Predictive analysis on histopathological images using metaheuristics and machine learning method

Aditi Ganapathi, Utsav Sharma,
Saumya Gupta and Arihan Deshwal

Amity University Uttar Pradesh,
Noida, India
Email: aditi.ganapathi.2000@gmail.com
Email: utsavsharma1006@gmail.com
Email: saumyagupta2025@gmail.com
Email: singharihan08@gmail.com

Surbhi Vijh

ASET, Amity University Uttar Pradesh,
Noida, India
Email: surbhivijh428@gmail.com

Sumit Kumar*

Amity University Uttar Pradesh,
Noida, India
Email: sumitkumarbsr19@gmail.com
*Corresponding author

Abstract: The computer-based diagnosis system using histopathology images has always been the centre of research and improvement. Automated histopathological image classification systems have advanced to greater heights as a result of development in digital histopathology for computer-aided diagnosis. Due to the complexity of these images, a high dimension feature map is generated that makes the process difficult. The usage of metaheuristics in the classification of these images has grown due to their remarkable results. To conduct predictive analysis on histopathology images using nature-inspired algorithms and machine learning techniques. The feature extraction methods are applied to determine the statistical and texture-based feature. Furthermore, whale optimisation algorithm (WOA), cat swarm optimisation (CSO), lion optimisation algorithm (LOA), adaptive particle swarm optimisation (APSO), golden eagle optimisation (GEO), hybrid LOA_CSO (H-LOA_CSO) are used as feature selection algorithm for acquiring optimal subset of features. The classification are performed using support vector machine (SVM) and artificial neural network (ANN) to determine the predictive analysis. The observations show that the H-LOA_CSO algorithm performed best with ANN, giving an accuracy of 98%, while APSO showed the worst performance with both ANN and SVM with an accuracy of 72%.

Keywords: nature-inspired algorithms; machine learning algorithms; histopathological images; image enhancement; feature extraction; feature selection; classification ; prediction analysis.

Reference to this paper should be made as follows: Ganapathi, A., Sharma, U., Gupta, S., Deshwal, A., Vijh, S. and Kumar, S. (xxxx) 'Predictive analysis on histopathological images using metaheuristics and machine learning method', *Int. J. Bioinformatics Research and Applications*, Vol. X, No. Y, pp.xxx–xxx.

Biographical notes: Aditi Ganapathi recently earned her Bachelor of Technology (BTech) degree from Amity University, Noida, a highly regarded institution recognised for its emphasis on innovative education and research. Throughout her academic journey, she developed a robust understanding of engineering principles, supported by hands-on experience through various projects, internships, and workshops. Amity University's comprehensive curriculum, which integrates technical expertise with essential soft skills, has thoroughly prepared her for the challenges of the professional world. Her accomplishment is a testament to her dedication, hard work, and passion for learning, qualities that will undoubtedly serve her well as she embarks on her career. Her time at Amity University has laid a solid foundation for success, equipping her with the knowledge and skills needed to excel in her chosen field.

Utsav Sharma recently earned his Bachelor of Technology (BTech) degree from Amity University, Noida, a highly regarded institution recognised for its emphasis on innovative education and research. Throughout his academic journey, he developed a robust understanding of engineering principles, supported by hands-on experience through various projects, internships, and workshops. Amity University's comprehensive curriculum, which integrates technical expertise with essential soft skills, has thoroughly prepared him for the challenges of the professional world. His accomplishment is a testament to his dedication, hard work, and passion for learning, qualities that will undoubtedly serve him well as she embarks on her career. His time at Amity University has laid a solid foundation for success, equipping her with the knowledge and skills needed to excel in his chosen field.

Saumya Gupta recently earned her Bachelor of Technology (BTech) degree from Amity University, Noida, a highly regarded institution recognised for its emphasis on innovative education and research. Throughout her academic journey, she developed a robust understanding of engineering principles, supported by hands-on experience through various projects, internships, and workshops. Amity University's comprehensive curriculum, which integrates technical expertise with essential soft skills, has thoroughly prepared her for the challenges of the professional world. Her accomplishment is a testament to her dedication, hard work, and passion for learning, qualities that will undoubtedly serve her well as she embarks on her career. Her time at Amity University has laid a solid foundation for success, equipping her with the knowledge and skills needed to excel in her chosen field.

Arihan Deshwal recently earned his Bachelor of Technology (BTech) degree from Amity University, Noida, a highly regarded institution recognised for its emphasis on innovative education and research. Throughout his academic journey, he developed a robust understanding of engineering principles, supported by hands-on experience through various projects, internships, and workshops. Amity University's comprehensive curriculum, which integrates technical expertise with essential soft skills, has thoroughly prepared him for the challenges of the professional world. His accomplishment is a testament to his dedication, hard work, and passion for learning, qualities that will undoubtedly serve him well as she embarks on her career. His time at Amity University has laid a solid foundation for success, equipping her with the knowledge and skills needed to excel in his chosen field.

