


Perspectives on combining Nonlinear Laser Scanning Microscopy and Bag-of-Features data classification strategies for automated disease diagnostics

Stefan G. Stanciu¹  · Denis E. Tranca¹ · George A. Stanciu¹ · Radu Hristu¹ · Juan M. Bueno²

Received: 11 November 2015 / Accepted: 5 May 2016
© Springer Science+Business Media New York 2016

Abstract Nonlinear Laser Scanning Microscopy (NLSM) techniques have been demonstrated in the past two decades as powerful imaging tools for disease diagnostics (DD). Currently, in most DD related experiments the interpretation of NLSM data sets is performed by trained specialists. Such approaches are both time consuming and prone to errors due to inter- and intra-observer discrepancies. The Bag-of-Features (BoF) paradigm has demonstrated its potential usefulness with respect to automated data classification in the frame of multiple experiments, but its intersections with the field of NLSM are at this moment scarce, to say the least. In this paper we review recent progress on DD using NLSM, and discuss necessary steps and potential future perspectives for merging NLSM and BoF to achieve complex frameworks for automated DD with high sensitivity and specificity.

Keywords Nonlinear Laser Scanning Microscopy · Two-photon excitation fluorescence microscopy · Second harmonic generation microscopy · Fluorescence lifetime imaging · Coherent anti-stokes Raman microscopy · Bag-of-Features · Data classification · Data retrieval

This article is part of the Topical Collection on Laser technologies and laser applications.

Guest Edited by José Figueiredo, José Rodrigues, Nikolai A. Sobolev, Paulo André and Rui Guerra.

✉ Stefan G. Stanciu
stefan.stanciu@cmmip-upb.org

¹ Center for Microscopy-Microanalysis and Information Processing, University Politehnica of Bucharest, 313 Splaiul Independentei Bvd., 060042 Bucharest, Romania

² Laboratorio de Óptica, Universidad de Murcia, Campus Espinardo, 30100 Murcia, Spain

1 Introduction

Nonlinear Laser Scanning Microscopy (NLSM) techniques represent powerful tools for investigating tissue morphology, functionality and biochemical composition with high spatial and temporal resolution. Such techniques can be used for probing tissue specific endogenous signals, thus allowing for various medically relevant aspects to be assessed in the frame of both ex- and in-vivo assays. Moreover, in-vivo diagnostics and tumor delineation using NLSM can be done in tandem with laser surgery, which holds significant potential for accelerating surgical interventions for tumor resection, and for enhancing overall quality aspects of the medical act. To date, the analysis and interpretation of NLSM data sets have been successfully performed manually in a large number of experiments aimed at disease diagnostics, however it cannot be neglected that such manual evaluation approaches are both time consuming and prone to errors due to inter- and intra-observer discrepancies.

The Bag-of-Features (BoF) paradigm (O'Hara and Draper 2011), inspired from the Bag-of-Words (BoW) representation used in textual information retrieval, has been introduced to the field of computer vision about a decade ago. Since then, BoF approaches became well-established in the field because of their simplicity and performance. The potential of BoF methods for biomedical classification problems has been demonstrated so far in several experiments, but the methods reported to date rely mainly on spatial intensity information such as image gradients.

In this paper we recall the main principles behind four NLSM techniques that have been demonstrated as powerful tools for tissue characterization: Second Harmonic Generation Microscopy (SHG), Two-Photon Excitation Fluorescence Microscopy (TPEF), two-photon Fluorescence Lifetime Imaging (FLIM) and Coherent Anti-Stokes Raman Microscopy (CARS), we then briefly review important recent progress in tissue characterization by these four techniques, and discuss potential future steps and perspectives for merging NLSM and BoF in the purpose of achieving complex frameworks for automated DD with high sensitivity and specificity.

2 Contrast mechanisms of SHG, TPEF, FLIM and CARS and their significance for cell and tissue characterization

The field of NLSM is comprised of various techniques that rely on different contrast mechanisms. Depending on the contrast mechanism that is being used, complementary information categories can be investigated (even simultaneously). For example, Second Harmonic Generation Microscopy (Gauderon et al. 2000), represents a NLSM technique based on a nonlinear non-resonant and coherent process that has been demonstrated as a superb tool for imaging molecules with non-centrosymmetric structure, which exhibit a nonvanishing second-order susceptibility tensor $\chi^{(2)}$. Such molecules generate under the influence of an external electric field a nonlinear optical signal at exactly half the wavelength of the excitation source, which can be easily separated from the excitation light by using low-pass or band-pass filters positioned in front of the detector. Because this phenomenon occurs only in the focal volume, same as in the case of all other NLSM techniques, optical sectioning represents an inherent feature. One of the fundamental applications of SHG with respect to disease diagnostics (DD) consists in collagen imaging (Chen et al. 2012), which represents the main structural protein in the extracellular matrix

of animal tissues. Investigating collagen distribution with SHG in tissues enables a precise and non-invasive assessment of extracellular matrix modifications, which represent a hallmark of a wide range of pathologies, including cancer (Bonnans et al. 2014; Lu et al. 2012).

A different NLSM technique, Two-Photon Excitation Fluorescence Microscopy (So et al. 2000) is based on the nonlinear resonant and incoherent process that involves the simultaneous absorption of two photons whose combined energy is sufficient to induce an electronic transition to an excited electronic state. Excited by these two photons, a fluorophore acts in the same way as if excited by only one photon, emitting a single photon whose wavelength is only determined by its intrinsic characteristics, such as fluorophore type, chemical structure, etc. TPEF can thus be used for imaging various endogenous fluorophores such as nicotinamide adenine dinucleotide (NADH), flavins, tryptophan, elastin, melanin, lipofuscin, and others. Such possibilities offered by TPEF hold significant potential for DD as these make possible the non-invasive assessment of tissular information such as cell morphology, size variation of cell nuclei, blood vessel hyperplasia, or inflammatory reaction related aspects, which all represent indicators with significant importance for DD.

