Pixel-Based Image Classification using a Grey Wolf Optimised Support Vector Machine*

[1]M. B. Poku, [1]I. Yakubu, and [1]Y. Y. Ziggah
[1]University of Mines and Technology, Tarkwa, Ghana

Poku, M. B., Yakubu, I. and Ziggah, Y. Y. (2024), "Pixel Based Image Classification Approach using a Grey Wolf Optimised Support Vector Machine", *Ghana Mining Journal,* Vol. 24, No. 1, pp. 12-26.

Abstract

Support Vector Machine (SVM) is one of the most effective machine learning algorithms widely employed for classification tasks. SVMs perform well in high-dimensional spaces, making them suitable for applications with a large number of features. This capability is crucial in tasks like image classification, where each pixel can represent a feature. Its effectiveness has made it a preferred choice among remote sensing experts. However, the performance of the SVM is highly dependent on the appropriate selection of the best combination of hyperparameters. Thus, optimisation is an essential step for maximising classification accuracy. This paper explores a metaheuristic optimisation algorithm, the Grey Wolf Optimisation Algorithm (GWO), to optimise the performance of the SVM by fine-tuning the optimal combination of hyperparameters that can improve the accuracy of the SVM. With an accuracy of 92%, the GWO-optimised SVM confirms its superiority compared to the standalone SVM, which obtained an accuracy of 89%. The findings of this research highlight the potential of metaheuristic algorithms in improving the effectiveness of machine learning algorithms for image classification tasks.

Keywords: Support Vector Machine, Hyperparameter Optimisation, Image Classification, Metaheuristic, Grey Wolf Algorithm.

1 Introduction

Satellite image classification is a fundamental component in remote sensing and geospatial analysis, holding paramount importance in land use and cover mapping (Alqurashi and Kumar, 2013; Fan *et al*., 2023). It involves categorising a satellite image's pixels according to the type of land cover they represent using their spectral properties (Rudrapal and Subhedar, 2015). This field of remote sensing technology has generated attention in a variety of applications, leading to the development of several algorithms (Huang *et al*., 2002).

Traditionally, land cover classification predominantly relied on statistical methods such as Maximum Likelihood Classification (MLC) (Ahmad and Quegan, 2012), Minimum Distance to Mean (MDM), and Markov Random Field (MRF) (Yang *et al*., 2013), among others. However, the advent of machine learning (ML) has revolutionised this field by introducing more versatile and powerful algorithms. Researchers have since examined numerous supervised classification algorithms, including Random Forest (Shihab *et al*., 2020; Piao *et al*., 2021; Amini *et al*., 2022; Thakur and Panse, 2022; Cengiz *et al*., 2023), Decision Trees (Hamad, 2020), K-Nearest Neighbour (Thakur and Panse, 2022), Artificial Neural Networks (Hamad, 2020; e Silva *et al*., 2020; Shihab *et al*., 2020), and Support Vector Machines (Pal and Mather, 2005; Prasad *et al*., 2017; Cengiz *et al*., 2023). These supervised ML algorithms have shown exceptional performance in handling the intricate patterns in remote sensing data. The algorithms' inherent ability to capture non-linear relationships and adapt to complex feature spaces makes them strong candidates for land cover classification. Unlike traditional classifiers like the MLC, these algorithms can work with both balanced and imbalanced datasets. This makes them better at dealing with the classification uncertainties associated with traditional models (Yuh *et al*., 2023).

Among the many machine learning algorithms employed for satellite image classification, the Support Vector Machine (SVM) holds a distinct position.SVM's capacity to delineate complex decision boundaries, handle high-dimensional data, and exhibit robust generalisation performance makes it a preferred choice among remote sensing experts(Abbas and Jaber, 2020; Basheer *et al.*, 2022; Tamirat *et al.*, 2023). While SVM offers substantial potential, its performance heavily relies on the appropriate selection of hyperparameters (Mantovani *et al*., 2015). These hyperparameters determine the SVM's ability to generalise from training data to unseen samples, impacting its classification accuracy and robustness (Ramasamy *et al.*, 2021). Therefore, it is essential to optimise the hyperparameters of the SVM to maximise its effectiveness for satellite image classification tasks.

Traditional methods for hyperparameter optimisation, such as grid and random searches, have been commonly used to fine-tune machine learning model hyperparameters. However, these

**Manuscript received February 21, 2024*
Revised version accepted June 25, 2024
https://doi.org/10.4314/gm.v22i1.3

techniques tend to converge slowly and are prone to getting stuck in local minima, resulting in lower precision (Tamimi *et al*, 2017). Additionally, they demand significant computational resources, particularly in high-dimensional search spaces, which reduces their efficiency (Li *et al.*, 2015).

An intuitive approach is the use of metaheuristic algorithms. Inspired by natural processes, these algorithms offer a more efficient and effective way to search for optimal hyperparameters. One such algorithm is Grey Wolf Optimisation (GWO), a relatively new metaheuristic approach that draws inspiration from the social hierarchy and hunting behaviour of grey wolves. GWO exhibits superior exploration-exploitation capabilities and has successfully optimised complex functions (Negi *et al.*, 2021).

Consequently, this work proposes an SVM-GWO model for land cover classification. The principal aim is to utilise the GWO algorithm to fine-tune the selection of SVM hyperparameters. The contributions of this paper to existing literature are:

- Investigate the performance of GWO for optimising the hyperparameters of SVM for land cover classification; and
- Evaluate the performance of the proposed SVM-GWO model with conventional SVM

1.1 Justification

In hyperparameter optimisation for image classification, the selection of an appropriate metaheuristic algorithm plays a pivotal role in determining the efficiency and effectiveness of the optimisation process (Kuo *et al.*, 2018). Over the last few years, advancements in metaheuristic algorithms have motivated numerous researchers to apply them to various optimisation problems (Ali *et al.*, 2017). Metaheuristic algorithms offer several advantages over conventional methods, including their derivative-free nature, lack of constraints on problem formulation, and adaptability to a diverse range of real-world problems(Tomar *et al.*,2023). Several metaheuristic algorithms have been applied to image classification tasks. For example, a hybridised Genetic Algorithm Particle Swarm Optimisation (GAPSO) was used to effectively tune the hyperparameters of the SVM (Cheng and Bao, 2014)

Among various metaheuristic algorithms, Grey Wolf Optimisation (GWO) stands out as a superior choice, offering compelling advantages that make it a better choice for the complex nature of image classification (Askarzadeh, 2013).GWO's inspiration from the social hierarchy and hunting behaviour of grey wolves in nature provides a unique foundation, enabling the emulation of cooperative and competitive dynamics observed in the natural world. This inherent ability contributes to GWO's exceptional balance between exploration and exploitation (Hatta *et al.*, 2019), a crucial aspect for effectively navigating high-dimensional search spaces encountered in hyperparameter optimisation for image classification. Notably, GWO excels in avoiding premature convergence to local minima, a common challenge in conventional optimisation methods and some metaheuristic algorithms. With fewer tuning parameters, specifically population and iterations, GWO stands out for its simplicity compared to other algorithms, which require more intricate parameter adjustments (Mirjalili *et al.*, 2014). Its adaptability to complex functions, particularly in the context of non-linear relationships inherent in image classification tasks, further underscores GWO's efficacy. The reasons stated above justify the choice of the GWO to optimise the SVM in this paper.

1.2 Review of Related Works

Google Scholar Advanced Search and Science Direct were adopted to identify the various works that had been conducted. In order to access the vast knowledge and methodologies adopted in image classifications, the search focused on manuscripts from 2010 to 2023, even though a few articles from earlier years were included based on their relevance to the subject. The keywords that were used to filter manuscripts included "Image Classification" or "Pixel Based Classification" and "Metaheuristic Algorithms" or "Hyperparameter Optimisation". Also, the searches were limited to peer-reviewed journals. For the screening stage, abstracts of the selected articles were analysed to ascertain if the write-up matched up to the focus of this paper. Articles that were unrelated to the research question were thus excluded. This section of the paper delves into the evolving landscape of pixel-based image classification, focusing on the hybridisation of machine learning and optimisation algorithms.

