Towards a portable platform integrated with multi-spectral non-contact probes for delineating normal and breast cancer tissue based on near-infrared spectroscopy

Uttam M. Pal, *Member, IEEE,* Anil Vishnu GK, Gayatri Gogoi, Saeed Rila, Saahil Shroff, Gokul AM, Manoj Varma, Vishnu Kurpad, Deb Baruah, Pronami Borah, Jayant S. Vaidya, Hardik J. Pandya, *Member, IEEE*

Abstract—**Currently, the confirmation of diagnosis of breast cancer is made by microscopic examination of an ultra-thin slice of a needle biopsy specimen. This slice is conventionally formalin-fixed and stained with hematoxylin-eosin and visually examined under a light microscope. This process is labor-intensive and requires highly skilled doctors (pathologists). In this paper, we report a novel tool based on near-infrared spectroscopy (Spectral-IRDx) which is a portable, non-contact, and cost-effective system and could provide a rapid and accurate diagnosis of cancer. The Spectral-IRDx tool performs absorption spectroscopy at near-infrared (NIR) wavelengths of 850 nm, 935 nm, and 1060 nm. We measure normalized detected voltage (V_{dn}) with the tool in 10 deparaffinized breast biopsy tissue samples, 5 of which were cancer (C) and 5 were normal (N) tissues. The difference in V_{dn} at 935 nm and 1060 nm between cancer and normal tissues is statistically significant with *p*-values of 0.0038 and 0.0022 respectively. Absorption contrast factor (N/C) of 1.303, 1.551, and 1.45 are observed for 850 nm, 935 nm, and 1060 nm respectively. The volume fraction contrast (N/C) of lipids and collagens are reported as 1.28 and 1.10 respectively. Higher absorption contrast factor (N/C) and volume fraction contrast (N/C) signifies higher concentration of lipids in normal tissues as compared to cancerous tissues, a basis for delineation. These preliminary results support the envisioned concept for non-invasive and non-carcinogenic NIR-based breast cancer diagnostic platform, which will be tested using a larger number of samples.**

Index Terms— **Near-infrared spectroscopy, Breast cancer, Biopsy tissue, Lipids, Collagens, Non-invasive.**

I. Introduction

BREAST cancer is the most common cancer accounting for almost 11.6% of all cancer incidences and 6.6% of cancer-related deaths globally in 2018 [1]. Current methods for diagnosis of breast cancer, whether in a symptomatic or a screening situation are clinical breast examination, imaging with X-ray mammography [2], [3], and ultrasound [4] as well as tissue diagnosis using needle biopsy [5]. The most common tool used for screening of breast cancer is X-ray mammography, which has been reported to have sensitivity and specificity of 77% and 97% respectively [2]. However, X-Ray mammography's sensitivity and specificity reduce significantly to 67% and 89% for dense breasts [2], [3]. Moreover, X-Ray mammography is expensive, hospital-based, involves radiation risks, and its efficacy in patients less than 40 years is not yet established [6]. Contrast-enhanced Magnetic Resonance Imaging (MRI) is reported to have the highest sensitivity between 93% and 100% [7]; but is costly, bulky, and hospital-based. Ultrasonography requires a skilled operator and gives high false positives [8] compared to X-ray mammography. Microwave based imaging modalities [9], [10] are also proposed, however large number of clinical trials are required before implementing in the field.

Women living in limited-resource settings, and geographically remote areas do not have easy access to high-quality diagnostic facilities traditionally designed for secondary or tertiary referral centres. Furthermore, with rising healthcare costs even in the developed world, there is a need for a diagnostic technique that is robust, cost-effective, and easy to interpret [11]. The concept of the diagnostic process including triple assessment at the secondary health center is illustrated in Fig. 1.

Portable systems have been studied that can delineate between normal and cancerous tissues based on electric properties [12], mechanical properties [13], combined electro-mechanical properties [14], and combined electro-thermo-mechanical properties [15] of tissues. However, these systems
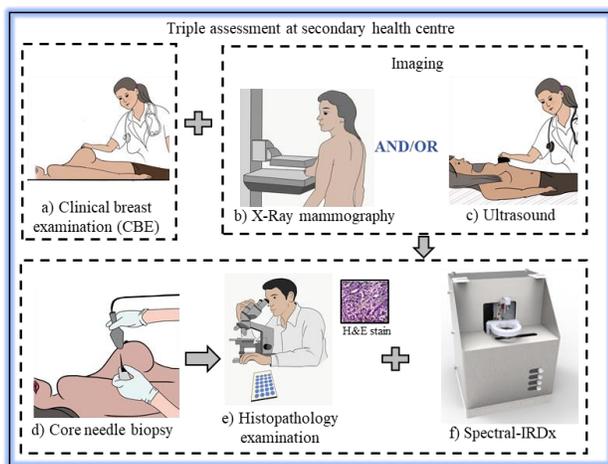
Fig. 1: Conceptual diagnostic process: The patient either visits primary health centre or screening centre for an initial assessment. If suspected, the patient visits the secondary health centre to undergo triple assessment: (a) clinical assessment by surgeons, (b) X-Ray mammography, and/or (c) Ultrasound based on the age group. In case of a suspicious lesion on a/b/c, (d) core needle biopsy is performed and examined by (e) Pathologist using standard histopathological hematoxylin-eosin (H&E) stain tests to confirm cancer, (f) The Spectral-IRDx is a platform to augment the histological analysis.

are based on physical contact with the biological tissue, which reduces the durability of the sensor and increases the expenditure in the product lifecycle. Hence, there is a need for a non-contact-based technique to delineate between normal and cancerous biopsy tissues.

Infrared spectroscopy has been used to optically characterize the different constituents of the tissue [16]. The operating wavelength of interest for near-infrared spectroscopy is between 700 nm and 1100 nm. The different constituents of tissue such as oxygenated hemoglobin, deoxygenated hemoglobin, lipids, collagens, and water have their unique characteristic absorption peaks within this range [17]. The difference in absorption of oxygenated and deoxygenated hemoglobin at 850 nm forms the basis for measuring its relative concentration. Likewise, the absorption for lipids peak at 930 nm [17]–[19] and collagens at 850 nm [17], [20], which forms the basis for measuring their relative concentration.

Additionally, the absolute absorption coefficient of lipid and collagen at 1030 nm is reported to be similar [20]. The quantification of such tissue constituents can be leveraged to demarcate and differentiate between normal and cancerous tissue. Several literatures report a higher lipid concentration in the normal tissues as compared to cancerous tissues [11], [17], [21]–[24].

The first study on exploring the absorption properties using infrared spectroscopy on breast tissue was reported in 1991 by Key et al. [25], where freshly excised tissues of women undergoing mastectomy or lumpectomy were collected. Peters et al. [26] specifically reported the delineation between normal and cancerous tissues based on the optical absorption at visible and infrared wavelengths. However, these systems used either halogen or Xeon arc lamp as a source, thereby making the system bulky, fragile, and expensive. Several new systems have now been developed as described below.