Surbhi Vijh contributing her enormous academics and research works in the area of image processing, machine learning, metaheuristics and soft computing. She has published various research papers in international journals (SCI, Scopus indexed) and international conference.

Sumit Kumar is currently working as an Associate Professor in Amity University, Noida, Uttar Pradesh India. He received his PhD degree from NIT Jamshedpur Jharkhand India. He has more than 20 years of teaching and research experience at repute college and university of India. His area of research is software testing, meta-heuristic algorithms, image processing.

1 Introduction

Several stochastic and deterministic methods have been employed in diagnostic systems for the identification and classification of cancerous cells over the past few years. The stochastic computing methods have been successfully applied to a wide range of real-world issues in a number of different fields (Kaur and Sharma, 2017).

The microscopic examination of cell morphology combined with situated molecular data is referred to as pathology. The work in the field originate from the early 1990s but are just few in number, likely because digital technology has recently widely used in pathology (Bartels et al., 1992; Thiran and Macq, 1996). Numerous cancer detection and classification applications, such as prostate, lung, liver, cervix, breast, and brain cancer classification have been presented as a result of recent developments in digital pathology.

Automated cytology image analysis, another well-known imaging approach, has received significant attention in pathological image analysis. The process is used for both disease screening and biopsy, it frequently comes from the least invasive biopsies. Histopathology slides are prepared in a way that preserves the underlying tissue structure by providing more understanding of disease and its impact on tissues. As a result, automated histopathological image analysis has been widely applied using computational image processing techniques and cutting-edge machine learning methodologies (Gurcan et al., 2009).

The task of identifying the priority areas in the image has several challenges, such as noise removal. The staining of the biopsy samples results in noise, and the uneven distribution of the stain typically results in difficulties in processing the stained material. The second challenge arises to segment the nuclei or cells in the image. The complexity of the image sceneries (such as contacting and overlaying cells) and the clutter make the process difficult (e.g., stain artefacts). The third challenge arises in selecting the features that represent a cell or tissue in the job of quantifying cellular or tissue-level properties. The attributes should offer distinct quantitative measurements to detect cancer automatically. The system evaluation step of the diagnosis process is the also the significant problem. If the generated detail is not carried out appropriately, there may be a significant level of biasness due to the restricted amount of available data.

Nature-inspired algorithms are most widely used in the field of computerised diagnoses. Metaheuristic optimisation algorithms are increasingly become popular in engineering applications due to their ability to solve a wide variety of problems across different fields and their ability to avoid local optima. These techniques were inspired by natural phenomena like humans, insects, birds, animals, water, and the cosmos. The characteristics including autonomous, distributed, emergent, adaptive, and self-organised behaviour of insects, birds, animals, and people are the inspiration for nature-inspired computing (Selvaraj et al., 2014).

In whale optimisation algorithm (WOA) method, the unique way that humpback whales hunt, known as the bubble-net feeding method (Watkins and Schevill, 1979). Hunting krill or small fish near the surface is preferred by humpback whales, has been seen to involve the formation of unique bubbles along a circular or '9'-shaped course (Mirjalili and Lewis, 2016).

One of the most significant swarm intelligence model is particle swarm optimisation (PSO). The PSO directs the particles to look for globally optimal solutions by using a straightforward process that imitates swarm behaviour in fish schooling and avian flocking (Kennedy and Eberhart, 1995). PSO has advanced quickly in recent years because of its simplicity, and it has had numerous successful applications in resolving actual optimisation issues (Eberhart and Shi, 2001; Li and Engelbrecht, 2007). In contrast to PSO, the adaptive particle swarm optimisation (APSO) method has a faster convergence speed and can conduct a global search across the entire search space. A real-time evolutionary state estimation approach used to identify one of the four defined evolutionary stages of exploration, exploitation, convergence, and jumping out in each generation (Zhan et al., 2009).

Cat swarm optimisation (CSO) is created by observing the behaviour of cats, and it is made up of two sub-models that mimic those behaviours: the tracing mode and the searching mode (Chu et al., 2006). Grey wolves (*Canis lupus*) serve as the inspiration for grey wolf optimiser (GWO). The leadership structure and hunting strategy of grey wolves in nature are modelled by the GWO algorithm. Additionally, GWO incorporates the three primary hunting techniques: seeking out prey, encircling prey, and attacking prey (Mirjalili et al., 2014). Another population-based nature-inspired algorithm is the lion optimisation algorithm (LOA). Of all the wild cat species, lions are the most sociable and exhibit both high levels of cooperation and rivalry (Mccomb et al., 1993). For the purpose of creating an optimisation technique, some of these lion traits have been mathematically modelled (Yazdani and Jolai, 2016).