The advantages offered by TPEF can be exploited as well for time-resolved imaging in fluorescence lifetime imaging (FLIM) implementations that rely on tunable mode-locked solid state lasers providing picosecond or femtosecond pulses. FLIM enables the spatial distribution mapping of the excited state lifetimes corresponding to endogenous or exogenous fluorophores, the physical phenomena measured being the fluorescence intensity decay following excitation. Measuring the lifetime of a fluorescent signal, in addition to its intensity, allows monitoring the biophysical environments of the responsible fluorophore, independent of its concentration or of the excitation intensity. Therefore, although the contrast mechanisms of both FLIM and TPEF are based on fluorescence, these two can be regarded as complementary techniques because the FLIM signals depend on factors that define the local environment of the interrogated fluorophore such as ion and oxygen concentrations, temperature, pH, or on additional aspects such as its conformational or binding state (Suhling et al. 2015). Thus, the same fluorophore can exhibit different emission lifetimes depending on its environment, an aspect that has been successfully exploited to date in the frame of multiple DD oriented experiments which exploit metabolic contrast.

Another NLSM technique that holds high potential for DD is Coherent anti-Stokes Raman scattering (CARS) microscopy, which generates images based on the vibrational signatures of molecules, independent of their fluorescent properties or their centrosymetry. For the CARS phenomenon to occur, a target molecule needs to be irradiated using two short-pulse laser beams, the pump beam and the Stokes beam, with frequencies tuned such that their difference corresponds to a vibration of the target. When this illumination scenario takes place, coherently vibrating molecules in the probe volume scatter the probe beam, resulting in a coherent signal with a higher frequency than that of the probe beam (Pezacki et al. 2011). CARS is thus sensitive to the same vibrational signatures of molecules which are of interest in Raman spectroscopy, typically the nuclear vibrations of chemical bonds, but the signals exploited in CARS have much higher intensities than those originating from spontaneous Raman scattering allowing fast-rate imaging. This contrast mechanism makes CARS a perfect tool for probing specific molecular information in tissues corresponding to a wide range of bio-molecular compounds relevant for DD, such as lipids (Baenke et al. 2013).

3 Disease diagnostics using Nonlinear Laser Scanning Microscopy: recent examples

NLSM techniques can be used to collect non-invasively, in a label free manner, 2D images that recapitulate essential histology features that are typically used by pathologists for tissue characterization. An important study that successfully demonstrated this aspect has been conducted by (Makino et al. 2012), who employed combined TPEF/SHG imaging for creating an atlas of images of the entire gastrointestinal tract of a mouse, from the esophagus to the anus, including the organs of the hepatobiliary system and pancreas. In our opinion, this experiment represents a landmark study as the authors demonstrate an excellent correlation between TPEF/SHG images and histopathology images collected with classical optical microscopy on excised and Hematoxylyn and Eosin (H&E) stained tissue. Besides 2D intensity information with high potential for morphological analysis, TPEF and SHG are capable to yield additional information categories specific to the underlying contrast mechanism. A prominent example in this regard consists in the recent experiment reported by (Tokarz et al. 2015), where the authors use a combination of polarization related metrics to gather additional information on the extracellular matrix collagen, complementary to the concentration and spatial distribution measures. More precisely, the authors use a polarization-in \leftrightarrow polarization-out SHG microscopy approach to characterize classical papillary thyroid carcinoma and follicular variant papillary thyroid carcinoma. For this, the authors generate a support image set by using different combinations of orientation angles by rotating a half-wave plate placed directly in front of the excitation objective lens, and a polarizer (analyzer) located after the collection objective before the detector, resulting in excitation and SHG polarization orientation ranges between -90° and 90° . Based on this support image set the authors calculate the second-order nonlinear optical susceptibility tensor component ratio, $\chi^{(2)ZZZ'}/\chi^{(2)ZXX'}$, and the degree of linear polarization (DOLP) observing significant changes of these two metrics between tumorous and non-tumorous thyroid tissues. Tokarz et al. (2015) conclude that increased $\chi^{(2)ZZZ'}/\chi^{(2)ZXX'}$ most likely indicates a greater collagen fibril disorder, which correlates well with the decreased DOLP values obtained for cancerous thyroid tissue. Collagen orientation was assessed also in the case of H&E stained samples, which are typically used for histopathological exam, by (Birk et al. 2014). The authors of this experiment have studied sixteen H&E stained slides from colon biopsies through SHG imaging, four for each of the following categories: normal colonic mucosa, polyps with low grade dysplasia, polyps with high grade dysplasia, and malignant tumors. The investigation of the H&E slides was done in two stages. In a first one the SHG reflective intensity was assessed for each slide by dividing the sum of all intensities by the total number of pixels, while in the second stage the anisotropy parameter was calculated by: $\beta = (I_{||} - I_{\perp})/(I_{||} + 2I_{\perp})$ where $I_{||}$ and I_{\perp} correspond to SHG intensity detected when the analyzing polarizer is oriented parallel and perpendicular to the laser polarization. The results of the work performed by (Birk et al. 2014) demonstrate that SHG can be successfully used not only for assessing disease relevant tissue configurations in label-free samples but also in H&E samples that are routinely used for histopathology exam. Thus SHG could play an important role in consolidating the diagnostic provided by a trained pathologist. More precisely, the experiment demonstrated that a significantly higher SHG intensity is observed for the H&E stained slides corresponding to malignant samples which could be linked to higher collagen content. Moreover, a higher anisotropy value was observed for the case of H&E stained slides corresponding to the normal tissue in comparison to the cases of low grade dysplasia

