpi.com/2673-4532/3/3/20

[37] N. Blake, R. Gaifulina, L. D. Griffin, I. M. Bell, and G. M. H. Thomas, "Machine learning of Raman spectroscopy data for classifying cancers: A review of the recent literature," *Diagnostics*, vol. 12, no. 6, p. 1491, Jun. 2022. [Online]. Available: https://www.mdpi.com/2075-4418/12/6/1491

[38] Z. Li, Z. Li, Q. Chen, A. Ramos, J. Zhang, J. P. Boudreaux, R. Thiagarajan, Y. Bren-Mattison, M. E. Dunham, A. J. McWhorter, X. Li, J.-M. Feng, Y. Li, S. Yao, and J. Xu, "Detection of pancreatic cancer by convolutional-neural-network-assisted spontaneous Raman spectroscopy with critical feature visualization," *Neural Netw.*, vol. 144, pp. 455–464, Dec. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0893608021003567

[39] F. Lussier, V. Thibault, B. Charron, G. Q. Wallace, and J.-F. Masson, "Deep learning and artificial intelligence methods for Raman and surface-enhanced Raman scattering," *TrAC Trends Anal. Chem.*, vol. 124, Mar. 2020, Art. no. 115796. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0165993619305783

[40] L. Pan, P. Zhang, C. Daengngam, S. Peng, and M. Chongcheawchamnan, "A review of artificial intelligence methods combined with Raman spectroscopy to identify the composition of substances," *J. Raman Spectrosc.*, vol. 53, no. 1, pp. 6–19, Jan. 2022, doi: 10.1002/jrs.6225.

[41] C. Papers. *Connected Papers—A Visual Tool to Help Researchers and Applied Scientists Find and Explore Papers*. Accessed: Mar. 27, 2023. [Online]. Available: https://www.connectedpapers.com/

[42] J. Carmicheal, C. Hayashi, X. Huang, L. Liu, Y. Lu, A. Krasnoslobodtsev, A. Lushnikov, P. G. Kshirsagar, A. Patel, M. Jain, Y. L. Lyubchenko, Y. Lu, S. K. Batra, and S. Kaur, "Label-free characterization of exosome via surface enhanced Raman spectroscopy for the early detection of pancreatic cancer," *Nanomed., Nanotechnol., Biol. Med.*, vol. 16, pp. 88–96, Feb. 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1549963418305598

[43] M. Aslam, F. Rajbdad, S. Azmat, Z. Li, J. P. Boudreaux, R. Thiagarajan, S. Yao, and J. Xu, "A novel method for detection of pancreatic ductal adenocarcinoma using explainable machine learning," *Comput. Methods Programs Biomed.*, vol. 245, Mar. 2024, Art. no. 108019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169260724000154

[44] A. Uthamacumaran, S. Elouatik, M. Abdouh, M. Berteau-Rainville, Z.-H. Gao, and G. Arena, "Machine learning characterization of cancer patients-derived extracellular vesicles using vibrational spectroscopies: Results from a pilot study," *Int. J. Speech Technol.*, vol. 52, no. 11, pp. 12737–12753, Sep. 2022, doi: 10.1007/s10489-022-03203-1.

[45] L. Zhang, C. Li, D. Peng, X. Yi, S. He, F. Liu, X. Zheng, W. E. Huang, L. Zhao, and X. Huang, "Raman spectroscopy and machine learning for the classification of breast cancers," *Spectrochimica Acta A, Mol. Biomolecular Spectrosc.*, vol. 264, Jan. 2022, Art. no. 120300. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1386142521008775

[46] H. Li, S. Wang, Q. Zeng, C. Chen, X. Lv, M. Ma, H. Su, B. Ma, C. Chen, and J. Fang, "Serum Raman spectroscopy combined with multiple classification models for rapid diagnosis of breast cancer," *Photodiagnosis Photodynamic Therapy*, vol. 40, Dec. 2022, Art. no. 103115. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S157210002200401X

[47] Q. Zeng, C. Chen, C. Chen, H. Song, M. Li, J. Yan, and X. Lv, "Serum Raman spectroscopy combined with convolutional neural network for rapid diagnosis of HER2-positive and triple-negative breast cancer," *Spectrochimica Acta A, Mol. Biomolecular Spectrosc.*, vol. 286, Feb. 2023, Art. no. 122000. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1386142522011489

[48] A. M. Fuentes, A. Narayan, K. Milligan, J. J. Lum, A. G. Brolo, J. L. Andrews, and A. Jirasek, "Raman spectroscopy and convolutional neural networks for monitoring biochemical radiation response in breast tumour xenografts," *Sci. Rep.*, vol. 13, no. 1, p. 1530, Jan. 2023, doi: 10.1038/s41598-023-28479-2.

[49] A. M. Fuentes, K. Milligan, M. Wiebe, A. Narayan, J. J. Lum, A. G. Brolo, J. L. Andrews, and A. Jirasek, "Stratification of tumour cell radiation response and metabolic signatures visualization with Raman spectroscopy and explainable convolutional neural network," *Analyst*, vol. 149, no. 5, pp. 1645–1657, 2024, doi: 10.1039/d3an01797d.

[50] X. Ma, H. Xiong, J. Guo, Z. Liu, Y. Han, M. Liu, Y. Guo, M. Wang, H. Zhong, and Z. Guo, "Label-free breast cancer detection and classification by convolutional neural network-based on exosomes surface-enhanced Raman scattering," *J. Innov. Opt. Health Sci.*, vol. 16, no. 2, Mar. 2023, Art. no. 2244001, doi: 10.1142/s1793545822440011.

[51] D. Ma, L. Shang, J. Tang, Y. Bao, J. Fu, and J. Yin, "Classifying breast cancer tissue by Raman spectroscopy with one-dimensional convolutional neural network," *Spectrochimica Acta A, Mol. Biomolecular Spectrosc.*, vol. 256, Jul. 2021, Art. no. 119732. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1386142521003085

[52] Q. Li, Z. Zhang, and Z. Ma, "Raman spectral pattern recognition of breast cancer: A machine learning strategy based on feature fusion and adaptive hyperparameter optimization," *Heliyon*, vol. 9, no. 7, Jul. 2023, Art. no. e18148. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405844023053562

[53] I. A. Bratchenko, L. A. Bratchenko, Y. A. Khristoforova, A. A. Moryatov, S. V. Kozlov, and V. P. Zakharov, "Classification of skin cancer using convolutional neural networks analysis of Raman spectra," *Comput. Methods Programs Biomed.*, vol. 219, Jun. 2022, Art. no. 106755. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0169260722001419

[54] D. C. Araújo, A. A. Veloso, R. S. de Oliveira Filho, M.-N. Giraud, L. J. Raniero, L. M. Ferreira, and R. A. Bitar, "Finding reduced Raman spectroscopy fingerprint of skin samples for melanoma diagnosis through machine learning," *Artif. Intell. Med.*, vol. 120, Oct. 2021, Art. no. 102161. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0933365721001548

[55] X. Qiu, X. Wu, X. Fang, Q. Fu, P. Wang, X. Wang, S. Li, and Y. Li, "Raman spectroscopy combined with deep learning for rapid detection of melanoma at the single cell level," *Spectrochimica Acta A, Mol. Biomolecular Spectrosc.*, vol. 286, Feb. 2023, Art. no. 122029. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1386142522011775

[56] M. Wu, S. Wang, S. Pan, A. C. Terentis, J. Strasswimmer, and X. Zhu, "Deep learning data augmentation for Raman spectroscopy cancer tissue classification," *Sci. Rep.*, vol. 11, no. 1, p. 23842, Dec. 2021, doi: 10.1038/s41598-021-02687-0.