The contribution of the study is to conduct the predictive analysis on machine learning-based model considering breast, lung, spleen and kidney histopathology images. The various nature-inspired algorithms, were used as feature selection algorithms on a set of extracted features. Further, the classification is performed to determine the quality metrics such as accuracy, precision, recall, and F1-score. The results obtained from this research will help to identify the most efficient algorithm for the purpose of quantitative analysis on histopathological images.

2 Related work

Cancer detection using histopathology images has been subjected to a wide study, with experimentation conducted on each step of the process (Hasan et al., 2018). H&E-stained histopathology images have been concluded to be highly effective for computer-aided diagnosis of cancer for breast, lung, and spleen histology images, respectively (Khalid et al., 2019).

Spanhol et al. (2016a) conducted a study on automated breast cancer diagnosis using histopathology images that shows Gaussian filtering was the most commonly used filtering technique in histopathology research. Saha et al. (2017) analysed the performance of cancer classification with and without the use of Gaussian filter to remove noise and concluded that noise removal provided better results. Gandomkar et al. (2018a, 2018b) compared the results obtained on performing mean, median, and Gaussian filtering based on accuracy and F1-score. The observation show that Gaussian filtering was the superior method with an accuracy and F1-score of 88.8% and 0.71 respectively, in comparison to median filtering, which achieved 87.7% accuracy and 0.68 F1-score.

Doyle et al. (2012) evaluated the effectiveness of GLCM, GLDM, and GLRLM as feature extraction techniques in the detection of prostate cancer using digital histopathology images. The examination are performed on dataset of 126 prostate biopsies, GLRLM had the best overall accuracy of 76.8%, followed by GLCM with 75.4% and GLDM with 72.9%. Saha et al. (2018) evaluated various feature extraction techniques, such as GLCM, GLDM, and LBP to be performed on the classification of breast histopathology images. According to their findings, GLDM had the second-best accuracy of 93.23%, while GLCM had the highest accuracy of 94.58%. Similarly, Nagarajan et al. (2016) reported the superiority of GLCM over GLDM in a texture analysis study for splenic images. Tashk et al. (2019) performed a comparative analysis of GLCM, GLDM, and GLRLM method with support vector machine (SVM) for cancer classification using histopathology images. The observation shows that the features extracted through GLRLM provided the best accuracy score of 95.7% with the SVM classifier.

Ahmad et al. (2021) used WOA for feature selection in a breast histology dataset of 306 images. The SVM method gave an accuracy of 96.67%, surpassing k-NN and Naïve Bayes classification results. In a study performed by Karami et al. (2019) on cancer detection using genetic algorithm (GA), firefly algorithm (FA), artificial bee colony (ABC), and cuckoo search (CS) on leukaemia, lung, breast, and colon cancer datasets, it was observed that PSO outperformed the other algorithms in terms of accuracy and F1-score on three of the four datasets when combined with SVM. Similarly, Sharifzadeh and Sabzpoushan (2020) derived that COA outperformed LOA. However, in the comparative evaluation conducted by Wang et al. (2019) on cancer classification using breast histopathology images, LOA was shown to be the best-performing algorithm among PSO and ABC with an accuracy of 95.65%. Alomari et al. (2021) used GWO, PSO, and GA to classify cancer in a dataset of 2,200 breast, lung, liver, and prostate histopathology images. After using GLCM, GLRLM, and GLDM to extract features and using SVM, k-NN, and decision tree classifiers. The examination shows that GWO performed the best with 98.60% accuracy for binary classification and 94.73% accuracy for multi-class classification.

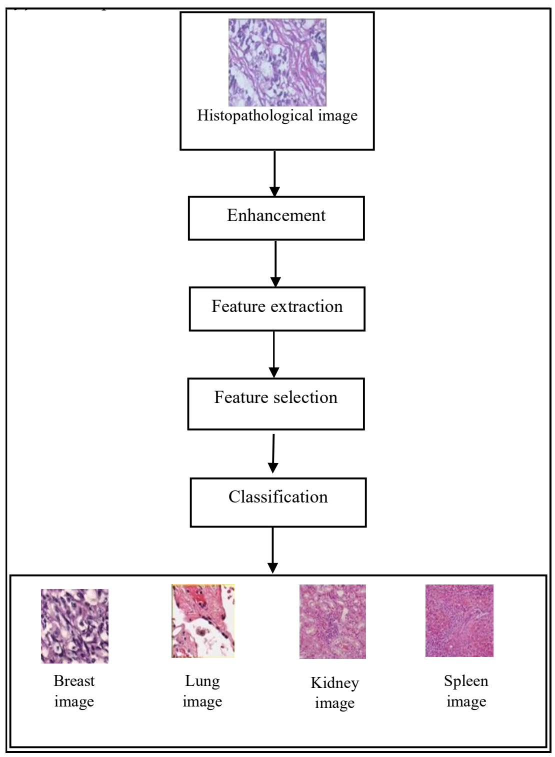
Litjens et al. (2017) conducted a comparative study of machine learning techniques for cancer diagnosis and concluded that SVM, kNN, and ANN had the most widespread application as classification algorithms in the histopathology images sphere. Cruz-Roa et al. (2017) showed that SVM is the superior displaying an accuracy of 90.1% and 90.8% respectively. However, the performance of a supervised machine learning algorithm while classifying cancer highly depends on the dataset, the feature extraction techniques and the feature selection methods (Baid et al., 2019).

