DOI Link: https://doi.org/10.61586/wG5Kx
Vol.40, Issue.7, Part.1, July 2025, PP.2-23

CANCER DIAGNOSIS WITH IMAGE FILTER-INTEGRATED ARTIFICIAL INTELLIGENCE ALGORITHMS: INNOVATIVE POSSIBILITIES FOR MELANOMA DETECTION WITH A HIGH DEGREE OF ACCURACY

Ahmet Kara
Istanbul Commerce University, Department of Economics, Turkey
kara2010@ticaret.edu.tr

ORCID ID: 0000-0002-0162-8137

Received May 2025 Accepted June 2025 Published July 2025

ABSTRACT

This paper makes use of image filter-generated high dimensional data with artificial intelligence algorithms for an innovative and accurate diagnosis of cancer.

The paper describes two methods, the first of which employs image filters to extract, from images, hundreds (and even thousands) of quantified features representing high dimensional data that can be fed into selection and classification algorithms to accurately diagnose cancer. The method is applicable to many different cancer-related cases. We have used this method of image filter-integrated artificial intelligence algorithms in the context of a data set to achieve up to 100 % accuracy. This method, which is static and hence focuses on cases at a point in time, could be generalized to a dynamic setting by collecting data at different points in time and evaluating them algorithmically so as to construct a trajectory describing cancer progression over time. Obtaining such a cancer trajectory would facilitate the decisions for optimal treatment and/or interventions to slow down cancer progression or help eradicate cancer altogether.

Overall, methods are flexible enough to handle a broad range of cases with various levels of complexity and intricacy as well as varying scale and scope associated with different types, subtypes, degrees and stages of cancer.

KEY WORDS: Cancer diagnosis, image filters, Melanoma, artificial intelligence algorithms, innovative possibilities.

1. INTRODUCTION

The state-of-art methods in artificial intelligence contain a rich spectrum of possibilities that can be innovatively tailored and recombined for the purpose of diagnosing cancer in an efficient and accurate manner. A plethora of ever-developing artificial intelligence algorithms have created an interactively progressing productive ground leading to an extensive mode of inquiry, facilitating, with an increasing speed and precision, the static and dynamic identification, classification and analysis of the types, subtypes, degrees and stages of cancer. There are many works in the literature that illustrate the extent of this inquiry, exploring topics ranging from the use of machine learning/artificial intelligence algorithms in cancer detection in general (Ilhan et al. 2021; Mahmood et al. 2021, Dadzie et al. (2025), Kenner et al. 2020; Mahmood et al. 2020; Alabi et al. 2021; Coudray et al. 2018) to the employment of specific algorithms such as deep learning in particular (Kanavati et al. 2021; Iizuka et al. 2020; Ghosh et al. 2021; Kumar et al. 2020; Phan et al. 2020; Zhu et al. 2021), Verma & Yadav (2025), Kreouzi et al. (2025), Unnisa et al. (2025) and Mio et al. (2025). The literature contains works exploring the use of artificial intelligence in melanoma research as well. Among these are Tschandl (2021) which presents the issue of artificial intelligence for melanoma diagnosis in a general manner, Johannet et. al. (2021), which deals with machine learning algorithms to predict immunotherapy response in patients with advanced melanoma, Esteva et. al. (2017), Brinker (2019), Fink (2020), and Haenssle (2018), which study issues associated with skin cancer/melanoma with deep neural networks and Barata et. al. (2023) which investigates artificial intelligence-based decision support in skin cancer with a reinforcement learning. Some works such as Yee, Rosendahl & Aoude (2024), examine the role of artificial intelligence in the management of melanoma from a more general clinical, pathological and radiological perspective.

Image-focused intricacies associated with the topic are also explored in the literature in some detail. Li et. al. (2021) tackles with the issue of automated diagnosis and localization of melanoma from skin histopathology slides using deep learning. Brinker et. al. (2019) undertakes head-to-head dermoscopic melanoma image classification tasks. Bi (2017), analyzes dermoscopic image segmentation via multistage fully convolutional networks. Xie (2021) studies whole-slide melanoma histology images using convolutional neural network. Hohn (2021) combines convolutional neural network-based histologic whole slide image analysis and patient data to improve skin cancer classification. Surveys/reviews of the topic, such as the ones presented by Jothi & Rajam (2017) on automated cancer diagnosis from histopathology images and Li et. al. (2022) on computer-aided whole-slide image analysis involving extraction, segmentation, classification and detection are informative as well.

In this paper, we will present and use methodologically sound and versatile combinations of various existing algorithms that can be flexibly, innovatively and efficiently adjusted to the cancer diagnosis and analysis. The first combination we

will draw attention to involves “filtering”, “selection” and “classification” algorithms, the joint structure of which we will call “the basic triadic/three-layered/static method”. The second structured combination adds a “simulation and forecasting” component to the basic method and will be called “the extended/dynamic method.” The details and innovative potential of the methods will be explained in the second-to-fourth sections. For a compact elaborated statement of the main issue of the paper, we should indicate that the procedure below leads to the construction of sufficiently explanatory quantitative features, a subset of which could be selected and put into use for a highly accurate tumor classification tasks. The described simulation exercise has the potential to dynamically capture the evolution of tumors.

In the context of a case study/a data set chosen for this paper, we will use the basic triadic method to first employ image filters to extract hundreds/thousands of features from a set of tissues, then use selection algorithms to determine the best features (i.e., the features that are most suitable to the proper classification of the tissues) and finally make recourse to classification algorithms to accurately classify them. The end results, in general, are shown to be a high degree of accuracy which may not be achieved through manual procedures. The results verify the great potential of the hybrid methods we use for distinguishing between benign and malignant lesions. The methods could also be used to differentiate cancers at different levels of development as well as identify the nature of the cancerous tissues. In other words, not only do the methods enable us to identify types and subtypes of cancer, they could also serve as successful identification instruments for cancer grading and staging. This could open the doors to the early diagnosis of cancer which is of great significance to overall survival rate.