FTIR (Fourier Transform Infrared) based spectroscopy has been touted as a potential technique to analyze the molecular and structural data of the tissue and aid the clinician performing histopathology test to differentiate between normal and cancer cases [27], [28]. Kumar et al. [27] review the applicability and challenges of the Fourier Transform Infrared (FTIR) spectroscopy and imaging as a potential tool to differentiate between normal and cancer tissues based on the molecular differences. The review also documents the processes of biopsy extraction, paraffinizing, sectioning, deparaffinizing, and comparing the IR spectroscopy data with histopathology data. The challenges with FTIR spectroscopy involve uniform sample preparation protocol, misinterpretation of data due to contamination, and difficulty in handling samples with aqueous solution. FTIR spectroscopy technique has been proposed to detect various cancer including breast, lung, oral, ovarian, and colon cancer [28].

Research in infrared spectroscopy reports the use of commercial Fourier Transform Infrared (FTIR) and Raman spectroscopy tools to delineate between normal and cancerous samples [29], [30]. Fernandez et al. [29] reported the use of Perkin-Elmer's Spotlight 300 imaging system to perform the FTIR spectroscopy on breast biopsy tissue. Depciuch et al. [30] reported the use of Bruker's VERTEX 70v to perform the FTIR spectroscopy and Thermo Scientific's Smart Raman DXR to perform the Raman spectroscopy on paraffinized and deparaffinized breast biopsy tissue. However, the commercially available FTIR and Raman spectroscopy systems are not explicitly designed for positioning of breast biopsy tissue and are prohibitively expensive. The recent progress involves the development of new specialized systems based on optical-infrared hybrid spectroscopy to analyze the deparaffinized biopsy tissue samples to create the absorption maps and compare them with histopathology results [31]. However, further research needs to be carried out to emphasize the delineation between normal and cancer tissues based on the constituent's contrast ratio (N/C) which acts as a biomarker.

Diffuse Optical Imaging (DOI) technique, which involves the propagation of photons through dozens of millimetres of breast tissue to quantify the cancer biomarkers are proposed as a potential rapid diagnostic tool to aid the clinician alongside the gold standard X-Ray mammography [11], [32], [33]. Cost-effective continuous-wave operation based systems have been proposed to delineate between cancer and adjacent normal regions based on the concentration of biomarkers such as hemoglobin, lipid, collagen, and oxygen saturation [21], [34]. However, more research needs to be carried out at the tissue level to understand further the correlation between different biomarkers and tissue types (normal and cancer).

In this work, we report the development of a portable, non-contact, cost-effective, and robust NIR spectroscopy tool to delineate normal and cancerous breast biopsy tissues and quantify the contrast ratio (Normal/Cancer) of different chromophores such as lipids and collagens, based on the near-infrared spectroscopy at 850 nm, 935 nm, and 1060 nm.

II. Design And Methodology

In this section, we discuss the theoretical basis for the study, the preparation of the biopsy tissue for measurement, and the system-level description of the Spectral-IRDx, including its mechanical, electronics, and optics module.

A. Theory and mathematical model

The absorption contrast factor (ACF) for different wavelengths are experimentally quantified by measuring the normalized detected voltage (V_{dn}); Section III discusses the quantification of ACF in detail. In order to evaluate the volume fraction of different constituents of tissue, we consider the Modified Beer-Lambert Law (MBBL) for optically thick (2 ± 0.0368 mm) and optically homogenous biological tissue [35]–[38]. The near-IR absorption contribution is comparatively significant as compared to the variation of sample thickness (SD = ± 0.0368 mm) and hence the variation of sample thickness is neglected. The MBBL can be written as:

$$OD(\lambda) = -log\left(\frac{I}{I_0}\right) = \mu_a(\lambda).DPF(\lambda).d + G \qquad (1)$$

where $OD(\lambda)$ is the wavelength-dependent optical density for normal and cancer tissue samples, I_0 is the source intensity, and I is the fraction of the light intensity arriving at the detector, d is the tissue sample thickness, λ is the operating wavelength, DPF is the differential pathlength factor, G is the attenuating factor from both geometrical configuration and scattering coefficient, and $\mu_a(\lambda)$ is the wavelength-dependent total absorption coefficient of the normal and cancer tissue samples. In order to cancel the scattering factor G, differential form of MBBL is considered [35]–[38]:

$$\Delta OD(\lambda) = -log\left(\frac{I}{I_0}\right) = \Delta\mu_a(\lambda).DPF(\lambda).d \qquad (2)$$

where Δ represent the change in the variable with respect to different state or wavelength [39]. As the experiments are carried out on the deparaffinized tissues where the state of the tissue don't change with time, the Δ difference is evaluated with a respect to a reference wavelength (λ_r), also known as differential MBBL approach [40]–[42]. For simplicity, it is assumed that the geometry configuration of the tissues and the scattering losses at different wavelengths are identical.
The closed-form solution of the steady-state photon diffusion equation in an optically homogenous tissue can be written as [43]:

$$\frac{I}{I_0} = \frac{1}{4\pi D}\frac{1}{d}e^{-\mu_{eff}d} \qquad (3)$$

where D is diffusion coefficient and $\mu_{eff} = \sqrt{3\mu_a\mu_s'}$ is the effective absorption coefficient. For simplicity, the definition of DPF from Fantini et al. [36] for an infinite slab model is assumed, the DPF_{inf} can be re-written in the form of the absorption coefficient $\mu_a(\lambda)$ and reduced scattering coefficient $\mu_{s'}(\lambda)$ [37], [44]:

$$DPF_{inf}(\lambda) = \frac{1}{d}\frac{\partial}{\partial\mu_a}\left[log\left(\frac{I}{I_0}\right)\right] = \frac{\partial}{\partial\mu_a}\mu_{eff} = \frac{\sqrt{3\mu_s'(\lambda)}}{2\sqrt{\mu_a(\lambda)}} \qquad (4)$$

The reduced scattering coefficient is defined as $\mu_{s'}(\lambda) = a(\lambda/\lambda_0)^{-b}$, where a is the scattering coefficient, b is the scattering power, and λ_0 is the reference wavelength for dimensionless evaluation [18]. Additionally, the total

absorption coefficient can be written in the form of volume fraction and the absorption coefficient for each of the constituents of the sample tissue as follows [18]:

$$\mu_a(\lambda) = \sum_{i=1}^{N} f_{vi}\,\mu_{ai}(\lambda) \qquad (5)$$

where f_{vi} volume fraction of the ith constituent, $\mu_{ai}(\lambda)$ is the wavelength dependent absorption coefficient of the ith constituent, and N is the total number of constituents. The experiment with FFPE (formalin-fixed, paraffin-embedded) samples involves biopsy tissue extraction, paraffinizing, sectioning, and deparaffinization. This preparation for deparaffinized tissue is reported to be able to preserve tissue constituents such as lipids and collagens (structurally proteins) [27], [45]. In this study, we assume the IR absorption contribution dominated by two chromophores in the deparaffinized biopsy tissue, namely lipids and collagens. Hence, the total absorption can be re-written in the form of volume fraction and absorption coefficient of lipids and collagens respectively. The system operating at two wavelengths (λ_1 and λ_2) gives two equations with two unknowns($f_{v.lipid}$ and $f_{v.collagen}$):

$$\mu_a(\lambda_1) = f_{v.lipid}\mu_{a.lipid}(\lambda_1) + f_{v.collagen}\mu_{a.collagen}(\lambda_1)$$
$$\mu_a(\lambda_2) = f_{v.lipid}\mu_{a.lipid}(\lambda_2) + f_{v.collagen}\mu_{a.collagen}(\lambda_2) \qquad (6)$$