Generally, there are two image classification methods: unsupervised and supervised. The unsupervised classification methods do not have a priori information on the dataset. Classes are formed based on the pixel characteristics of the dataset itself (Wang *et al.*, 2017). Some examples include K-Means Clustering (Venkateswaran *et al.*, 2013; Wang *et al.*, 2013; Vishwanath *et al.*, 2016), the Iterative Self-Organising Data Analysis Training Algorithm (ISODATA) (Abbas *et al.*, 2016), the Fuzzy Clustering Algorithm (HongLei *et al.*, 2013), and Random Forest Theory (PeerBhay *et al.*, 2015). However, these techniques require prior knowledge about the dataset for supervised classification. Research reveals that supervised techniques yield better results than unsupervised ones (Hasmadi *et al.*, 2009). Thus, supervised classification techniques have been utilised by many researchers

**Manuscript received January 19, 2024*
Revised version accepted June 23, 2024
https://doi.org/10.4314/gm.v22i1.1

in the field of remote sensing. Many research works have been conducted in the literature on remote sensing image classification techniques and optimisation algorithms for image classification techniques.

For instance, Elmanai *et al*. (2013) researched the best classifier for multispectral data by comparing three classifiers. In their research, they concluded that the SVM, which obtained an accuracy of 88.35%, was the best technique for classifying multispectral data as compared to the K-Means and the Minimum Distance classifiers, which obtained accuracies of 51.27% and 85.01%, respectively. Research was also conducted by Mondal *et al*. (2012) to assess the accuracy of the Support Vector Machine and Maximum Likelihood Classifiers. They applied the classifiers to detect and assess land cover change along the Birupa basin in Odisha. The spatial similarity and area statistics were assessed between the two classifiers, and the results obtained demonstrated that the SVM classifier provides better results than MLC.

Zhu *et al*. (2019) examined the performance of three intelligence algorithms: artificial bee colony (ACO), particle swarm optimisation (PSO) and genetic algorithms (GA). The research evaluated the performance of the three algorithms on SVM. Three hyperspectral images were used, and the optimisers' convergence rate, sample size, feature selection, parameter settings and classification accuracies were compared. The GA was observed to be more robust regarding sample size and feature selection.

Xue *et al*. (2014) also introduced a novel hyperspectral image classification method, HA-PSO-SVM, which combines harmonic analysis (HA), particle swarm optimisation (PSO), and support vector machine (SVM). Initially, HA transforms pixels into the frequency domain, enhancing features for classification. Subsequently, PSO optimises SVM parameters (penalty parameter C and kernel parameter), improving classification performance.

Wang *et al*. (2017) presented a classification model for remotely sensed images using the optimal SVM and modified binary ant colony algorithm. The results obtained were compared with an optimal SVM optimised with the binary-coded particle swarm optimisation (BPSO), binary-coded ant colony optimisation (BACO), and binary-coded cuckoo search (BCS) algorithms. A general observation was that swarm intelligence and evolutionary algorithms perform better at increasing the performance accuracy of SVM.

The literature review illustrates the efficiency of metaheuristics in optimising hyperparameter selection. However, while existing studies have demonstrated the effectiveness of metaheuristic algorithms in optimising hyperparameters of ML models, the exploration of new metaheuristics remains a vital area of research. The No Free Lunch (NFL) theorem, proposed by Ho and Pepyne (2002), posits that no single metaheuristic algorithm can outperform all others across a wide range of problem domains. Consequently, researchers are continually developing new metaheuristic algorithms, variants, and hybrid techniques that may potentially surpass existing methods or offer unique advantages As a result, scientists are continually creating new metaheuristic algorithms, variations, and hybrid approaches that could outperform current ones. (Nsiah *et al.*, 2023).

2 Study Area and Methodology

2.1 Study Area

The Greater Accra Region, which houses Accra, the capital city of Ghana, is one of Ghana's most rapidly urbanising regions. The region serves as the country's capital hub and houses its most prominent seaport and international airport, making it an important international gateway. The region has attracted many investors, resulting in an influx of immigrants, primarily searching for greener pastures. This influx of immigrants has resulted in the need for space, mainly for settlement, thus resulting in the rapid change in the land cover and land use of the area. The land cover of the study area is a heterogeneous mix of natural and man-made features, including water bodies, vegetation, bare lands, and built-up areas. For the training and testing of the model, 2 440 ground truth points were randomly picked using the Google Earth Pro software. The sample points were split into training and test datasets in the 0.8: 0.2 ratio. Fig. 1 shows the map of the study area.

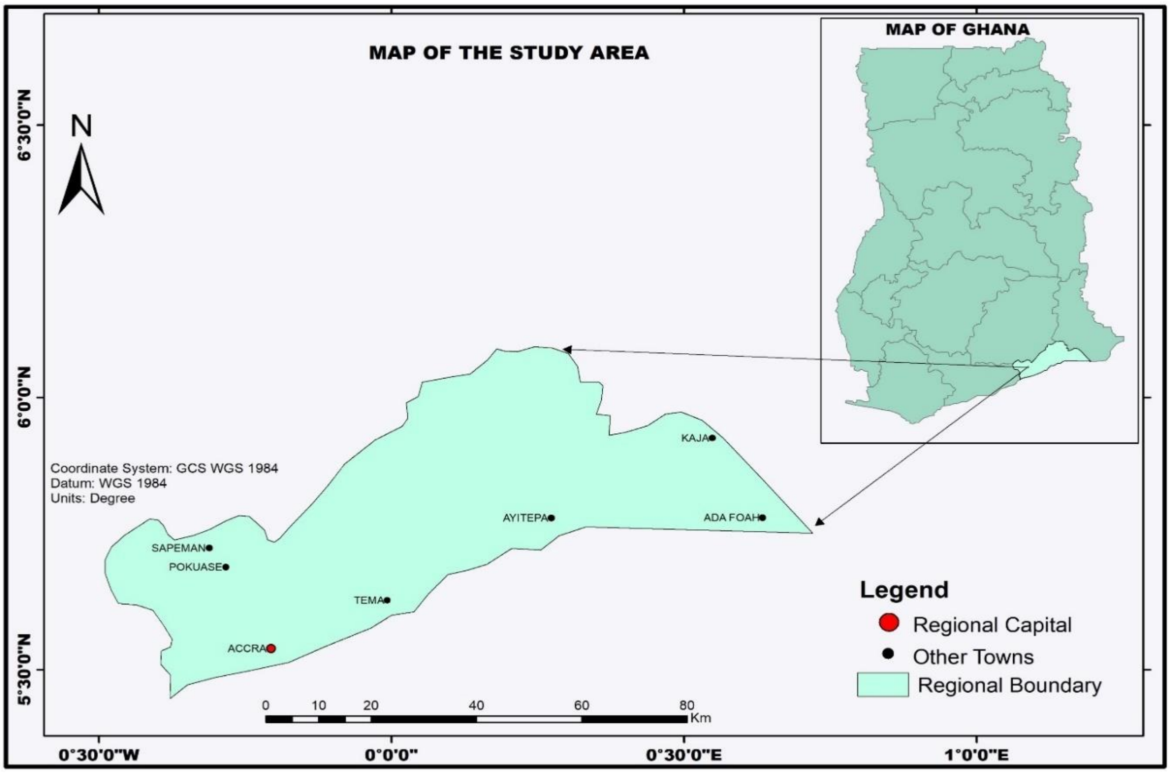

Fig. 1 Map of Greater Accra Region

2.2 Methods Used

This section elaborates on the various methods employed in the study. This includes the data used, training sample generation, parameter tuning, model formulation, assessment and classification. A flowchart of the adopted methodology is provided in Fig. 2.