and high grade dysplasia (for which the lowest anisotropy value was recorded). These results are linked to the alignment of the collagen fibers which is progressively disturbed along cancer progression, as shown also in (Tokarz et al. 2015). In another recent study that highlights the importance of NLSM with respect to DD, (Li et al. 2015) evaluate the feasibility of using combined TPEF/SHG imaging for distinguishing the layers of the bowel wall, including mucosa, submucosa, muscularis propria, and serosa, whose label-free investigation either in-vivo or ex-vivo holds as well significant potential for understanding and diagnosing colorectal cancers. In their experiment, the authors analyzed freshly excised tissue with nonlinear spectral analysis of TPEF signals, and conducted as well a detailed morphological analysis using SHG and TPEF intensity data. Their results demonstrate that the nonlinear spectral analysis of TPEF spectra can lead to the correct identification of the constituents of colorectal tissues and to label-freely identify the four-layer microstructures of fresh colorectal tissue. Colorectal tissue has been investigated by NLSM also in an experiment presented by (Li et al. 2016), where the authors investigate normal tissue and tissue affected by rectal adenoma-carcinoma excised from patients enrolled in a study at the Fujian Medical University Union Hospital. In this experiment, the main focus of attention was given to the thorough investigation of the collagen changes that occur in cancerous tissue in the course of neoadjuvant treatment consisting in radiochemotherapy. Modifications in the collagen structure of colorectal tissue were quantitatively described using an intensity ratio of SHG over TPEF signals, which was obtained by measuring the nonlinear emission spectrum. Their quantitative analysis showed that the intensity ratio of SHG to TPEF in desmoplasia and in stromal fibrosis is consistently increased in comparison to the case of normal tissue. Moreover, the collagen orientation index in the case of normal tissue was found to be much higher than the one calculated for tissue affected by desmoplasia and stromal fibrosis, which represents direct evidence that the collagen orientation gradually becomes more disordered during the development of rectal carcinoma undergoing neoadjuvant treatment. Colorectal tissue was also investigated by NLSM by (Cicchi et al. 2013) in a work in which the authors discriminate between healthy mucosa, adenomatous polyp, and adenocarcinoma in ex-vivo fresh colorectal biopsies using TPEF data sets. Cellular/nuclear ratios, as well as asymmetry aspects, were assessed based on TPEF intensity maps in the frame of an extensive morphological characterization experiment while functional differences were observed in terms of metabolic activity by measuring the relative abundance of NADH and flavin adenine dinucleotide (FAD) through the selective acquisition of their fluorescence. The ratio between these two endogenous metabolic biomarkers, also known as the metabolic redox ratio, plays a key role in the detection and assessments of cellular metabolism which is an important indicator for cancer diagnostics, as it is believed that increased cellular metabolism in cancer cells may cause the route of energy production to shift from the oxidative phosphorylation pathway to the glycolytic pathway (Gulledge and Dewhirst 1995). For calculating the metabolic redox ratio, (Cicchi et al. 2013) use two excitation wavelengths and two corresponding spectral windows for detection, in order to minimize the cross-talk between NADH and FAD fluorescence signals. Based on this approach, they observe an altered metabolic condition not only in adenocarcinomas but also in adenomatous polyps, demonstrating that the altered metabolism, precursor of cancer development, can be detected optically. The same metabolic redox ratio was also used by (Thomas et al. 2014) to investigate the cellular metabolism of epidermal cells. Their experiment led to the conclusion that the NADH/FAD ratio in epidermal cells is higher in skin with preclinical hyperplasia and clinically visible early skin tumors, when compared to normal skin. This represents an important discovery with respect to cancer diagnostics, as it leads

to the conclusions that elevated cellular metabolism in the early preclinical stage of hyperplasia can be regarded as an important indicator for the addressed pathology.

Metabolic contrast was also considered by (Chakraborty et al. 2016) who performed an experiment meant to differentiate and quantify the amounts of free and protein-bound NADH and FAD with respect to monitoring the cellular energetics in Parkinson's disease (PD). More precisely, the authors used two-photon FLIM for functional mapping of cellular redox states in normal and pathological neuronal cells, based on a protocol that serves as a model system of PD in laboratory conditions. NADH/NAD⁺ and FADH₂/FAD, redox ratios of neuronal cells, were evaluated in terms of the ratios of free-to-protein-bound NADH and FAD. The performed experiments brought solid experimental proof that FLIM data derived from interrogating the two considered autofluorescent coenzymes via the proposed methodology can be used to differentiate between normal and diseased cells, as FLIM based metrics turned out to be good indicators with respect to metabolic activity and changes in metabolic state of neuronal cells. In conclusion to their experiment, the authors claimed that similar FLIM based studies focused on tissue model or in-vivo animal models could significantly advance our current understanding of PD causes and progression, and could also lead to novel label-free PD diagnostic tools, which could operate in-vivo. A study that represents a consistent step forward towards developing brain oriented in-vivo diagnostic methodologies based on fluorescence data consists in the recent milestone experiment reported by (Kantelhardt et al. 2016), where consecutive patients with cranial tumors were enrolled in a study that incorporated TPEF and FLIM imaging of ex-vivo tumor samples, but also intraoperative investigations using a cutting-edge MPTflexTM Multiphoton Tomograph (JenLab GmbH, Germany) based on a tuneable NIR femtosecond laser. In their experiments focused on the investigation of ex-vivo high grade glioma tissues, the authors found TPEF images to show a chaotic architecture of tumor cells, as could also be observed with conventional histology. FLIM provided complementary information demonstrating image areas of long fluorescence lifetime (up to 2000 ps) interspersed with image areas of shorter fluorescence lifetime (1000 ps), these latter corresponding to necrotic tissue. TPEF investigations of ex-vivo glioblastoma (malignant glioma) tissues demonstrated images in which typical cytological features of the considered pathology, such as hypercellularity, pleomorphism, necrosis or microvascular proliferation were identified; FLIM images allowed the differentiation of necrotic regions, with absent nuclei, which showed a specifically low fluorescent lifetime. Further on, human glioblastoma was also imaged in-situ, by positioning the imaging system over the craniotomy after it was sterilized. The intraoperative in-situ TPEF intensity and FLIM images collected on solid tumor tissue resembled the images collected ex-vivo, showing heterogeneous cells with loss of any structured histo-architecture, which is typical for high grade glioma. FLIM images collected on the intact arachnoid membrane/pia covering the tumor revealed longer fluorescence lifetimes in comparison to those obtained in the solid tumor. As a conclusion to their experiment, the authors point out that combined TPEF & FLIM imaging holds potential to be used for intraoperative optical biopsy and guidance of resection in neurosurgery. In a different study aimed at FLIM imaging of human tissues, (Giacomelli et al. 2015) demonstrate what they claim to be the highest resolution TPEF-FLIM images generated in freshly-excised human tissue. These were collected with their setup which exploits direct temporal sampling of fluorescent decays phase-locked to a high harmonic of mode-locked laser repetition rate (DTS-FLIM). Their DTS-FLIM investigations performed on healthy tissue specimens originated from partial mastectomy demonstrated consistent lifetime ratios for normal fat and stroma within widely separated areas of the same specimen, but also across different specimens. According to the authors, these

