

یک مطالعه مقایسه‌ای برای بررسی استفاده از الگوریتم‌های ماشین بردار پشتیبان و نزدیک‌ترین نقطه همسایگی در بهبود تشخیص تومورهای ریه جهت پرتودرمانی

INC29-1243

جواد طیبی، احمد اسماعیلی ترشابی*، محمد مهدی ابراهیمی بشلی

گروه مهندسی هسته‌ای، دانشکده علوم و فناوری‌های نوین، دانشگاه تحصیلات تکمیلی صنعتی و فناوری پیشرفته، کدپستی: ۷۶۳۱۸۸۵۳۵۶، کرمان، ایران

چکیده:

آمار مرگ و میر ناشی از سرطان ریه در حال افزایش است. در اغلب کشورهای جهان، از شایع شدن آن در بین جامعه آماری سالم، نگرانی را به وجود آورده است که به یک مشکل بزرگ در میان کادر درمانی کشورها تبدیل شده است. تشخیص به موقع سرطان ریه می‌تواند تا حدود زیادی جلوی شیوع این بیماری را بگیرد. سرطان ریه از طریق تصویربرداری پزشکی ریه قابل تشخیص است. انتخاب نوع درمان مناسب با آن بیماری، یکی از اهداف تصویربرداری از ریه است. بهبود و کیفیت تصاویر سی‌تی‌اسکن برای افزایش دقت و قدرت تفکیک در محیط ریوی، از کارهای مهم پیش‌پردازش تصاویر پزشکی است. درمان سرطان ریه ضروری است و اگر درمان نشود، احتمال رخداد متاستاز زیاد می‌شود. در این مقاله سعی شده است با پردازش تصاویر پزشکی و طبقه‌بندی آن‌ها، از طریق ماشین بردار پشتیبان و نزدیک‌ترین نقطه همسایگی متناسب با تصاویر گرفته‌شده به منظور پیشرفت در روند پرتودرمانی، تصاویر پزشکی را پیش‌پردازش و طبقه‌بندی کنیم.

کلیدواژه‌ها: سرطان ریه، سی‌تی‌اسکن ریه، پردازش تصویر، طبقه‌بندی کننده‌ها، پرتودرمانی

A Comparative Study to Investigate the Use of Support Vector Machine and Nearest Neighbor Point Algorithms in Improving the Diagnosis of Lung Tumors for Radiotherapy

Javad Tayebi, Ahmad Esmaili Torshabi*, Mohammad Mahdi Ebrahimi Besheli

Faculty of Sciences and Modern Technologies, Graduate University of Advanced Technology, Postal Code: 7631885356, Kerman, Iran.

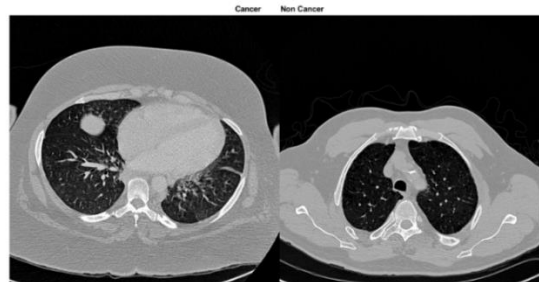
Abstract:

The number of deaths caused by lung cancer is increasing. In most countries of the world, its spread among the healthy population has caused concern. This has become a big problem among the medical staff of the countries. Early detection of lung cancer can largely prevent the spread of this disease. Lung cancer can be diagnosed through medical imaging of the lung. Choosing the right type of treatment for that disease is one of the goals of lung imaging. Improving the quality of CT scan images to increase accuracy and resolution in the pulmonary environment is one of the important preprocessing tasks of medical images. Lung cancer treatment is essential, and if it is not treated, the probability of metastasis increases. In this article, an attempt has been made to preprocess and classify medical images by processing medical images and classifying them, using the support vector machine and the nearest neighborhood point corresponding to the images taken in order to progress in the radiation therapy process.

