

## تشخیص سرطان پستان با استفاده از طیف‌سنجی رامان

زهره دهقانی بیدگلی<sup>۱</sup>  
 محمدحسین میران بیگی<sup>۲</sup>  
 رسول ملک‌فر<sup>۳</sup>  
 احسان‌اله کبیر<sup>۴</sup>  
 طاهره خامه‌چیان<sup>۵</sup>

<sup>۱</sup>گروه مهندسی پزشکی، دانشکده مهندسی برق، دانشگاه تربیت مدرس، تهران، ایران  
<sup>۲</sup>گروه مهندسی پزشکی، دانشکده مهندسی برق، دانشگاه تربیت مدرس، تهران، ایران  
<sup>۳</sup>گروه فیزیک، دانشکده علوم پایه، دانشگاه تربیت مدرس، تهران، ایران  
<sup>۴</sup>گروه الکترونیک، دانشکده مهندسی برق، دانشگاه تربیت مدرس، تهران، ایران  
<sup>۵</sup>گروه پاتولوژی، دانشگاه علوم پزشکی کاشان، کاشان، ایران

نویسنده مسئول: محمدحسین میران بیگی، تلفن: ۰۹۱۱۱۰۳۳۸۱  
 نشانی الکترونیک: miran@focus-me.com

### چکیده

**مقدمه:** امروزه، سرطان پستان یکی از شایع‌ترین انواع سرطان می‌باشد. روشی که همواره جهت تشخیص نهایی به‌کار می‌رود، آزمایش پاتولوژی است که از مشکلاتی همچون تهاجمی بودن، زمان پاسخ طولانی و نتایج وابسته به فرد رنج می‌برد. لذا، امروزه استفاده از تکنیک‌های نوری از جمله طیف‌سنجی رامان در تشخیص این بیماری بسیار مورد توجه قرار گرفته است.

**روش بررسی:** در مطالعه حاضر استفاده از روش طیف‌سنجی رامان در تشخیص سرطان پستان در سه سطح نرمال، تومور خوش‌خیم و بدخیم (سرطان) مورد بررسی قرار گرفته است. سیستم طیف‌سنجی رامان به‌کاررفته در این مطالعه از لیزری با طول موج ۵۳۲ نانومتر استفاده می‌نماید. با استفاده از پردازش طیف‌های رامان با تبدیل موجک (جهت حذف نویز و فلورسانس پس‌زمینه) و نرمال‌سازی SNV و سپس استخراج دو مؤلفه اصلی اول طیف‌ها و استفاده از طبقه‌بندهای SVM و MLP به تمایز سه گروه سرطانی، نرمال و خوش‌خیم اقدام نموده‌ایم.

**یافته‌ها:** با استفاده از پردازش‌های ذکرشده و طبقه‌بندی روی دو مؤلفه اصلی نخست طیف‌ها توانستیم به صحت ۹۵/۳ درصد در طبقه‌بندی سه گروه سرطانی، نرمال و خوش‌خیم به‌ترتیب با حساسیت ۹۳/۷ درصد، ۸۵/۷ درصد و ۱۰۰ درصد و صحت پیش‌بینی ۸۸/۲ درصد، ۹۲/۳ درصد و ۱۰۰ درصد دست یابیم.

**نتیجه‌گیری:** این مطالعه پتانسیل طیف‌سنجی رامان را در تشخیص سرطان پستان با صحت بالا نشان می‌دهد.

**واژه‌های کلیدی:** طیف‌سنجی رامان، سرطان پستان، تبدیل موجک، تحلیل مؤلفه اصلی

### مقدمه

امروزه، یکی از بیماری‌هایی که جوامع بشری را با مشکلات فراوانی در بخش بهداشت و درمان، مددکاری و بهزیستی مواجه ساخته است، انواع مختلف سرطان‌ها است. سرطان پستان یکی از شایع‌ترین سرطان‌ها در بین خانم‌ها و مسئول مرگ‌ومیر هزاران نفر می‌باشد. تشخیص زودهنگام این بیماری کمک شایانی به درمان بیمار خواهد کرد. با توجه به این امر مطالعات بسیاری در جهت تشخیص و درمان هرچه سریع‌تر این بیماری انجام شده است. اما، روشی که همواره جهت تشخیص نهایی به‌کار می‌رود، آزمایش پاتولوژی است. این روش با وجود توانایی‌هایی که در زمینه تشخیص انواع بیماری‌ها دارد، دارای مشکلاتی همچون تهاجمی بودن (نیاز به نمونه‌برداری)، زمان پاسخ نسبتاً طولانی و نتایج وابسته به فرد که به تجربه و تخصص پاتولوژیست بستگی دارد، می‌باشد [۱]. لذا در سال‌های گذشته، تکنیک‌های مختلفی برای حل این مشکلات پیشنهاد شده است. در این میان روش‌های نوری از آنجایی که می‌توانند به‌صورت غیر تهاجمی و غیر مخرب به‌کار گرفته شوند، بسیار مورد توجه قرار گرفته‌اند. از جمله این روش‌ها که در زمینه تشخیص تغییرات مولکولی و در نتیجه تشخیص سریع و صحیح ضایعات بافتی قابلیت بسیاری از خود