Having presented, in Section 2, the structure of the methods to be employed, the results obtained through the application of the first method, and hence the tables of accuracy rates for various image filter-classification algorithm combinations are mentioned in the third section and displayed at the end of the paper. A discussion of the results will be presented; some concluding remarks will be spelled out and various fruitful extensions will be indicated in the fourth part of the paper.

2. MATERIAL AND METHODS

The multi-layered method we will make use of for cancer diagnosis has a combinatorially rich architecture which gets empowered with every new algorithm incorporated into it. Let’s explain this architecture which has several building blocks, namely, filtering, selection, classification and simulation/forecasting, a relational sketch of which is as follows (Figure 1).

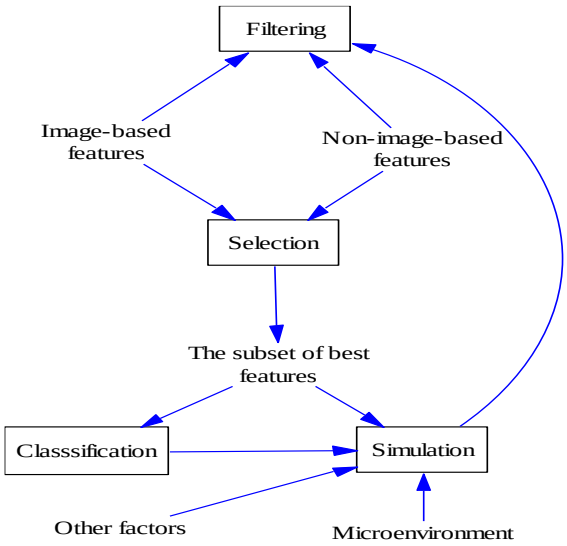

Figure 1. A sketch of the multi-layered method of filtering-selection-classification-simulation

2.1 FILTERING: Cancerous tissues have certain physical, chemical and biological properties/features[1] which could be decomposed into two main categories: image-based features and non-image-based features. For the purposes of this article, we will focus on image-based features involving geometrically or physically describable morphological and other properties such as color, tone, size, shape/texture and related intricate details concerning curvature, kinks, (dis)continuities, degree of fuzziness, irregularity, asymmetry, possible rupture and non-convexity, broken or diffracted edges or segments, structures of interconnectedness as well as convoluted differentiation, overlapping patterns etc. Image filters are designed to extract a wide spectrum of mathematically-represented (quantifiable) features from the images, leading to a very rich, high dimensional dataset with invaluable properties. Filter-based/scrutinized mathematical representation of morphological, topological and other features of images could bring into light hidden patterns or structures embodied in images. Patterns revealed through such filters make it possible to more accurately diagnose the cancerous tissues, thus facilitating the proper discrimination of benign tissues from the cancerous ones or the discrimination of tissues at advanced cancerous stages from

[1] For some features of malignant tumors, see (Prabhu et al. 2022).

the early-stage ones. Features extracted through image filters make it possible to distinguish between different types or subtypes of cancer as well.

Cancerous tissues appear to display a range of features including various kinds of heterogeneity, irregularity and possible complex differentiation and connectedness, the patterns of which could be captured by nano, micro or macro level images that could be effectively used in distinguishing cancerous tissues/lesions from the non-cancerous ones. Clearly, if the existing image filter algorithms (or algorithms that can be designed for specific image analyses) enable us to extract, from images, quantifiable (real-valued) features that properly represent the intricate details of those images embodying cancer-signifying information, we can use those quantifiable features, via further algorithms, to accurately identify whether a tissue is cancerous or not. Similarly, properly designed image filters may help identify features that facilitate the detection of early signs for cancer tendencies before the critical or metastatic stage, allowing a time-specific classification of cancer-prone tissues.

In this paper, we will make use of a subset of the image filter algorithms built in the machine learning program, "Waikato Environment for Knowledge Analysis" (WEKA), which are named, in a suggestive manner, as "Auto Color Correlogram Filter", "Binary Patterns Pyramid Filter", "Color Layout Filter", "Edge Histogram Filter", "Fuzzy Color and Texture Histogram Filter" (FCTH Filter), "Fuzzy Opponent Histogram Filter", "Gabor Filter", "JPEG Coefficient Filter", "Pyramid Histogram Oriented Gradients Filter" and "Simple Color Histogram Filter". Some of these filters extract color information from the images, some extract texture information and some do both. The procedures through which these algorithms extract information/features from images are beyond the scope of this paper. Without getting into the mathematical descriptions of these algorithms, we will incorporate the filter-extracted quantified features into a selection algorithm which we will point out next.

2.2 SELECTION: Through image-filters, we can extract/construct hundreds (and sometimes thousands) of quantified features representing the images. Features may direct us in different directions and using all those features in the classification process need not always lead to the best results. Instead, it would often be better to construct a subset of features that are most suitable for classification purposes. For that reason, we will use a "feature selection algorithm" to select the best features for the classification target. There are a number of algorithms ranging from "Cfs Subset Eval" to "Correlation Attribute Eval" that can be used for selection purposes. We will choose Cfs Subset Eval algorithm, which is based on the predictive abilities of the individual features[2], for the task of selecting the best subset of the image-based features.