results indicate that the lifetimes of the fluorescent components of normal breast tissue are well regulated, and that the lifetime ratio increases with cellularity. Increased cellularity, which takes places in breast carcinoma or ductal carcinoma in-situ, could therefore result in tumor specific FLIM signatures, thus exhibiting significant potential for the in vivo diagnostics of such pathologies.

In recent years CARS was also demonstrated in multiple experiments as an important tool for DD (Kong et al. 2015; Schie et al. 2015). The potential of CARS with respect to DD stands in its capabilities of collecting signals that arise from lipidic C-H bond stretches. The roles of lipids in biology are diverse and hold deep implications for a wide range of pathologies (Baenke et al. 2013; Kraemer et al. 2013; Nordestgaard and Varbo 2014). Dysregulations in lipid synthesis or metabolism can result in diseases such as diabetes, neurodegeneration, cancers and many others. A recent example on how CARS data collected on the CH-stretching region can be used for DD consists in the experiment reported in (Huang et al. 2015). There, CARS data was used for discriminating between gastrocnemius tissue biopsies collected from healthy patients and from patients diagnosed with Peripheral Artery Disease (PAD), a common circulatory problem which can ultimately lead to leg dysfunction. Sets of spectra from 12 tissue specimens corresponding to 12 different patients (4 healthy, 4 with moderate PAD and 4 with severe PAD) were recorded, and after a pre-processing step meant to minimize multiplicative error, a Principal Component Analysis Scheme was performed to reduce the data dimensions to a smaller number of Principal Components that incorporate the most important information from the spectra. Further on, a classification method known as principal component—discriminant function analysis (PC-DFA) (Hamasha et al. 2013) was applied for classifying the collected spectra. This led to achieving an overall classification accuracy of $\sim 79\%$, $\sim 7\%$ lower compared to the case when spontaneous Raman spectra were used in a similar classification scenario. Further on, in order to evaluate the sensitivity and specificity of the proposed methodology for classification of CARS spectra, these were divided into “healthy” and “diseased” classes. In this binary classification scenario the PC-DFA framework based on CARS data proved to be very sensitive ($\sim 95\%$) on identification of disease severity, which indicates that future efforts in this direction could lead to a valuable tool for clinical diagnostics of PAD. Less than a decade ago, CARS was demonstrated for the first time as a useful tool for imaging unstained brain tissue (Evans et al. 2007). Since then, many other experiments continuously consolidated the idea that CARS, either by itself or in combination with other optical techniques, can make possible the in-vivo diagnostic of various brain related pathologies. For example, in the experiment of (Uckermann et al. 2014) different human tumors were induced in an orthotopic mouse model and imaged along with biopsies collected from human patients using a CARS configuration tuned to probe vibrational signatures of lipids. Among many of the important results reported in this study, the authors observed that in both mouse model and in human biopsies glioblastoma tumors were characterized by CARS signal intensities lower than those corresponding to healthy tissues. In both considered cases, CARS intensities collected for tumor, infiltration zone and normal tissue were found to fall in separated ranges, allowing thus a precise delineation of tumor borders, which is extremely important with respect to the success of the medical act aimed to surgically remove the pathologic tissue in its entirety, thus minimizing the risk of recidive. We end this mini-review oriented towards highlighting recent experiments that stand as direct proof of the potential that NLSM techniques hold with respect to DD, by providing a representative example on the significant value exhibited by CARS also in clinical settings. In an experiment reported by (Breunig et al. 2013) CARS data (together with TPEF and SHG data) were collected in-

vivo from the forearms of male and female volunteers and patients at the at the Charite Hospital in Berlin using an adapted variant of the DermaInspect system (JenLab GmbH., Germany). A comparison of TPEF/SHG intensity data to CARS data revealed that the lipid distribution imaged in-vivo by CARS on healthy and dysplastic tissues does not merely replicate the cellular structure of the tissular layer but provides additional information of pathological relevance. All these experiments represent convincing evidence for the huge potential that NLSM techniques hold with respect to DD. This considerable potential can be amplified by combining NLSM data acquisition with computer vision and other machine intelligence techniques to accelerate the speed at which analysis tasks are performed, but also to generate outputs that are not biased by human subjectivity and error.

4 Combining the Bag-of-Features paradigm and Nonlinear Laser Scanning Microscopy for automated diseases diagnostics

The BoF paradigm (O'Hara and Draper 2011) has been used to date in a wide range of image classification methods and scenarios, e.g. (Lazebnik et al. 2006; Cao et al. 2010; Xu et al. 2010; Zhou et al. 2013; Eweiji et al. 2015; Nanni and Melucci 2016). BoF has its origins in the BoW text categorization methods used for document classification. In such BoW approaches, a document is classified as belonging to a particular category based on a normalized histogram of word counts. BoF methods, also known as Bag-of-Visual-Words, adapt this text categorization approach to a visual categorization one by replacing the dictionary of textual words with a dictionary of visual ones, usually referred to as “visual features”. A typical BoF method, schematically illustrated in Fig. 1, requires the assignment of a BoF representation to an image, which consists in a histogram that indicates the number of occurrences of the visual words that take part in a dictionary in the respective image. The dictionary (a.k.a. codebook) can be either assembled manually so as to include relevant terms with respect to the approached classification problem, or can be built automatically by running a clustering algorithm over a large set of visual features in order to divide them into distinct groups and to identify the representative of each group (e.g. cluster mean). Given a novel training or test image, visual features are detected in it and assigning them their nearest matching terms from the visual vocabulary results in a normalized histogram of the quantized features detected in the image. Based on such BoF representations (a.k.a term vector), a target image can be classified as belonging to a specific class by comparing its term vector to the term vectors of images included apriori in training classes. Such comparisons are typically performed in an automatic fashion, using different types of classifiers such as k-Nearest Neighbor (k-NN) classifiers, Support Vector Machines (SVM), etc.