نشان داده است، می‌توان به طیف‌سنجی رامان اشاره کرد. تا به امروز مطالعات بسیاری با استفاده از طیف‌سنجی رامان در تشخیص انواع سرطان‌ها صورت گرفته است [۲-۵]. این مطالعات طیف وسیعی از بافت‌ها را همچون پوست [۶]، پستان [۴، ۷ و ۸]، بافت‌های دستگاه گوارش [۴]، مغز [۳]، کلیه [۲]، حلق و بینی [۵]، رحم و دهانه رحم [۹ و ۱۰] شامل می‌شود.

طیف‌سنجی رامان یک روش طیف‌سنجی براساس پراکندگی غیر الاستیک نور تک‌رنگ است که معمولاً از یک منبع نور لیزر تأمین می‌شود [۱۰ و ۱۱]. هنگامی که نور تک‌رنگ لیزر به یک نمونه برخورد می‌کند، بخشی از نور پراکنده می‌شود. در نور پراکنده‌شده علاوه بر فرکانس نور فرودی (پراکندگی رایلی) تابش‌هایی در فرکانس متفاوت نیز مشاهده می‌شود که پراکندگی رامان نامیده می‌شود. میزان اختلاف فرکانس نور تابشی و پراکنده‌شده به فرکانس نوسان پیوندهای اتمی موجود در ماده تحت مطالعه بستگی دارد. لذا، می‌توان با ثبت میزان پراکندگی در هر فرکانس به ساختار مولکولی ماده تحت مطالعه پی برد. بدین ترتیب طیف رامان می‌تواند به‌عنوان اثر انگشت نمونه به‌کار گرفته شود و از این

تکنیک در شناسایی مواد مختلف از جمله نمونه‌های بیولوژیک استفاده نمود.

از نظر بافت‌شناسی، نئوپلاسم‌ها (تومورها) با تجمع و بی‌نظمی سلولی مشخص می‌شوند اما، این ویژگی‌های عمومی سلول‌های نئوپلاستیک موجب تغییرات خاص در مقادیر یا ترکیب اسیدنوکلئیک، پروتئین، چربی و کربوهیدرات در بافت نیز می‌شود. به‌طور خلاصه می‌توان گفت چندین نشانگر مولکولی در غشا، سیتوپلاسم، هسته و فضای بین سلولی وجود دارند که می‌توانند شاخص نئوپلازی باشند. مولکول‌های نشانگر شامل پروتئین‌ها، چربی‌ها و اسیدهای نوکلئیک هستند و تغییر در خود نشانگر یا تعامل آن با مولکول‌های دیگر مورد انتظار است. تغییرات سلولی و مولکولی که در سرطان رخ می‌دهند، منجر به تفاوت در طیف‌های رامان به‌دست‌آمده از بافت‌های سالم و سرطانی می‌شود [۱۲ و ۱۳].

هدف از این مطالعه تشخیص سرطان پستان از بافت نرمال و تومور خوش‌خیم با استفاده از طیف‌سنجی رامان با صحت بالا می‌باشد لذا، در این مقاله به بررسی سرطان پستان در سه سطح نرمال، تومور خوش‌خیم و سرطانی با استفاده از طیف‌سنجی رامان پرداخته شده است.