Further, many improvised version of nature inspired optimisation methods are developed which can be utilised for optimisation purpose such as improved sand-CSO (Zhang et al., 2023), modified golden eagle optimisation (GEO) (Vijh et al., 2021), SALP swarm optimisation (Xie et al., 2023), etc.

3 Methods

The methodology involves a series of the following steps to categorise the histopathological images: image enhancement, feature extraction, feature selection, and classification. The process is illustrated in Figure 1 and elaborated in this section.

Figure 1 Flowchart of the proposed methodology (see online version for colours)



3.1 *Selection of dataset*

For this study, the total of 400 histopathology images consisting of breast, lung, kidney and spleen. The breast histopathology images dataset (Janowczyk and Madabhushi, 2016) were selected from Kaggle. Initially, the dataset comprised 162 whole-mount slide images of invasive ductal carcinoma (IDC) specimens that were examined at 40× magnification. BreakHis (Spanhol et al., 2016a) is a collection of 7,909 H&E-stained

breast tumour tissue samples captured at a 20× magnification. Lung, spleen and spleen images were collected from the ADL dataset (Srinivas et al., 2014).

3.2 Image pre-processing

To minimise arbitrary noise and to enhance the contrast in the image, the collection of images was subject to the Gaussian filtering technique in order to reduce noise from the image. The method is recognised for its ability to preserve the edges along with the fine details of the images. The working of the Gaussian filter involves applying a kernel that is determined using the Gaussian distribution function. The kernel comprises of a set of weights which determine the contribution of each pixel in the image to the final output image. The continuous function is used in Gaussian filter as shown in equation (1)

$$G(x, y) = \left(1/\left(2 * \pi * \sigma^2\right)\right) * \left(e^{-((x^2+y^2)/(2*\sigma^2))}\right) \tag{1}$$

In the above formula, $G(x, y)$ represents the value of the Gaussian function at point (x, y), and σ represents the standard deviation of the Gaussian distribution.

3.3 Feature extraction

For this study, grey-level co-occurrence matrix (GLCM), grey-level difference method (GLDM), and grey-level run length matrix (GLRLM) were chosen. GLCM computes the spatial dependence of the grey values of an image while GLDM calculates the difference known as grey level difference (GLD) between the values of the adjacent pixels within every square region. These GLD values are utilised to create a matrix known as the GLDM matrix, which contains the information of the texture of the image.

GLRLM is a matrix that is used to measure the distribution of runs of pixels with the same grey-level value in an image. GLRLM counts the runs of pixels for same grey-level value in each block and stores this count in the matrix.

3.4 Feature selection

Six nature-inspired algorithms were employed to select features from the ones extracted in the above step, by using APSO, WOA, CSO, LOA, GEO and hybrid loin and cat swarm optimisation (H-LOA_CSO).

3.4.1 Adaptive particle swarm optimisation

Each particle in swarm represents a potential solution and is identified by its location as well as velocity within the search space. A series of computation incorporate a self-adaptive technique that modifies the search radius in response to the movements of the particles is employed to modify the positions of each particle. By calculating the local search space and continuously modifying its search radius according to the variance of the particle positions, the self-adaptive system seeks to strike a balance between exploration and exploitation. Based on both the swarm's best solution and each particle's current best solution, each particle's velocity is updated. The particle velocities are updated by the algorithm randomly using an inertia weight that decreases linearly.

3.4.2 Whale optimisation algorithm

The WOA metaheuristic optimisation algorithm models its hunting behaviour after the social behaviour of humpback whales. The locations of whales in the field of search are used by WOA to indicate a population of possible solutions. Each whale has three different behaviour modes that it can use to determine where it should be in the search area: seek, encircle, and attack. In the seek method, the whale roams the search space at random to discover new regions. The whale positions itself around the victim with the encircle mode, which is the best method thus far. In attack mode, the whale advances on the target in search of a better answer. The calculations are performed to include a randomly linearly increasing function, the whale's location is adjusted.