2.2.1 Data Used

The dataset used in this study was a multispectral Landsat Image of the year 2021, obtained from the USGS Earth Explorer website (https://earthexplorer.usgs.gov). For the year under consideration, the Landsat 8 satellite imagery was obtained.. Landsat 8 provides high-quality satellite imagery with various spectral bands that are valuable for remote sensing and land cover classification tasks. For this study, six out of the eleven bands were selected. The bands were selected based on their relevance to the classification task in this study. Table 1 shows the specific bands that were selected and their relevance to the classification task. Radiometric and reflectance corrections were performed to correct all atmospheric errors. To make the data more convenient for data analysis and manipulation, the bands of the array structure were stacked along the third axis, creating a 3D array with bands as the third dimension. The 3D array was then reshaped into a 2D array, where every row signifies a pixel, and each column signifies a different band. This format rendered the data more convenient for analysis and manipulation.

Table 1 Specific Bands Selected for the Classification Task

SELECTED BANDS	RELEVANCE
Band 1 - Coastal/Aerosol (0.43 - 0.45 µm)	Useful for mapping shallow water and differentiating land from water, as well as assessing coastal and aerosol properties
Band 2 - Blue (0.45 - 0.51 µm)	Helps in distinguishing soil and vegetation, and is essential for water body mapping and monitoring
Band 3 - Green (0.53 - 0.59 µm)	Useful for assessing plant vigour and is a key component in vegetation analysis.
Band 4- Red Band (0.64 - 0.67 µm)	Essential for differentiating vegetation types and health due to its sensitivity to chlorophyll absorption.
Band 5- Near Infrared (NIR) (0.85 - 0.88 µm)	Able to distinguish between different land cover types

	based on their reflectance properties.
Band 8 - Panchromatic (0.50 - 0.68 µm)	Provides high spatial resolution images which can be used to sharpen multispectral bands. It enhances the detail and clarity of the image.

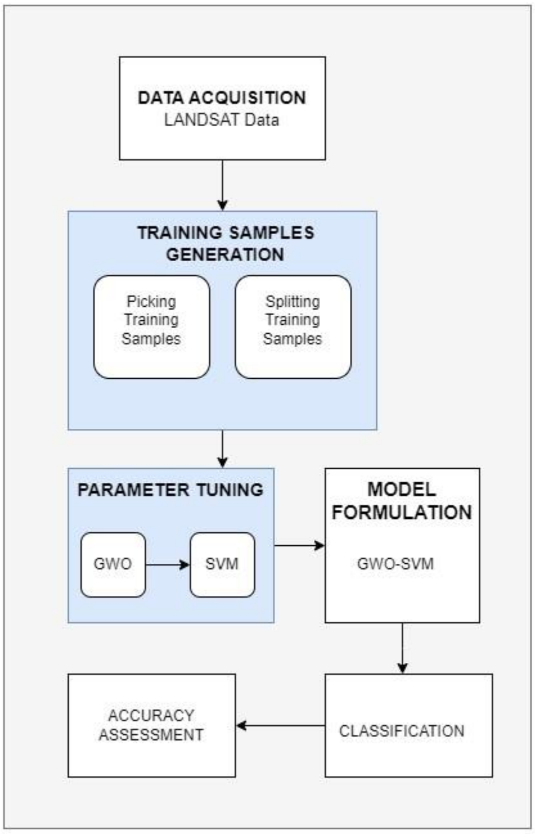

Fig. 2 Flowchart of Adopted Methodology

2.2.2 Training Samples Generation

The training data were randomly generated using Google Earth Pro. The generated data were then sampled to extract their band values. A total of 2440 points were generated and split into training (80 %) and testing (20 %). This ratio was chosen due to its preferred usage among machine learning researchers. The training data provided the model with input data examples and their corresponding correct outputs or labels. These datasets are used to enable the model to learn patterns and relationships between input features and output labels. The purpose of the testing data was to evaluate the performance of a trained machine-learning model. It consisted of a different set of sample points from the training dataset. The testing dataset aims to assess how well the models can perform on unseen data. This helps understand if the model can generalise its predictions to new and unseen data. A sample of the training data is represented in Fig. 3.

	fid	Name	id	rvalue_1	rvalue_2	rvalue_3	rvalue_4	rvalue_5	rvalue_6
1	1283	s	4	15347	14246	12166	11397	9807	8911
2	1284	s	4	14500	14006	11565	10884	9538	8915
3	1285	s	4	15702	14732	12092	11399	9903	9090
4	1286	s	4	16228	14719	12059	11379	9833	9145
5	1287	s	4	14782	14036	11898	11168	9791	9142
6	1528	v	3	12513	17839	10612	10497	8951	8173
7	1529	v	3	11977	17367	10380	10390	8773	7923
8	1530	v	3	12873	16290	10758	10498	8817	8083
9	1531	v	3	11363	16267	10101	10171	8353	7534
10	1532	v	3	10033	18454	9247	9566	7965	7222
11	1533	v	3	10013	17477	9425	9736	8126	7369
12	1534	v	3	9755	19100	9212	9590	8062	7365
13	1535	v	3	10777	17484	9643	9979	8382	7691
14	1520	v	3	10729	20494	9980	10493	8445	7499
15	1521	v	3	10799	19400	10009	10465	8492	7524
16	1522	v	3	9952	12548	10212	10208	8541	7571
17	1523	v	3	13684	16169	10792	10581	8930	7977
18	1524	v	3	11792	17186	10317	10442	8578	7580
19	1525	v	3	11945	17267	10457	10493	8715	7786
20	1526	v	3	12268	18696	10389	10594	8711	7807
21	1527	v	3	11643	16723	10134	10263	8486	7575

Fig. 3 Sample Training Data

2.2.3 Grey Wolf Optimiser

The Grey Wolf Optimisation algorithm (GWO), formulated by Mirjalili *et al*. (2014), draws inspiration from grey wolves' hierarchical structure and hunting tactics. The Grey Wolf Optimiser (GWO) is a nature-inspired optimisation algorithm based on grey wolves' social hierarchy and hunting behaviour. It mimics the cooperative and competitive interactions among wolves to search for optimal solutions efficiently. GWO effectively navigates complex search spaces by balancing exploration and exploitation, making it suitable for solving various optimisation problems.

Grey wolves are well-known for their social nature and typically live in packs that comprise 5 to 12 wolves. The social hierarchy of grey wolves is characterised by a structured organisation reflecting their cooperative and familial dynamics. The alpha wolves stand at the top of the hierarchy, typically composed of a dominant breeding pair – an alpha male and an alpha female. These leaders wield significant influence, making critical decisions for the pack. Subordinate to the alphas are the beta wolves, serving as supportive members who contribute to maintaining order within the pack. They may assume leadership roles in the absence of the alphas. The omega wolf, occupying the lowest rank, experiences greater social stress and plays a vital role in diffusing tensions within the pack, contributing to overall social harmony. In some descriptions of wolf hierarchies, delta wolves may exist as intermediaries between higher-ranking and lower-ranking members(Wong *et al*., 2014).

The Grey Wolf Optimisation (GWO) algorithm strategically emulates the hunting behaviour of grey wolves, orchestrating its optimisation process through four distinct phases: searching for prey, encircling the prey, hunting, and attacking. In this algorithmic framework, each wolf symbolises an initialised solution within the hyperspace of the given problem, drawing an analogy between potential solutions and prey in the wolves' natural hunting environment. The wolves are classified based on their fitness, with the most adept solution identified as the alpha wolf, followed by the beta and delta wolves, while all other solutions are collectively denoted as Omega(Rezaei *et al.*, 2018).

The algorithm unfolds its hunting strategy as the Alpha, Beta, and Delta wolves take the lead in the pack, guiding the exploration of potential solutions within the problem's hyperspace. The Omega wolf follows suit, representing additional solutions in the search space. As the wolf pack collectively identifies promising prey, the algorithm seamlessly transitions through the phases: searching for prey, encircling the prey, hunting, and attacking (Mirjalili *et al.*, 2014).