[2]The WEKA program provides descriptions/brief accounts of these algorithms.

The list of features we consider for overall analysis need not consist only of image-based features. We can enlarge the list of finally-chosen features for overall analysis by adding, to the image-based subset, those non-image-based features the current medical knowledge considers to be important for diagnostic purposes.

2.3. CLASSIFICATION: Following the steps of extracting features from images (filtering) and forming a best subset of features (selection), we can choose algorithms to identify the nature of the tissues/lesions (classification). That is to say, through classification algorithms based on image-filtered and properly selected features, we can determine whether a tissue is cancerous or not or we can classify, with a reasonably high degree of accuracy, the types, subtypes, degrees or stages of cancer. The (supervised) classification algorithms we can use in the diagnostic process ranges from "multiobjective evolutionary fuzzy classifier" to "deep learning", the performances (or successes) of which are likely to depend, among other things, on the particular targeted classification problem. The set of criteria that are used to measure the performances/successes of these algorithms include many metrics such as "accuracy rate", "sensitivity", "specificity", "F measure" and "AUC" (area under the ROC curve).

The performance evaluation problem of the classification algorithms or the overall chain of diagnosis-targeting algorithms is a particularly complicated or sensitive matter in the context of our paper, precisely because we use a hybrid method that incorporates image filters and selection algorithms beforehand. In other words, the classification procedures we use in this paper are image filter-integrated and best feature-selected artificial intelligence algorithms, the performance of which resides not only in the classification process itself but also in the filtering and selection processes and perhaps combination of all three processes. We conjecture that image-filtering and feature-selection improve the performance of the classification algorithms they are integrated with, but the dependence of the classification task on the filtering and selection tasks renders the classification performance sensitive to the filtering and selection performances. Certain combinations of filtering, selection and classification may well prove to be highly successful for the identification of certain types of cancer. The question of which combinations of image filters and classification algorithms work better for the identification of what types of cancer is, we think, a fruitful and sophisticated mathematical topic with powerful practical implications/consequences for cancer diagnosis.[3] In the third ("the results") section of this paper, without delving into the mathematical details, we will try to find combinations of image-filters and classification algorithms that yield a high degree of accuracy.

[3] Different features may be crucial to detecting different types of cancer. The features extracted through image filters may need to be classified in terms of their instrumental role in cancer diagnosis. Color-centered filters may be more successful in some cases than others. Similarly, filters extracting some topological features or shape-related features would be more conducive to successful diagnosis in cases where these features play a more distinguishing role. So a study of filters and taxonomy of them might be useful.

2.4. SIMULATION/FORECASTING OF FUTURE TRENDS: The above-described triadic method based on filtering, selection and classification is, as we will show in the context of a concrete case in the next section, very successful at cancer diagnosis at a point in time. This is, if the phrase is appropriate, "a success in static contexts". Can this success in static contexts be generalized to dynamic contexts? The answer to this question is, I think, a "qualified yes". The answer is qualified because of the multidimensionally intricate cases with complicated feedback relations that might render the cancer trajectories difficult to predict over time. Despite the challenging difficulties, artificial intelligence methods, if properly combined with appropriate simulation paradigms, could produce a significant success in dynamic contexts as well. For an illustrative description, let us begin with a relatively simple case (example) where we can diagnose whether a tissue is cancerous at different points in time, using the triadic method explained above. Using the obtained estimates at different points in time, we can form a time series implicitly characterizing the cancer trajectory over time. We can then use artificial intelligence algorithms to undertake a time series forecasting of the future trends. This can be done with different programs including the WEKA mentioned above. The forecasting algorithms could also enable us to incorporate the effects of different factors on the trajectory, assuming that the data about those factors are available. This way we can, for instance, forecast the effects of possible interventions (involving drugs, radiation etc.) on the cancer trajectories. Some versions of these types of forecasting could also be carried out via ARIMA (Auto Regressive Integrated Moving Average)-type models which could be run with computer programs such as EVIEWS.

Artificial intelligence-based possibilities in dynamic contexts are not limited to the time series forecasting. We can construct simulation models where we can incorporate artificial intelligence algorithms - models that enable us to formulate the feedback relations characterizing cancer processes and take into account a multiplicity of possibly interconnected or interwoven factors, delays or interdependencies in the microenvironment. Such multiple factors and the web of interactions among those factors could be modeled with the methods of "system dynamics" or "agent-based modeling" or a combination of both. These simulation paradigms present possibilities which cancer practitioners can take advantage of in dealing with or analyzing cancer-related phenomena such as metastatic potential and recurrence, not to mention the possibilities involving the timing and magnitude of intervention to limit cancer progression. And as such, the simulative framework is highly suitable for innovative exploration and experimentation of a wide range of dynamic structures in cancer research.

In sum, the main propositions and tasks of the paper has four subcomponents through which significant progress could be made towards cancer diagnosis. First, via the filtering component, image filters greatly facilitate the extraction of the patterns that could characterize the various types of tumors. Second, via the optimal selection component, an optimal subset of extracted quantitative variables

embodying the patterns in question could be formed and used for accurate classification of tumors. Third, via the classification component, a number of classification algorithms could be employed to identify tumors. Fourth, via the simulation component, the static results could be extended to dynamic contexts where the progress of tumors could be effectively analyzed.