[57] L. Zhang, Y. Zhou, B. Wu, S. Zhang, K. Zhu, C.-H. Liu, X. Yu, and R. R. Alfano, "A handheld visible resonance Raman analyzer used in intraoperative detection of human glioma," *Cancers*, vol. 15, no. 6, p. 1752, Mar. 2023.

[58] D. Vrazhnov, A. Mankova, E. Stupak, Y. Kistenev, A. Shkurinov, and O. Cherkasova, "Discovering glioma tissue through its biomarkers' detection in blood by Raman spectroscopy and machine learning," *Pharmaceutics*, vol. 15, no. 1, p. 203, Jan. 2023.

[59] A. Quesnel, N. Coles, C. Angione, P. Dey, T. M. Polvikoski, T. F. Outeiro, M. Islam, A. A. Khundakar, and P. S. Filippou, "Glycosylation spectral signatures for glioma grade discrimination using Raman spectroscopy," *BMC Cancer*, vol. 23, no. 1, pp. 1–15, Feb. 2023.

[60] T. Hollon et al., "Artificial-intelligence-based molecular classification of diffuse gliomas using rapid, label-free optical imaging," *Nature Med.*, vol. 29, no. 4, pp. 828–832, Apr. 2023.

[61] M. Ma, X. Tian, F. Chen, X. Ma, W. Guo, and X. Lv, "The application of feature engineering in establishing a rapid and robust model for identifying patients with glioma," *Lasers Med. Sci.*, vol. 37, no. 2, pp. 1007–1015, Mar. 2022.

[62] Z. Jin, Q. Yue, W. Duan, A. Sui, B. Zhao, Y. Deng, Y. Zhai, Y. Zhang, T. Sun, G. Zhang, L. Han, Y. Mao, J. Yu, X. Zhang, and C. Li, "Intelligent SERS navigation system guiding brain tumor surgery by intraoperatively delineating the metabolic acidosis," *Adv. Sci.*, vol. 9, no. 7, Mar. 2022, Art. no. 2104935.

[63] N. Iturrioz-Rodríguez, D. De Pasquale, P. Fiaschi, and G. Ciofani, "Discrimination of glioma patient-derived cells from healthy astrocytes by exploiting Raman spectroscopy," *Spectrochimica Acta A, Mol. Biomolecular Spectrosc.*, vol. 269, Mar. 2022, Art. no. 120773.

[64] C. Chen, Y. Ma, M. Zhu, Z. Yan, X. Lv, C. Chen, and F. Tian, "A new method for Raman spectral analysis: Decision fusion-based transfer learning model," *J. Raman Spectrosc.*, vol. 54, no. 3, pp. 314–323, Mar. 2023.

[65] X. Tian, C. Chen, C. Chen, Z. Yan, W. Wu, F. Chen, J. Chen, and X. Lv, "Application of Raman spectroscopy technology based on deep learning algorithm in the rapid diagnosis of glioma," *J. Raman Spectrosc.*, vol. 53, no. 4, pp. 735–745, Apr. 2022.

[66] M. Bukva, G. Dobra, J. Gomez-Perez, K. Koos, M. Harmati, E. Gyukity-Sebestyen, T. Biro, A. Jenei, S. Kormondi, P. Horvath, Z. Konya, A. Klekner, and K. Buzas, "Raman spectral signatures of serum-derived extracellular vesicle-enriched isolates may support the diagnosis of CNS tumors," *Cancers*, vol. 13, no. 6, p. 1407, Mar. 2021.

[67] L. J. Livermore, M. Isabelle, I. M. Bell, O. Edgar, N. L. Voets, R. Stacey, O. Ansorge, C. Vallance, and P. Plaha, "Raman spectroscopy to differentiate between fresh tissue samples of glioma and normal brain: A comparison with 5-ALA–induced fluorescence-guided surgery," *J. Neurosurgery*, vol. 135, no. 2, pp. 469–479, 2020.

[68] J. Ao, X. Shao, Z. Liu, Q. Liu, J. Xia, Y. Shi, L. Qi, J. Pan, and M. Ji, "Stimulated Raman scattering microscopy enables Gleason scoring of prostate core needle biopsy by a convolutional neural network," *Cancer Res.*, vol. 83, no. 4, pp. 641–651, Feb. 2023.

[69] T. Chen, A. Yavuz, and M. C. Wang, "Dissecting lipid droplet biology with coherent Raman scattering microscopy," *J. Cell Sci.*, vol. 135, no. 5, Mar. 2022, Art. no. jcs252353.

[70] D. Grajales, F. Picot, R. Shams, F. Dallaire, G. Sheehy, S. Alley, M. Barkati, G. Delouya, J.-F. Carrier, M. Birlea, D. Trudel, F. Leblond, C. Ménard, and S. Kadoury, "Image-guided Raman spectroscopy navigation system to improve transperineal prostate cancer detection. Part 2: In-vivo tumor-targeting using a classification model combining spectral and MRI-radiomics features," *J. Biomed. Opt.*, vol. 27, no. 9, Sep. 2022, Art. no. 095004.

[71] K. Milligan, X. Deng, R. Ali-Adeeb, P. Shreeves, S. Punch, N. Costie, J. M. Crook, A. G. Brolo, J. J. Lum, J. L. Andrews, and A. Jirasek, "Prediction of disease progression indicators in prostate cancer patients receiving HDR-brachytherapy using Raman spectroscopy and semi-supervised learning: A pilot study," *Sci. Rep.*, vol. 12, no. 1, p. 15104, Sep. 2022.

[72] K. Milligan, S. J. Van Nest, X. Deng, R. Ali-Adeeb, P. Shreeves, S. Punch, N. Costie, N. Pavey, J. M. Crook, D. M. Berman, A. G. Brolo, J. J. Lum, J. L. Andrews, and A. Jirasek, "Raman spectroscopy and supervised learning as a potential tool to identify high-dose-rate-brachytherapy induced biochemical profiles of prostate cancer," *J. Biophotonics*, vol. 15, no. 11, Nov. 2022, Art. no. e202200121.

[73] F. Picot, R. Shams, F. Dallaire, G. Sheehy, T. Trang, D. Grajales, M. Birlea, D. Trudel, C. Ménard, S. Kadoury, and F. Leblond, "Image-guided Raman spectroscopy navigation system to improve transperineal prostate cancer detection. Part 1: Raman spectroscopy fiber-optics system and in situ tissue characterization," *J. Biomed. Opt.*, vol. 27, no. 9, Sep. 2022, Art. no. 095003.