3.4.3 Cat swarm optimisation

CSO mimics the behaviour of cats in order to find the optimised solution to a problem by dividing it into searching mode and trailing mode. The searching mode defines behaviours like resting and looking for the subsequent location to move to, while the trailing mode is when the cats are looking for some targets. In the searching mode, k duplicates of the i^{th} cat are created of which one copy is retained, while the positions of the remaining $k - 1$ duplicates are arbitrarily changed by the addition or subtraction of the SRD. Here, SRD is seeking range of selected dimensions, which gives the mutative fraction for the chosen dimension. CDC is the counts of dimension to change; this gives the total dimensions that can be changes. In tracing mode, the cats tend to move towards the targets. Each cat is assigned its own velocity for each dimension when in trailing mode.

3.4.4 Lion optimisation algorithm

The initial phase of the LOA, a population-based meta-heuristic method, is to produce a random population across the solution space. The initial population is made up of a collection of randomly generated solutions known as lion. In the initial population, N% lions are chosen to be nomads, while the other lions (resident lions) are randomly divided into P subgroups known as the pride. The rest of the members of the pride are thought to be male, whereas S% of the members of the pride are thought to be female; this ratio is contrary to nomad lions.

The best-reached solution for each lion is referred to as the best-visited position, and it is gradually updated throughout the optimisation process.

3.4.5 Golden eagle optimisation

Each search agent selects a target prey from the flock's collective memory for each iteration. Then, for each golden eagle, attack and cruise vectors are determined in relation to the chosen prey. Based on the assault vector, the cruise vector is determined. The assault and cruise vector make up the golden eagles' displacement. Golden eagles exhibit a greater propensity to cruise early in the hunting flight and a greater propensity to attack late in the flight, which corresponds to greater exploration in the early iterations and greater exploitation in the late iterations in the suggested optimiser.

3.4.6 Hybrid loin and cat swarm optimisation

H-LOA_CSO is developed by considering the merit of both loin and cat swarm optimisation. Due to social framework, helpful action, and awareness, hybrid LOA CSO (H-LOA CSO) is favourable (Vijh et al., 2023). The fusion of two algorithm improves the exploration and exploitation stages and enables to provide better solution.

3.5 Classification

SVM and ANN method are considered as the two classification method for classification. SVM are effective for dealing with linear in addition to nonlinear data, but ANN try to simulate how the brain functions to tackle challenging situations. A hyperplane is used in SVM to segregate multidimensional data (Wenzhuo and Shuo, 2023). It makes use of kernel functions to systematically locate higher-dimensional support vector classifiers. A branch of computing called artificial neural networks (ANNs) imitates how the human brain processes information. By examining the provided dataset, ANN has the ability to identify patterns and other things independently.

4 Results

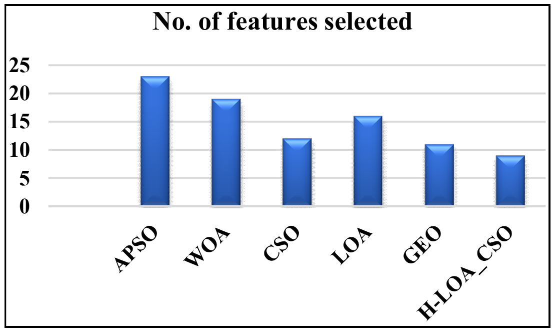
The experimental analysis conducts the predictive analysis of cancer presence in histopathology images using metaheuristics. The feature extraction and feature selection values acquired from algorithm are described in Table 1.

From Table 1, it is observed that GLCM extracted the highest number of features (24) while GLDM and GLRLM extracted 4 and 7 features each, thereby leading to a total of 35 extracted features. Among the feature selection algorithms, it is noted that APSO has selected 23 feature, is the maximum. WOA and LOA have selected 19 and 16 features, respectively. GEO, having selected 11 out of the 35 extracted features. H-LOA_CSO extracted nine relevant and least number of features. Figure 2 is a graphical representation of selected features.

Table 1 Selected features

Nature-inspired algorithm	*No. of features extracted*				*No. of features selected*
	GLCM	*GLDM*	*GLRLM*	*Total*	
APSO	24	4	7	35	23
WOA					19
CSO					12
LOA					16
GEO					11
H-LOA_CSO					9

Figure 2 Feature selection results (see online version for colours)



The predictive performance analysis of the accuracy, precision, recall, and F1-score of each meta-heuristic algorithm using SVM and ANN has been evaluated in Table 2 and Table 3, respectively.