Keywords: Lung cancer, CT lung scan, Image processing, Classifiers, Radiation therapy

۱. مقدمه

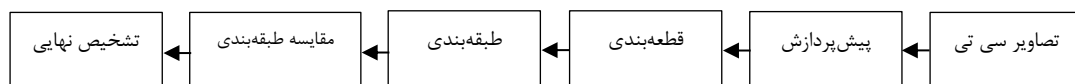
بدن ما برای تنفس، نیازمند دستگاهی به نام ریه یا شش است. محل این دستگاه مهم بدن، در بین دو سویی قلب و به صورت مخروطی شکل در این نواحی قرار گرفته است. در اکثر جانداران این عضو در اطراف قلب و در بدن انسان در حفره سینه درون سینه قرار دارد که وظیفه اصلی آن‌ها دریافت اکسیژن از هوا و منتقل کردن آن‌ها به دستگاه گردش خون است. در این محل، گلبول‌های قرمز خون اکسیژن را دریافت کرده و به بقیه سلول‌های اندام‌ها، اکسیژن می‌رساند. این عضو حیاتی در بدن انسان، مداوم با هر دم و بازدمی ارتباطی را با دنیای بیرون از بدن انسان دارد. در نتیجه مستعد بیماری خواهد شد. هر نوع بیماری و در هر اندازه‌ای، ریه انسان را تحت تأثیر خود قرار می‌دهد. سرطان ریه، به دلیل رشد بیش از حد و غیرقابل کنترل سلول‌ها در ناحیه ریه اتفاق می‌افتد. عوامل مختلفی می‌تواند در ایجاد سرطان مؤثر باشد. یکی از معروف‌ترین دلایل، آلودگی هوا و استعمال دخانیات است [1]. دود سیگار، به دلیل داشتن مواد سرطان‌زا، یکی از بسترهای مهم برای ایجاد سرطان در ریه را فراهم می‌کند. از دیگر عوامل می‌توان به عوامل ژنتیکی و گاز رادون [2] در ایجاد بیماری سرطان ریه، نام برد. علائم شخص مبتلا به سرطان ریه، از جمله سرفه‌های مداوم [3] همراه با خلط خونی، تنگی نفس [4]، کاهش وزن [5]، ضعف بدنی و تب در شخص است، که به سرطان ریه مبتلا است. بعد از تشخیص، از راه‌های معمول درمان بیماری سرطان ریه، می‌توان به شیمی‌درمانی، پرتودرمانی و جراحی اشاره کرد [6]. سرطان ریه هم در مردان و هم در زنان، در میان شایع‌ترین انواع سرطان‌ها قرار دارد [7]. اشکال مختلفی از سرطان ریه وجود دارد، که خود را به شکل‌های گوناگونی ظاهر می‌سازد، که قابلیت رشد در ریه را دارند. گاهی علامتی از خود برای شخص بیمار ندارد و این همان مسئله مهمی است، که دیر تشخیص داده می‌شود، یعنی زمانی که فرآیند رشد غده سرطانی افزایش یافته است. در ریه انواع مختلفی از توده‌ها و بافت‌ها وجود دارند، که با توجه به این که خوش خیم یعنی غیر سرطانی و یا پیش سرطانی (ندول‌ها) باشند، تقسیم بندی می‌شوند. ندول ریه، یک ندول متاستاز است، یعنی این که سلول‌های این توده سرطانی از حرکت سلول‌های توده سرطانی در اعضای دیگر پخش شده و وارد ریه شده است. ساده‌ترین مثال آن سرطان سینه است [8]، که توده سرطانی در ریه از پخش و انتقال سلول‌های سرطان سینه در قسمت‌های مختلف، وارد ریه شده است. تشخیص نوع بیماری و این که در کدام ناحیه از ریه قرار دارد، برای ما خیلی مهم است. برای اینکه اگر ما بخواهیم عمل جراحی، شیمی‌درمانی و یا پرتودرمانی انجام بدهیم، در نقطه مورد نظر باشد، و به نقاط دیگر آسیب جبران ناپذیری نزنند. این خود از مزایای بسیار مهم در تشخیص مکان بیماری در فضا و محیط ریه، به عنوان یک عضو حیاتی بدن است. برای تشخیص بیماری سرطان ریه، از تصویربرداری پزشکی استفاده می‌کنیم. در تصویربرداری پزشکی، با توجه به تصویری که از ریه‌ها، و از نمای کلی آن خواهیم داشت، تغییر و تحولی که داخل ریه در اثر بیماری ایجاد شده است را می‌توان به وضوح دید. گاهی از مواقع با توجه به نوپز و اغتشاشی که در تصویربرداری پزشکی ایجاد می‌شود، می‌توان از طریق تکنیک‌های مدرن در عصر حاضر، مانند پردازش تصاویر پزشکی [9]، تصویر خود را با وضوح بهتر و دقت بالا دید و آن را تحلیل و بررسی کرد. تفاوت یک ریه‌ی سالم با ریه ناسالم را می‌توان از تصویر زیر مشاهده کرد.