[74] X. Zhao, Q. Xu, Y. Lin, W. Du, X. Bai, J. Gao, T. Li, Y. Huang, Y. Yu, X. Wu, and J. Lin, "Label-free surface-enhanced Raman spectroscopy detection of prostate cancer combined with multivariate statistical algorithm," *J. Raman Spectrosc.*, vol. 53, no. 11, pp. 1861–1870, Nov. 2022.

[75] A.-A. Grosset et al., "Identification of intraductal carcinoma of the prostate on tissue specimens using Raman micro-spectroscopy: A diagnostic accuracy case–control study with multicohort validation," *PLOS Med.*, vol. 17, no. 8, Aug. 2020, Art. no. e1003281.

[76] Y. Ma, J. Chi, Z. Zheng, A. Attygalle, I. Y. Kim, and H. Du, "Therapeutic prognosis of prostate cancer using surface-enhanced Raman scattering of patient urine and multivariate statistical analysis," *J. Biophotonics*, vol. 14, no. 1, Jan. 2021, Art. no. e202000275.

[77] D. K. R. Medipally, D. Cullen, V. Untereiner, G. D. Sockalingum, A. Maguire, T. N. Q. Nguyen, J. Bryant, E. Noone, S. Bradshaw, M. Finn, M. Dunne, A. M. Shannon, J. Armstrong, A. D. Meade, and F. M. Lyng, "Vibrational spectroscopy of liquid biopsies for prostate cancer diagnosis," *Therapeutic Adv. Med. Oncol.*, vol. 12, Jan. 2020, Art. no. 175883592091849.

[78] W. Lee, A. T. M. Lenferink, C. Otto, and H. L. Offerhaus, "Classifying Raman spectra of extracellular vesicles based on convolutional neural networks for prostate cancer detection," *J. Raman Spectrosc.*, vol. 51, no. 2, pp. 293–300, Feb. 2020.

[79] F. Chen, C. Sun, Z. Yue, Y. Zhang, W. Xu, S. Shabbir, L. Zou, W. Lu, W. Wang, Z. Xie, L. Zhou, Y. Lu, and J. Yu, "Screening ovarian cancers with Raman spectroscopy of blood plasma coupled with machine learning data processing," *Spectrochimica Acta A, Mol. Biomolecular Spectrosc.*, vol. 265, Jan. 2022, Art. no. 120355.

[80] S. David, A. Plante, F. Dallaire, J. Tremblay, G. Sheehy, E. Macdonald, L. Forrest, M. Daneshmand, D. Trudel, B. C. Wilson, L. Hopkins, S. Murugkar, B. Vanderhyden, and F. Leblond, "Multispectral label-free Raman spectroscopy can detect ovarian and endometrial cancer with high accuracy," *J. Biophotonics*, vol. 15, no. 2, Feb. 2022, Art. no. e202100198.

[81] P. Giamougiannis, C. L. M. Morais, R. Grabowska, K. M. Ashton, N. J. Wood, P. L. Martin-Hirsch, and F. L. Martin, "A comparative analysis of different biofluids towards ovarian cancer diagnosis using Raman microspectroscopy," *Anal. Bioanal. Chem.*, vol. 413, no. 3, pp. 911–922, Jan. 2021.

[82] R. A. Hunter, M. Asare-Werehene, A. Mandour, B. K. Tsang, and H. Anis, "Determination of chemoresistance in ovarian cancer by simultaneous quantification of exosomes and exosomal cisplatin with surface enhanced Raman scattering," *Sens. Actuators B, Chem.*, vol. 354, Mar. 2022, Art. no. 131237.

[83] V. Moisoiu et al., "SERS-based differential diagnosis between multiple solid malignancies: Breast, colorectal, lung, ovarian and oral cancer," *Int. J. Nanomed.*, vol. 14, pp. 6165–6178, Aug. 2019.

[84] J. Perumal, A. Mahyuddin, G. Balasundaram, D. Goh, C. Y. Fu, A. Kazakeviciute, U. Dinish, M. Choolani, and M. Olivo, "SERS-based detection of haptoglobin in ovarian cyst fluid as a point-of-care diagnostic assay for epithelial ovarian cancer," *Cancer Manage. Res.*, vol. Volume 11, pp. 1115–1124, Jan. 2019.

[85] M. Paraskevaidi, K. M. Ashton, H. F. Stringfellow, N. J. Wood, P. J. Keating, A. W. Rowbottom, P. L. Martin-Hirsch, and F. L. Martin, "Raman spectroscopic techniques to detect ovarian cancer biomarkers in blood plasma," *Talanta*, vol. 189, pp. 281–288, Nov. 2018.

[86] X. Chang, M. Yu, R. Liu, R. Jing, J. Ding, J. Xia, Z. Zhu, X. Li, Q. Yao, L. Zhu, and T. Zhang, "Deep learning methods for oral cancer detection using Raman spectroscopy," *Vibrational Spectrosc.*, vol. 126, May 2023, Art. no. 103522.

[87] A. Ghosh, D. Chaudhuri, S. Adhikary, K. Chatterjee, A. Roychowdhury, A. K. Das, and A. Barui, "Deep reinforced neural network model for cyto-spectroscopic analysis of epigenetic markers for automated oral cancer risk prediction," *Chemometric Intell. Lab. Syst.*, vol. 224, May 2022, Art. no. 104548.

[88] H. J. Koster, A. Guillen-Perez, J. S. Gomez-Diaz, M. Navas-Moreno, A. C. Birkeland, and R. P. Carney, "Fused Raman spectroscopic analysis of blood and saliva delivers high accuracy for head and neck cancer diagnostics," *Sci. Rep.*, vol. 12, no. 1, p. 18464, Nov. 2022.

[89] K. Wang, Y. Qiu, C. Wu, Z. Wen, and Y. Li, "Surface-enhanced Raman spectroscopy and multivariate analysis for the diagnosis of oral squamous cell carcinoma," *J. Raman Spectrosc.*, vol. 54, no. 4, pp. 355–362, Apr. 2023.

[90] J. Xia, L. Zhu, M. Yu, T. Zhang, Z. Zhu, X. Lou, G. Sun, and M. Dong, "Analysis and classification of oral tongue squamous cell carcinoma based on Raman spectroscopy and convolutional neural networks," *J. Modern Opt.*, vol. 67, no. 6, pp. 481–489, Mar. 2020.

[91] H. Yan, M. Yu, J. Xia, L. Zhu, T. Zhang, Z. Zhu, and G. Sun, "Diverse region-based CNN for tongue squamous cell carcinoma classification with Raman spectroscopy," *IEEE Access*, vol. 8, pp. 127313–127328, 2020.

[92] M.-J. Jeng, M. Sharma, L. Sharma, T.-Y. Chao, S.-F. Huang, L.-B. Chang, S.-L. Wu, and L. Chow, "Raman spectroscopy analysis for optical diagnosis of oral cancer detection," *J. Clin. Med.*, vol. 8, no. 9, p. 1313, Aug. 2019.