3. RESULTS

We have used, with proper permission, a subset of the images of the DermNet (DermNet 2021). Images are pictures of some malignant/cancerous and some benign/non-cancerous skin lesions. The 50 % of the balanced set (subset) of 84 observations consists of images of malignant lesions and the rest of the 50 % consists of benign lesions. The data set is problem-free, removing the need for prepossessing. Since the data set is not of time series nature and consists of static images, we have used the above-mentioned basic triadic/static method with the proposed steps of filtering, selection and classification. Though the mathematics underlying some of these steps could be challenging to some readers, the computerized procedures associated with them are accessible to anyone with some familiarity with the topic. Among the computer programs that can be used to implement these procedures, we have used the freely available WEKA program where some image filters are already built in (or could be freely downloaded). There are some freely available online material and video courses that teach these procedures and describe the algorithms used in WEKA (Frank et al. 2016; Witten 2022a,b, Mayo 2022).

To implement the procedures, we have first prepared the data set in the ARFF format (Attribute Relation File Format). Then we have applied various image filters to extract mathematically representable (real-valued) features (variables) from the images. Among the hundreds (in some cases thousands) of features extracted from the application of a particular image filter or image filter-combinations, we have selected, through a feature selection algorithm (Cfs Subset Eval), the best features that are most suitable for classification. More specifically, with the Auto Color Correlogram Filter, we have extracted 1024 features and selected 18 of them. With the combination of the Auto Color Correlogram Filter and the Fuzzy Color and Texture Histogram Filter, we have extracted 1216 features and selected 22 of them. With the combination of the Auto Color Correlogram Filter, the Edge Histogram Filter and the Fuzzy Color and Texture Histogram Filter, we have extracted 1296 features and selected 37 of them. With the combination of the Auto Color Correlogram Filter, the Edge Histogram Filter, Fuzzy Color and Texture Histogram Filter and JPEG Coefficient Filter, we have extracted 1488 features and selected 40 of them. With the combination of the Auto Color Correlogram Filter, the Edge Histogram Filter, Fuzzy Color and Texture Histogram Filter, the JPEG Coefficient Filter and Fuzzy Opponent Histogram Filter, we have extracted 2064 features and selected 40 of them.

Having selected the best (and still a fairly high number of) features for classification, we have applied various classification algorithms to see how accurately they identify the malignant and benign lesions. The classification algorithms we have used are “Sequential Minimal Optimization/Support Vector Machines” (SMO), “Non-linear Support Vector Machines (Lib SVM)”, “Random Forest”, “MultiobjectiveEvolutionary Fuzzy Classifier”, “Deep Learning”, “Random Sub Space” and Simple Logistic”.

Before the classification task, however, we have used two methods/options to determine the training-and-testing-related fractions of the data set, which consists of independent observations. Basically, the image set is divided into two sets for training and testing purposes, which are, by virtue of the composition of the data set, independent samples, which could be reliably used in any employable option/method . In the percentage split option, a particular percentage is set aside randomly or via some other method, and the rest of the set is used for testing. In the cross-validation option, a series of training-testing splits are applied to the same set. For instance, in the 10-fold cross validation, the set is divided into ten sections. The first 10 % is used for testing and the rest (90%) for training. Then the second 10% is used for testing and the rest (90 %) for training and so on. This process is repeated until every one of the ten folds is used for testing. Then the average of the 10 trials is taken to find the overall scores for accuracy, sensitivity, specificity etc.

The particular metrics we will use to represent the performances of the classification algorithms are the accuracy rate and AUC value, which is the area under “Receiver Operating Characteristic” (ROC) curve. In Table 1a, Table 1b, Table 1c, Table 2 and Table 3, we provide the accuracy rates for various cases. Table 1a, Table 1b, Table 1c present a number of cases that give accuracy rates (the last column in the table) for particular combinations of filtering, selection and classification algorithms with particular train-test split of the data. Table 2 shows the accuracy rates obtained through a 90%-10% train-test split for particular combinations of image filter and classification algorithms. Similarly, Table 3 gives the accuracy rates obtained through 10-fold cross validation for particular combinations of image filter and classification algorithms.

Table 1a. Accuracy rates for various combinations of filtering, selection and classification algorithms with particular train-test split options.

	Filtering Algorithm	Selection Algorithm	Classification Algorithm	Train-test Split	Accuracy Rate
Case 1	Auto Color Correlogram	Cfs Subset Eval	Multiobjective Evolutionary Fuzzy Classifier	90%-10% split	100%
Case 2	Auto Color Correlogram	Cfs Subset Eval	Deep Learning	10-fold cross validation	88.09%
Case 3	Auto Color Correlogram+ Fuzzy Color and Texture Histogram	Cfs Subset Eval	Support Vector Machines	90%-10% split	100%
Case 4	Auto Color Correlogram+ Fuzzy Color and Texture Histogram	Cfs Subset Eval	Multiobjective Evolutionary Fuzzy Classifier	10-fold cross validation	69.04%
Case 5	Auto Color Correlogram+ Fuzzy Color and Texture Histogram+ Edge Histogram	Cfs Subset Eval	Random Forest	10-fold cross validation	96.42%
Case 6	Auto Color Correlogram+ Fuzzy Color and Texture Histogram+ Edge Histogram	Cfs Subset Eval	Non-linear Support Vector Machines	90%-10% split	100%

Table 1b. Accuracy rates for various combinations of filtering, selection and classification algorithms with particular train-test split options.

	Filtering Algorithm	Selection Algorithm	Classification Algorithm	Train-test Split	Accuracy Rate
Case 7	Auto Color Correlogram+ Fuzzy Color and Texture Histogram+ Fuzzy Color and Texture Histogram+ Edge Histogram	Cfs Subset Eval	Support Vector Machines	10-fold cross validation	92.85%
Case 8	Auto Color Correlogram+ Fuzzy Color and Texture Histogram+ Fuzzy Color and Texture Histogram+ Edge Histogram	Cfs Subset Eval	Random Forest	90%-10% split	100%
Case 9	Auto Color Correlogram+ Fuzzy Color and Texture Histogram+ Edge Histogram+ JPEG Coefficient+ Fuzzy Opponent Histogram	Cfs Subset Eval	Multiobjective Evolutionary Fuzzy Classifier	90%-10% split	100%

Table 1c. Accuracy rates for various combinations of filtering, selection and classification algorithms with particular train-test split options.