Hence, eqn. 2, 4, and 6 can be combined to be re-written as:

$$\begin{pmatrix} \frac{4(\Delta OD(\lambda_1))^2}{3\mu_s'(\lambda_1).d^2} \\ \frac{4(\Delta OD(\lambda_2))^2}{3\mu_s'(\lambda_2).d^2} \end{pmatrix} = \begin{pmatrix} \Delta\mu_{a.lipid}(\lambda_1) & \Delta\mu_{a.collagen}(\lambda_1) \\ \Delta\mu_{a.lipid}(\lambda_2) & \Delta\mu_{a.collagen}(\lambda_2) \end{pmatrix}\begin{pmatrix} f_{v.lipid} \\ f_{v.collagen} \end{pmatrix} \qquad (7)$$

where ΔOD is the difference in OD between operating wavelength (λ_1 and λ_2) and λ_r and $\Delta\mu_a$ is the difference in μ_a between operating wavelength (λ_1 and λ_2) and λ_r. The OD at the operating wavelength λ_1 and λ_2 are evaluated using I and I_0 through the electronic system as discussed in C.3 of this section. The $\mu_{a.lipid}$ and $\mu_{a.collagen}$ at wavelength λ_1, λ_2, and λ_r are known [19], [20], while $\mu_{s'}(\lambda_1)$ and $\mu_{s'}(\lambda_2)$ are assumed from [18], [46]. Solving Eqn. 7 we evaluate two unknown parameters, that is volume fraction of lipid ($f_{v.lipid}$) and volume fraction of collagen ($f_{v.collagen}$).

B. Sample Preparation

Breast biopsy tissues were obtained from Assam Medical College from patients after obtaining their informed consent. All the studies were performed in accordance with the institutional ethics committee of Assam Medical College with the ethical clearance certificate number AM/EC/1333 and of Indian Institute of Science with the ethical clearance certificate number 17-14012020. Baker et al. [47] proposed the sample preparation protocol of the FFPE tissue samples for performing FTIR spectroscopy. The same protocol with the help of an oncopathologist was further tailored to prepare the tissue sample for near-infrared spectroscopy measurements. A total of N = 10 breast tissue blocks were examined, from which N = 5 were cancer, and N = 5 were adjacent normal tissues. Each cancer-normal pair were obtained from the breast tissue of the same patient, excised in the course of their routine treatment. Formalin fixed paraffin embedded (FFPE) tissue blocks were prepared from these excised tissues. Uniform cylindrical tissue blocks of 5 mm diameter and 2 mm thickness were extracted
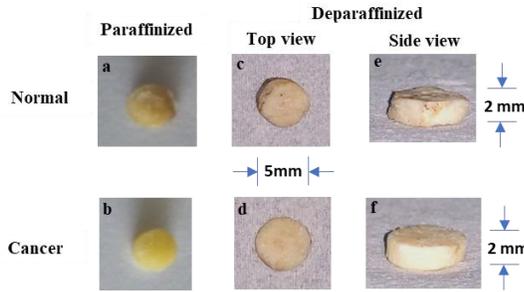
Fig. 2: Breast biopsy tissue. (a) Normal paraffinized tissue, (b) Cancerous paraffinized tissue, (c-d) Top view of normal and cancerous tissue after deparaffinization process, and (e-f) Side view of normal and cancerous tissue after deparaffinization process.

from the FFPE blocks as shown in Fig. 2a-b. These cylindrical blocks were first kept in a hot air oven at 70°C for two hours to melt the paraffin- embedded around the tissue. This paraffin is subsequently removed by dipping in Xylene for 4-5 mins followed by vortexing and centrifugation at 4500 rpm for 20 seconds to separate the paraffin wax from the tissue. The sample is then passed through a series of ethanol washes of decreasing concentrations of 100%, 95%, 70%, and 50% for 4 minutes each to remove the formalin fixation. Finally, the deparaffinized tissue sample is dipped in distilled water for 5 mins and stored in 1.5 mL PBS vials for testing and measurements, as shown in Fig. 2c-d (Top view) and Fig. 2e-f (Side view).

C. Spectral-IRDx Tool

The Spectral-IRDx comprises of mechanical, electronic, and optical modules. Each module is designed, taking into consideration the requirements for being portable, compact, cost-effective, and easy to use.

The acquisition cycle is categorized into three steps: 1) Sample loading, 2) Initialization, and 3) Measurement. The first step loads the sample holder from the original position (Fig. 3a) to the position where the operator can easily place the sample on the sample holder (Fig. 3b), taking about 90 s (T_1). The initialization step translates the sample holder and optical assembly to position the sample in the line of sight of the 850 nm LED and its corresponding OPT101(A) detector (Fig. 3c), taking about 30 s (T_2). The measurement step begins with the continuous-wave (CW) operation of 850 nm LED and acquisition of signals from the corresponding OPT101(A) detector (Fig. 3d-e). The acquisition is performed at an interval of 1 second, the full acquisition lasting for 60 seconds and a buffer of 10 seconds for switching to the next operating wavelength (Fig. 3e). After the completion of the acquisition cycle of 850 nm LED, the sample holder rotates by 45°, positioning the biopsy tissue perpendicular to the line of sight of the 935 nm LED and its corresponding OPT101(B) detector (Fig. 3f), while the same acquisition cycle is repeated for 1060 nm LED (Fig. 3g). The total measurement time (T_3) of 3 min 30 seconds consists of the acquisition cycle for the three operating wavelengths, each taking 70 s.

The acquisition time for the complete cycle, including loading sample, initialization, and measurements of single biopsy tissue, takes about 6 minutes ($T_1+T_2+T_3$).
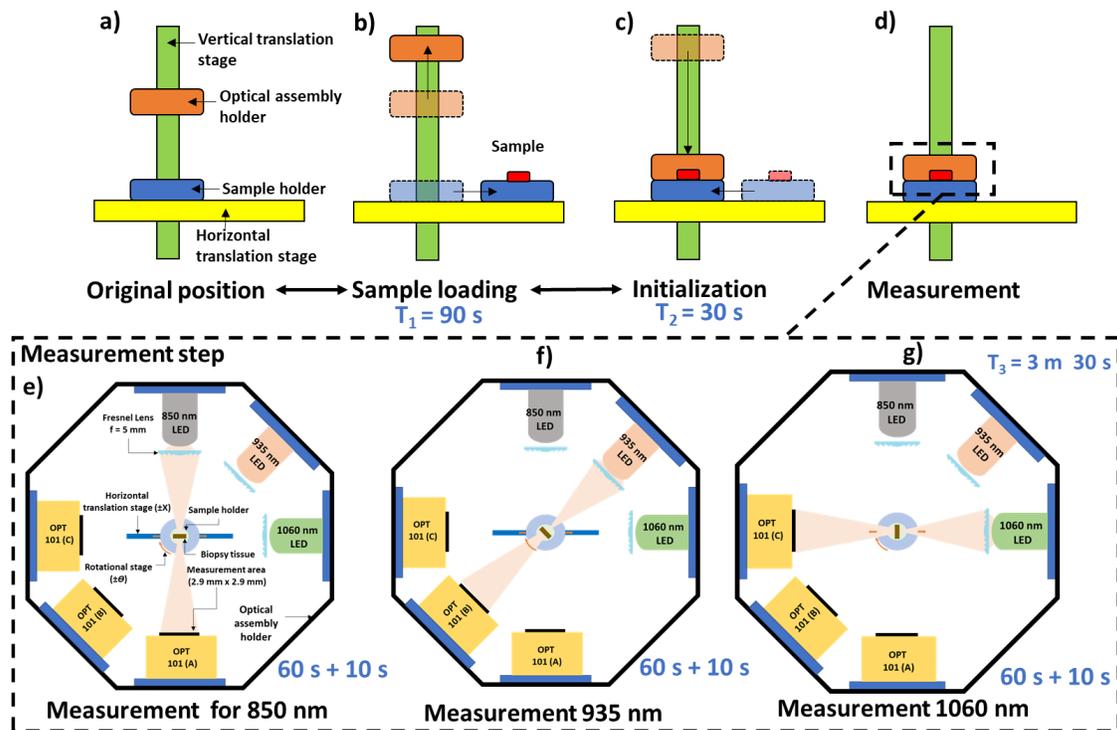