During the encircling phase, mathematically expressed through equations 1 and 2, the Alpha, Beta, and Delta wolves update their positions to surround the targeted prey. This phase emphasises exploitation, focusing the algorithm's efforts on refining and converging towards the identified solutions. The subsequent hunting phase intensifies the pursuit of optimal solutions, with the pack concentrating its search on the regions identified during the exploration phase. The attacking phase then refines the wolves' positions, optimising each solution's fitness (Hatta *et al*., 2019).

$$\vec{D} = \vec{C} \cdot \left|\vec{X}_p(t) - \vec{X}(t)\right| \quad (1)$$

$$\vec{X}(t+1) = \left|\vec{X}_p(t) - \vec{A} \cdot \vec{D}\right| \quad (2)$$

where $\vec{X}(t) and \vec{X}_p(t)$ denote the grey wolves and the prey's location, respectively, at the tth iteration. $\vec{D}$ denotes the position alteration element. $\vec{A}$ *and* $\vec{C}$ are coefficient vectors and are computed as shown in Equations (3) and (4).

$$\vec{A} = 2\vec{a} \cdot \vec{r}_1 - \vec{a} \quad (3)$$

$$\vec{C} = 2 \cdot r_2 \quad (4)$$

where $\vec{r}_1$ *and* r_2 are vectors with values between 0 and 1 that are generated randomly, and $\vec{a}$ is a moderating entity that linearly diminishes from 2 to 0.

In the GWO algorithm, the positions of the fittest solutions, that is, α, β, and δ, are updated first. Then, the other search agents (ω) are repositioned based on Equations (5), (6) and (7).

$$\vec{D}_\alpha = \left|\vec{C}_1 \cdot \vec{X}_\alpha - \vec{X}\right| \quad (5)$$

$$\vec{D}_\beta = \left|\vec{C}_2 \cdot \vec{X}_\beta - \vec{X}\right| \quad (6)$$

$$\vec{D}_\delta = \left|\vec{C}_3 \cdot \vec{X}_\delta - \vec{X}\right| \quad (7)$$

where $\vec{D}_\alpha$, $\vec{D}_\beta$, and $\vec{D}_\delta$ denotes the step size of ω with regards to α, β, and δ, with their respective as $\vec{X}_\alpha$, $\vec{X}_\beta$, and $\vec{X}_\delta$. $\vec{C}_1$, $\vec{C}_2$ and $\vec{C}_3$ are randomly initiated vectors and $\vec{X}$ is the current solution location.

After the distances are defined, $\vec{X}(t+1)$ which denotes the final position of the current solution is subsequently computed by Equations (8), (9), (10) and (11).

$$\vec{X}_1 = \vec{X}_\alpha - \vec{A}_1(\vec{D}_\alpha) \quad (8)$$

$$\vec{X}_2 = \vec{X}_\beta - \vec{A}_2(\vec{D}_\beta) \quad (9)$$

$$\vec{X}_3 = \vec{X}_\delta - \vec{A}_3(\vec{D}_\sigma) \quad (10)$$

$$\vec{X}(t+1) = \frac{\vec{X}_1 + \vec{X}_2 + \vec{X}_3}{3} \quad (11)$$

The GWO was adopted for this study due to the random adaptability of $\vec{A}$ and $\vec{C}$. These parameters enable the algorithm to achieve a balance when exploring and exploiting a search space. Thus, $\vec{A}$ initiates the exploration of search space and $\left|\vec{A}\right| > 1$ prompts candidate solutions to diverge from weaker prey in search of a fitter one. Similarly, candidate solutions converge toward the prey when $\left|\vec{A}\right| < 1$. $\vec{C} = [0,\ 2]$[0,2] which a random vector secondarily specifies weights for the prey considering its location from the wolf (Makhadmeh *et al.*, 2019). The pseudocode for the GWO algorithm is shown in Fig. 4.

PSEUDOCODE OF THE GREYWOLF OPTIMISATION ALGORITHM

Initialise the Greywolf population X_i (i= 1,2, 3……. n)
Initialise a, A and C
Calculate the fitness of each agent
Where X_a=best search, X_b=second best, X_c=third best
While t< Max number of iterations
For each search agent, update the position of the current search agent
End Update a, A and C
Calculate the fitness of all search agents
Update X_a, X_b and X_c
Return X_a

Fig. 4 Pseudocode of the GWO (Source: Mirjalili *et al.* (2014))

2.2.4 Support Vector Machine

Introduced in 1963 by Vapnik, the SVM is a classifier that linearly separates datasets using a decision boundary known as the hyperplane. The concept of the Support Vector Machine is to find a decision boundary that maximises the margin between the datasets closest to the boundary. The support vectors are the datasets lying closest to the decision boundary, also known as a hyperplane. SVMs have been employed effectively to solve many real-world problems, such as text recognition, regression, and image classification (Cervantes *et al.*, 2020). The Equation of the hyperplane is represented mathematically in Equation (12)

$$w.x + b = 0 \quad (12)$$

Where w and b are the normal vectors and the bias of the hyperplane, respectively. The points that satisfy Equation (12) are to be considered when defining the margin hyperplanes because they lie exactly on one of the margin borders. The distances for these points to the optimal hyperplane ($\omega \cdot x_i + b = 0$) are therefore $d_+ = d\_ = \frac{1}{\|w\|}$ and the margin width is equal to $\frac{2}{\|w\|}$. The optimisation of the margin width is obtained by solving the constrained quadratic optimisation problem presented in Equations (13) and (14).

$$minimise\ 1/\frac{2}{\|\omega\|} \quad (13)$$

$$under\ the\ constraints\ y = \omega \cdot x_i + b - 1 > 0 \quad (14)$$

The architecture of SVM is depicted in Fig. 5.

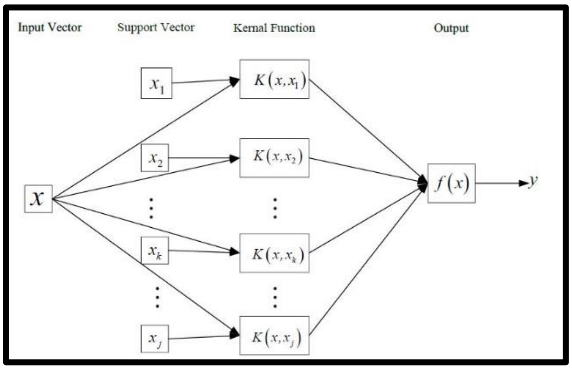

Fig. 5 SVM Architecture (Source: Adopted from Ruiz-Gonzalez *et al.*, 2014)

2.2.5 Hyperparameter Optimisation using GWO

This study optimised the hyperparameters, one of the key features of the support vector machine, using the grey wolf optimiser. These parameters are not directly learned from the data. However, they are set before the learning process begins. The hyperparameters are the regularisation parameter (C) and the choice of kernel function. The selection of the best c value and kernel for enhanced performance can be quite complex.

The regularisation parameter balances minimising training error and maintaining the model's simplicity to avoid overfitting. A small value of C allows for a larger margin and a smoother decision boundary, which promotes better generalisation but potentially increases bias. On the other hand, a large C value prioritises accurate classification of training samples, potentially leading to a more complex decision boundary and higher variance, hence increasing the risk of overfitting (Poku *et al.*, 2023).

The kernel function maps the input features into a higher-dimensional space, allowing SVM to handle non-linear decision boundaries. The shape of the decision boundary is determined by the type of kernel function used (linear, polynomial, radial basis function, among others*)* (Prajapati and Patle, 2010). The kernel function, thus, significantly affects the performance of the SVM model. Even though it was initially designed as a binary classifier, SVM can be modified to solve multiclass problems by converting the model into a higher dimensional using kernels (Cawley and Talbot, 2010).

Support Vector Machine (SVM) becomes an optimisation problem when trained on a dataset to find the optimal decision boundary that maximises the margin between different classes while minimising the classification error. Overall, SVM optimisation aims to find the parameters that minimise the classification error on the training data while maintaining good generalisation performance on unseen data.

The objective of the SVM is to find the decision boundary (hyperplane) that maximises the margin between different classes while minimising the classification error. This involves minimising the objective function, which quantifies the misclassification error and the complexity of the decision boundary. This is subject to constraints that ensure correct classification and sufficient margin between classes. Consequently, the objective function was defined to achieve the minimum fitness values.