### روش بررسی

تعداد ۳ نمونه سرطانی، ۳ نمونه نرمال و ۵ نمونه خوش‌خیم از بافت پستان از آزمایشگاه پاتولوژی بیمارستان شهید بهشتی کاشان به امانت گرفته شدند و در آزمایشگاه رامان پس از خروج از محلول فرمالین تحت طیف‌سنجی رامان قرار گرفتند. از هر نمونه تعداد ۳ تا ۶ طیف از نقاط مختلف آن اخذ شد و در مجموع تعداد ۱۶، ۱۴ و ۳۴ طیف به‌ترتیب از نمونه‌های سرطانی، نرمال و خوش‌خیم اخذ گردید و یک مجموعه شامل ۶۴ طیف تشکیل داده‌اند. طیف‌های به‌دست‌آمده پس از تبدیل به فرمت mat. در نرم‌افزار متلب مورد پردازش قرار گرفتند.

طیف‌های رامان با استفاده از سیستم Thermo Nicolet Almega واقع در آزمایشگاه فیزیک رامان دانشگاه تربیت مدرس اخذ شدند. این سیستم از هارمونیک دوم لیزر Nd:YLF با طول موج ۵۳۲ نانومتر استفاده می‌کند. توان لیزر به‌کاررفته ۵۰ میلی‌وات و شعاع لیزر در فاصله کانونی ۱۰۰ μm می‌باشد. طیف‌ها با ۳۲ بار اسکن در محدوده طیفی ۴۰۰۰-۴۰۰ cm<sup>-1</sup> با رزولوشن ۴ cm<sup>-1</sup> اخذ شده‌اند.

### پردازش

رزولوشن طیفی طیف‌های به‌دست‌آمده با استفاده از درون‌یابی اسپلاین به 1 cm<sup>-1</sup> ارتقاء یافت و سپس با استفاده از تبدیل موجک به حذف نویز و فلورسانس پس‌زمینه اقدام نمودیم.

بیان سلسله‌مراتبی مجموعه داده امکان آنالیز چنددقتی را فراهم می‌کند که تحت عنوان تبدیل موجک گسسته (DWT) شناخته می‌شود [۱۴].

K تابع ویولت  $\Phi_{kj}(x)$  که بیان فضای پایه را در سطح j تشکیل می‌دهد، از یک تابع منفرد به‌نام ویولت مادر  $\Psi(x)$  از طریق رابطه ۱ تشکیل می‌شود:

$$\phi_{kj}(x) = 2^{j/2} \Psi(2^j x - k) \quad (1)$$

تقریب yn و جزئیات zn در سطح j از فرمول‌های ۲ و ۳ محاسبه می‌شود:

$$z_n = \sum_{m=1}^{M_{high}} d_m f_{n-m} \quad (2)$$

$$y_n = \sum_{m=1}^{M_{low}} c_m f_{n-m} \quad (3)$$

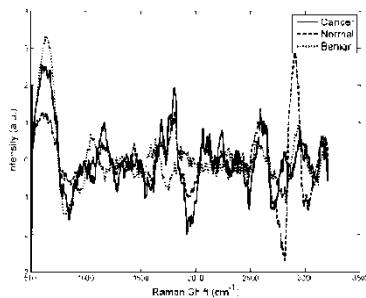
که  $n=2ji$  با  $i=1,2,\dots,N/2j$  ،  $cm$  ها  $M_{low}$  تا ضرایب فیلتر تجزیه پایین و  $dm$  ها  $M_{high}$  تا ضریب فیلتر تجزیه بالا هستند. داده‌ها با استفاده از جعبه‌ابزار ویولت نرم‌افزار Matlab 7.7 پردازش شدند.

با استفاده از بخش‌های تجزیه‌شده سیگنال می‌تواند در فرآیند معکوس بازسازی شود که به‌نام تبدیل معکوس ویولت گسسته (IDWT) شناخته می‌شود. اگر آخرین مؤلفه تقریب در فرآیند IDWT وارد نشود، هموارترین بخش سیگنال حذف می‌شود. در مورد طیف رامان، این مؤلفه عموماً بخش ناشی از فلورسانس زمینه است. به‌طور مشابه با حذف جزئیات مرتبه بالا نویز تا حدی حذف می‌شود.