شکل ۱. ریه سمت راست، نشان‌دهنده‌ی یک ریه سالم و ریه سمت چپ، دارای یک توده‌ی سرطانی، معروف به سرطان ریه است.

۲. روش کار

برای انجام این کار، ما از محیط برنامه‌نویسی متلب (نسخه ۲۰۲۱) استفاده کردیم. در ابتدای این کار، پیش‌پردازش تصاویر را با فرمت (TIFF) انجام دادیم. در این عملیات، بهبودهایی از لحاظ شدت روشنایی و اعمال وضوح بیشتر بر روی نواحی تار داده‌های تصویر، اعمال می‌کنیم. تا امکان استخراج صحیح و دقیق‌تر اطلاعات، میسر گردد. افزایش کیفی و تنظیم شدت افزایش تباين (کنتراست)، حذف نویز و حذف مؤلفه‌های غیرضروری از مجموعه اقدامات پیش‌پردازش این تصاویر می‌باشد. بعد از عملیات پیش‌پردازش به استخراج ویژگی^۲ [10] پرداختیم. استخراج ویژگی بخشی از فرایند کاهش ابعاد است که در آن مجموعه اولیه‌ای از داده‌های خام تقسیم‌شده و به گروه‌های قابل‌کنترل کاهش می‌یابند. مهم‌ترین ویژگی این مجموعه داده‌های بزرگ این است که تعداد زیادی متغیر دارند. برای پردازش این متغیرها، به منابع محاسباتی زیادی نیاز داریم. بنابراین استخراج ویژگی با انتخاب و ترکیب متغیرها در ویژگی‌ها، به بهترین ویژگی از آن مجموعه داده‌های بزرگ، کمک می‌کند. این کار باعث می‌شود، حجم داده‌های ما به‌طور مؤثری کاهش پیدا کند. طبقه‌بندی تصاویر^۳ [11]، کاری اساسی است، که سعی می‌کند یک تصویر را به‌طور کلی درک کند. هدف این است که تصویر را با اختصاص دادن آن به یک برجسب خاص طبقه‌بندی کند. به‌طور معمول، طبقه‌بندی تصویر به تصاویری اطلاق می‌شود، که در آن‌ها تنها یک شی ظاهر می‌شود، و مورد تجزیه و تحلیل قرار می‌گیرد. بعد از طبقه‌بندی‌ها برای این‌که به داده‌های آموزش‌یافته نوع و اندازه تومور را با شکل هندسی خاصی آموزش دهیم و تصاویر دارای سرطان را از غیر سرطان تشخیص بدهد، از فراکتال‌ها^۴ [12] استفاده می‌کنیم. فراکتال‌ها اشکال هندسی هستند که به‌صورت کلی، آرایشی هندسی و تکرارشونده دارند. یعنی با چند قطعه شدن آن قسمت، مشابه قسمت دیگری است. تصاویر دارای مجموعه‌ای از پیکسل‌ها هستند. با استفاده از ماتریس وابستگی فضای سطح خاکستری^۵ [13] برای پیکسل‌ها، بازه‌های مختلف شدت نور از سیاهی مطلق تا سفیدی، مانند ماتریس‌هایی در کنار هم قرار می‌گیرند. پیکسل‌هایی را که دارای رنگ خاکستری می‌باشند، به نوعی سطح خاکستری دارند، با ماتریس وابستگی فضای سطح خاکستری، می‌توان استخراج کرد. از تبدیل موجک^۶ [14] برای فشرده‌سازی تصاویر باکیفیت خوب استفاده شده است. همچنین تبدیل موجک هم در حوزه زمان و هم در حوزه فرکانس قدرت تفکیک بالایی دارد. اصول کار و عملکرد این، تبدیل موج یک تابع قابل تجزیه را به شکل مربع و دارای وزن خاص ایجاد می‌کند. تبدیل موجک دارای قابلیت قسمت‌بندی تابع قابل تجزیه به مقیاس‌های بزرگ و کوچک است. برای استخراج ویژگی‌های دلخواه از تکنیک‌های آماری بافت^۷ [15]، استفاده می‌شود. به منظور شناسایی جهت ناحیه یا صافی، بزرگ و کوچک بودن نقاطی از اجزای تصویر با این روش انجام شد. از این روش برای تجزیه و تحلیل بافت ریه به‌منظور مرزبندی بین اجزا و همچنین مرز آن‌ها استفاده می‌کنیم. تکنیک زرنیک^۸ [16]، آخرین عملیات ما در این تحقیق است. از طریق این تکنیک، تصاویر را در نقاطی که می‌خواهیم بازسازی یا اینکه مورد تجزیه و تحلیل قرار دهیم، استفاده می‌کنیم. در دیاگرام زیر، نحوه عملکرد این تکنیک‌ها برای تشخیص نهایی آورده شده است:



شکل ۲. دیاگرام مراحل مختلف کارهای انجام شده برای تصاویر سی تی اسکن به عنوان داده های ورودی.

¹ Tag Image File Format

² Feature Extraction

³ Classification of Images

^۴ Fractals

⁵ Gray-Level Co-occurrence Matrix

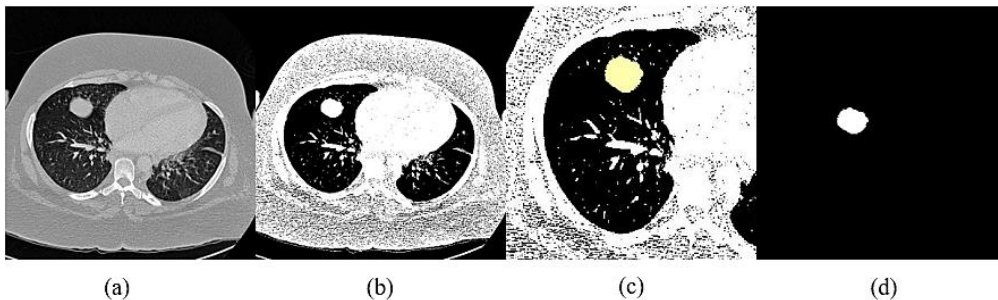
⁶Wavelet Transform

⁷ Statistical Texture

⁸Zernike

۳. یافته‌ها

بعد از فراخوانی تصویر، تکنیک‌های ذکر شده از جمله فراکتال، تبدیل موجک، ماتریس هم‌زمانی سطح خاکستری، روش‌های مبتنی بر زرنیک یا متغیرهای گشتاوری، به عنوان استخراج‌کننده ویژگی، استفاده شده است. برخی از ویژگی‌های آماری از بافت تصویر استخراج و سپس ویژگی‌های آماری با استفاده از تبدیل موجک انجام می‌شود و در مرحله بعد، برای طبقه‌بندی تصاویر، مورد استفاده قرار می‌گیرد. در ابتدا باید ویژگی‌ها را نرمالیزه کرد. با استفاده از مقادیر کمینه و بیشینه، قادر به نرمالیزه کردن ویژگی‌ها خواهیم بود با ترکیب کردن داده‌های سرطانی و غیر سرطانی، ۱۰۰ نمونه تصویر برای سرطان، ۱۱۰ نمونه تصویر برای غیرسرطان به منظور آموزش دادن و یاد دادن الگوریتم طبقه‌بندی کننده، استفاده می‌کنیم. پس از استخراج ویژگی‌های تصویر که به صورت دو بار تکرار است از ۸۳ ویژگی، شش ویژگی را به صورت دلخواه انتخاب کرده‌ایم. اختیار در انتخاب نوع و عدد ویژگی، بستگی به شخص و تحلیلگر دارد و سپس با طبقه‌بندی کننده های ^۹ (SVM) و ^{۱۰} (K-NN) داده‌ها را به ۱۰ قسمت تقسیم کرده و ۹ قسمت هر سری برای آموزش و یک قسمت، دقت را به ما می‌دهد. در اینجا تصاویر سی‌تی اسکن به عنوان ورودی گرفته شده است. تشخیص محل تومور و ایجاد ماسک برای بررسی اندازه‌ی تومور برای هر تصویر سی تی اسکن انجام شده است. یک تصویر دارای تومور سرطان برای پردازش، مطابق شکل زیر آمده است.