[93] H. Yan, M. Yu, J. Xia, L. Zhu, T. Zhang, and Z. Zhu, "Tongue squamous cell carcinoma discrimination with Raman spectroscopy and convolutional neural networks," *Vibrational Spectrosc.*, vol. 103, Jul. 2019, Art. no. 102938.

[94] L. Zhang, Y. Wu, B. Zheng, L. Su, Y. Chen, S. Ma, Q. Hu, X. Zou, L. Yao, Y. Yang, L. Chen, Y. Mao, Y. Chen, and M. Ji, "Rapid histology of laryngeal squamous cell carcinoma with deep-learning based stimulated Raman scattering microscopy," *Theranostics*, vol. 9, no. 9, pp. 2541–2554, 2019.

[95] M. Yu, H. Yan, J. Xia, L. Zhu, T. Zhang, Z. Zhu, X. Lou, G. Sun, and M. Dong, "Deep convolutional neural networks for tongue squamous cell carcinoma classification using Raman spectroscopy," *Photodiagnosis Photodynamic Therapy*, vol. 26, pp. 430–435, Jun. 2019.

[96] M. Sharma, Y.-C. Li, S. N. Manjunatha, C.-L. Tsai, R.-M. Lin, S.-F. Huang, and L.-B. Chang, "Identification of healthy tissue from malignant tissue in surgical margin using Raman spectroscopy in oral cancer surgeries," *Biomedicines*, vol. 11, no. 7, p. 1984, Jul. 2023.

[97] D. Chaudhuri, A. Ghosh, S. Raha, A. Chakraborty, K. Chatterjee, and A. Barui, "Machine learning algorithm ensembles for early oral cancer risk assessment using Raman cyto-spectroscopy," *Soft Comput.*, vol. 27, no. 19, pp. 13861–13875, Oct. 2023.

[98] G. W. Auner, S. K. Koya, C. Huang, B. Broadbent, M. Trexler, Z. Auner, A. Elias, K. C. Mehne, and M. A. Brusatori, "Applications of Raman spectroscopy in cancer diagnosis," *Cancer Metastasis Rev.*, vol. 37, pp. 691–717, Jan. 2018.

[99] D. Cialla-May, M. Schmitt, and J. Popp, "Theoretical principles of Raman spectroscopy," *Phys. Sci. Rev.*, vol. 4, no. 6, May 2019, Art. no. 20170040.

[100] R. R. Jones, D. C. Hooper, L. Zhang, D. Wolverson, and V. K. Valev, "Raman techniques: Fundamentals and frontiers," *Nanosc. Res. Lett.*, vol. 14, no. 1, pp. 1–34, Dec. 2019.

[101] K. S. Lee, Z. Landry, F. C. Pereira, M. Wagner, D. Berry, W. E. Huang, G. T. Taylor, J. Kneipp, J. Popp, and M. Zhang, "Raman microspectroscopy for microbiology," *Nature Rev. Methods Primers*, vol. 1, no. 1, p. 80, 2021.

[102] A. I. Pérez-Jiménez, D. Lyu, Z. Lu, G. Liu, and B. Ren, "Surface-enhanced Raman spectroscopy: Benefits, trade-offs and future developments," *Chem. Sci.*, vol. 11, no. 18, pp. 4563–4577, 2020.

[103] C. Krafft, J. Popp, P. Bronsert, and A. Miernik, "Raman spectroscopic imaging of human bladder resectates towards intraoperative cancer assessment," *Cancers*, vol. 15, no. 7, p. 2162, Apr. 2023.

[104] H. Høgset, C. C. Horgan, J. P. K. Armstrong, M. S. Bergholt, V. Torraca, Q. Chen, T. J. Keane, L. Bugeon, M. J. Dallman, S. Mostowy, and M. M. Stevens, "In vivo biomolecular imaging of zebrafish embryos using confocal Raman spectroscopy," *Nature Commun.*, vol. 11, no. 1, p. 6172, Dec. 2020.

[105] C. Corden, R. Boitor, P. K. Dusanjh, A. Harwood, A. Mukherjee, D. Gomez, and I. Notingher, "Autofluorescence-Raman spectroscopy for ex vivo mapping colorectal liver metastases and liver tissue," *J. Surgical Res.*, vol. 288, pp. 10–20, Aug. 2023.

[106] B. Matsumoto, *Cell Biological Applications of Confocal Microscopy*. Amsterdam, The Netherlands: Elsevier, 2003.

[107] A. Szaniawska and A. Kudelski, "Applications of surface-enhanced Raman scattering in biochemical and medical analysis," *Frontiers Chem.*, vol. 9, May 2021, Art. no. 664134.

[108] R. Pilot, R. Signorini, C. Durante, L. Orian, M. Bhamidipati, and L. Fabris, "A review on surface-enhanced Raman scattering," *Biosensors*, vol. 9, no. 2, p. 57, 2019.

[109] S. Kruszewski, "Enhancement mechanisms in the sers phenomenon," *Proc. SPIE*, vol. 3320, pp. 281–293, Jul. 1998.

[110] K. Yuan, B. Jurado-Sánchez, and A. Escarpa, "Nanomaterials meet surface-enhanced Raman scattering towards enhanced clinical diagnosis: A review," *J. Nanobiotechnol.*, vol. 20, no. 1, pp. 1–28, Dec. 2022.

[111] C. Lin, Y. Li, Y. Peng, S. Zhao, M. Xu, L. Zhang, Z. Huang, J. Shi, and Y. Yang, "Recent development of surface-enhanced Raman scattering for biosensing," *J. Nanobiotechnol.*, vol. 21, no. 1, p. 149, May 2023.

[112] A. Alfonso-García, R. Mittal, E. S. Lee, and E. O. Potma, "Biological imaging with coherent Raman scattering microscopy: A tutorial," *J. Biomed. Opt.*, vol. 19, no. 7, Mar. 2014, Art. no. 071407.

[113] S. Weng, X. Xu, J. Li, and S. T. C. Wong, "Combining deep learning and coherent anti-Stokes Raman scattering imaging for automated differential diagnosis of lung cancer," *J. Biomed. Opt.*, vol. 22, no. 10, p. 1, Oct. 2017.

[114] R. Galli, O. Uckermann, A. Temme, E. Leipnitz, M. Meinhardt, E. Koch, G. Schackert, G. Steiner, and M. Kirsch, "Assessing the efficacy of coherent anti-Stokes Raman scattering microscopy for the detection of infiltrating glioblastoma in fresh brain samples," *J. Biophotonics*, vol. 10, no. 3, pp. 404–414, Mar. 2017.

[115] E. A. McCullagh, S. Poleg, D. Stich, R. Moldovan, and A. Klug, "Coherent anti-Stokes Raman spectroscopy (CARS) application for imaging myelination in brain slices," *J. Visualized Exp.*, no. 185, Jul. 2022, Art. no. e64013.

[116] G. I. Petrov, R. Arora, and V. V. Yakovlev, "Coherent anti-Stokes Raman scattering imaging of microcalcifications associated with breast cancer," *Analyst*, vol. 146, no. 4, pp. 1253–1259, 2021.