2.2.6 Accuracy Metrics

Four standard metrics namely; Recall, Precision, F1-Score, and Overall Accuracy were used to evaluate the effectiveness of the different classification algorithms. Equation (15) illustrates how the recall, which quantifies a model's completeness, is calculated as the ratio of positively identified targets to positive targets. The ratio of the number of positively identified targets to the total number of targets recognised as positive is known as Precision, which measures how accurate or precise a model is. Equation (16) provides the mathematical expression for Precision. Equation (17) is the mathematical representation of Accuracy, which is the percentage of accurately detected targets to all detected targets. Equation (18) depicts that F1 is the harmonic mean of accuracy and recall.

$$\text{Recall} = \frac{TP}{TP+FN} \tag{15}$$

$$\text{Precision} = \frac{TP}{TP+FP} \tag{16}$$

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \tag{17}$$

$$\text{F1-score} = \frac{2*Precision*Recall}{Precision+Recall} \tag{18}$$

where TP is the true positive, FP is the false positive, FN represents the false negative, and TN is the true negative.

3 Results and Discussion

3.1 Experimental Design

The experiment was conducted on a Windows Operating System using Python as the programming language, utilising several libraries, including Scikit Learn, Numpy, Label Encoder, Pandas, Osgeo, and Mealpy (Thieu and Mirjalili, 2023). The image dataset was structured as an objective function for the Grey Wolf Optimiser (GWO). The GWO parameters were initialised to find the optimal hyperparameter combinations that would yield the highest accuracy for the Support Vector Machine (SVM) model. The specific GWO settings and parameters are detailed in Table 2.

Table 2 GWO Parameter Settings

GWO PARAMETER	VALUE
Fitness Function	Objective Function
Lower Bound	[0, 0.1]
Upper Bound	[3.99, 1000]
Epoch Size	50
Population Size	50

3.2 Performance Comparison

The robustness of the GWO-SVM model was ascertained by conducting a comparative analysis with the standalone SVM using recall, precision, F-1 score and overall accuracy. The GWO-SVM and SVM models were trained and tested using the same data. The findings of the evaluation are discussed in the subsequent sections.

3.2.1 Comparison between the GWO-SVM and SVM

The SVM classifier was employed with default parameter settings, while the GWO-SVM underwent a hyperparameter tuning process. After 50 iterations, GWO selected a regularisation parameter of 52.5 and an RBF kernel as the optimal configuration.

The results shown in Table 3 and Fig. 6 demonstrate improvements in the classification performance of GWO-SVM compared to the standalone SVM. GWO-SVM achieved a precision of 0.91, surpassing SVM's precision of 0.84. Both classifiers exhibited similar recall values of 0.82, indicating consistent performance in identifying positive instances. The F1-Score, a measure of the balance between precision and recall, was higher for GWO-SVM (0.85) than SVM (0.83). Overall accuracy showed a significant improvement, with GWO-SVM achieving 0.9192, outperforming SVM's 0.8932.

The higher precision of GWO-SVM suggests that predicting the positive class is more accurate than SVM. This is critical in applications where the misclassification of positive instances is costly. The consistent recall values indicate that both classifiers perform equally well in identifying positive instances. However, the superior F1-Score of GWO-SVM emphasises its better balance between precision and recall, making it a more reliable classifier. The substantial improvement in overall accuracy further solidifies the superiority of GWO-SVM, as it makes more accurate predictions across both positive and negative classes. Table 3 displays the classification performance of the developed models. Therefore, based on this comprehensive evaluation, GWO-SVM emerges as a promising classifier, excelling in making precise positive predictions while maintaining a strong balance between precision and recall.

Table 3 Accuracy Metrics for SVM and GWO-SVM

CLASSIFIER	PRECISION	RECALL	F1-SCORE	OVERALL ACCURACY
SVM	0.84	0.82	0.83	0.8932
GWO-SVM	0.91	0.82	0.85	0.9192

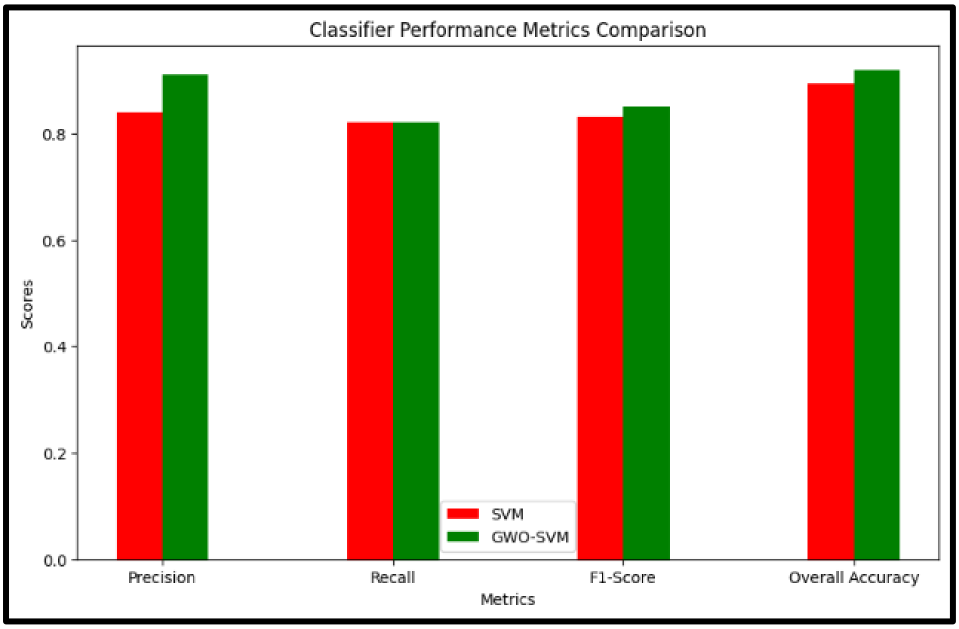

Fig. 6 Comparison between the Four Accuracy Metrics

3.3.2 Runtime

The computational efficiency was also observed. The GWO spent less than 7 secs for each epoch to complete its search. A graph showing the runtime of the GWO is depicted in Fig. 7. Table 4 also shows the runtime used by both the standalone SVM and the GWO-optimised SVM.

Table 4 Runtime of SVM and GWO-SVM

CLASSIFIER	**TIME USED (MINS)**
SVM	28
GWO-SVM	4.73

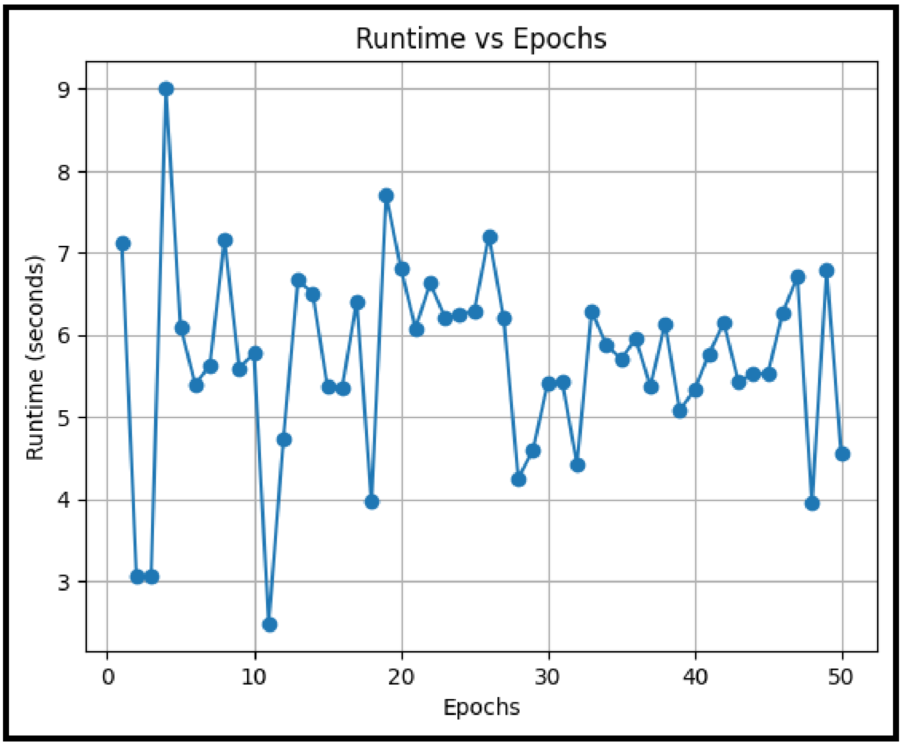

Fig. 7 Graph of Runtime for GWO

3.3.2 Land Cover Maps

The resulting land cover maps were visually compared between the standalone SVM and the GWO-SVM models. It could be observed that both models generally classified the various land cover types. However, it could be observed that some parts of the forest cover types were classified as bareland areas by the standalone SVM. The resulting images are shown in Figs. 8 and 9.