A first experiment to exploit the versatility of the BoF paradigm for the classification of NLSM data sets was aimed at liver fibrosis diagnostic in the case of a Wistar rat model using TPEF data sets collected ex-vivo by scanning the liver surface (Stanciu et al. 2014). This experiment is aligned to the general trend of BoF methods aimed at biomedical image classification problems, which consists in using image features based on spatial intensity information (Caicedo et al. 2009; Tamaki et al. 2013; Sadek et al. 2015). In this regard, the Scale Invariant Feature Transform (SIFT) (Lowe 2004) and adapted variants are currently considered as default choices for implementing BoF frameworks aimed at classification of image data sets because of their efficiency in detecting image regions of high invariance to affine transformations (a.k.a keypoints), and in describing these using robust descriptors. In

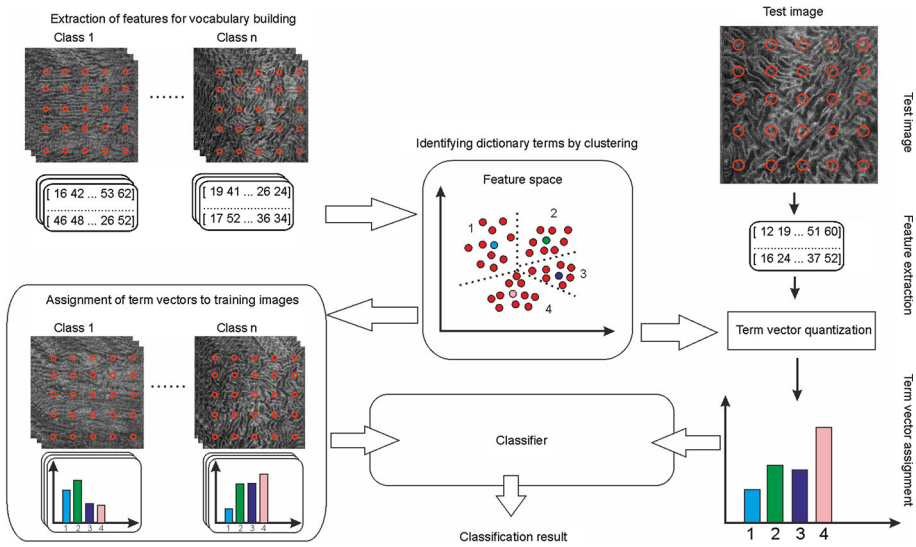


Fig. 1 A typical BoF framework for image classification [adapted from (Stanciu et al. 2014)]. A codebook feature space is created by extracting image features (usually at fixed grid locations) in a set of images dedicated to vocabulary building. A clustering algorithm can be used to partition the resulted feature space into k regions and to identify the centroids of these regions, which will represent the codebook terms. The codebook (a.k.a. dictionary) allows for BoF representations, *term vectors*, to be assigned to training and test images. The *term vector* is a histogram that indicates the number of occurrences of the codebook terms in an image. An image is classified as belonging to a specific class by assigning a term vector to it, and running this term vector through a classifier that compares it with term vectors assembled for images with ground-truth tagging, which constitute the training classes. Example images provided for vocabulary building, training and test in this example represent images collected by SHG on pig cornea

SIFT, this is achieved by constructing a keypoint descriptor which is a histogram representation that combines local gradient orientations and magnitudes from a certain neighborhood around a keypoint. More precisely, the SIFT descriptor is in fact a 3D histogram of gradient location and orientation, where location is quantized into a 4×4 location grid and the gradient angle is quantized into 8 orientations, one for each of the cardinal directions. The resulting descriptor, which represents the “image feature” (or “visual word”) in SIFT based BoF approaches, is a normalized vector with the dimension of 128 elements. SIFT, and other local feature descriptors such as SURF (Bay et al. 2008), HOG (Dalal and Triggs 2005), BRISK (Leutenegger et al. 2011), ORB (Rublee et al. 2011) or FREAK (Alahi et al. 2012), do a great job at processing spatial intensity information in a BoF-friendly manner, but the great majority of such techniques are developed by the computer vision community taking as reference the specifics of natural images. These techniques can be of course used in association with image data originating from any other source than a digital camera, e.g. microscopes or medical imaging instruments, but this is done at the cost of unpredictable performance due to image content specifics that differ from those characteristic to natural images. For example, in two previous experiments (Stanciu et al. 2010; Stanciu et al. 2011) it was shown that the detection and matching of SIFT and SURF features is sensitive to image content modifications related to the adjustment of typical Laser Scanning Microscopy (LSM) acquisition parameters, such as pinhole aperture, photomultiplier gain or laser beam power. Moreover, in (Stanciu et al.