شکل ۳. قسمت (a) تصویر سی‌تی اسکن ریه، قسمت (b) بهبود کنتراست در دقت و تفکیک تصویر، قسمت (c) تشخیص و تعیین محل تومور در ناحیه سمت راست ریه و قسمت (d) ایجاد ماسک بر روی ناحیه ریه به جهت نمایش شکل و اندازه تومور.

⁹Support Vector Machine

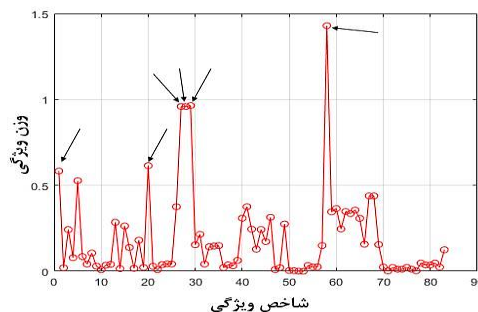
¹⁰ K-Nearest Neighbor

در شکل شماره 3، قسمت (c) با نشان دادن محل تومور و قسمت (d) اندازه تومور، به ما در تشخیص بیماری در مرحله اول و انتخاب نوع بیماری در مرحله دوم کمک می‌کند. پرتودرمانی یکی از درمان‌های رایج برای این بیماری‌ها است. توجه به این که از چه جهتی به بدن پرتو تابانده شود تا مؤثر واقع شود و همچنین به حداقل رساندن مقدار آسیب به بدن در حین پرتودرمانی، از اهداف مهم پردازش تصاویر پزشکی ریه است. برای نمونه‌های مختلف ورودی از تصاویر سی‌تی‌اسکن ریه، با استفاده از طبقه‌کننده‌های ماشین بردار پشتیبان و نزدیک‌ترین نقطه همسایگی نمونه‌هایی از تصاویر دارای تومور سرطانی تشخیص داده شده است. این طبقه‌بندی کننده‌ها کمک بزرگی در بررسی داده‌های سرطانی با اندازه‌های مختلف و تشخیص محل تومور می‌کنند. حال نتایج ما با استفاده از اعتبار سنجی متقابل داده‌ها^{۱۱} و با توجه به تعداد حلقه‌ها، به ازای ۱ تا ۱۰ محاسبه شده است. مقادیر دقت تشخیص طبقه‌بندی کننده‌های ماشین بردار پشتیبان و نزدیک‌ترین نقطه همسایگی، برای (K-fold) های مختلف در جدول زیر آمده است.