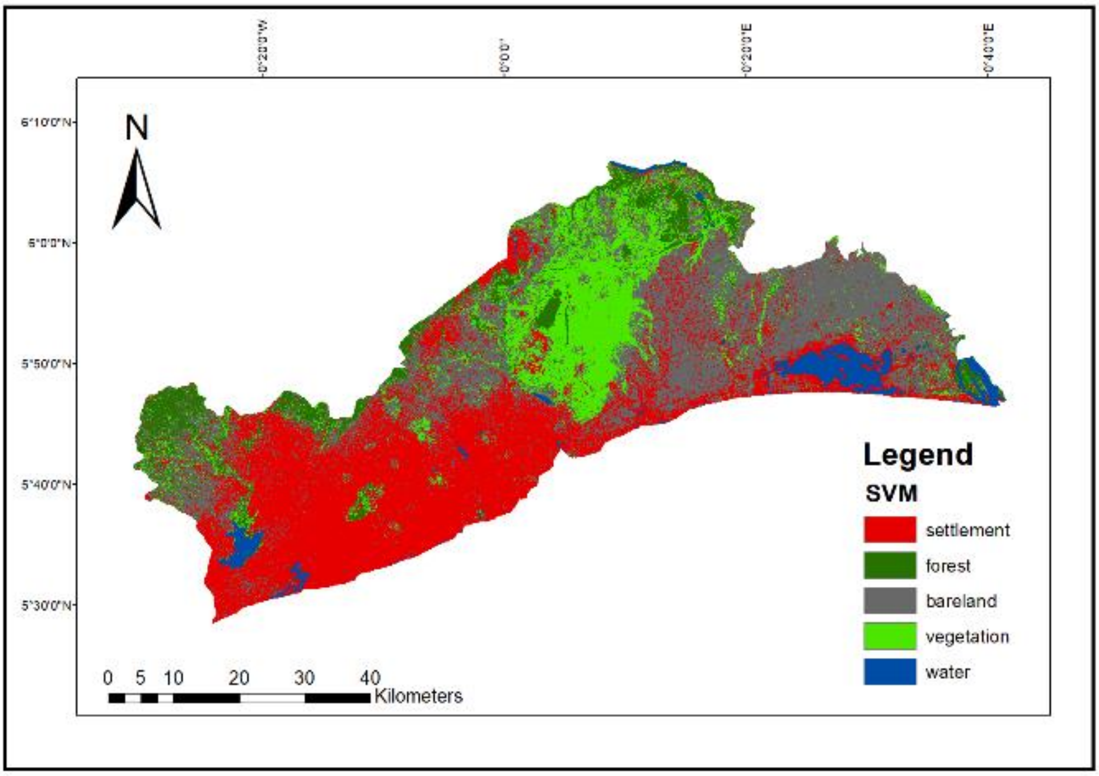

Fig. 8 SVM-derived LULC Map

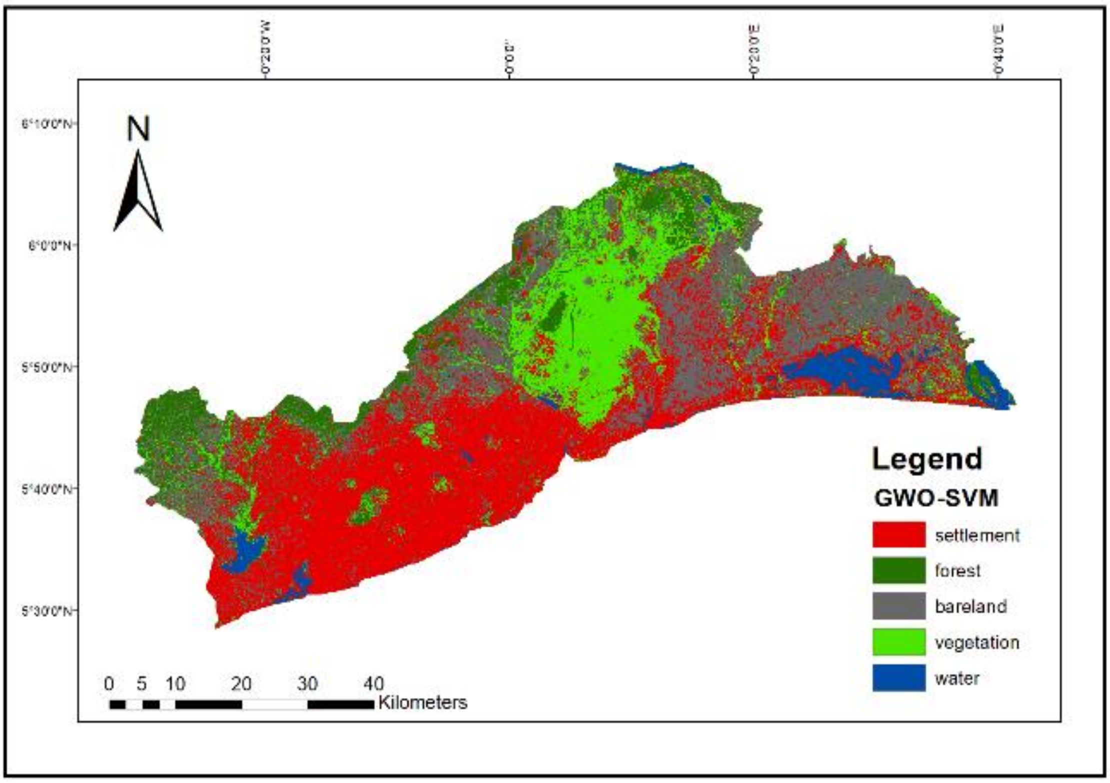

Fig. 9 GWO-SVM-derived LULC Map

4.4 Research Implication

Applying the Grey Wolf Optimisation (GWO) Algorithm to Support Vector Machine (SVM) for pixel-based image classification has yielded promising results and significant research implications. The successful integration of GWO with SVM showcases its potential as an effective and efficient approach for optimising hyperparameters in the context of image classification tasks, enhancing accuracy and efficiency in the classification of land cover types. The superior performance of this hybrid approach implies that it can serve as a valuable tool for remote sensing applications, environmental monitoring, and land cover mapping. Furthermore, the research implication of this study extends beyond pixel-based image classification. The effective integration of GWO with SVM may inspire researchers to investigate its potential in other machine-learning applications, such as object detection, feature selection, and anomaly detection. In conclusion, using GWO-SVM in pixel-based image classification opens new avenues for improving the performance and efficiency of machine learning models. The research implications of this study contribute to advancing the field of machine learning optimisation techniques, stimulating further research and innovation in developing sophisticated algorithms for a wide range of applications in image analysis.