2013) it was shown that image deconvolution, which is routinely used in combination with laser scanning microscopies for eliminating unwanted contributions of the point spread function, affect the matching of gradient based local feature descriptors. All these represent direct evidence that the application of feature detection/description techniques developed for general use to LSM images, although straightforward, raises a series of concerns with respect to the robustness and consistency of the results. In order to address such issues, more efforts are required for either adjusting popular computer vision techniques developed for natural images so that these take into account specific properties of LSM images, or for developing completely novel methods for local feature detection/description that are tailored to the specifics of LSM imaging data. Moreover, as recalled by the brief review of recent progress for DD using NLSM provided in Sect. 3, NLSM data sets host not only morphology related information available as 2D image data but also other information categories that are equally important for DD. Exploiting for classification purposes only 2D image data with BoF strategies that rely solely on image features such as SIFT, SURF, FREAK or others, leaves out important additional content harbored in additional information categories, leading thus to an incomplete exploitation of available data. In our view, important steps that are required to address this issue consist in the development of novel strategies to extract BoF-friendly data features from NLSM information categories complementary to imaging data, which at this moment represents an unexplored domain. This would play a very important role with respect to the better penetration of BoF in the realm of NLSM data classification. Among such NLSM information categories that should be considered in novel pattern recognition, content modeling and machine intelligence approaches, fluorescence emission spectra, fluorescence lifetime spectra, TPEF/SHG signal ratio, SHG dependence on polarization or CARS spectra are noteworthy candidates to be taken into account. Features extracted from these categories could play important roles in DD oriented classification scenarios based on the BoF paradigm, a claim that we make based on what has been discussed in Sects. 2 and 3. With respect to classification of fluorescence or CARS spectra, strategies similar to the “Bag of Peaks” method (Brelstaff et al. 2009), where the authors characterize nuclear magnetic resonance traces on the basis of the occurrences of particular peaks in the collected data sets, could represent solid solutions. Also, BoF strategies developed for Music Information Retrieval, e.g. (Fu et al. 2011; Nanni et al. 2016) seem to embed as well many concepts useful for exploitation of NLSM data. For example, Bag-of-Frames approaches, such as ones discussed in (Aucouturier et al. 2007), rely on audio pattern recognition by representing audio sequences as a statistical distribution of contained local spectral features. Modalities of detecting and modeling spectral features relevant for audio analysis rely on processing 1D data sequences, which is also the topic of interest when it comes to NLSM information categories such as fluorescence emission spectra, fluorescence lifetime spectra or CARS spectra. Thus, placing efforts on identifying concepts from the realm of audio data classification that could find application to NLSM data classification could turn out to be very valuable, since in this area of audio recognition a huge volume of work has taken place so far. Another important step that we also consider important with respect to the approached subject consists in the development of novel BoF strategies for exploiting in parallel NLSM data features based on different information categories. Such different categories could be either multiple 2D information categories, multiple 1D information categories, or mixed 1D and 2D categories. In this regard two types of strategies for exploiting in parallel multiple feature classes associated to multiple information categories were briefly proposed in (Stanciu et al. 2015), one based on the Multiple Dictionary BoF (a.k.a MDBoW) model (Aly et al. 2011) and a second one based on running in parallel independent BoF

frameworks for different information categories embedded in the same NLSM data set, and merging the results using machine learning methods for multiple expert decision, e.g. (Raykar et al. 2009). In our view such combined approaches could represent elegant solutions for automatically classifying complex NLSM data sets in the purpose of diagnosing a wide range of pathologies with high sensitivity and specificity.

5 Conclusions

NLSM data sets host multiple information categories whose importance for DD has been demonstrated in a wide range of experiments during the past two decades. Novel automated strategies for NLSM data classification based on the BoF paradigm could play a key-role in aiding or even replacing classical approaches for DD based on morphological interpretation of tissue by standard histopathology. For developing such strategies a concerted effort is required for developing novel feature extraction and description methods tailored to the specifics of various NLSM information categories, such as fluorescence emission spectra, fluorescence lifetime spectra, TPEF/SHG ratio, SHG intensity dependence on polarization, CARS spectra, etc. In this effort an intense collaboration between scientists conducting research in high resolution imaging, optics, spectroscopy, computer vision, machine learning and statistical analysis will be crucial. Methods for automated NLSM data classification yielded by such multidisciplinary collaborations could play an important role with respect to the sustainability of healthcare systems across the world by improving DD in terms of speed and accuracy.

Acknowledgments The presented work was partially supported by the PN-II-RU-TE-2014-4-1803 and PN-II-PT-PCCA-2011-3.2-1162 Research Grants, funded by the Romanian Executive Agency for Higher Education, Research, Development and Innovation Funding (UEFISCDI). The corresponding author acknowledges as well the financial support of the SOP HRD, financed from the European Social Fund and the Romanian Government under the contract number POSDRU/159/1.5/S/137390/. The contribution of J.M. Bueno was supported by the Spanish SEIDI through the research grants FIS2013-41237-R and FIS2015-71933-REDT.

References

- Alahi, A., Ortiz, R., Vandergheynst, P.: Freak: fast retina keypoint. In: 2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 510–517. IEEE (2012)
- Aly, M., Munich, M., Perona, P.: Multiple dictionaries for bag of Words Large Scale Image Search. In: IEEE International Conference on Image Processing (ICIP) (2011)
- Aucouturier, J.-J., Defreville, B., Pachet, F.: The bag-of-frames approach to audio pattern recognition: a sufficient model for urban soundscapes but not for polyphonic music. *J. Acoust. Soc. Am.* **122**(2), 881–891 (2007)
- Baenke, F., Peck, B., Miess, H., Schulze, A.: Hooked on fat: the role of lipid synthesis in cancer metabolism and tumour development. *Dis. Models Mech.* **6**(6), 1353–1363 (2013)
- Bay, H., Ess, A., Tuytelaars, T., Van Gool, L.: Speeded-up robust features (SURF). *Comput. Vis. Image Underst.* **110**(3), 346–359 (2008)
- Birk, J.W., Tadros, M., Moezardalan, K., Nadyarnykh, O., Forouhar, F., Anderson, J., Campagnola, P.: Second harmonic generation imaging distinguishes both high-grade dysplasia and cancer from normal colonic mucosa. *Dig. Dis. Sci.* **59**(7), 1529–1534 (2014)
- Bonnans, C., Chou, J., Werb, Z.: Remodelling the extracellular matrix in development and disease. *Nat. Rev. Mol. Cell Biol.* **15**(12), 786–801 (2014)
- Brelstaff, G., Bicego, M., Culeddu, N., Chessa, M.: Bag of peaks: interpretation of nmr spectrometry. *Bioinformatics* **25**(2), 258–264 (2009)