Fig. 3: Acquisition cycle of the Spectral-IRDx a) Original position, b) Loading sample (90 s): Operator places the sample on the sample holder, c) Initialization step (30 s): The sample holder along with the optical assembly holder initializes for the measurement position, d) Measurement step (3 m 30 s): Involves the measurement for the operating wavelength of 850 nm, 935 nm, and 1060 nm. The light source from e) 850 nm, f) 935 nm, and g) 1060 nm LED is focused by the Fresnel lens, which interacts with centre of the biopsy tissue and gets detected by its corresponding OPT 101(A/B/C).
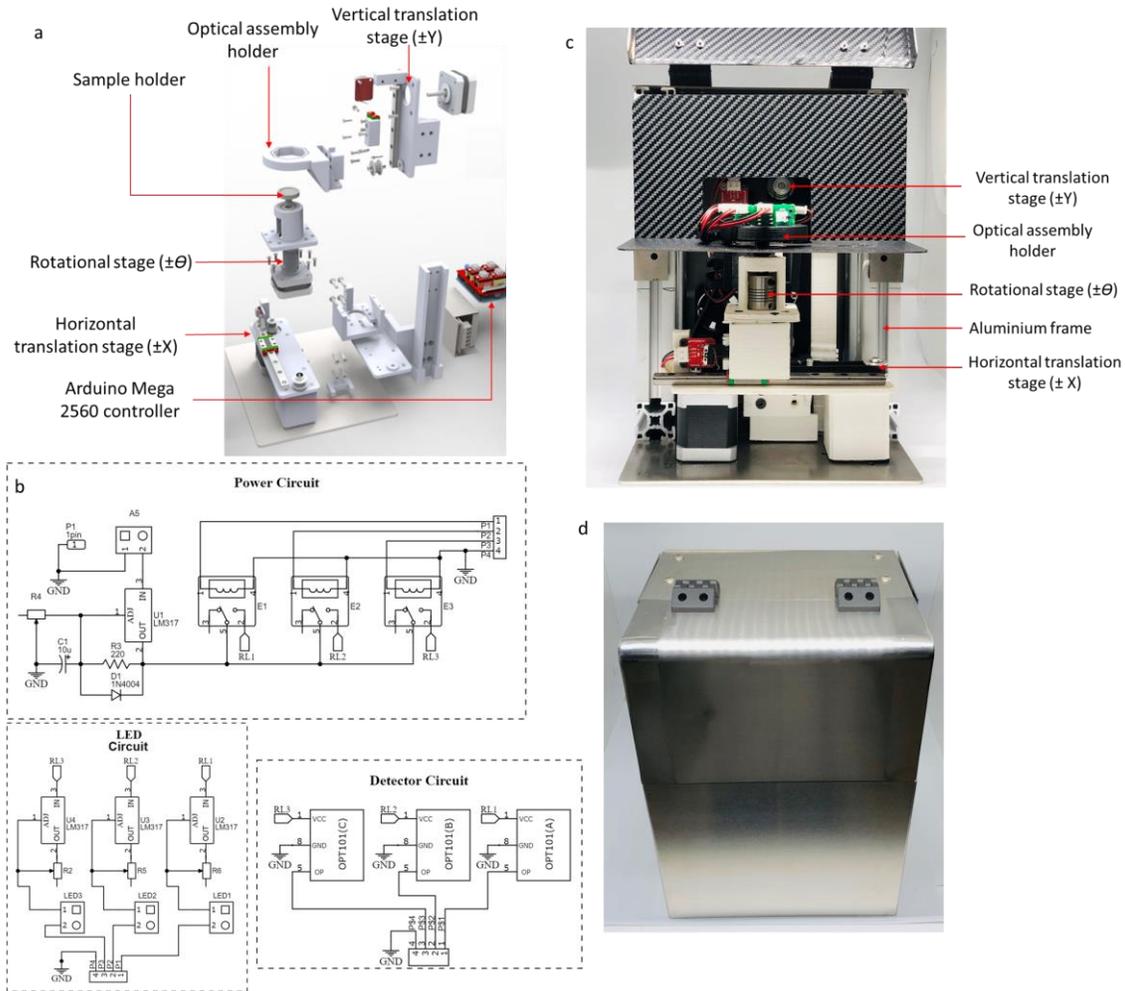
Fig. 4: (a) Exploded view of the Spectral-IRDx showcasing the space optimization to deliver compact and portable system, (b) Source-detector driving circuit. The constant current controlled signal is delivered to the 850 nm, 935 nm, and 1060 nm LEDs using the LED circuit, while a regulated 9V input is delivered to the OPT101s using the detector circuit. (c) Spectral-IRDx without the front casing. The figure showcases the three translation stages, optical assembly holder, and the mechanical components of the system, (d) The exterior view of Spectral-IRDx with the front flap closed.

1) Mechanical module

The Spectral-IRDx consists of three NEMA 17 motors that control the horizontal ($\pm$X), vertical ($\pm$Y), and rotational motion ($\pm\Theta$), as shown in the exploded view Fig. 4a. The sample holder is associated with the horizontal ($\pm$X) and rotational ($\pm\Theta$) translation stages, while the optical assembly holder is associated with the vertical ($\pm$Y) translation stage. Arduino Mega 2560 microcontroller controls the motion of the three stages using NEMA 17 motors and the crash sensors. The crash sensors are integrated at the end of the horizontal (-X direction) and vertical railing (+Y direction), to give the end stop feedback to the controller.

The black infrared absorbing layers are attached to all the inner surfaces of the system to reduce the infrared scattering. The structures of the Spectral-IRDx are 3D printed using polylactic acid (PLA) material, while the aluminum frame is used to make the system durable, as shown in Fig. 4c. The outer casing is built using stainless-steel material to reduce corrosion and increase the longevity of the system, as shown in Fig. 4d. The system is completely closed during the acquisition process to minimize the influx of ambient light that may get measured as noise. The access within the system is only permitted during the first step of loading sample, where the operator can open the flap window and place the biopsy tissue. The total weight of the system is 4.8 kg with the dimensions of 20.3 cm (length) x 18 cm (breadth) x 26.8 cm (height).