Table 2 SVM performance metrics using nature-inspired algorithm

Classifier	*Measures (overall)*	*APSO*	*WOA*	*CSO*	*LOA*	*GEO*	*H-LOA_CSO*
Support vector machine (SVM)	Accuracy	0.72	0.79	0.86	0.83	0.95	0.97
	Precision	0.60	0.73	0.80	0.74	0.88	0.91
	Recall	0.63	0.75	0.85	0.84	0.98	0.98
	F1-score	0.67	0.76	0.79	0.81	0.90	0.92

Table 3 ANN performance metrics using nature-inspired algorithm

Classifier	*Measures (overall)*	*APSO*	*WOA*	*CSO*	*LOA*	*GEO*	*H-LOA_CSO*
Artificial neural networks (ANN)	Accuracy	0.72	0.89	0.91	0.95	0.95	0.98
	Precision	0.63	0.83	0.88	0.94	0.97	0.99
	Recall	0.62	0.84	0.90	0.94	0.95	0.97
	F1-score	0.68	0.85	0.92	0.93	0.97	0.98

On observing the overall performance of the nature-inspired algorithms with the SVM and the ANN classifier, it is seen that ANN has yielded significantly better results across all the metrics considered. H-LOA_CSO method with ANN classifier is found to be the best-performing combination, yielding the highest metrics with 98% accuracy, 99% precision, 97% recall, and 98% F1-score. In contrast, APSO-SVM has shown the least predictive capacity, with 72% accuracy, 60% precision, 63% recall, and 67% F1-score.

While the difference in predictive performance is less for GEO and APSO with ANN and SVM, there is a visibly large forward leap for WOA, CSO, and LOA. ANN were found to be the better classifier by about 10% in all the metrics under consideration.

The tabulated results have been formulated as bar charts for pictorial analysis in Figures 3 and 4 are depicted.

Figure 3 Overall performance metrics with SVM (see online version for colours)

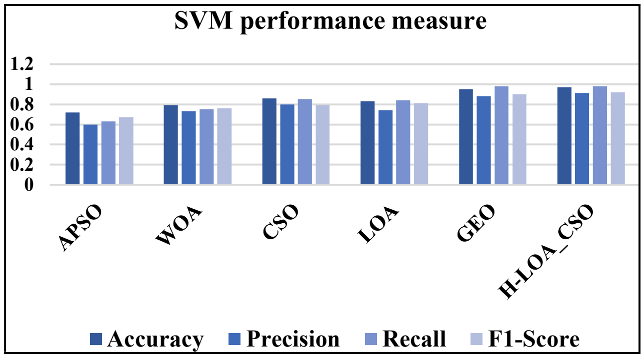
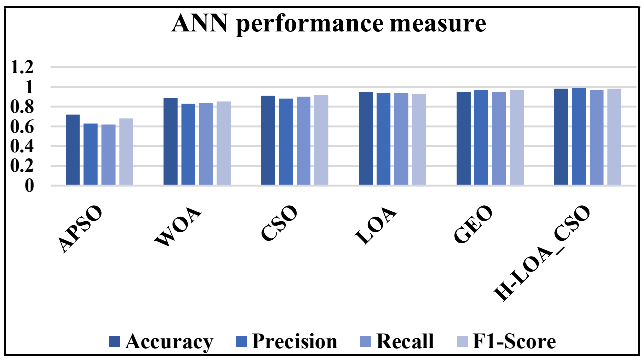


Figure 4 Overall performance metrics with ANN (see online version for colours)



H-LOA_CSO has resulted in the best performance overall, the algorithm has selected the least number of features (9), whereas APSO had chosen the highest number of features (23), has shown the worst performance. GEO has selected 11 features and shown good performance outcome. WOA having selected 19 features thus has yielded the penultimate worst performance. CSO and LOA selected 12 and 16 features each, and in the SVM chart, it is observed that CSO shows better results. However, these results have been flipped as per the ANN chart, where LOA, despite selecting more features than CSO, has the better performance.

5 Conclusions and future scope

Histopathology is a key component of contemporary medicine, allowing doctors to decipher the riddles of cancer by delving into the tiny realm of tissue samples, thereby improving patient outcomes and saving numerous lives. An enormous amount of research has been done for computerised diagnosis system during the past years. The automated diagnosis model has enormous promise for widespread application in providing advanced treatment at early stage of disease. The nature inspired algorithms are utilised because of their versatility, adaptability, and simplicity and solving real-world problem across many fields specially in medical imaging effectively. In medical field, classification algorithms are crucial because they enable the effective analysis and comprehension of intricate medical data, assisting in the precise diagnosis and treatment of a variety of diseases, improving patient outcomes, and expanding medical knowledge. From the study, it is concluded that H-LOA_CSO gave the highest classification algorithms by providing the accuracy of 97% with SVM and 98% with ANN, while APSO showed the lowest accuracy at 72% with both classifiers. To further optimise the results, hybrid algorithms will be designed in future to incorporate the best features of multiple algorithms.

References

Ahmad, I., Iqbal, N. and Khan, A. (2021) 'Cancer classification in histopathology images using whale optimization algorithm', *International Journal of Advanced Computer Science and Applications*, Vol. 12, No. 3, pp.1–10.

Alomari, O.A., AlRyalat, S.A. and Alkafaween, E. (2021) 'Cancer classification in histopathology images using genetic wolf optimization', *International Journal of Computational Intelligence Systems*, Vol. 14, No. 1, pp.1497–1507.