5. Conclusion and Future Works

This study has optimised the SVM for pixel-based image classification on heterogeneous land types. The grey wolf optimiser conducted an exhaustive search within the hyperspace, and the results obtained from the Grey Wolf Optimised-SVM yielded the highest accuracy metric. Four evaluation metrics, precision, recall, F1-score, and overall accuracy, are considered for evaluating the proposed model. The performance assessment confirmed that the proposed hybrid classification technique was superior to the standalone SVM. Therefore, the optimised SVM could significantly improve land cover mapping accuracy, especially for complex land cover characteristics. Whilst this research has addressed the significance of optimisation techniques in improving classification accuracies, future works will explore beyond spectral features, considering textural, statistical (such as mean, median, standard deviation, *etc*) and various topographical indices (such as normalised differential vegetative index, built-up index, modified normalised differential water index, soil adjusted vegetative index, *etc*). This will provide richer information for classification, improving model robustness and accuracy. In the context of hyperparameter tuning, future research shall focus on hybrid methods. By combining the strengths of multiple models, these methods can correct the weaknesses of each model, leading to a more robust and accurate prediction

References

Abbas, A.W., Minallh, N., Ahmad, N., Abid, S.A.R. and Khan, M.A.A. (2016), "K-Means and ISODATA Clustering Algorithms for Landcover Classification using Remote Sensing". *Sindh University Research Journal-SURJ (Science Series)*, Vol 48, No 2.

Abbas, Z and Jaber, H. S.(2020), "Accuracy Assessment of Supervised Classification Methods for Extraction Land Use Maps using Remote Sensing and GIS Techniques" *IOP Conference Series: Materials Science Engineering*,745

Ahmad, A. and Quegan, S. (2012), "Analysis of Maximum Likelihood Classification on Multispectral Data", *Applied Mathematical Sciences*, Vol 6, No 129, pp 6425-6436

Ali, M., El-Hameed, M.A. and Farahat, M.A. (2017), "Effective Parameters' Identification for Polymer Electrolyte Membrane Fuel Cell Models using Grey Wolf Optimiser", *Renewable Energy*, Vol. 111, pp.455-462.

Alqurashi, A. and Kumar, L. (2013) "Investigating the Use of Remote Sensing and GIS Techniques to Detect Land Use And Land Cover Change: A Review" *Advances in Remote Sensing*.

Amini, S., Saber, M., Rabiei-Dastjerdi, H. and Homayouni, S. (2022), "Urban Land Use and Land Cover Change Analysis Using Random Forest Classification of Landsat Time Series", *Remote Sensing*, Vol, 14, No 11, pp 2654.

Askarzadeh, A. (2013), "Parameter Estimation of Fuel Cell Polarisation Curve using BMO Algorithm", *International Journal of Hydrogen Energy*, *38*(35), pp.15405-15413.

Basheer, S., Wang, X., Farooque, A. A., Nawaz, R. A., Liu, K., Adekanmbi, T. and Liu, S. (2022), "Comparison of Land Use Land Cover Classifiers Using Different Satellite Imagery and Machine Learning Techniques", *Remote Sensing*, Vol 14, No. 19, pp 4978.

Cawley, G. C and Talbot, N. L. C. (2010), "On Over-fitting in Model Selection and Subsequent Selection Bias in Performance Evaluation, *Journal of Machine Learning Research*, Vol 11, pp 2079-2107.

Cengiz, A. V. C. I., Budak, M., Yagmur, N. and Balcik, F. (2023). Comparison between Random Forest and Support Vector

Machine Algorithms for LULC Classification" *International Journal of Engineering and Geosciences*, Vol 8, No1, pp 1-10.

Cervantes, J., Garcia-Lamont, F., Rodríguez-Mazahua, L. and Lopez, A. (2020), "A Comprehensive Survey on Support Vector Machine Classification: Applications, Challenges and Trends", *Neurocomputing*, Vol 408, pp.189-215.

Cheng, L. and Bao, W. (2014), "Remote Sensing Image Classification Based On Optimised Support Vector Machine", *TELKOMNIKA Indonesian Journal of Electrical Engineering*, Vol 12, No 2, pp.1037-1045.

e Silva, L.P., Xavier, A.P.C., da Silva, R.M. and Santos, C.A.G. (2020) "Modelling Land Cover Change Based on an Artificial Neural Network for a Semiarid River Basin in Northeastern Brazil", *Global Ecology and Conservation*, Vol 21, pp e00811.

Elmanai, H. Loghmani, M.A. and Naceur, M. S. (2013), "SVM for Remote Sensing Classification", *International Conference on Control, Engineering and Information Technology*, Vol. 2, pp 68-72.

Fan, X., Chen, L., Xu, X., Yan, C., Fan, J. and Li, X. (2023), "Land Cover Classification of Remote Sensing Images Based on Hierarchical Convolutional Recurrent Neural Network", *Forests*, Vol 14, No 9, pp 1-27

Hasmadi, M., Pakhriazad, H.Z. and Shahrin, M.F. (2009), "Evaluating Supervised and Unsupervised Techniques for Land Cover Mapping using Remote Sensing Data", *Geografia: Malaysian Journal of Society and Space*, Vol *5*, No 1, pp.1-10

Hatta, N.M, Zain, A. M., Sallehuddin, R., Shayfull, Z. and Yusoff, Y. (2019), "Recent Studies on Optimisation Method of Grey Wolf Optimiser (GWO): A Review (2014–2017)", *Artificial Intelligence Review*, Vol 52, pp.2651-2683.

Hamad, R. (2020) "An Assessment of Artificial Neural Networks, Support Vector Machines and Decision Trees for Land Cover Classification Using Sentinel-2A Data." *Applied Ecology and Environmental Sciences*, Vol. 8, No 6, pp 459-464.

HongLei, Y., JunHuan, P., BaiRu, X. and DingXuan, Z. (2013), "Remote Sensing Classification using Fuzzy C-Means Clustering with Spatial Constraints Based on Markov Random Field", *European Journal of Remote Sensing*, Vol 46, No 1, pp.305-316.

Ho Y. C and Pepyne D. L (2002), "Simple Explanation of the No-free-lunchTheorem and its Implications"., *Journal of Optimisation Theory and Application*, Vol 115, No 3, pp 49–570.

Huang, C., Davis, L. S. and Townshend, J. R. G. (2002), "An Assessment of SVM for Land Cover Classification", *International Journal of Remote Sensing*, Vol 23, No 4, pp. 725-749

Kuo, R.J., Huang, S.L., Zulvia, F.E., and Liao, T.W. (2018), "Artificial Bee Colony-based Support Vector Machines with Feature Selection and Parameter Optimisation for Rule Extraction" *Knowledge Information Systems*, Vol 55, pp. 253–274.

Li, B., Deng, C., and Li, S. (2015), "High Resolution Remote Sensing Image Classification based on Particle Swarm Optimization And Support Vector Machine", *Computer Modelling and New Technologies*, Vol 19, No 2B, pp 22-26.

Makhadmeh, S. N., Khader, A. T., Al-Betar, M. A., and Naim, S. (2019), "Multi-Objective Power Scheduling Problem in Smart Homes using Grey Wolf Optimiser" *Journal of Ambient Intelligence and Humanised Computing*, Vol 10, pp 3643-3667.

Mantovani, R.G., Rossi, A.L., Vanschoren, J., Bischl, B. and De Carvalho, A.C., (2015), "Effectiveness of Random Search in SVM Hyper-parameter Tuning". I*nternational Joint Conference on Neural Networks (IJCNN) IEEE*, pp 1-8.

Mirjalili, S., Mirjalili, S. M. and Lewis, A. (2014), "Grey Wolf Optimiser", *Advances in Engineering Software*, vol. 69, 2014, pp. 46-61.

Mondal, A., Kundu, S., Chandniha, S.K., Shukla, R. and Mishra, P.K. (2012), "Comparison of Support Vector Machine and Maximum Likelihood Classification Technique using Satellite Imagery", *International Journal of Remote Sensing and GIS*, Vol 1, No 2, pp.116-123.

Negi, G., Kumar, A., Pant, S. and Ram, M. (2021). "GWO: A Review and Applications", *International Journal of System Assurance Engineering and Management*, 12, pp.1-8.