2) Electronics module

An SMPS is used to convert the AC power supply into $\pm$12 V DC supply to drive the Computer Numerical Control (CNC) shield for motor control and source-detector driver circuit. Arduino Mega 2560 controls both the CNC shield and the driver circuit. The CNC shield allows the full control over the clockwise and anti-clockwise rotation of the three NEMA 17 motors, while Arduino Mega 2560 controlling the electromagnetic relays (RL1, RL2, and RL3) are used to switch the operation between the LEDs and its corresponding OPT101s.

The driver circuit consists of two sections for driving LEDs and Silicon photodiode (Si-PD) based detectors, as shown in Fig. 4b. The LEDs and Si-PD detectors are powered using LM317 regulator IC, which regulates the SMPS output of 12 V with noise into a stable 9 V signal. The regulated 9 V signal is

then converted to a constant current supply using LM317 to drive the LEDs. The potentiometer for the LED driving circuit is adjusted to drive the LED between 5-10 mA as shown in the Eqn. 8. The R_{pot} in the LED driver circuit is set to 200 Ω to drive the LED with ~ 6.25 mA. The regulated 9 V from LM 317 is supplied directly to the Vcc connection of OPT101 for its operation.

$$I_{out} = \frac{1.25}{R_{pot}} \qquad (8)$$

3) Optics module

The optics module comprises of three source LEDs and its corresponding Si-PD OPT101 detectors. The peak wavelength of LEDs is selected to match the absorption peak of lipids and collagens. SFH4350 LED from OSRAM with a peak operating wavelength of 850 nm and spectral full width half maximum (FWHM) of 30 nm is used to match higher absorption of collagens as compared to lipids [17], [20] at 850 nm. OP133 LED from TT Electronics, having a peak operating wavelength of 935 nm and FWHM of 50 nm is used to match higher absorption of lipids at 930 nm [17], [20]. While EOLD-1060-525 LED from OSA Optolight having the peak operating wavelength of 1060 nm and FWHM of 80 nm is used to approximately match the absorption peak of both lipids and collagens at 1030 nm [17], [20]. All the LEDs are selected to deliver radiant power ~ 5 mW at the operating condition with a narrow beam angle of less than 20º to improve the capacity to interact at the core of the tissue sample. A converging beam of light is created using Fresnel lens, which has a focal length of 5 mm. The converging beam of light focus the light source towards the centre of the tissue sample and hence effective optical path length (OPL) for each of the sample measurement was assumed to be similar. This approach also reduces the placement error caused due to the shift in the placement of the tissue with each measurement. Lasers which provide narrower beamwidth and bandwidth were found to be comparatively costly at the required near-infrared wavelengths, increasing the overall development price of the system, and hence were replaced by narrow beam angle LEDs in this design. OPT101 from Texas Instrument (TI) detects this light beam using Si-PD

TABLE I
SPECTRAL-IRDX SYSTEM SPECIFICATIONS

Requirements	Specification
Power supply	230 V (50 Hz)
System dimensions	20.3 cm (length) x 18 cm (breadth) x 26.8 cm (height)
Weight	4.8 Kg
Operation type	Continuous wave (CW)
Operating wavelength	850 nm, 935 nm, and 1060 nm
Source	LED
Detector	Si Photodetector with integrated amplifier (OPT101)
Measurement time	~6 minutes
Portable	Yes
Tissue sample dimension	5 mm diameter and 2 mm thickness
Tissue type	Deparaffinized FFPE tissue samples
Controller	Arduino Mega 2560
Approximate cost	$350
Acceptable sensitivity	> 90% *
Acceptable specificity	> 90% *

* Based on references [11], [27], [28]

integrated with an inbuilt amplifier. The OPT101 is a universally accepted detector for designing biomedical optical tools due to its durability and robustness [48]. The OPT101 detects visible and near IR light (between the range of 400 nm and 1100 nm) with varying sensitivity and hence, the detected voltage at 850 nm, 935 nm, and 1060 nm from the OPT101 were normalized with a common reference value. Moreover, the OPT101 follows a linear relationship between the light intensity and the detected voltage, as shown in Eqn. 9 [48]. This linear relationship makes it possible to characterize the intensity of light as a representative detector voltage.

$$OD(\lambda) = -log\left(\frac{I}{I_0}\right) = -log\left(\frac{U}{U_0}\right) = -log(V_{dn}) \qquad (9)$$

where U and U_0 are the detected voltage with and without placing the sample respectively, and V_{dn} is the normalized detected voltage (=U/U_0). The complete system specifications of Spectral IRDx are summarized in Table I.

III. RESULTS AND DISCUSSION

Transmittance based near-infrared spectroscopy measurements were performed at the operating wavelengths of 850 nm, 935 nm, and 1060 nm for a total of 10 tissue specimens, of which 5 were known to be cancer and 5 were adjacent normal tissues. The measurement cycle of 60 datapoints was repeated for $n = 3$ times, while reporting the mean for each data point. The operating wavelengths of 935 nm and 1060 nm is used for the delineation between the normal and cancer tissues. The combination of operating wavelengths of 850 nm and 935 nm is used to quantify the volume fraction of lipid and collagen. Fig. 5a-b shows typical results from conventional hematoxylin-eosin (H&E) staining histology tests for two suspected cancerous biopsy tissue samples, later confirmed as positive by the oncopathologist.

Fig. 5c show that the normal samples on an average reported lower normalized voltage (V_{dn}) at 850 nm, 935 nm, and 1060 nm, signifying higher absorption as compared to cancerous samples. Fig. 5c compares the average of the normalized detected voltage along with the standard deviation for normal and cancerous samples for the operating wavelengths of 850 nm, 935 nm, and 1060 nm respectively. Statistical analysis were performed on the data obtained using the two-tailed unequal covariance *t-test* with Welch's correction. Data in Fig. 5c were grouped into two groups, namely normal (N = 5) and cancerous (N = 5) for the operating wavelength of 935 nm and 1060 nm. The measurements at an operating wavelength of 935 nm report slight delineation (*p*-value = 0.0038), while the measurement at 1060 nm report significant delineation (*p*-value = 0.0022) between the normal and cancer tissue samples. In comparison, the contact-based probes reported the bio-impedance analysis with a *p*-value ~ 0.00001667 [12], mechanical tissue properties analysis with a *p*-value ~ 0.000001 [13], and electro-thermo-mechanical analysis with a *p*-value ~ 0.00002 [15].