[117] F. Lyng, E. Gazi, and P. Gardner, "Preparation of tissues and cells for infrared and Raman spectroscopy and imaging," in *Biomedical Applications of Synchrotron Infrared Microspectroscopy, RSC Analytical Spectroscopy Monographs* (RSC Analytical Spectroscopy Series), vol. 11. Royal Society of Chemistry, 2010, pp. 147–185.

[118] H. J. Butler, L. Ashton, B. Bird, G. Cinque, K. Curtis, J. Dorney, K. Esmonde-White, N. J. Fullwood, B. Gardner, and P. L. Martin-Hirsch, "Using Raman spectroscopy to characterize biological materials," *Nature Protocols*, vol. 11, no. 4, pp. 664–687, 2016.

[119] S. Gao, Y. Lin, X. Zhao, J. Gao, S. Xie, W. Gong, Y. Yu, and J. Lin, "Label-free surface enhanced Raman spectroscopy analysis of blood serum via coffee ring effect for accurate diagnosis of cancers," *Spectrochimica Acta A, Mol. Biomolecular Spectrosc.*, vol. 267, Feb. 2022, Art. no. 120605.

[120] A. L. Fradkov, "Early history of machine learning," *IFAC-PapersOnLine*, vol. 53, no. 2, pp. 1385–1390, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2405896320325027

[121] E. Çallı, E. Sogancioglu, B. van Ginneken, K. G. van Leeuwen, and K. Murphy, "Deep learning for chest X-ray analysis: A survey," *Med. Image Anal.*, vol. 72, Aug. 2021, Art. no. 102125. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1361841521001717

[122] S. Gassenmaier, T. Küstner, D. Nickel, J. Herrmann, R. Hoffmann, H. Almansour, S. Afat, K. Nikolaou, and A. E. Othman, "Deep learning applications in magnetic resonance imaging: Has the future become present?" *Diagnostics*, vol. 11, no. 12, p. 2181, Nov. 2021.

[123] K. Gupta and V. Bajaj, "Deep learning models-based CT-scan image classification for automated screening of COVID-19," *Biomed. Signal Process. Control*, vol. 80, Feb. 2023, Art. no. 104268.

[124] F. León-Bejarano, M. O. Méndez, M. G. Ramírez-Elías, and A. Alba, "Improved vancouver Raman algorithm based on empirical mode decomposition for denoising biological samples," *Appl. Spectrosc.*, vol. 73, no. 12, pp. 1436–1450, Dec. 2019.

[125] J. Zhao, H. Lui, D. I. McLean, and H. Zeng, "Automated autofluorescence background subtraction algorithm for biomedical Raman spectroscopy," *Appl. Spectrosc.*, vol. 61, no. 11, pp. 1225–1232, Nov. 2007.

[126] N. B. Gallagher, "Savitzky–Golay smoothing and differentiation filter," Eigenvector Res. Incorporated, Tech. Rep., 2020. [Online]. Available: https://eigenvector.com

[127] I. J. Goodfellow, J. Pouget-Abadie, M. Mirza, B. Xu, D. Warde-Farley, S. Ozair, A. Courville, and Y. Bengio, "Generative adversarial networks," *Commun. ACM*, vol. 63, no. 11, pp. 139–144, 2014.

[128] A. Maćkiewicz and W. Ratajczak, "Principal components analysis (PCA)," *Comput. Geosci.*, vol. 19, no. 3, pp. 303–342, 1993. [Online]. Available: https://www.sciencedirect.com/science/article/pii/009830049390090R

[129] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Trans. Inf. Theory*, vol. IT-13, no. 1, pp. 21–27, Jan. 1967.

[130] C. Cortes and V. Vapnik, "Support-vector networks," *Mach. Learn.*, vol. 20, no. 3, pp. 273–297, Sep. 1995.

[131] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*. New York, NY, USA: ACM, Aug. 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[132] W. Mcculloch and W. Pitts, "A logical calculus of the ideas immanent in nervous activity," *Bull. Math. Biol.*, vol. 52, nos. 1–2, pp. 99–115, 1990.

[133] J. Schmidhuber, "Deep learning in neural networks: An overview," 2014, *arXiv:1404.7828*.

[134] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, Nov. 1997.

[135] S. Narkhede, "Understanding AUC-ROC curve," *Towards Data Sci.*, vol. 26, no. 1, pp. 220–227, 2018.

[136] A. P. Bradley, "The use of the area under the ROC curve in the evaluation of machine learning algorithms," *Pattern Recognit.*, vol. 30, no. 7, pp. 1145–1159, Jul. 1997. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0031320396001422

[137] S. Kunjan, T. S. Grummett, K. J. Pope, D. M. W. Powers, S. P. Fitzgibbon, T. Bastiampillai, M. Battersby, and T. W. Lewis, "The necessity of leave one subject out (LOSO) cross validation for EEG disease diagnosis," in *Brain Informatics*, M. Mahmud, M. S. Kaiser, S. Vassanelli, Q. Dai, and N. Zhong, Eds. Cham, Switzerland: Springer, 2021, pp. 558–567.

[138] R. L. Siegel, K. D. Miller, N. S. Wagle, and A. Jemal, "Cancer statistics, 2023," *CA, Cancer J. Clinicians*, vol. 73, no. 1, pp. 17–48, Jan. 2023, doi: 10.3322/caac.21763.

[139] L. Zhang, X. Zou, J. Huang, J. Fan, X. Sun, B. Zhang, B. Zheng, C. Guo, D. Fu, L. Yao, and M. Ji, "Label-free histology and evaluation of human pancreatic cancer with coherent nonlinear optical microscopy," *Anal. Chem.*, vol. 93, no. 46, pp. 15550–15558, Nov. 2021, doi: 10.1021/acs.analchem.1c03861.

[140] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke, and A. Rabinovich, "Going deeper with convolutions," 2014, *arXiv:1409.4842*.

[141] H. Liu, J. Li, and L. Wong, "A comparative study on feature selection and classification methods using gene expression profiles and proteomic patterns," *Genome Informat.*, vol. 13, pp. 51–60, Jan. 2002.

[142] K. Kira and L. A. Rendell, "The feature selection problem: Traditional methods and a new algorithm," in *Proc. Nat. Conf. Artif. Intell.*, vol. 2, 1992, pp. 129–134.

[143] J. Kowal, M. Tkach, and C. Théry, "Biogenesis and secretion of exosomes," *Current Opinion Cell Biol.*, vol. 29, pp. 116–125, Aug. 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S095506741400057X

[144] Y. H. Soung, T. Nguyen, H. Cao, J. Lee, and J. Chung, "Emerging roles of exosomes in cancer invasion and metastasis," *BMB Rep.*, vol. 49, no. 1, pp. 18–25, Jan. 2016. [Online]. Available: https://www.scopus.com/inward/record.uri?eid=2-s2.0-84958167805&doi=10.5483%2fBMBRep.2016.49.1.239&partnerID=40&md5=e60f6c451bdab021327a4c2a7175adc0

[145] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Mach. Learn.*, vol. 46, no. 1, pp. 389–422, 2002.