جدول ۱. مقادیر طبقه‌بندی کننده‌های ماشین بردار پشتیبان و نزدیک‌ترین نقطه همسایگی با (K-fold) های مختلف

K-fold	1	2	3	4	5	6	7	8	9	10
SVM	100%	100%	100%	95%	100%	100%	100%	100%	95%	100%
K-NN	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

با استخراج بهترین ویژگی‌ها (بااهمیت) با توجه به تصویر نمودار پایین، ماکزیمم داده‌های ما در نوک قله‌های نمودار بهترین ویژگی‌های نمودار محسوب می‌شوند. ۸۳ ویژگی در شکل زیر وجود دارد. که از آن‌ها ۶ ویژگی بر نوک قله‌ها بهترین نوع ویژگی هستند. وزن‌های ویژگی به فرمی، بهینه شده^{۱۲} از سوی توابع عملگر در محیط برنامه نویسی متلب به دست آمده است.



شکل ۴. استخراج ویژگی‌های بااهمیت

جدول ۲. مقادیر وزن ویژگی‌های بااهمیت بر حسب شاخص ویژگی

شاخص ویژگی بااهمیت	۱	20	27	28	29	58
وزن ویژگی‌های بااهمیت	0/56۶۹	0/66۸۸	0/99۲۸	0/99۲۸	0/97۴۵	1/38۵۷

۴. نتیجه‌گیری

¹¹ K-fold Cross Validation

¹² Optimized

تشخیص به موقع سرطان ریه بسیار اهمیت دارد. و این تشخیص با تصویربرداری پزشکی از ریه میسر می‌شود. پردازش تصاویر پزشکی ریه برای افزایش دقت در تشخیص بیماری، مکان و اندازه تومور ریه، جهت پرتودرمانی کمک بزرگی به ما می‌کند. طبقه‌بندی تصاویر دارای سرطان ریه کمک جانبی به پزشک می‌کند. پزشک می‌تواند با خطای کمتری، تصاویر پزشکی ریه را مورد تجزیه و تحلیل قرار دهد. تعیین ناحیه‌ی پرتودهی به تومور سرطانی و پی بردن به اندازه و حجم تومور، به منظور بازدهی بیشتر و به حداقل رساندن آسیب ناشی از پرتو درمانی به بافت سالم بدن، از اهداف این مقاله بوده است. ما با دقت بالایی برای طبقه بندی تصاویر دارای سرطان و غیر سرطان با دقت ۹۵٪ درصد به بالا (تا ۱۰۰٪) برای (K-fold) های از ۱۰ تا ۱۰، دست یافتیم. در حالی که (E.S.N. Joshua و همکاران) [17] برای طبقه بندی تصاویر سی تی اسکن ریه، به منظور تشخیص سرطان ریه با طبقه بندی کننده‌های ماشین‌بردار پشتیبان و نزدیک‌ترین نقطه همسایگی، بدون ذکر کردن تعداد (K-fold) ها و یا استفاده از آن‌ها به ترتیب با دقت ۶۲/۵٪ و ۷۲/۲۴٪ دست یافتند.