The absorption contrast factor (ACF) ratio between normal and cancer (N/C) is defined in Eqn. 10. The ACF (N/C) for the operating wavelength of 850 nm, 935 nm, and 1060 nm were reported as 1.303, 1.551, and 1.45 respectively. The higher ACF

(N/C) represents higher absorption in normal tissues as compared to cancerous tissues, a basis for delineation. Moreover, it can be inferred from ACF (N/C) for different wavelengths, the absorption at 935 nm is higher (~1.551) as compared to 850 nm (~1.303) and 1060 nm (~1.45), due to the lipid absorption peak overlapping at 930 nm.

$$
\text{Absorption contrast factor} \left(\frac{N}{C}\right) = \left(\text{Average voltage contrast} \left(\frac{N}{C}\right)\right)^{-1} = \sum_{i=1}^{5} \left(\frac{\left(\frac{U}{U_0}\right)_C}{\left(\frac{U}{U_0}\right)_N}\right)_i \quad (10)
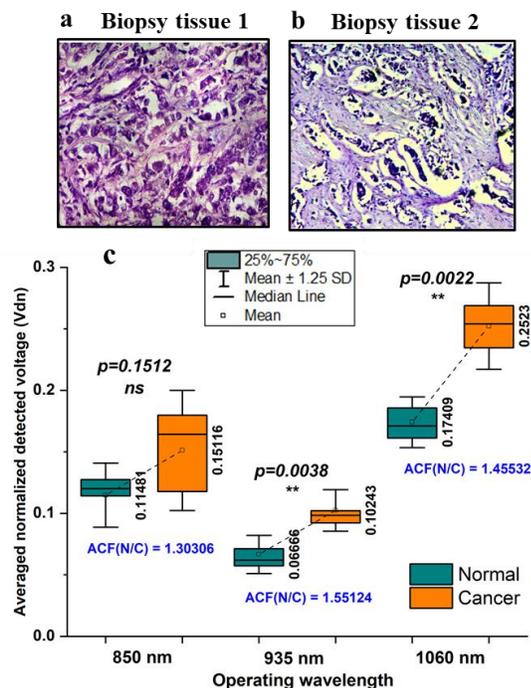$$

Fig. 5. (a-b) Histopathology tests (H&E) for two cancerous biopsy tissue samples, (c) Pooled data comparison: Average normalized detected voltage (Vdn) for normal and cancerous tissues with operating wavelength of 850 nm, 935 nm, and 1060 nm showcasing slight demarcation (*p-value = 0.0038)* at 935 nm and significant demarcation (*p*-value = 0.0022) at 1060 nm along with the Absorption contrast factor ACF(N/C).

Since the normal and cancer tissue samples are extracted from the same patient, the comparison of Vdn data points for each patient and for the operating wavelengths of 935 nm and 1060 nm is shown in Fig. 6a and Fig. 6b respectively. It can be observed from the measurement data points that the difference in Vdn between normal and cancer tissues is significantly different (p-value < 0.0001) as compared to the inter-personal distribution of each of the sample dataset. The Vdn value for the clustered normal and cancer tissues for the operating wavelengths of 935 nm and 1060 nm is reported in Fig. 6c and Fig. 6d respectively. From Fig. 6c, it is observed that for the operating wavelength of 935 nm, a significant deviation in Vdn is observed for the samples 1-2 and 4-5, while the region between the mean and third quartile of the B3 sample overlaps with the region of the mean and first quartile of the C4 sample. This may be due to the different physiological parameters such as BMI, age, and weight of the patients. From Fig. 6d, it is observed that for the operating wavelength of 1060 nm, a significant delineation (p-value = 0.0022) in Vdn is observed.

The volume fraction contrast (N/C) evaluates the ratio of biomarker (lipid or collagen) concentration of the normal tissue to the cancer tissue. To quantify the volume fraction contrast (N/C) of lipids and collagens, the operating wavelength of 850 nm and 935 nm were considered, assuming the maximum difference of absorption between the lipids and collagens. At 850 nm wavelength, the absolute absorption coefficient of collagens is significantly higher (~ 4 times) than lipids, while at 935 nm, the absolute absorption coefficient of lipids is significantly higher (~ 4 times) than collagens, as shown in Table II.. As the CW measurement cannot separate the contribution between absorption and scattering effects, the scattering variables for normal and cancer regions were assumed to be similar [21]. The reduced scattering coefficient for both normal and cancer tissues were evaluated by assuming the breast tissue with a = 31.8 cm^{-1}, b = 2.741, and reference wavelength (λ_0) as 500 nm [18], [37].

After solving Eqn. 7, using 1060 nm as a reference wavelength (λ_r), the volume fraction contrast (N/C) for lipids and collagens was obtained as 1.28 and 1.10 respectively, signifying higher lipid concentration in normal as compared to cancerous tissue, a basis for delineation. Wang et. al. [22] reported the lipid contrast factor (N/C) of ~1.5, while Cerussi et al. [23] reported a reduced lipid content (~20%) in the cancer region as compared to the adjacent normal region, leading to lipid contrast factor (N/C) ~1.25. Anderson et al. [21] reported a reduced lower lipid concentration (8%) in cancer regions as compared to the normal tissue regions, leading to lipid contrast factor (N/C) ~1.09. The study performed by Anderson et al. was able to statistically differentiate the normal and cancer regions based on the lipid contrast factor with the *p*-value = 0.002.

As these data are derived from 5 matched pairs of samples from 5 cancer patients, we recognized that they are preliminary. However, the results indicate a definite signal and are promising enough for us to expand the study to a larger sample and could potentially pave the way towards a clinical trial to test the new instrument in the clinical environment.

TABLE II
ABSORPTION COEFFICIENT FOR LIPID AND COLLAGEN

Reference	Constituents	Absorption Coefficient m^{-1}		
		850 nm	935 nm	1060 nm
Van Veen et al. [25]	Lipid	0.637	11.75	5.705
Sekar et al. [26]	Collagen	2.82	3.01	5.4

IV. COST ANALYSIS

The entire cost of the prototype including the source LED, detector OPT101, electronic components, mechanical components, fabrication, and display system amounts to ~ $350. In comparison, the commercial systems are priced >$10,000 [29], [30].

V. CONCLUSION

We have conceived, developed, and manufactured a portable system that performs infrared spectroscopy technique to aid the clinician to differentiate between the normal and cancer biopsy tissues with results comparable to the current Histopathology tests (H&E) which is considered as gold standard. The IR

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TBCAS.2020.3005971, IEEE Transactions on Biomedical Circuits and Systems

> REPLACE THIS LINE WITH YOUR PAPER IDENTIFICATION NUMBER (DOUBLE-CLICK HERE TO EDIT) <　　　8
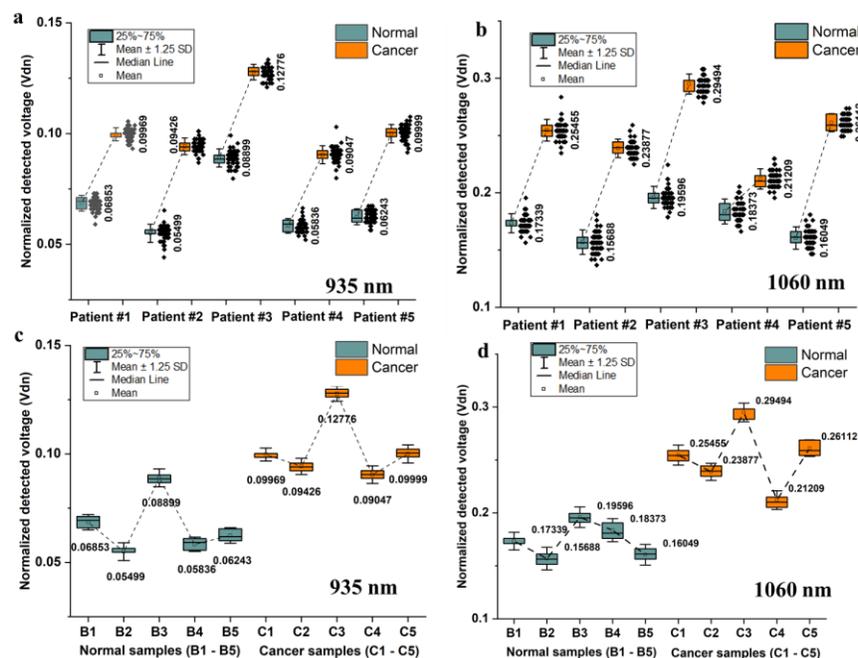
Fig. 6. Vdn comparison between cancer and normal tissues from each patient with operating wavelength of (a) 935 nm and (b) 1060 nm showcase significant statistical difference (p-value < 0.0001). Clustered sample comparison of Vdn with operating wavelength of (c) 935 nm and (d) 1060 nm, for a total of N = 10 samples (N = 5 normal, N = 5 cancer).