[146] B. Huang, H. Huang, S. Zhang, D. Zhang, Q. Shi, J. Liu, and J. Guo, "Artificial intelligence in pancreatic cancer," *Theranostics*, vol. 12, no. 16, pp. 6931–6954, 2022.

[147] R. V. Selvarani and P. S. H. Jose, "A label-free marker based breast cancer detection using hybrid deep learning models and Raman spectroscopy," *Trends Sci.*, vol. 20, no. 4, p. 6299, Jan. 2023. [Online]. Available: https://tis.wu.ac.th/index.php/tis/article/view/6299

[148] A. G. Rivenbark, S. M. O'Connor, and W. B. Coleman, "Molecular and cellular heterogeneity in breast cancer: Challenges for personalized medicine," *Amer. J. Pathol.*, vol. 183, no. 4, pp. 1113–1124, 2013.

[149] C. Qiu, W. Zhang, Y. Zhou, H. Cui, Y. Xing, F. Yu, and R. Wang, "Highly sensitive surface-enhanced Raman scattering (SERS) imaging for phenotypic diagnosis and therapeutic evaluation of breast cancer," *Chem. Eng. J.*, vol. 459, Mar. 2023, Art. no. 141502. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1385894723002334

[150] A. Savitzky and M. J. E. Golay, "Smoothing and differentiation of data by simplified least squares procedures," *Anal. Chem.*, vol. 36, no. 8, pp. 1627–1639, Jul. 1964.

[151] Z. Cheng, H. Li, C. Chen, X. Lv, E. Zuo, X. Xie, Z. Li, P. Liu, H. Li, and C. Chen, "Application of serum SERS technology based on thermally annealed silver nanoparticle composite substrate in breast cancer," *Photodiagnosis Photodynamic Therapy*, vol. 41, Mar. 2023, Art. no. 103284. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1572100023000133

[152] H. Li, T. Ning, F. Yu, Y. Chen, B. Zhang, and S. Wang, "Raman microspectroscopic investigation and classification of breast cancer pathological characteristics," *Molecules*, vol. 26, no. 4, p. 921, Feb. 2021. [Online]. Available: https://www.mdpi.com/1420-3049/26/4/921

[153] S. Akbar, M. I. Majeed, H. Nawaz, N. Rashid, A. Tariq, W. Hameed, S. Shakeel, G. Dastgir, R. Z. A. Bari, M. Iqbal, A. Nawaz, and M. Akram, "Surface-enhanced Raman spectroscopic (SERS) characterization of low molecular weight fraction of the serum of breast cancer patients with principal component analysis (PCA) and partial least square-discriminant analysis (PLS-DA)," *Anal. Lett.*, vol. 55, no. 10, pp. 1588–1604, Jul. 2022, doi: 10.1080/00032719.2021.2017948.

[154] Early Breast Cancer Trialists' Collaborative Group, "Effect of radiotherapy after breast-conserving surgery on 10-year recurrence and 15-year breast cancer death: Meta-analysis of individual patient data for 10801 women in 17 randomised trials," *Lancet*, vol. 378, no. 9804, pp. 1707–1716, 2011. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0140673611616292

[155] S. J. Van Nest, L. M. Nicholson, L. DeVorkin, A. G. Brolo, J. J. Lum, and A. Jirasek, "Raman spectroscopic signatures reveal distinct biochemical and temporal changes in irradiated human breast adenocarcinoma xenografts," *Radiat. Res.*, vol. 189, no. 5, p. 497, Feb. 2018, doi: 10.1667/rr15003.1.

[156] P. Shreeves, J. L. Andrews, X. Deng, R. Ali-Adeeb, and A. Jirasek, "Nonnegative matrix factorization with group and basis restrictions," 2021, *arXiv:2107.00744*.

[157] S. Carr, C. Smith, and J. Wernberg, "Epidemiology and risk factors of melanoma," *Surgical Clinics North Amer.*, vol. 100, no. 1, pp. 1–12, Feb. 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0039610919301239

[158] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun, "Dermatologist-level classification of skin cancer with deep neural networks," *Nature*, vol. 542, no. 7639, pp. 115–118, Feb. 2017.

[159] A. K. Sharma, S. Tiwari, G. Aggarwal, N. Goenka, A. Kumar, P. Chakrabarti, T. Chakrabarti, R. Gono, Z. Leonowicz, and M. Jasinski, "Dermatologist-level classification of skin cancer using cascaded ensembling of convolutional neural network and handcrafted features based deep neural network," *IEEE Access*, vol. 10, pp. 17920–17932, 2022.

[160] H. Ye, U. Kruger, T. Wang, S. Shi, J. Norfleet, and S. De, "Raman spectroscopy accurately classifies burn severity in an ex vivo model," *Burns*, vol. 47, no. 4, pp. 812–820, Jun. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0305417920305052

[161] Y. Kanemura, M. Kanazawa, S. Hashimoto, Y. Hayashi, E. Fujiwara, A. Suzuki, T. Ishii, M. Goto, H. Nozaki, T. Inoue, and H. Takanari, "Assessment of skin inflammation using near-infrared Raman spectroscopy combined with artificial intelligence analysis in an animal model," *Analyst*, vol. 147, no. 12, pp. 2843–2850, 2022, doi: 10.1039/d2an00193d.

[162] I. A. Bratchenko, L. A. Bratchenko, A. A. Moryatov, Y. A. Khristoforova, D. N. Artemyev, O. O. Myakinin, A. E. Orlov, S. V. Kozlov, and V. P. Zakharov, "In vivo diagnosis of skin cancer with a portable Raman spectroscopic device," *Experim. Dermatol.*, vol. 30, no. 5, pp. 652–663, May 2021, doi: 10.1111/exd.14301.

[163] Y. A. Khristoforova, I. A. Bratchenko, O. O. Myakinin, D. N. Artemyev, A. A. Moryatov, A. E. Orlov, S. V. Kozlov, and V. P. Zakharov, "Portable spectroscopic system for in vivo skin neoplasms diagnostics by Raman and autofluorescence analysis," *J. Biophotonics*, vol. 12, no. 4, Apr. 2019, Art. no. e201800400, doi: 10.1002/jbio.201800400.

[164] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," in *Proc. Int. Conf. Adv. Neural Inf. Process. Syst.*, vol. 30, 2017, pp. 3146–3154.

[165] S. M. Lundberg, G. Erion, H. Chen, A. DeGrave, J. M. Prutkin, B. Nair, R. Katz, J. Himmelfarb, N. Bansal, and S.-I. Lee, "From local explanations to global understanding with explainable AI for trees," *Nature Mach. Intell.*, vol. 2, no. 1, pp. 56–67, Jan. 2020.

[166] X. Qiu, T. He, X. Wu, P. Wang, X. Wang, Q. Fu, X. Fang, S. Li, and Y. Li, "Combining fiber optical tweezers and Raman spectroscopy for rapid identification of melanoma," *J. Biophotonics*, vol. 15, no. 12, pp. 1–10, Dec. 2022.