	Filtering Algorithm	Selection Algorithm	Classification Algorithm	Train-test Split	Accuracy Rate
Case 10	Auto Color Correlogram+ Fuzzy Color and Texture Histogram+ Edge Histogram+ JPEG Coefficient+ Fuzzy Opponent Histogram	Cfs Subset Eval	Deep Learning	10-fold cross validation	91.66%

Table 2. Accuracy rates for various image filter and classification algorithms with 90%-10% train-test split of the data.

Image Filter . / Class. Alg.	Auto Color Correlogram	Auto Color Correlogram+ Fuzzy Color and Texture Histogram	Auto Color Correlogram+ Fuzzy Color and Texture Histogram+ Edge Histogram	Auto Color Correlogram+ Fuzzy Color and Texture Histogram+ Fuzzy Color and Texture Histogram+ Edge Histogram	Auto Color Correlogram+ Fuzzy Color and Texture Histogram+ Edge Histogram+ JPEG Coefficient+ Fuzzy Opponent Histogram
Support Vector Machines	100%	100%	87.5%	100%	100%
Random Forest	100%	100%	87.5%	100%	100%
Multiobjective Evolutionary Fuzzy Classifier	100%	100%	100%	100%	100%
Deep Learning	87.5%	100%	87.5%	100%	87.5%
Random Sub Space	100%	87.5%	75%	87.5%	87.5%
Simple Logistic	100%	100%	75%	100%	87.5 %

Table 3. Accuracy rates for various image filter and classification algorithms with 10-fold cross-validation.

Image Filter . / Class. Alg.	Auto Color Correlogram	Auto Color Correlogram+ Fuzzy Color and Texture Histogram	Auto Color Correlogram+ Fuzzy Color and Texture Histogram+ Edge Histogram	Auto Coor Correlogram+ Fuzzy Color and Texture Histogram+ Fuzzy Color and Texture Histogram+ Edge Histogram	Auto Color Correlogram+ Fuzzy Color and Texture Histogram+ Edge Histogram+ JPEG Coefficient+ Fuzzy Opponent Histogram
Support Vector Machines	89.28%	90.47%	89.28%	92.85%	91.66%
Random Forest	89.28%	91.66%	96.42%	92.85%	92.85%
Multiobjective Evolutionary Fuzzy Classifier	78.57%	69.04%	80.05%	83.33%	84.52%
Deep Learning	88.09%	89.28%	92.85%	91.66%	91.66%
Random Sub Space	85.71%	82.14%	85.71%	82.14%	84.52%
Simple Logistic	89.09%	88.09%	91.66%	92.85%	91.66%

Tables indicate up to 100 % accuracy, which is indeed quite high. Nevertheless, depending on the particular combinations of algorithms used, the accuracy rate varies. For instance, as shown in Table 1a, the combination of algorithms in Case 6 yield 100% accuracy whiles the combination in Case 4 yield 69.04 % accuracy. This is of course expected because not all algorithm combinations would be equally suitable for the identification of a particular set (or subset) of tumors. Hence, choices for image filters and choices for particular combinations of image filters and classification algorithms appear to make some difference. This is by virtue of the fact that the performance of the filters has a lot to do with the color and shape/texture of the tumors. Some tumors may be distinctive in the color spectrum they display while some may be more distinguished by the textural patterns they exhibit. Thus color-based filters may perform better in properly identifying color-characterized lesions while texture-based filter may do better in extracting quantized features for the shape-distinguished lesions. A combination of filters could, as shown in the tables, be used as well.

We have used a subset of tumor/lesion images with particular algorithm combinations and get the results displayed in Tables 1a, 1b, 1c, 2 and 3. Had we used a different set or subset of images, we could have obtained different results. This tumor set-specific result does not, however, diminish/undermine the strength of the method precisely because of the fact that, for any set or subset of tumors, we can almost always find algorithm combinations and features sets that yield a sufficiently high degree of accuracy, regardless of how complex the structures or properties of tumors are. If one combination does not work, we can keep trying another one till we find a successful one. Furthermore, if we have some knowledge about other properties of the tumors, we can add some of these “non-image-based” properties to the feature set, which could have a great potential of improving the accuracy of the classification. In short, the power of the method we use resides in its flexibility and broad applicability.

In addition to the accuracy rate, we can use AUC values as a performance criterion, which yields similar results. For the purpose of illustration, we will only give the AUC values associated with the classification algorithms in Table 3 Column 3. With the selected filters, the associated AUC values for Support Vector Machines, Random Forest, Random Sub Space, Deep Learning, Simple Logistic and Multi Objective Evolutionary Fuzzy Classifier are 0.893, 0.976, 0.810, 0.979, 0.921 and 0.974, respectively. The associated ROC curves are given in Figure 2 below.