spectroscopic analysis was performed on cancerous and adjacent normal tissues from N = 5 patients, with a total of N = 10 tissue sample. The difference in normalized detected voltage (Vdn) at 935 nm and 1060 nm between cancer and normal tissues was reported to be statistically significant with p-values of 0.0038 and 0.0022 respectively. Additionally, the volume fraction contrast (N/C) of lipid (~1.28) reports higher lipids in normal tissue as compared to cancerous tissue. The cost is significantly lower than other commercial devices, which are also not specifically designed for positioning of a breast biopsy tissue. The Spectral-IRDx tool is designed to be portable, robust, non-contact, and easy-to-interpret. It is, therefore, suitable to be used in larger studies to formally compare its outputs with conventional histopathological diagnosis of breast cancer. This could potentially have a significant impact. Furthermore, if the Spectral-IRDx system is found to approach the accuracy of pathological examination of the biopsy specimen, the principles could be used for a device for in-vivo use in patients suspected of breast cancer through non-invasive imaging of the breast.

VI. AUTHORS' CONTRIBUTIONS

Conceptualization: UP, AV, GG, PB and HJP
Tool design and fabrication: SR, UP, AV, and HJP
Experiments: UP, AV, SS, GA, 0047G, and HJP
Analysis: UP, AV, MV, GG, and HJP
Interpretation: UP, AV, MV, GG, JSV, and HJP
Funding acquisition: HJP
Investigation: UP, AV, MV, GG, JSV, and HJP
Methodology: UP, AV, DB, GG, PB, JSV, and HJP
Project Administration: GG, JSV, and HJP
Supervision: MV, VK, GG, JSV, and HJP
Validation: MV, VK, DB, and GG

Visualization: UP, AV, and HJP
Manuscript Preparation: UP, AV, MV, JSV, and HJP

VIII. REFERENCES

[1] The Global Cancer Observatory, "Age standardized (World) incidence rates, breast, all ages," 2018.

[2] P. Skaane, S. Hofvind, and A. Skjennald, "Randomized trial of screen-film versus full-field digital mammography with soft-copy reading in population-based screening program: Follow-up and final results of Oslo II study," *Radiology*, vol. 244, no. 3, pp. 708–717, Sep. 2007.

[3] P. A. Carney *et al.*, "Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography," *Ann. Intern. Med.*, vol. 138, no. 3, pp. 168–175, Feb. 2003.

[4] P. B. Gordon, "Ultrasound for breast cancer screening and staging.," *Radiol. Clin. North Am.*, vol. 40, no. 3, pp. 431–41, May 2002.

[5] R. J. Jackman, K. W. Nowels, M. J. Shepard, S. I. Finkelstein, and F. A. Marzoni, "Stereotaxic large-core needle biopsy of 450 nonpalpable breast lesions with surgical correlation in lesions with cancer or atypical hyperplasia," *Radiology*, vol. 193, no. 1, pp. 91–95, 1994.

[6] D. L. Miglioretti *et al.*, "Radiation-Induced Breast Cancer Incidence and Mortality From Digital Mammography Screening: A Modeling Study.," *Ann. Intern. Med.*, vol. 164, no. 4, pp. 205–14, Feb. 2016.

[7] S. J. Lord *et al.*, "A systematic review of the effectiveness of magnetic resonance imaging (MRI) as an addition to mammography

and ultrasound in screening young women at high risk of breast cancer.," *Eur. J. Cancer*, vol. 43, no. 13, pp. 1905–17, Sep. 2007.

[8] M. Al-Foheidi, M. M. Al-Mansour, and E. M. Ibrahim, "Breast cancer screening: Review of benefits and harms, and recommendations for developing and low-income countries," *Med. Oncol.*, vol. 30, no. 2, 2013.

[9] A. Santorelli, E. Porter, E. Kang, T. Piske, M. Popović, and J. D. Schwartz, "A Time-Domain Microwave System for Breast Cancer Detection Using a Flexible Circuit Board," *IEEE Trans. Instrum. Meas.*, vol. 64, no. 11, pp. 2986–2994, Nov. 2015.

[10] M. R. Casu *et al.*, "A COTS-Based Microwave Imaging System for Breast-Cancer Detection," *IEEE Trans. Biomed. Circuits Syst.*, vol. 11, no. 4, pp. 804–814, Aug. 2017.

[11] U. M. Pal *et al.*, "Optical spectroscopy-based imaging techniques for the diagnosis of breast cancer: A novel approach," *Appl. Spectrosc. Rev.*, pp. 1–27, Apr. 2020.

[12] H. J. Pandya *et al.*, "Towards an automated MEMS-based characterization of benign and cancerous breast tissue using bioimpedance measurements," *Sensors Actuators, B Chem.*, vol. 199, pp. 259–268, 2014.

[13] H. J. Pandya, W. Chen, L. A. Goodell, D. J. Foran, and J. P. Desai, "Mechanical phenotyping of breast cancer using MEMS: A method to demarcate benign and cancerous breast tissues," *Lab Chip*, vol. 14, no. 23, pp. 4523–4532, Dec. 2014.

[14] H. J. Pandya, K. Park, and J. P. Desai, "Design and fabrication of a flexible MEMS-based electro-mechanical sensor array for breast cancer diagnosis," *J. Micromechanics Microengineering*, vol. 25, no. 7, Jul. 2015.

[15] H. J. Pandya, K. Park, W. Chen, L. A. Goodell, D. J. Foran, and J. P. Desai, "Toward a Portable Cancer Diagnostic Tool Using a Disposable MEMS-Based Biochip," *IEEE Trans. Biomed. Eng.*, vol. 63, no. 7, pp. 1347–1353, Jul. 2016.

[16] G. Bellisola and C. Sorio, "Infrared spectroscopy and microscopy in cancer research and diagnosis," *American Journal of Cancer Research*, vol. 2, no. 1. E-Century Publishing Corporation, pp. 1–21, 2012.

[17] P. Taroni *et al.*, "Non-invasive optical estimate of tissue composition to differentiate malignant from benign breast lesions: A pilot study," *Sci. Rep.*, vol. 7, no. January, pp. 1–12, 2017.

[18] S. L. Jacques, "Optical properties of biological tissues: A review," *Phys. Med. Biol.*, vol. 58, no. 11, 2013.

[19] R. L. P. van Veen, H. J. C. M. Sterenborg, A. Pifferi, A. Torricelli, and R. Cubeddu, "Determination of VIS- NIR absorption coefficients of mammalian fat, with time- and spatially resolved diffuse reflectance and transmission spectroscopy," in *OSA Annual BIOMED Topical Meeting posted on (http://omlc.ogi.egu/spectra/fat/ by S A Prahl)*, 2004.