[167] C.-H. Liu, B. Wu, L. A. Sordillo, S. Boydston-White, V. Sriramoju, C. Zhang, H. Beckman, L. Zhang, Z. Pei, L. Shi, and R. R. Alfano, "A pilot study for distinguishing basal cell carcinoma from normal human skin tissues using visible resonance Raman spectroscopy," *J. Cancer Metastasis Treatment*, vol. 2019, p. 4, Jan. 2019, doi: 10.20517/2394-4722.2018.55.

[168] S. A. Fox, A. A. Shanblatt, H. Beckman, J. Strasswimmer, and A. C. Terentis, "Raman spectroscopy differentiates squamous cell carcinoma (SCC) from normal skin following treatment with a high-powered CO_2 laser," *Lasers Surg. Med.*, vol. 46, no. 10, pp. 757–772, Oct. 2014.

[169] G. P. Guy, S. R. Machlin, D. U. Ekwueme, and K. R. Yabroff, "Prevalence and costs of skin cancer treatment in the U.S., 2002–2006 and 2007–2011," *Amer. J. Preventive Med.*, vol. 48, no. 2, pp. 183–187, Feb. 2015.

[170] J. Ferlay, M. Colombet, I. Soerjomataram, T. Dyba, G. Randi, M. Bettio, A. Gavin, O. Visser, and F. Bray, "Cancer incidence and mortality patterns in europe: Estimates for 40 countries and 25 major cancers in 2018," *Eur. J. Cancer*, vol. 103, pp. 356–387, Nov. 2018.

[171] W. Tang, W. Fan, J. Lau, L. Deng, Z. Shen, and X. Chen, "Emerging blood–brain-barrier-crossing nanotechnology for brain cancer theranostics," *Chem. Soc. Rev.*, vol. 48, no. 11, pp. 2967–3014, 2019.

[172] G. Menna, G. Piaser Guerrato, L. Bilgin, G. M. Ceccarelli, A. Olivi, and G. M. Della Pepa, "Is there a role for machine learning in liquid biopsy for brain tumors? A systematic review," *Int. J. Mol. Sci.*, vol. 24, no. 11, p. 9723, Jun. 2023.

[173] C. M. Rogers, P. S. Jones, and J. S. Weinberg, "Intraoperative MRI for brain tumors," *J. Neuro-Oncol.*, vol. 151, no. 3, pp. 479–490, Feb. 2021.

[174] C. Orillac, W. Stummer, and D. A. Orringer, "Fluorescence guidance and intraoperative adjuvants to maximize extent of resection," *Neurosurgery*, vol. 89, no. 5, pp. 727–736, 2021.

[175] J. L. Roland, C. D. Hacker, and E. C. Leuthardt, "A review of passive brain mapping techniques in neurological surgery," *Neurosurgery*, vol. 88, no. 1, pp. 15–24, 2021.

[176] H. Fabelo, M. Halicek, S. Ortega, M. Shahedi, A. Szolna, J. Piñeiro, C. Sosa, A. O'Shanahan, S. Bisshopp, C. Espino, M. Márquez, M. Hernández, D. Carrera, J. Morera, G. Callico, R. Sarmiento, and B. Fei, "Deep learning-based framework for in vivo identification of glioblastoma tumor using hyperspectral images of human brain," *Sensors*, vol. 19, no. 4, p. 920, Feb. 2019.

[177] S. Chen, H. Zhang, X. Yang, X. Shao, T. Li, N. Chen, Z. Chen, W. Xue, J. Pan, and S. Liu, "Raman spectroscopy reveals abnormal changes in the urine composition of prostate cancer: An application of an intelligent diagnostic model with a deep learning algorithm," *Adv. Intell. Syst.*, vol. 3, no. 4, Apr. 2021, Art. no. 2000090.

[178] M. Haroon, M. Tahir, H. Nawaz, M. I. Majeed, and A. A. Al-Saadi, "Surface-enhanced Raman scattering (SERS) spectroscopy for prostate cancer diagnosis: A review," *Photodiagnosis Photodynamic Therapy*, vol. 37, Mar. 2022, Art. no. 102690.

[179] I. Thompson, "Prevalence of prostate cancer among men with a prostate-specific antigen level <= 4.0 ng per milliliter," *New England J. Med.*, vol. 351, no. 14, p. 1470, May 2239.

[180] M. Sekhoacha, K. Riet, P. Motloung, L. Gumenku, A. Adegoke, and S. Mashele, "Prostate cancer review: Genetics, diagnosis, treatment options, and alternative approaches," *Molecules*, vol. 27, no. 17, p. 5730, Sep. 2022.

[181] J.-L. Descotes, "Diagnosis of prostate cancer," *Asian J. Urol.*, vol. 6, no. 2, pp. 129–136, 2019.

[182] J. I. Epstein, M. J. Zelefsky, D. D. Sjoberg, J. B. Nelson, L. Egevad, C. Magi-Galluzzi, A. J. Vickers, A. V. Parwani, V. E. Reuter, S. W. Fine, J. A. Eastham, P. Wiklund, M. Han, C. A. Reddy, J. P. Ciezki, T. Nyberg, and E. A. Klein, "A contemporary prostate cancer grading system: A validated alternative to the Gleason score," *Eur. Urol.*, vol. 69, no. 3, pp. 428–435, Mar. 2016.

[183] S. Lheureux, M. Braunstein, and A. M. Oza, "Epithelial ovarian cancer: Evolution of management in the era of precision medicine," *CA, Cancer J. Clinicians*, vol. 69, no. 4, pp. 280–304, Jul. 2019.

[184] U. Menon et al., "Ovarian cancer population screening and mortality after long-term follow-up in the U.K. collaborative trial of ovarian cancer screening (UKCTOCS): A randomised controlled trial," *Lancet*, vol. 397, no. 10290, pp. 2182–2193, Jun. 2021.

[185] K. El Bairi, S. Afqir, and M. Amrani, "Is HE4 superior over CA-125 in the follow-up of patients with epithelial ovarian cancer?" *Current Drug Targets*, vol. 21, no. 10, pp. 1026–1033, Jul. 2020.

[186] T. Muinao, H. P. Deka Boruah, and M. Pal, "Multi-biomarker panel signature as the key to diagnosis of ovarian cancer," *Heliyon*, vol. 5, no. 12, Dec. 2019, Art. no. e02826.

[187] R. E. Bristow, D. R. Gossett, D. R. Shook, M. L. Zahurak, R. S. Tomacruz, D. K. Armstrong, and F. J. Montz, "Recurrent micropapillary serous ovarian carcinoma: The role of secondary cytoreductive surgery," *Cancer*, vol. 95, no. 4, pp. 791–800, Aug. 2002.

[188] S. Abati, C. Bramati, S. Bondi, A. Lissoni, and M. Trimarchi, "Oral cancer and precancer: A narrative review on the relevance of early diagnosis," *Int. J. Environ. Res. Public Health*, vol. 17, no. 24, p. 9160, Dec. 2020.