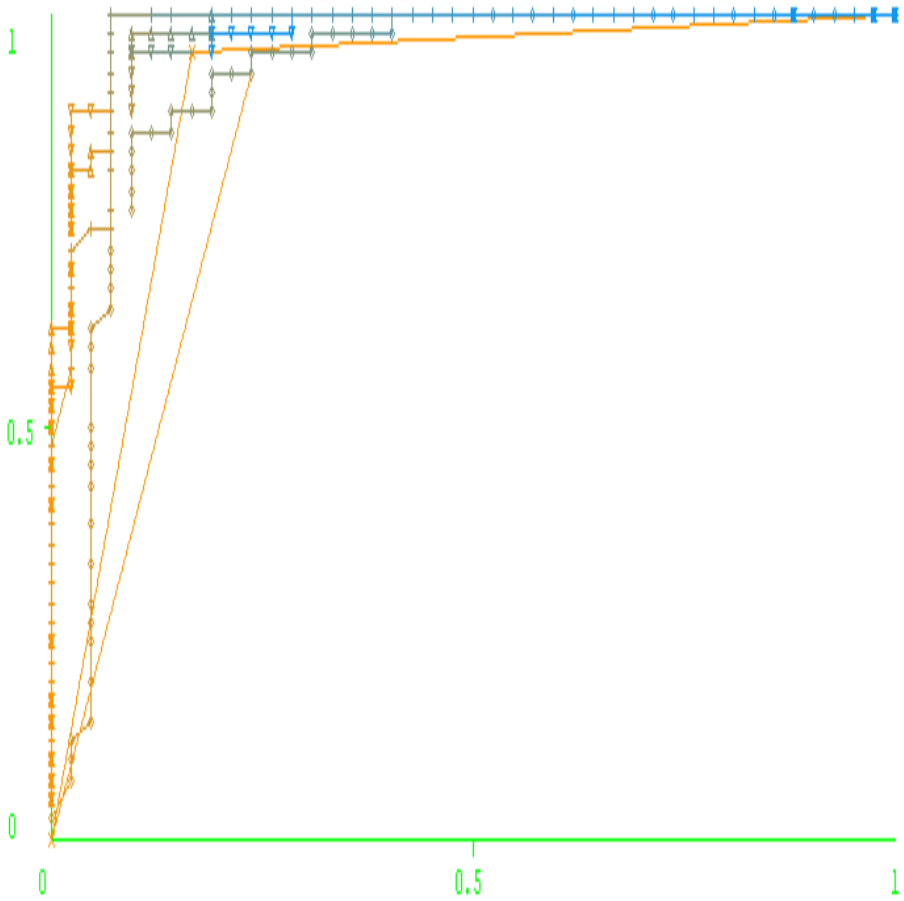

Figure 2. ROC curves obtained through Support Vector Machines, Random Forest, Random Sub Space, Deep Learning, Simple Logistic and Multi Objective Evolutionary Fuzzy Classifier integrated with Auto Color Correlogram, Fuzzy Color and Texture Histogram and Edge. 10-fold cross validation is used as a train-test split method.

x: Support Vector Machines (SMO)
+: Random Forest
◊ : Random Sub Space
Δ: Deep Learning
∇: Simple Logistic
x: Multi Objective Evolutionary Fuzzy Classifier

Accuracy rates and AUC values expectedly point out parallel performance for filter-classification algorithm combinations. Nevertheless, the joint performance of filters and classification algorithms is a highly complicated, mathematically challenging issue. Here, a practitioner would be advised to make recourse to trial and error and see which combinations perform better.

4. DISCUSSION, EXTENSIONS AND CONCLUDING REMARKS

The multilayered method presented in this paper has broad applications with far-reaching consequences in a number of respects. First, the scale of the applications is very flexible, ranging from nano scale to macro scale. Depending on the nature of the task to be undertaken, we can construct a data set of nano, micro or macro scale skin images from which we can extract a high number of quantifiable features (through filtering), which could be combined with non-image features, so as to proceed with selection, classification and even simulation. The extractable features could be too many and hence the data could be strikingly high dimensional enabling the representation of a wide range of heterogeneity and complexity that can be encountered in real cases, confirming the versatility of the method employed here. Possibilities for applied works in different areas are practically endless. For instance, extracting information (features) from nano-scale images of the cell network structure of a malignant tissue could help distinguish it, with sometimes painstaking precision, from the benign one through the procedure outlined in this paper. Second, the scope of the applications is very wide and flexible as well, ranging from the dichotomous tissue classification involving the presence or absence of cancer to multiplex classification of types, subtypes, degrees and stages of cancer. Not only is the method flexible enough to enable classification among the subtypes of a particular type, it is also versatile enough to perform classifications across different types, degrees and stages of cancer. Third, the method can handle both static (at a point in time) and dynamic (over time) analysis of tissues. The static analysis is already illustrated above. As described in the simulation part of the paper above, the static analysis could be extended to a dynamic setting by collecting data at different points in time and evaluating them algorithmically so as to construct a sketch of the cancer or cancer-prone trajectory of the tissues over time. The evolution of the saddle dimensions of the images could give an idea of where the trajectory is heading. Using the results of classification at different points in time, as captured by the derived trajectory, we can forecast the cancer progression and optimally intervene to slow down the progression of cancer. The constructed model will enable us to simulate the effects of various types of intervention so as to determine the optimal course of treatment.

The main idea signifying the productivity and accuracy of artificial intelligence algorithms in cancer research this paper exemplifies is also supported by some recent research such as Dadzie et al. (2025), Mio et al. (2025), Unnisa et al. (2025), Kreouzi et al. (2025) and Verma & Yadav (2025).

The work presented here has a fascinatingly wide array of extensions, some of which revolve around the additional constructible variables/measures that can be included in the set of features (the high dimensional dataset) to be used for classification. I will point out three of such measures drawn from system theory and physics.