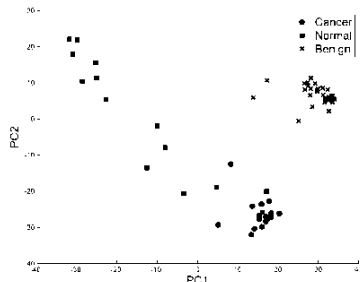
در تبدیل موجک با موجک مادر db4، تجزیه به ۸ سطح، حذف تقریبات (به‌عنوان فلورسانس) و استفاده از پنج سطح جزئیات در بازسازی طیف‌ها به حذف فلورسانس پس‌زمینه و نویز اقدام نمودیم. در نهایت، نرمال‌سازی SNV باتوجه به رابطه ۱ روی طیف‌ها صورت گرفته است.

$$y_{SNV} = \frac{y - \text{mean}(y)}{\text{std}(y)} \quad (1)$$

در رابطه ۱  $y_{SNV}$  طیف نرمال‌شده،  $y$  طیف نرمال‌نشده،  $\text{mean}(y)$  میانگین تمام نقاط طیفی و  $\text{std}(y)$  انحراف معیار



شکل ۲: میانگین طیف‌های به‌دست‌آمده پس از پیش‌پردازش در سه گروه سرطانی، نرمال و خوش‌خیم



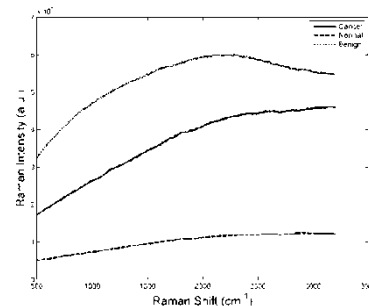
شکل ۳: تصویر پراکندگی مؤلفه‌های اول و دوم طیف‌ها پس از اعمال پیش‌پردازش در سه گروه سرطانی، نرمال و خوش‌خیم

پس از استخراج مؤلفه‌های اصلی، دو مؤلفه اصلی اول و دوم به‌منظور طبقه‌بندی به‌کار گرفته شد و با استفاده از طبقه‌بندهای SVM<sup>۱</sup> و شبکه عصبی پرسپترون چندلایه<sup>۲</sup> (MLP) به طبقه‌بندی نمونه‌ها با استفاده از اعتبارسنجی ده‌لایه<sup>۳</sup> در سه گروه اقدام گردید. در طبقه‌بندی SVM از سه طبقه‌بند SVM دوکلاسه به‌روش یک کلاس دربرابر بقیه<sup>۴</sup> با هسته چندجمله‌ای مرتبه سه استفاده گردید. ماتریس سردرگمی به‌دست‌آمده توسط این طبقه‌بند در جدول نشان داده شده است. در طبقه‌بندی با استفاده از MLP از یک شبکه عصبی دارای یک لایه میانی با هفت نرون استفاده شد. این شبکه همچنین دارای دو نرون ورودی برای دو ویژگی (دو مؤلفه اول طیف‌ها) و سه نرون خروجی به‌ازای سه گروه (سرطانی و خوش‌خیم) می‌باشد. ماتریس سردرگمی طبقه‌بندی MLP نیز دقیقاً مشابه مورد SVM به‌دست آمده است که در جدول نشان داده شده است. در این جدول ناحیه نشان داده شده با زمینه سفید درواقع همان ماتریس سردرگمی است که نشان‌دهنده توزیع تشخیص‌های درست و نادرست روی سه گروه است. در این ماتریس گروه‌های واقعی روی

نقاط طیفی می‌باشد.

سه نمونه از طیف‌های خام اندازه‌گیری در سه گروه در شکل ۱ نشان داده شده است. طیف‌های به‌دست‌آمده پس از انجام پردازش‌های ذکرشده در شکل ۲ نشان داده شده است. در پایان، با استفاده از آنالیز PCA دو مؤلفه اول طیف‌ها استخراج گردیدند. PCA یا آنالیز مؤلفه‌های اصلی یک فرآیند ریاضی است که با استفاده از یک تبدیل متعامد، مجموعه‌ای از مشاهده‌ها از متغیرهای احتمالاً وابسته را به مجموعه‌ای از متغیرهای به‌طور خطی ناهمبسته (مؤلفه‌های اصلی) تبدیل می‌کند [۱۵]. تعداد مؤلفه‌های اصلی برابر یا کمتر از تعداد متغیرهای اولیه است. این تبدیل به‌گونه‌ای انجام می‌شود که اولین مؤلفه اصلی بیشترین واریانس ممکن (تغییرات ممکن در مجموعه داده) را دارا است. مؤلفه‌های بعدی درحالی‌که بر مؤلفه‌های قبلی عمود هستند (با مؤلفه‌های قبلی ناهمبسته هستند) به‌ترتیب دارای بیشترین واریانس ممکن می‌باشند. به‌این‌ترتیب اولین مؤلفه دارای بیشترین واریانس است و به‌ترتیب این واریانس در مؤلفه‌های بعدی کاهش می‌یابد. در صورتی که مجموعه داده دارای توزیع توأماً نرمال باشد، مؤلفه‌های اصلی مستقل خواهند بود. تحلیل مؤلفه‌های اصلی می‌تواند برای کاهش ابعاد داده مورد استفاده قرار بگیرد، به‌این‌ترتیب مؤلفه‌هایی از مجموعه داده که بیشترین تأثیر را در واریانس دارند، حفظ می‌کند.