[189] J. A. Woolgar and A. Triantafyllou, "A histopathological appraisal of surgical margins in oral and oropharyngeal cancer resection specimens," *Oral Oncol.*, vol. 41, no. 10, pp. 1034–1043, Nov. 2005.

[190] G. La Rosa, G. Gattuso, E. Pedullà, E. Rapisarda, D. Nicolosi, and M. Salmeri, "Association of oral dysbiosis with oral cancer development," *Oncol. Lett.*, vol. 19, no. 4, pp. 3045–3058, Mar. 2020.

[191] S. Bond-Taylor, A. Leach, Y. Long, and C. G. Willcocks, "Deep generative modelling: A comparative review of VAEs, GANs, normalizing flows, energy-based and autoregressive models," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 11, pp. 7327–7347, Nov. 2022.

[192] B. Yelmen and F. Jay, "An overview of deep generative models in functional and evolutionary genomics," *Annu. Rev. Biomed. Data Sci.*, vol. 6, no. 1, pp. 173–189, Aug. 2023, doi: 10.1146/annurev-biodatasci-020722-115651.

[193] R. Mehmood, R. Bashir, and K. J. Giri, "Deep generative models: A review," *Indian J. Sci. Technol.*, vol. 16, no. 7, pp. 460–467, Feb. 2023.

[194] F. Xu, H. Uszkoreit, Y. Du, W. Fan, D. Zhao, and J. Zhu, "Explainable AI: A brief survey on history, research areas, approaches and challenges," in *Natural Language Processing and Chinese Computing*, J. Tang, M.-Y. Kan, D. Zhao, S. Li, and H. Zan, Eds. Cham, Switzerland: Springer, 2019, pp. 563–574.

[195] T. Hulsen, "Explainable artificial intelligence (XAI): Concepts and challenges in healthcare," *AI*, vol. 4, no. 3, pp. 652–666, Aug. 2023. [Online]. Available: https://www.mdpi.com/2673-2688/4/3/34

[196] R. Caruana, Y. Lou, J. Gehrke, P. Koch, M. Sturm, and N. Elhadad, "Intelligible models for HealthCare: Predicting pneumonia risk and hospital 30-day readmission," in *Proc. 21st ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*. New York, NY, USA: Association for Computing Machinery, Aug. 2015, pp. 1721–1730, doi: 10.1145/2783258.2788613.

[197] C. J. Rowlands, S. Varma, W. Perkins, I. Leach, H. Williams, and I. Notingher, "Rapid acquisition of Raman spectral maps through minimal sampling: Applications in tissue imaging," *J. Biophotonics*, vol. 5, no. 3, pp. 220–229, Mar. 2012.

[198] K. Kong, C. Kendall, N. Stone, and I. Notingher, "Raman spectroscopy for medical diagnostics—From in-vitro biofluid assays to in-vivo cancer detection," *Adv. Drug Del. Rev.*, vol. 89, pp. 121–134, Jul. 2015.

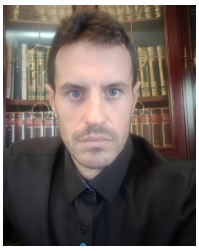
[199] A. Gupta, P. Matta, and B. Pant, "Graph neural network: Current state of art, challenges and applications," *Mater. Today, Proc.*, vol. 46, pp. 10927–10932, Jan. 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S2214785321010543

[200] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin, "Attention is all you need," 2017, *arXiv:1706.03762*.

[201] A. Waqas, A. Tripathi, R. P. Ramachandran, P. Stewart, and G. Rasool, "Multimodal data integration for oncology in the era of deep neural networks: A review," 2023, *arXiv:2303.06471*.

PIETRO MANGANELLI CONFORTI received the bachelor's degree in computer and automatic engineering and the master's degree in artificial intelligence and robotics from the Sapienza University of Rome, in 2019 and 2022, respectively, where he is currently pursuing the Ph.D. degree with the Department of Computer, Control, and Management Engineering "Antonio Ruberti." In 2022, he started the Ph.D. Program in computer engineering with the Sapienza University of Rome, where he is undergoing his research training in collaboration with the AlcorLab and Instituto Superiore di Sanità. His main areas of research focus on the application of machine learning and deep learning methodologies to tasks involving computer vision and signal analysis, with a focus on model explainability, data preprocessing, and applications in the medical field.

GIANMARCO LAZZINI received the Ph.D. degree in physics from the University of Parma, Italy. His research interests include surface physics and numerical modeling of the interactions between textured surfaces and bacteria. His current activity is focused on the employment of Raman spectroscopy coupled with machine learning for cancer diagnosis.



PAOLO RUSSO received the B.S. degree in telecommunication engineering from Università degli Studi di Cassino, Italy, in 2008, and the M.S. degree in artificial intelligence and robotics and the Ph.D. degree in computer science from the Sapienza University of Rome, Italy, in 2016 and 2020, respectively.

He is an Assistant Researcher with the Alcor-Lab, DIAG Department, Sapienza University of Rome. He currently teaches algorithms and data structures and vision and perception. From 2018 to 2019, he was a Researcher with the Italian Institute of Technology (IIT), Tourin, Italy. His main research interests include deep learning, computer vision, generative adversarial networks, and reinforcement learning. More information can be found on website https://www.paolorusso.org together with comprehensive publication lists.

MARIO D'ACUNTO received the Ph.D. degree in nanotribology from the University of Pisa, Italy, in 1999. He is a permanent Researcher with the Institute of Biophysics, Pisa, Italy. During the course of his scientific career, he has made pioneering contributions in the areas of nanotribology, nanooptics, nanophotonics, scanning probe microscopy, essentially atomic force microscopy, near-field optical scanning microscopy, and nonlinear dynamical systems. His scientific achievements include the first model ever proposed for onset wear on atomic scale, near-field plasmonic behavior of metal nanoparticles interacting with biological systems, Raman scattering and imaging, and super-resolution methods for image analysis in microscopy. He has given about 15 invited plenary talks at international conferences. He is the author or coauthor of about 130 indexed papers, cited ca. 1300 with an Hindex is 20 (Google Scholar, visited 11 March 2023), (eight papers cited more than 50 times), one book titled *Nature-Inspired Computation*, where, among various natural systems, nanophotonic paradigms stimulating and leading to computation skills are deeply described, and 12 monographic papers (book chapters). In 2012, he founded with Ovidio Salvetti the NanoICT Laboratory as a facility inside CNR. He is an Active Member of the International Technical Community; and the Chair, the Co-Chair, an Organizer, and a member of scientific committees of many national and international conferences on the foundations of light-matter interaction at nanoscale and material sciences. He has been a fellow of the European COST Action MP1302 Nanospectroscopy as Italian Delegate, from 2013 to 2017. He has been an Italian Delegate for Italy–Japan Bilateral Project on Fluctuation-Mediated Phonon-Assisted Nanophotonic Fabrication and Applications and proposed as a National Leader for the same project, since 2016. More information can be found on the following http://www.pi.ibf.cnr.it/?page_id=2115 together with comprehensive publication lists.

Open Access funding provided by 'Università degli Studi di Roma "La Sapienza" 2' within the CRUI CARE Agreement