Baid, U., Choudhary, R. and Raja, K. (2019) 'A comparative study of feature extraction techniques for classification of breast cancer histopathological images', *Journal of Medical Systems*, Vol. 43, No. 8, pp.1–9.

Bartels, P.H., Thompson, D., Bibbo, M. and Weber, J.E. (1992) 'Bayesian belief networks in quantitative histopathology', *Analytical and Quantitative Cytology and Histology*, Vol. 14, No. 6, pp.459–473.

Chu, S.C., Tsai, P.W. and Pan, J.S. (2006) 'Cat swarm optimization', in *PRICAI 2006: Trends in Artificial Intelligence: 9th Pacific Rim International Conference on Artificial Intelligence, Proceedings*, Springer, Berlin, Heidelberg, Guilin, China, 7–11 August, Vol. 9, pp.854–858.

Cruz-Roa, A., Basavanhally, A., González, F.A., Gilmore, H., Feldman, M., Ganesan, S. and Madabhushi, A. (2017) 'Breast cancer histopathological image classification using convolutional neural networks and ensemble learning', *IEEE Transactions on Multi-Scale Computing Systems*, Vol. 3, No. 4, pp.208–218.

Doyle, S., Agner, S., Madabhushi, A., Feldman, M. and Tomaszewski, J.E. (2012) 'Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features', *IEEE Transactions on Biomedical Engineering*, Vol. 59, No. 3, pp.649–656.

Eberhart, R.C. and Shi, Y.H. (2001) 'Particle swarm optimization: developments, applications, and resources', in *Proceedings of the IEEE Congress on Evolutionary Computation*, Seoul, Korea, pp.81–86.

Gandomkar, Z., Hoogi, A. and Chang, P. (2018a) 'Impact of different types of noise on performance of convolutional neural networks for CT datasets', *Radiology*, Vol. 286, No. 1, pp.938–945.

Gandomkar, Z., Hoogi, A., Lu, L., Wong, K.P. and Samuelson, F.W. (2018b) 'Computer aided detection and diagnosis of lung cancer: a review', *Journal of Medical Systems*, Vol. 42, No. 7, p.129.

Gurcan, M.N., Boucheron, L.E., Can, A., Madabhushi, A., Rajpoot, N.M. and Yener, B. (2009) 'Histopathological image analysis: a review', *IEEE Reviews in Biomedical Engineering*, Vol. 2, pp.147–171.

Hasan, M., Rahman, M.M. and Khandakar, A. (2018) 'Histopathological image analysis for cancer detection and diagnosis using machine learning: a review', *Journal of Medical Systems*, Vol. 42, No. 11, p.226.

Janowczyk, A. and Madabhushi, A. (2016) 'Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases', *Journal of Pathology Informatics*, Vol. 7, No. 1, p.29.

Karami, J., Mohammadi, M. and Safdari, R. (2019) 'A comparative study of nature-inspired feature selection methods for cancer diagnosis using gene expression data', *Journal of Medical Signals and Sensors*, Vol. 9, No. 2, pp.69–79.

Kaur, P. and Sharma, M. (2017) 'A survey on using nature inspired computing for fatal disease diagnosis', *International Journal of Information System Modeling and Design*, Vol. 8, No. 2, pp.70–91.

Kennedy, J. and Eberhart, R. (1995) 'Particle swarm optimization', in *Proceedings of ICNN '95 – International Conference on Neural Networks*, IEEE, November, Vol. 4, pp.1942–1948.

Khalid, M., Arsalan, M., Mahmood, T. and Iqbal, N. (2019) 'Histopathological image analysis for breast cancer using transfer learning and improved deep convolutional neural network', *Journal of Medical Systems*, Vol. 43, No. 10, p.317.

Li, X. and Engelbrecht, A.P. (2007) 'Particle swarm optimization: an introduction and its recent developments', in *Proceedings of the 9th annual conference companion on Genetic and evolutionary computation*, July, pp.3391–3414.

Litjens, G., Kooi, T., Bejnordi, B.E., Setio, A.A.A., Ciompi, F., Ghafoorian, M., van der Laak, J.A., van Ginneken, B. and Sánchez, C.I. (2017) 'A comprehensive review of deep learning for image analysis in cancer diagnosis, segmentation, and treatment', *Medical Image Analysis*, Vol. 42, pp.60–88.

Mccomb, K., Pusey, A., Packer, C. and Grinnell, J. (1993) 'Female lions can identify potentially infanticidal males from their roars', *Proceedings of the Royal Society of London, Series B: Biological Sciences*, Vol. 252, No. 1333, pp.59–64.

Mirjalili, S. and Lewis, A. (2016) 'The whale optimization algorithm', *Advances in Engineering Software*, Vol. 95, pp.51–67.

Mirjalili, S., Mirjalili, S.M. and Lewis, A. (2014) 'Grey wolf optimizer', *Advances in Engineering Software*, Vol. 69, pp.46–61.