مؤلفه‌های اول و دوم طیف‌ها در سه کلاس سرطانی، نرمال و خوش‌خیم در نمودار پراکندگی شکل ۳ نمایش داده شده است. همان‌طور که از مقایسه شکل ۱ و شکل ۲ مشاهده می‌شود، طیف‌های خام به‌شدت متأثر از فلورسانس پس‌زمینه نمونه‌ها می‌باشد که در طول موج لیزر به‌کار رفته (۵۳۲ نانومتر) بسیار قوی است درحالی‌که پس از انجام پردازش‌های مذکور پیک‌های رامان به‌خوبی آشکار شده‌اند. نمونه‌های مربوط به سه گروه در فضای دو مؤلفه اصلی به‌خوبی متمایز شده‌اند این امر در شکل ۳ به‌وضوح مشاهده می‌شود.



شکل ۴: سه نمونه از طیف‌های خام اندازه‌گیری‌شده در سه گروه سرطانی، نرمال و خوش‌خیم

<sup>1</sup> Support Vector Machine

<sup>2</sup> Multi Layer Perceptron

<sup>3</sup> 10-fold Cross Validation

<sup>4</sup> One Vs Rest

### نتیجه گیری

مطالعه حاضر نشان می‌دهد که طیف‌سنجی رامان به خوبی می‌تواند جهت تشخیص سرطان در پستان استفاده گردد. همچنین با وجود اینکه لیزر به کاررفته برای طیف‌سنجی در این مطالعه دارای طول موج ۵۳۲ نانومتر می‌باشد و نمونه‌های بیولوژیک در این طول موج فلورسانس زیادی از خود نشان می‌دهند که سیگنال رامان را می‌پوشاند، اما با استفاده از تبدیل موجک امکان دستیابی به ویژگی‌های اصلی رامان به دست آمد. آنچه در این مطالعه به دست آمد، توانسته است نتایج به دست آمده توسط آقای بروزک-پلوسکا و همکاران در مورد تشخیص سرطان پستان در شرایطی مشابه را که همانا حساسیت ۶۲ درصد، ۸۳ درصد و ۷۲ درصد به ترتیب برای گروه‌های سرطانی، نرمال و خوش خیم می‌باشد [۱۶]، تا حد زیادی بهبود بخشد.

در این مطالعه نشان داده شد که با استفاده از تکنیک طیف‌سنجی رامان می‌توان به صحت بالایی از تشخیص در درجه بندی سرطان دست یافت. این درحالی است که در اینجا برخلاف روش متداول هیستوپاتولوژی، تمام تشخیص‌ها به صورت اتوماتیک انجام می‌شود. این امر علاوه بر کاهش زمان تشخیص، احتمال بروز خطاهای انسانی و وابستگی نتایج به تجربه فردی را نیز کاهش می‌دهد. به علاوه این روش از آنجاکه مبتنی بر نمونه برداری نوری از بافت می‌باشد، می‌تواند به صورت غیرتهاجمی نیز برای بررسی‌های بافتی به کار گرفته شود.

ستون‌ها و گروه پیش‌بینی شده توسط طبقه‌بند روی سطرها نشان داده شده است. به این ترتیب عناصر روی قطر اصلی نشان‌دهنده تعداد تشخیص‌های صحیح در هر گروه می‌باشد. ردیف آخر این جدول نشان‌دهنده حساسیت تشخیص در هر گروه می‌باشد که از تقسیم تعداد نمونه‌های طبقه‌بندی شده به صورت صحیح به کل نمونه‌های آن گروه به دست می‌آید و همچنین ستون آخر جدول نشان‌دهنده صحت پیش‌بینی به ازای هر گروه می‌باشد که با تقسیم تعداد نمونه‌هایی که به درستی به یک گروه تعلق گرفته‌اند به تمام نمونه‌های تعلق یافته به آن گروه به دست می‌آید.