- Breunig, H.G., Weinigel, M., Bückle, R., Kellner-Höfer, M., Lademann, J., Darvin, M.E., Sterry, W., König, K.: Clinical coherent anti-Stokes Raman scattering and multiphoton tomography of human skin with a femtosecond laser and photonic crystal fiber. *Laser Phys. Lett.* **10**(2), 025604 (2013)
- Caicedo, J.C., Cruz, A., Gonzalez, F.A.: Histopathology image classification using bag of features and kernel functions. In: *Artificial Intelligence in Medicine*. pp. 126–135. Springer (2009)
- Cao, Y., Wang, C., Li, Z., Zhang, L., Zhang, L.: Spatial-bag-of-features. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3352–3359. IEEE (2010)
- Chakraborty, S., Nian, F.-S., Tsai, J.-W., Karmenyan, A., Chiou, A.: Quantification of the metabolic state in cell-model of Parkinson's disease by fluorescence lifetime imaging microscopy. *Sci. Rep.* **6**, 19145 (2016)
- Chen, X., Nadiarynkh, O., Plotnikov, S., Campagnola, P.J.: Second harmonic generation microscopy for quantitative analysis of collagen fibrillar structure. *Nat. Protoc.* **7**(4), 654–669 (2012)
- Cicchi, R., Sturiale, A., Nesi, G., Kapsokalyvas, D., Alemanno, G., Tonelli, F., Pavone, F.S.: Multiphoton morpho-functional imaging of healthy colon mucosa, adenomatous polyp and adenocarcinoma. *Biomed. Opt. Express* **4**(7), 1204–1213 (2013)
- Dalal, N., Triggs, B.: Histograms of oriented gradients for human detection. In: *Computer Vision and Pattern Recognition, 2005. CVPR 2005*. In: Conference on IEEE Computer Society, pp. 886–893. IEEE (2005)
- Evans, C.L., Xu, X., Kesari, S., Xie, X.S., Wong, S.T., Young, G.S.: Chemically-selective imaging of brain structures with CARS microscopy. *Opt. Express* **15**(19), 12076–12087 (2007)
- Eweivi, A., Cheema, M.S., Bauckhage, C.: Action recognition in still images by learning spatial interest regions from videos. *Pattern Recogn. Lett.* **51**, 8–15 (2015)
- Fu, Z., Lu, G., Ting, K.M., Zhang, D.: Music classification via the bag-of-features approach. *Pattern Recogn. Lett.* **32**(14), 1768–1777 (2011)
- Gauderon, R., Lukins, P., Sheppard, C.: Second-harmonic generation imaging. In: *Optics and Lasers in Biomedicine and Culture*. pp. 66–69. Springer (2000)
- Giacomelli, M.G., Sheikine, Y., Vardeh, H., Connolly, J.L., Fujimoto, J.G.: Rapid imaging of surgical breast excisions using direct temporal sampling two photon fluorescent lifetime imaging. *Biomed. Opt. Express* **6**(11), 4317–4325 (2015)
- Gulledge, C., Dewhirst, M.: Tumor oxygenation: a matter of supply and demand. *Anticancer Res.* **16**(2), 741–749 (1995)
- Hamasha, K., Mohaidat, Q.I., Putnam, R.A., Woodman, R.C., Palchadhuri, S., Rehse, S.J.: Sensitive and specific discrimination of pathogenic and nonpathogenic *Escherichia coli* using Raman spectroscopy—a comparison of two multivariate analysis techniques. *Biomed. Opt. Express* **4**(4), 481–489 (2013)
- Huang, X., Irmak, S., Lu, Y., Pipinos, I., Casale, G., Subbiah, J.: Spontaneous and coherent anti-Stokes Raman spectroscopy of human gastrocnemius muscle biopsies in CH-stretching region for discrimination of peripheral artery disease. *Biomed. Opt. Express* **6**(8), 2766–2777 (2015)
- Kantelhardt, S.R., Kalasauskas, D., König, K., Kim, E., Weinigel, M., Uchugonova, A., Giese, A.: In vivo multiphoton tomography and fluorescence lifetime imaging of human brain tumor tissue. *J. Neuro-Oncol.* **127**(3), 473–482 (2016)
- Kong, K., Kendall, C., Stone, N., Notingher, I.: Raman spectroscopy for medical diagnostics—From in vitro biofluid assays to in vivo cancer detection. *Adv. Drug Deliv. Rev.* **89**, 121–134 (2015)
- Krahmer, N., Farese, R.V., Walther, T.C.: Balancing the fat: lipid droplets and human disease. *EMBO Mol. Med.* **5**(7), 973–983 (2013)
- Lazebnik, S., Schmid, C., Ponce, J.: Beyond bags of features: Spatial pyramid matching for recognizing natural scene categories. In: 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, pp. 2169–2178. IEEE (2006)
- Leutenegger, S., Chli, M., Siewgart, R.Y.: BRISK: Binary robust invariant scalable keypoints. In: 2011 IEEE International Conference on Computer Vision (ICCV), pp. 2548–2555. IEEE (2011)
- Li, L., Chen, Z., Wang, X., Liu, X., Jiang, W., Zhuo, S., Guan, G., Chen, J.: Visualization of tumor response to neoadjuvant therapy for rectal carcinoma by nonlinear optical imaging. *IEEE J. Sel. Top. Quantum Electron.* **22**(3), 6800206 (2016)
- Li L, Li H, Chen Z, Zhuo S, Feng C, Yang Y, Guan G, Chen J Layer-resolved colorectal tissues using nonlinear microscopy. *Lasers Med. Sci.* 1–9 (2015)
- Lowe, D.G.: Distinctive image features from scale-invariant keypoints. *Int. J. Comput. Vision* **60**(2), 91–110 (2004)
- Lu, P., Weaver, V.M., Werb, Z.: The extracellular matrix: a dynamic niche in cancer progression. *J. Cell Biol.* **196**(4), 395–406 (2012)
- Makino, T., Jain, M., Montrose, D.C., Aggarwal, A., Sterling, J., Bosworth, B.P., Milsom, J.W., Robinson, B.D., Shevchuk, M.M., Kawaguchi, K.: Multiphoton tomographic imaging: a potential optical biopsy