First, malignant tumors are likely to have polymorphic structures with "rough" shapes that need not have integer dimensions. "Fractal dimension" (or fractional dimension") could be used as a measure of roughness, which could serve to distinguish tumors/lesions with intricately rough structures.[4] For instance, the fractal dimensions of proteins secreted by cancer cells could be indicative of their surface-specific irregularities, which could be captured, quantized and be included in the high dimensional feature set to be used for classification. Second, tumors are likely to be "disorderly" structures for which the law of increasing entropy would hold. Thus using an entropy measure to characterize the degree of disorder of the tumors could be a distinguishing feature as well. Third, the cancerous trajectories of some tumors may display chaotic tendencies for which "Liapunov exponent"[5] could be a reasonably appropriate and fairly practical measure. There are computer programs or algorithms/procedures to estimate all these three measures (i.e., fractal dimension, entropy and Liapunov exponent). We conjecture that the inclusion of these measures into the feature set could greatly enhance the power of classification especially in cases of structurally complex tumors. The generation of these measures could be done with the help of separate computer programs or alternatively new filters could be designed so as to extract quantitative estimates of these measures to be used for classification. Developing such new filters would be a time-saving short-cut and a valuable contribution to the theory and practice of high- dimensional-data-centered and artificial-intelligence-based cancer diagnosis.

The extraordinarily rich application spectrum and new and extended possibilities associated with the hybrid combinations of image filter integrated artificial intelligence algorithms with high dimensional data are worthy of future research.[6]

[4] A technical definition of the term "fractal" provided by Mandelbrot is as follows (Mandelbrot 1982): A fractal is a set whose Hausdorff dimension strictly exceeds its topological dimension Though Hausdorff dimension and topological dimension are standard concepts in the field of topology, for practitioners, it suffices to have a basic understanding of a fractal (for which Falconer's work (Falconer 2013) would be useful) and use a computerized estimate of its dimension.

[5]Liapunov exponent of a dynamical system (such as a malignant tumor) is a real number representing the rate at which infinitesimally close trajectories separate from one another.

[6] An earlier version of this article has appeared as a preprint (Kara2022).

CONFLICT OF INTEREST
The author has no conflicts of interest.

ETHICS STATEMENT
The paper does not require an ethics statement

DATA AVAILABILITY STATEMENT
The data used in the paper is publicly/freely available for non-commercial/scholarly purposes at DermNet.

REFERENCES

Alabi, R.O., Youssef, O., Pirinen, M., Elmusrati, M., Makitie, A.A., Leivo, I. & Almangush, A. (2021). Machine learning in oral squamous cell carcinoma: current status, clinical concerns and prospects for future-a systematic review. *Artif Intell Med*, 115, 102060.

Barata, C., Rotemberg, V., Codella, N.C.F., Tschandl, P., Rinner, C., Akay, B.N., et al. (2023). A reinforcement learning model for AI-based decision support in skin cancer. *Nat Med*, 29, 1941–1946.

Bi, L., Kim, J., Ahn, E., Kumar, A., Fulham, M. & Feng, D. (2017). Dermoscopic image segmentation via multistage fully convolutional networks. *IEEE Trans Biomed Eng*, 64, 2065–2074.

Brinker, T.J., Hekler, A., Enk, A.H., Klode, J., Hauschild, A., Berking, C., et al. (2019). Collaborators. Deep learning outperformed 136 of 157 dermatologists in a head-to-head dermoscopic melanoma image classification task. *Eur J Cancer*, 113, 47–54.

Coudray, N., Ocampo, P.S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., Moreira, A.L., Razavian, N. & Tsirigos, A. (2018). Classification and mutation prediction from non–small cell lung cancer histopathology images using deep learning. *Nat Med*, 24(10), 1559–1567.

Dadzie, A.K., et al. (2025). Artificial intelligence in the diagnosis of uveal melanoma: Advances and applications. *Exp Biol Med*, Vol. 250. Article no. 10444.

DermNet. (2021). https://dermnetnz.org/image-catalog ue/lesion-tumour-and-cancer-images.

Esteva, A., Kuprel, B., Novoa, R.A., Ko, J., Swetter, S.M., Blau, H.M., et al. (2017). Dermatologist-level classification of skin cancer with deep neural networks. *Nature*, 542, 115–118.

Fink, C., Blum, A., Buhl, T., Mitteldorf, C., Hofmann-Wellenhof, R., Deinlein, T., et al. (2020). Diagnostic performance of a deep learning convolutional neural network in the differentiation of combined naevi and melanomas. *J Eur Acad Dermatol Venereol*, 34, 1355–1361.

Frank, E., Hall, M.A. & Witten, I.H. (2022). The WEKA Workbench.Online appendix for "Data Mining: Practical Machine Learning Tools and Techniques". https://www.cs.waikato.ac.nz/ml/weka/Witten_et_al_2016_appendix.pdf

Ghosh, S., Bandyopadhyay, A., Sahay, S., Ghosh, R., Kundu, I. & Santosh, K. (2021). Colorectal histology tumor detection using ensemble deep neural network. *Eng App Artif Intell*, 100, 104202.

Haenssle, H. A., Fink, C., Schneiderbauer, R., Toberer, F., Buhl, T., Blum A, et al. (2018). Reader study level-I and level-II Groups. Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists. *Ann Oncol*, 29, 1836–1842.

Hohn, J., Krieghoff-Henning, E., Jutzi, T.B., von Kalle, C., Utikal, J.S., Meier, F., et al. (2021). Combining CNN-based histologic whole slide image analysis and patient data to improve skin cancer classification. *Eur J Cancer*, 149, 94–101.

Iizuka, O., Kanavati, F., Kato, K., Rambeau, M., Arihiro, K., Tsuneki, M. (2020). Deep learning models for histopathological classification of gastric and colonic epithelial tumours, *Sci Rep*, 10(1), 1–11.