۵. مراجع

- 1) De Ruyscher, D., Faivre-Finn, C., Nackaerts, K., Jordan, K., Arends, J., Douillard, J. Y., ... & Peters, S. (2020). Recommendation for supportive care in patients receiving concurrent chemotherapy and radiotherapy for lung cancer. *Annals of Oncology*, 31(1), 41-49.
- 2) Riudavets, M., Garcia de Herreros, M., Besse, B., & Mezquita, L. (2022). Radon and lung cancer: current trends and future perspectives. *Cancers*, 14(13), 3142.
- 3) Thakerar, A., Simadri, K., Alexander, M., & Fullerton, S. (2020). Paroxetine for the treatment of intractable and persistent cough in patients diagnosed with cancer. *Journal of Oncology Pharmacy Practice*, 26(4), 803-808.
- 4) Mamun, M., Farjana, A., Al Mamun, M., & Ahammed, M. S. (2022, June). Lung cancer prediction model using ensemble learning techniques and a systematic review analysis. In *2022 IEEE World AI IoT Congress (AIIoT)* (pp. 187-193). IEEE.
- 5) Ryan, A. M., Prado, C. M., Sullivan, E. S., Power, D. G., & Daly, L. E. (2019). Effects of weight loss and sarcopenia on response to chemotherapy, quality of life, and survival. *Nutrition*, 67, 110539.
- 6) Dhuguru, J., & Skouta, R. (2020). Role of indole scaffolds as pharmacophores in the development of anti-lung cancer agents. *Molecules*, 25(7), 1615.
- 7) Munawar, J., Khan, M. S., Syeda, S. E. Z., Nawaz, S., Janjhi, F. A., Haq, H. U., ... & Bilal, M. (2022). Metal-organic framework-based smart nanoplatfoms with multifunctional attributes for biosensing, drug delivery, and cancer theranostics. *Inorganic Chemistry Communications*, 110145.
- 8) Peart, O. (2017). Metastatic breast cancer. *Radiologic technology*, 88(5), 519M-539M.
- 9) Shakeel, P. M., Burhanuddin, M. A., & Desa, M. I. (2022). Automatic lung cancer detection from CT image using improved deep neural network and ensemble classifier. *Neural Computing and Applications*, 1-14.

- 10) Nageswaran, S., Arunkumar, G., Bisht, A. K., Mewada, S., Kumar, J. N. V. R., Jawarneh, M., & Asenso, E. (2022). Lung cancer classification and prediction using machine learning and image processing. *BioMed Research International*, 2022.
- 11) Murugesan, M., Kaliannan, K., Balraj, S., Singaram, K., Kaliannan, T., & Albert, J. R. (2022). A hybrid deep learning model for effective segmentation and classification of lung nodules from CT images. *Journal of Intelligent & Fuzzy Systems*, 42(3), 2667-2679.
- 12) Karri, S., Aviel-Ronen, S., & Firer, M. A. (2022). Fractal and textural imaging identify new subgroups of patients with colorectal cancer based on biophysical properties of the cancer cells. *Pathology-Research and Practice*, 238, 154040.
- 13) Althubiti, S. A., Paul, S., Mohanty, R., Mohanty, S. N., Alenezi, F., & Polat, K. (2022). Ensemble learning framework with GLCM texture extraction for early detection of lung cancer on CT images. *Computational and Mathematical Methods in Medicine*, 2022.
- 14) Xu, Y., Wang, Y., & Razmjoo, N. (2022). Lung cancer diagnosis in CT images based on Alexnet optimized by modified Bowerbird optimization algorithm. *Biomedical Signal Processing and Control*, 77, 103791.
- 15) Corrias, G., Micheletti, G., Barberini, L., Suri, J. S., & Saba, L. (2022). Texture analysis imaging “what a clinical radiologist needs to know”. *European Journal of Radiology*, 146, 110055.
- 16) Shaffie, A., Soliman, A., Eledkawy, A., van Berkel, V., & El-Baz, A. (2022). Computer-assisted image processing system for early assessment of lung nodule malignancy. *Cancers*, 14(5), 1117.
- 17) Joshua, E. S. N., Chakkravarthy, M., & Bhattacharyya, D. (2020). An Extensive Review on Lung Cancer Detection Using Machine Learning Techniques: A Systematic Study. *Rev. d'Intelligence Artif.*, 34(3), 351-359.