- tool for detecting gastrointestinal inflammation and neoplasia. *Cancer Prev. Res.* **5**(11), 1280–1290 (2012)
- Nanni, L., Costa, Y.M., Lumini, A., Kim, M.Y., Baek, S.R.: Combining visual and acoustic features for music genre classification. *Expert Syst. Appl.* **45**, 108–117 (2016)
- Nanni, L., Melucci, M.: Combination of projectors, standard texture descriptors and bag of features for classifying images. *Neurocomputing* **173**, 1602–1614 (2016)
- Nordestgaard, B.G., Varbo, A.: Triglycerides and cardiovascular disease. *Lancet* **384**(9943), 626–635 (2014)
- O'Hara, S., Draper, B.A.: Introduction to the bag of features paradigm for image classification and retrieval. arXiv preprint [arXiv:1101.3354](https://arxiv.org/abs/1101.3354) (2011)
- Pezacki, J.P., Blake, J.A., Danielson, D.C., Kennedy, D.C., Lyn, R.K., Singaravelu, R.: Chemical contrast for imaging living systems: molecular vibrations drive CARS microscopy. *Nat. Chem. Biol.* **7**(3), 137–145 (2011)
- Raykar, V.C., Yu, S., Zhao, L.H., Jerebko, A., Florin, C., Valadez, G.H., Bogoni, L., Moy, L.: Supervised learning from multiple experts: whom to trust when everyone lies a bit. In: Proceedings of the 26th Annual international conference on machine learning, pp. 889–896. ACM (2009)
- Rublee, E., Rabaud, V., Konolige, K., Bradski, G.: ORB: an efficient alternative to SIFT or SURF. In: 2011 IEEE International Conference on Computer Vision (ICCV), pp. 2564–2571. IEEE (2011)
- Sadek, I., Sidibé, D., Meriaudeau, F.: Automatic discrimination of color retinal images using the bag of words approach. In: SPIE Medical Imaging 2015, pp. 94141 J-94141 J-94148. International Society for Optics and Photonics
- Schie, I.W., Krafft, C., Popp, J.: Applications of coherent Raman scattering microscopies to clinical and biological studies. *Analyst* **140**(12), 3897–3909 (2015)
- So, P.T., Dong, C.Y., Masters, B.R., Berland, K.M.: Two-photon excitation fluorescence microscopy. *Annu. Rev. Biomed. Eng.* **2**(1), 399–429 (2000)
- Stanciu, S.G., Coltuc, D., Tranca, D.E., Stanciu, G.A.: Matching DSIFT descriptors extracted from CSLM images. *Engineering* **5**(10), 199–202 (2013)
- Stanciu, S.G., Hristu, R., Boriga, R., Stanciu, G.A.: On the suitability of SIFT technique to deal with image modifications specific to Confocal Scanning Laser Microscopy. *Microsc. Microanal.* **16**(05), 515–530 (2010)
- Stanciu, S.G., Hristu, R., Stanciu, G.A.: Influence of confocal scanning laser microscopy specific acquisition parameters on the detection and matching of speeded-up robust features. *Ultramicroscopy* **111**(5), 364–374 (2011)
- Stanciu, S.G., Hristu, R., Tranca, D.E., Stanciu, G.A.: Bags of features for classification of Laser Scanning Microscopy data. In: 2015 17th International Conference on Transparent Optical Networks (ICTON), pp. 1–4. IEEE (2015)
- Stanciu, S.G., Xu, S., Peng, Q., Yan, J., Stanciu, G.A., Welsch, R.E., So, P.T., Csucs, G., Yu, H.: Experimenting liver fibrosis diagnostic by two photon excitation microscopy and bag-of-features image classification. *Sci. Rep.* **4**, 4636 (2014)
- Suhling, K., Hirvonen, L.M., Levitt, J.A., Chung, P.-H., Tregidgo, C., Le Marois, A., Rusakov, D.A., Zheng, K., Ameer-Beg, S., Poland, S.: Fluorescence lifetime imaging (FLIM): basic concepts and some recent developments. *Med. Photonics* **27**, 3–40 (2015)
- Tamaki, T., Yoshimuta, J., Kawakami, M., Raytchev, B., Kaneda, K., Yoshida, S., Takemura, Y., Onji, K., Miyaki, R., Tanaka, S.: Computer-aided colorectal tumor classification in NBI endoscopy using local features. *Med. Image Anal.* **17**(1), 78–100 (2013)
- Thomas, G., van Voskuilen, J., Truong, H., Song, J.-Y., Gerritsen, H.C., Sterenborg, H.: In vivo nonlinear spectral imaging as a tool to monitor early spectroscopic and metabolic changes in a murine cutaneous squamous cell carcinoma model. *Biomed. Opt. Express* **5**(12), 4281–4299 (2014)
- Tokarz, D., Cisek, R., Golaraei, A., Asa, S.L., Barzda, V., Wilson, B.C.: Ultrastructural features of collagen in thyroid carcinoma tissue observed by polarization second harmonic generation microscopy. *Biomed. Opt. Express* **6**(9), 3475–3481 (2015)
- Uckermann, O., Galli, R., Tamosaityte, S., Leipnitz, E., Geiger, K.D., Schackert, G., Koch, E., Steiner, G., Kirsch, M.: Label-free delineation of brain tumors by coherent anti-stokes Raman scattering microscopy in an orthotopic mouse model and human glioblastoma. *PLoS One* **9**(9), e107115 (2014)
- Xu, S., Fang, T., Li, D., Wang, S.: Object classification of aerial images with bag-of-visual words. *Geosci. Remote Sens. Lett. IEEE* **7**(2), 366–370 (2010)
- Zhou, L., Zhou, Z., Hu, D.: Scene classification using a multi-resolution bag-of-features model. *Pattern Recogn.* **46**(1), 424–433 (2013)