Ilhan, B., Guneri, P. & Wilder-Smith, P. (2021). The contribution of artificial intelligence to reducing the diagnostic delay in oral cancer. *Oral Oncol*, 116, Article Number: 105254.

Johannet, P., Coudray, N., Donnelly, D.M., Jour, G., Illa-Bochaca, I., Xia, Y., et al. (2021). Using machine learning algorithms to predict immunotherapy response in patients with advanced melanoma. *Clin Cancer Res*, 27, 131–140.

Jothi, J.A.A., Rajam, V.M.A. (2017). A survey on automated cancer diagnosis from histopathology images. *Artif Intell Rev*, 48 (1), 31-81.

Kanavati, F., Toyokawa, G., Momosaki, S., Takeoka, H., Okamoto, M., Yamazaki, K., Takeo, S., Iizuka, O. & Tsuneki, M. (2021). A deep learning model for the classification of indeterminate lung carcinoma in biopsy whole slide images. *Sci Rep*, 11 (1), 1–14.

Kara, A. (2022). Cancer diagnosis with image filter-integrated artificial intelligence algorithms: Innovative possibilities. Research Square. Preprint posted online October 28, 2022. DOI: https://doi.org/10.21203/rs.3.rs-2193871/v1.

Kenner B., Chari, S. T., Kelsen, D. et al. (2021). Artificial intelligence and early detection of pancreatic cancer: 2020 summative review. *Pancreas*, 50(3), 251-279.

Kreouzi, M. et al. (2025). Deep Learning for Melanoma Detection: A Deep Learning Approach to Differentiating Malignant Melanoma from Benign Melanocytic Nevi. *Cancers*, Vol. 17 (1). Article No. 28

Kumar, A., Singh, S.K., Saxena, S., Lakshmanan, K., Sangaiah, A.K., Chauhan, H., Shrivastava, S. & Singh, R.K. (2020). Deep feature learning for histopathological image classification of canine mammary tumors and human breast cancer. *Inf Sci*, 508, 405–421.

Li, X.T., Li, C., Rahaman, M.M. et al. (2022).
A comprehensive review of computer-aided whole-
slide image analysis: from datasets to feature extraction, segmentation, classificati on and detection approaches. *Artif Intell Rev* 55(6): 4809-4878.

Li, T., Xie, P., Liu, J., Chen, M., Zhao, S., Kang, W., et al. (2021). Automated diagnosis and localization of melanoma from skin histopathology slides using deep learning: a multicenter study. *J Healthc Eng*, 5972962.

Mahmood, H., Shaban, M., Rajpoot, N., et al. (2021). Artificial intelligence-based methods in head and neck cancer diagnosis: an overview. *Br J Cancer*, 124, 1934-1940.

Mahmood, H., Shaban, M., Indave, B., Santos-Silva, A., Rajpoot, N., Khurram, S. (2020). Use of artificial intelligence in diagnosis of head and neck precancerous and cancerous lesions: A systematic review. *Oral Oncol*, 110, 104885.

Mayo M. (2022). Image classification (in the online course Advanced data mining with WEKA).
https://www.youtube.com/watch?v=Lhw_XcGCTFg&list=PLm4W7_iX_v4Msh-7lDOpSFWHRYU_6H5Kx

Mio, W. et al. (2025). Establishment of an intelligent analysis system for clinical image features of melanonychia based on deep learning image segmentation. *Comput Med Imaging Graph*, Vol. 123, Article no. 102543.

Phan, D.-V., Chan, C.-L., Li, A.-A., Chien, T.-Y., Nguyen, V.-C. (2020). Liver cancer prediction in a viral hepatitis cohort: A deep learning approach. *Int. J. Cancer*, 147, 2871–2878. https://doi.org/10.1002/ijc.33245.

Tschandl, P. (2021). Artificial intelligence for melanoma diagnosis. *Ital J Dermatol Venerol*, 156, 289–299.

Unnisa, Z. et al. (2025). Impact of fine-tuning parameters of convolutional neural network for skin cancer detection. *Sci Rep*, Vol. 15. Article No. 14779.

Verma, N. & Yadav, D.K. (2025). A comprehensive review on step-based skin cancer detection using machine learning and deep learning methods. *Arch Comput Methods Eng*, DOI: 10.1007/s11831-025-10275-y.

Xie, P., Zuo, K., Liu, J., Chen, M., Zhao, S., Kang, W., et al. (2021). Interpretable diagnosis for whole-slide melanoma histology images using convolutional neural network. *J Healthc Eng*, 8396438.

Witten, I.H. (2022a). More data mining with WEKA. Online course. https://www.youtube.com/watch?v=iqQn6YfyGs0&list=PLm4W7_iX_v4OMSgc8xowC2h70s-unJKCp

Witten, I.H. (2022b). Advanced data mining with WEKA. Online course. https://www.youtube.com/watch?v=Lhw_XcGCTFg&list=PLm4W7_iX_v4Msh-7lDOpSFWHRYU_6H5Kx

Yee, J., Rosendahl, C., Aoude, L.G. (2024). The role of artificial intelligence and convolutional neural networks in the management of melanoma: a clinical, pathological, and radiological perspective. *Melanoma Res*, 34(2), 96-104.

Zhu, H.-B., Xu, D., Ye, M., et al. (2021). Deep learning-assisted magnetic resonance imaging prediction of tumor response to chemotherapy in patients with colorectal liver metastases. *Int. J. Cancer*, 148, 1717–1730. https://doi.org/10.1002/ijc.33427