جدول ۱: ماتریس سردرگمی به دست آمده با استفاده از طبقه‌بندهای MLP و SVM

		Target Class			
		C	N	B	
Output Class	U	۵	۲	۰	۰/۸۸۲
	Z	۱	۱۲	۰	۰/۹۲۳
	W	۰	۰	۳۴	۱
		۰/۹۳۷	۰/۸۵۷	۱	۰/۹۵۳

همان‌طور که در جدول ۱ نشان داده شده است، با استفاده از پردازش‌های ذکر شده و طبقه‌بندی روی دو مؤلفه اصلی نخست طیف‌ها توانستیم به صحت ۹۵/۳ درصد در طبقه‌بندی سه گروه سرطانی، نرمال و خوش خیم به ترتیب با حساسیت ۹۳/۷ درصد، ۸۵/۷ درصد و ۱۰۰ درصد و صحت پیش‌بینی ۸۸/۲ درصد، ۹۲/۳ درصد و ۱۰۰ درصد دست یابیم.

### References

1. Crow P, Stone N, Kendall CA, Persad RA, Wright MPJ. Optical diagnostics in urology: current applications and future prospects. *BJU International* 2003; 92: 400-7.
2. Zhuang Z, Li N, Guo Z, Zhu M, Xiong K, Chen S. Study of molecule variations in renal tumor based on confocal micro-Raman spectroscopy. *Journal of Biomedical Optics* 2013; 18: 031103-031103.
3. Zhou Y, Liu CH, Sun Y, Pu Y, Boydston-White S, Liu Y, Alfano RR. Human brain cancer studied by resonance Raman spectroscopy. *Journal of Biomedical Optics* 2012; 17: 116021-116021.
4. Tu Q, Chang C. Diagnostic applications of Raman spectroscopy. *nanomedicine* 2012; 8: 545-58.
5. Li SX, Chen QY, Zhang YJ, Liu ZM, Xiong HL, Guo ZY, Mai HQ, Liu SH. Detection of nasopharyngeal cancer using confocal Raman spectroscopy and genetic algorithm technique. *Journal of Biomedical Optics* 2012; 17.
6. Philipsen PA, Knudsen L, Gniadecka M, Ravnbak MH, Wulf HC. Diagnosis of malignant melanoma and basal cell carcinoma by in vivo NIR-FT Raman spectroscopy is independent of skin pigmentation. *Photochemical & Photobiological Sciences* 2013; 12: 770-6.
7. Brozek-Pluska B, Musial J, Kordek R, Bailo E, Dieing T, Abramczyk H. Raman spectroscopy

- and imaging: applications in human breast cancer diagnosis. *Analyst*, 2012.
8. Hu C, Wang J, Zheng C, Xu S, Zhang H, Liang Y, Bi L, Fan Z, Han B, Xu W. Raman spectra exploring breast tissues: Comparison of principal component analysis and support vector machine-recursive feature elimination. *Medical physics* 2013; 40: 063501.
  9. Duraipandian S, Zheng W, Ng J, Low JJH, Ilancheran A, Huang Z. Near-infrared-excited confocal Raman spectroscopy advances in vivo diagnosis of cervical precancer. *Journal of Biomedical Optics* 2013; 18: 067007-067007.
  10. Lewis IR. *Handbook of raman spectroscopy*: Marcel Dekker, 2001.
  11. Twardowski J, Anzenbacher P. *Raman And IR Spectroscopy in Biology and Biochemistry*: Ellis Horwood, 1994.
  12. Mahadevan-Jansen A, Richards-Kortum R. Raman spectroscopy for the detection of cancers and precancers. *Journal of biomedical optics* 1996; 1: 31-70.
  13. Mahadevan-Jansen A, Richards-Kortum R. Raman Spectroscopy For Cancer Detection: A Review. in *Proceedings Of The 19th Int. Conf. Of The IEEE EMBS* 1997; 2722-8.
  14. Mallat S. *A wavelet tour of signal processing*: Academic Press, 1998.
  15. Svensson O, Josefson M, Langkilde FW. Reaction monitoring using Raman spectroscopy and chemometrics. *Chemometrics and Intelligent Laboratory Systems* 1999; 49: 49-66.
  16. Brozek-Pluska B, Placek I, Kurczewski K, Morawiec Z, Tazbir M, Abramczyk H. Breast Cancer Diagnostics By Raman Spectroscopy. *Journal of molecular liquids* 2008; 141: 145-8.