[20] S. K. V. Sekar *et al.*, "Diffuse optical characterization of collagen absorption from 500 to 1700 nm," *J. Biomed. Opt.*, vol. 22, no. 1, p. 015006, 2017.

[21] P. G. Anderson *et al.*, "Broadband optical mammography: Chromophore concentration and hemoglobin saturation contrast in breast cancer," *PLoS One*, vol. 10, no. 3, pp. 1–23, 2015.

[22] J. Wang *et al.*, "In vivo quantitative imaging of normal and cancerous breast tissue using broadband diffuse optical tomography," *Med. Phys.*, vol. 37, no. 7, pp. 3715–3724, 2010.

[23] A. Cerussi, N. Shah, D. Hsiang, A. Durkin, J. Butler, and B. J. Tromberg, "In vivo absorption, scattering, and physiologic properties of 58 malignant breast tumors determined by broadband diffuse optical spectroscopy," *J. Biomed. Opt.*, vol. 11, no. 4, p. 044005, 2006.

[24] A. S. Sergio Fantini, "Near-Infrared Optical Mammography for Breast Cancer Detection with Intrinsic Contrast," *Ann Biomed Eng.*, vol. 40, no. 2, pp. 398–407, 2013.

[25] H. Key, E. R. Davies, V. G. Peters, D. R. Wyman, M. R. Arnfield, and R. P. Mathew, "Optical attenuation characteristics of breast tissues at visible and near-infrared wavelengths and near-infrared wavelengths," *Phys. Med. Biol.*, vol. 36, no. 5, pp. 579–590, 1991.

[26] V. G. Peters, D. R. Wyman, M. S. Patterson, and G. L. Frank, "Optical properties of normal and diseased human breast tissues in the visible and near infrared," *Phys. Med. Biol.*, vol. 35, no. 9, pp. 1317–1334, 1990.

[27] S. Kumar, A. Srinivasan, and F. Nikolajeff, "Role of Infrared Spectroscopy and Imaging in Cancer Diagnosis," *Curr. Med. Chem.*, vol. 25, no. 9, pp. 1055–1072, 2018.

[28] K.-Y. Su and W.-L. Lee, "Fourier Transform Infrared Spectroscopy as a Cancer Screening and Diagnostic Tool: A Review and Prospects," *Cancers (Basel).*, vol. 12, no. 1, p. 115, Jan. 2020.

[29] D. C. Fernandez, R. Bhargava, S. M. Hewitt, and I. W. Levin, "Infrared spectroscopic imaging for histopathologic recognition," *Nat. Biotechnol.*, vol. 23, no. 4, pp. 469–474, 2005.

[30] J. Depciuch *et al.*, "Comparing paraffined and deparaffinized breast cancer tissue samples and an analysis of Raman spectroscopy and infrared methods," *Infrared Phys. Technol.*, vol. 76, no. May 2018, pp. 217–226, 2016.

[31] M. Schnell *et al.*, "All-digital histopathology by infrared-optical hybrid microscopy," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 117, no. 7, pp. 3388–3396, Feb. 2020.

[32] B. J. Tromberg, B. W. Pogue, K. D. Paulsen, A. G. Yodh, D. A. Boas, and A. E. Cerussi, "Assessing the future of diffuse optical imaging technologies for breast cancer management.," *Med. Phys.*, vol. 35, no. 6, pp. 2443–51, Jun. 2008.

[33] C. G. Hadjipanayis, H. Jiang, D. W. Roberts, and L. Yang, "Current and future clinical applications for optical imaging of cancer: From intraoperative surgical guidance to cancer screening," *Semin. Oncol.*, vol. 38, no. 1, pp. 109–118, 2011.

[34] P. G. Anderson *et al.*, "Optical mammography: bilateral breast symmetry in hemoglobin saturation maps," *J. Biomed. Opt.*, vol. 21, no. 10, p. 101403, 2016.

[35] A. Sassaroli and S. Fantini, "Comment on the modified Beer-Lambert law for scattering media," *Phys. Med. Biol.*, vol. 49, no. 14, pp. 255–257, 2004.

[36] S. Fantini *et al.*, "Non-invasive optical monitoring of the newborn piglet brain using continuous-wave and frequency-domain spectroscopy," *Phys. Med. Biol.*, vol. 44, no. 6, pp. 1543–1563, 1999.

[37] M. Bhatt, K. R. Ayyalasomayajula, and P. K. Yalavarthy, "Generalized Beer–Lambert model for near-infrared light propagation in thick biological tissues," *J. Biomed. Opt.*, vol. 21, no. 7, p. 076012, 2016.

[38] L. Kocsis, P. Herman, and A. Eke, "The modified Beer-Lambert law revisited," *Phys. Med. Biol.*, vol. 51, no. 5, 2006.

[39] F. Scholkmann *et al.*, "A review on continuous wave functional near-infrared spectroscopy and imaging instrumentation and methodology," *Neuroimage*, vol. 85, pp. 6–27, 2014.

[40] P. B. Benni *et al.*, "Validation of the CAS neonatal NIRS system by monitoring VV-ECMO patients: Preliminary results," *Adv. Exp. Med. Biol.*, vol. 566, pp. 195–201, 2005.

[41] Y. Hoshi, O. Hazeki, Y. Kakihana, and M. Tamura, "Redox behavior of cytochrome oxidase in the rat brain measured by near-infrared spectroscopy," *J. Appl. Physiol.*, vol. 83, no. 6, pp. 1842–1848, 1997.

[42] Y. Hashimoto, K. Watanabe, K. Yumoto, Y. Sekine, H.-K. Kim, and T. Otoshi, "Near Infrared Quadruple Wavelength Spectrophotometry of the Rat Head," *J. Japan Soc. Air Pollut.*, vol. 24, no. 1, pp. 45–51, 1989.

[43] A. Zhang, D. Piao, C. F. Bunting, and B. W. Pogue, "Photon diffusion in a homogeneous medium bounded externally or internally by an infinitely long circular cylindrical applicator I Steady-state theory," *J. Opt. Soc. Am. A*, vol. 27, no. 3, p. 648, Mar. 2010.

[44] A. M. Chiarelli, D. Perpetuini, C. Filippini, D. Cardone, and A. Merla, "Differential pathlength factor in continuous wave functional near-infrared spectroscopy: reducing hemoglobin's cross talk in high-density recordings," *Neurophotonics*, vol. 6, no. 03, p. 1, 2019.

[45] H. J. Butler *et al.*, "Shining a light on clinical spectroscopy: Translation of diagnostic IR, 2D-IR and Raman spectroscopy towards the clinic," *Clin. Spectrosc.*, vol. 1, p. 100003, Dec. 2019.

[46] J. L. Sandell and T. C. Zhu, "A review of in-vivo optical properties of human tissues and its impact on PDT," *J. Biophotonics*, vol. 4, no. 11–12, pp. 773–787, Nov. 2011.

[47] M. J. Baker *et al.*, "Using Fourier transform IR spectroscopy to analyze biological materials," *Nat. Protoc.*, vol. 9, no. 8, pp. 1771–1791, Jul. 2014.

[48] T. Li, F. Zhong, B. Pan, Z. Li, C. Huang, and Z. Deng, "A brief review of OPT101 sensor application in near-infrared spectroscopy instrumentation for intensive care unit clinics," *Sensors (Switzerland)*, vol. 17, no. 8, pp. 1–13, 2017.