Nsiah R.A, Mantey S., and Ziggah Y.Y. (2023), Building Segmentation from UAV Orthomosaics using Unet-Resnet-34 Optimised with Grey Wolf Optimisation Algorithm, *Smart Construction and Sustainable Cities*, Vol 1, No. 21, pp. 1-18.

Pal, M. and Mather, P.M. (2005), "Support Vector Machines for Classification in Remote Sensing", *International Journal of Remote Sensing*, Vol 26, No 5, pp 1007-1011.

Peerbhay, K.Y., Mutanga, O. and Ismail, R. (2015), "Random Forests Unsupervised Classification: The Detection and Mapping of Solanum Mauritianum Infestations in Plantation Forestry using Hyperspectral Data", *IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing*, Vol 8, pp.3107–3122.

Piao, Y., Jeong, S., Park, S. and Lee, D.(2021), "Analysis of Land Use and Land Cover Change using Time-series Data and Random Forest in North Korea" *Remote Sensing*, Vol 13, No 17, pp 3501.

Poku, M.B., Issaka, Y. and Ziggah, Y.Y. (2023) Assessing the Performance of Machine Learning Algorithms for Urban Land Cover Classification using Multispectral Satellite Imagery. *Ghana Journal of Technology,* Vol 7, No 2, pp.10-18.

Prajapati, GL and Patle, A. (2010), "On Performing Classification using SVM With Radial Basis and Polynomial Kernel Functions", *3rd International Conference on Emerging Trends in Engineering and Technology*, pp. 512-515).

Prasad, S.V.S., Savithri, T.S. and Krishna, I.V.M., (2017), "Comparison of Accuracy Measures for RS Image Classification using SVM and ANN classifiers", *International Journal of Electrical and Computer Engineering*, Vol *7*, No 3, pp 1180.

Ramasamy, L.K., Kadry, S., Nam, Y. and Meqdad, M.N. (2021), "Performance analysis of sentiments in Twitter dataset using SVM models", *International Journal of Electrical and Computer Engineering (IJECE),* Vol 11, No 3, pp.2275-2284.

Rezaei, H., Bozorg-Haddad, O. and Chu, X., (2018), Grey Wolf Optimization (GWO) Algorithm, *Advanced Optimization by Nature-inspired Algorithms,* pp.81-91.

Rudrapal D. and Subhedar, M.S. (2015), "Land Cover Classification using Support Vector Machines", *International Journal of Engineering, Research and Technology*, Vol. 4, No. 9, 6pp.

Ruiz-Gonzalez, R., Gomez-Gil, J., Gomez-Gil, F.J. and Martínez-Martínez, V. (2014), "An SVM-Based Classifier for Estimating the State of Various Rotating Components in Agro-Industrial Machinery with a Vibration Signal Acquired from a Single Point on the Machine Chassis", *Sensors*, *14*(11), pp.20713-20735.

Shihab, T. H., Al-Hameedawi, A. N., and Hamza, A. M. (2020), "Random Forest (RF) and Artificial Neural Network (ANN) Algorithms for LULC mapping". *Engineering and Technology Journal*, Vol 38, No 4A, Pp 510-514.

Tamimi, E., Ebadi, H. and Kiani, A. (2017), "Evaluation of Different Metaheuristic Optimization Algorithms in Feature Selection and Parameter Determination in SVM Classification", *Arabian Journal of Geosciences,* Vol 10, N0 478, pp 1-19

Tamirat, H., Argaw, M. and Tekalign, M. (2023), "Support Vector Machine-Based Spatiotemporal Land Use Land Cover Change Analysis in a Complex Urban and Rural Landscape of Akaki River Catchment, a suburb of Addis Ababa, Ethiopia",Vol 9, No 11, pp

Thieu, V. N And Mirjalili, S. (2023), "MEALPY: An Open-Source Library for Latest Meta-Heuristic Algorithms in Python", *Journal Of Systems Architecture*.

Tomar, V., Bansal, M. and Singh, P.(2023), "Metaheuristic Algorithms for Optimization: A Brief Review" *Engineering Proceedings*, Vol 59, No, 1

Venkateswaran, K., Kasthuri, N., Balakrishnan, K. and Prakash, K. (2013) "Performance Analysis of K-Means Clustering for Remotely Sensed Images," *International Journal of Computer Applications* Vol.84, No 12, pp 23-27.

Vishwanath, N., Ramesh, B. and Sreenivasa Rao, P. (2016), "Unsupervised Classification of Remote Sensing Images using K-Means Algorithm", *International Journal of Latest Trends in Engineering and Technology (IJLTET),* Vol 7, No. 2, pp.548-552.

Wang, M., Wan, Y., Ye, Z. and Lai, X. (2017), "Remote Sensing Image Classification Based on the Optimal SVM and Modified Binary Ant Colony Optimisation Algorithm", *Information Sciences*, Vol 402, pp 50-68.

Wang, C., Guo, Z., Wang, S., and Ma, C. (2013), "Improving Remote Sensing Image Segmentation for land Cover Classification using Combined K-Means and Minimum Spanning Tree Algorithm.

Wong, L. I., Sulaiman, M. H., Mohamed, M. R. and Hong,M. S. (2014), "Grey Wolf Optimiser for Solving Economic Dispatch Problems", 2014 IEEE International Conference on Power and Energy (PECon), Kuching, Malaysia, pp. 150-154,

Xue, Z.; Du, P.; Su, H. (2014), "Harmonic Analysis for Hyperspectral Image Classification Integrated with PSO Optimised SVM", *IEEE Journal of Selected Topics in Applied Earth Observation and Remote Sensing* Vol 7, pp. 2131–2146.

Yang H., Peng J., Xia B. and Zhang D. (2013) "Remote Sensing Classification Using Fuzzy C-means Clustering with Spatial Constraints Based on Markov Random

Field", *European Journal of Remote Sensing*, Vol 46, No.1, pp.305-316.

Yuh, Y. G., Tracz, W., Matthews, H. D., Turner, S.E. (2023), "Application of Machine Learning Approaches for Land Cover Monitoring in Northern Cameroon", *Ecological Informatics*, Vol 74, pp 1-15.

Zhu, X., Li, N. and Pan, X. (2019), "Optimisation Performance of Three Different Group Intelligent Algorithms on SVM for Hyperspectral Image Classification", *Remote Sensing*, Vol 11, No 734, pp 1-20.

Authors

Maame Boama Poku is a PhD candidate at the Geomatic Engineering Department of the University of Mines and Technology (UMaT). She holds a BSc in Geomatic Engineering from the University of Mines and Technology, Tarkwa, Ghana. Her research interests include Geographic Information Systems, Remote Sensing Applications, Urban Planning and Artificial Intelligent Techniques.

Yakubu Issaka is an Associate Professor of Geomatic Engineering and a Consulting Engineer currently working at UMaT, Tarkwa-Ghana. He holds BSc in Geomatic Engineering from KNUST, Kumasi-Ghana, MPhil and PhD in Geomatic Engineering from the (UMaT), Tarkwa-Ghana. He holds an MSc in Land Policy and Administration from the University of Cape Coast. He also has a Certificate in Disaster Risk Management and Environmental Assessment for Spatial Planning from the ITC-University of Twente, The Netherlands. He is a member of the GHIE, GhIS, IAENG, ISPRS, a fellow of WAIMM. His research and consultancy work cover; Land Policy and Administration, Mine Surveying, Infrastructural Mapping, Cadastral and Engineering Surveying, Global Navigation Satellite System (GNSS) applications, Geo-Information for Disaster Management, Geo-Information for Transport Management.

Y. Y. Ziggah is a Senior Lecturer at the Geomatic Engineering Department of the University of Mines and Technology (UMaT). He holds a BSc in Geomatic Engineering from Kwame Nkrumah University of Science and Technology, Kumasi, Ghana. He obtained his Master of Engineering degree and Doctor of Philosophy in Geodesy and Survey Engineering from China University of Geosciences (Wuhan). His research interests include 2D/3D coordinate transformation, height systems, gravity field modelling, geodetic deformation modelling, and application and development of computational intelligent techniques in Geodesy.