Nagarajan, R., Rajeswari, R. and Raja, N.S. (2016) 'Texture analysis for classification of splenic histopathology images', *Journal of Medical Systems*, Vol. 40, No. 8, p.190.

Saha, M., Chakraborty, C. and Nasipuri, M. (2017) 'Comparative study of texture analysis methods for cancer detection in histopathological images', in *2017 IEEE Region 10 Symposium (TENSYMP)*, IEEE, pp.1–6.

Saha, M., Chakraborty, C. and Nasipuri, M. (2018) 'Development of a computer-aided diagnostic system for detection and classification of lung nodules in computed tomography images', *Journal of Medical Systems*, Vol. 42, No. 6, p.p109.

Selvaraj, C., Kumar, R.S. and Karnan, M. (2014) 'A survey on application of bio-inspired algorithms', *International Journal of Computer Science and Information Technologies*, Vol. 5, No. 1, pp.366–370.

Sharifzadeh, M. and Sabzpoushan, S.H. (2020) 'A comparative study of classifiers for breast cancer diagnosis based on fine needle aspiration biopsy', *Journal of Medical Signals and Sensors*, Vol. 10, No. 2, pp.125–133.

Spanhol, F., Oliveira, L.S., Petitjean, C. and Heutte, L. (2016a) 'A dataset for breast cancer histopathological image classification', *IEEE Transactions on Biomedical Engineering*, Vol. 63, No. 7, pp.1455–1462.

Spanhol, F.A., Oliveira, L.S., Petitjean, C., Heutte, L. and Fernandes, A.A. (2016b) 'Breast cancer histopathological image classification using convolutional neural networks', in *International Joint Conference on Neural Networks*, IEEE, pp.2560–2567.

Srinivas, U., Mousavi, H.S., Monga, V., Hattel, A. and Jayarao, B. (2014) 'Simultaneous sparsity model for histopathological image representation and classification', *IEEE Transactions on Medical Imaging*, Vol. 33, No. 5, pp.1163–1179.

Tashk, A., Gity, M. and Soltanian-Zadeh, H. (2019) 'A comparative study of feature extraction and selection for cancer classification in histopathology images', *Computer Methods and Programs in Biomedicine*, Vol. 178, pp.97–108.

Thiran, J.P. and Macq, B. (1996) 'Morphological feature extraction for the classification of digital images of cancerous tissues', *IEEE Transactions on Biomedical Engineering*, Vol. 43, No. 10, pp.1011–1020.

Vijh, S., Kumar, S. and Saraswat, M. (2021) 'Efficient feature selection method for histopathological images using modified golden eagle optimization algorithm', in *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, IEEE, September, pp.1–5.

Vijh, S., Saraswat, M. and Kumar, S. (2023) 'Automatic multilevel image thresholding segmentation using hybrid bio-inspired algorithm and artificial neural network for histopathology images', *Multimedia Tools and Applications*, Vol. 82, No. 4, pp.4979–5010.

Wang, H., Liu, X., Li, S. and Li, J. (2019) 'An improved approach to cancer classification based on particle swarm optimization and artificial bee colony', *Journal of Healthcare Engineering*, Vol. 2019, pp.1–11.

Watkins, W.A. and Schevill, W.E. (1979) 'Aerial observation of feeding behaviour in four baleen whales: *Eubalaena glacialis*, *Balaenoptera borealis*, *Megaptera novaeangliae*, and *Balaenoptera physalus*', *Journal of Mammalogy*, Vol. 60, No. 1, pp.155–163.

Wenzhuo, Y. and Shuo, L. (2023) 'Optimizing parameters of support vector machines using an enhanced whale optimization algorithm', in *2023 2nd Asia Conference on Electrical, Power and Computer Engineering (EPCE)*, IEEE, April, pp.151–157.

Xie, X., Xia, F., Wu, Y., Liu, S., Yan, K., Xu, H. and Ji, Z. (2023) 'A novel feature selection strategy based on SALP swarm algorithm for plant disease detection', *Plant Phenomics*, Vol. 5, p.0039, DOI: 10.34133/plantphenomics.0039.

Yazdani, M. and Jolai, F. (2016) 'Lion optimization algorithm (LOA): a nature-inspired metaheuristic algorithm', *Journal of Computational Design and Engineering*, Vol. 3, No. 1, pp.24–36.

Zhan, Z.H., Zhang, J., Li, Y. and Chung, H.S.H. (2009) 'Adaptive particle swarm optimization', *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, Vol. 39, No. 6, pp.1362–1381.

Zhang, Y., Dai, J., Huang, Q. and Luo, Y. (2023) 'An improved sand cat swarm optimization and its application in aluminum melting process', in *2023 5th International Conference on Intelligent Control, Measurement and Signal Processing (ICMSP)*, IEEE, May, pp.380–384.