Introduction

Pneumonia is a common and sometimes life-threatening lung infection that causes major illness and death around the world.[1] Lower respiratory infections remain a leading cause of mortality across ages and regions.[2] Community-acquired pneumonia (CAP) ranges from mild disease managed as an outpatient to severe illness with respiratory failure and sepsis that requires intensive care.[3] Recent reviews summarize this clinical spectrum and its implications for care.[5] Clinicians typically combine symptoms and signs with chest imaging and basic laboratory tests to diagnose CAP.[3, 4] Quality standards emphasize early chest radiography and timely antibiotics when pneumonia is suspected.[6] Chest X-ray can be unavailable, inconclusive, or even negative early in the disease.[7] Computed tomography may reveal CAP when the X-ray is negative and can support earlier decisions in the emergency department.[8, 9] These gaps sustain diagnostic uncertainty, which can both delay treatment and promote unnecessary antibiotics. There is a need for data-driven tools that use routine clinical information to make diagnosis earlier and more reliable.

Blood biomarkers of the host response are widely used in suspected pneumonia, including complete blood count (CBC) indices and acute-phase reactants such as C-reactive protein (CRP) and procalcitonin (PCT).[10, 11] PCT-guided antibiotic strategies have been studied extensively in acute respiratory infections.[12] Leukocytosis with neutrophil predominance and raised CRP levels are common in bacterial CAP.[13] Higher CRP has also been linked to disease severity and worse outcomes.[14, 15] Ratios derived from the CBC capture inflammation in a compact form. The neutrophil-to-lymphocyte ratio (NLR) shows prognostic value in hospitalized CAP.[16, 18] The platelet-to-lymphocyte ratio (PLR) has similar associations in derivation–validation cohorts.[17] Combinations such as NLR, MLR, and PLR can also help differentiate etiologies and predict outcomes.[19] Beyond single markers, machine-learning (ML) models can integrate multiple variables to predict pneumonia or its complications.[20, 24] Models that rely only on clinical and laboratory data have reported strong diagnostic performance.[21] In severe CAP within the ICU, ML-based mortality prediction has also shown promise.[22, 23] Even so, many studies emphasize imaging-derived features from chest radiography or CT.[25, 26] Other studies restrict attention to a small, hand-picked set of lab variables, which may overlook useful signals in the full panel.

A modest blood panel becomes high-dimensional once we include derived indices, ratios, and interactions.[27] Such high dimensionality increases the risk of overfitting, variance, and reduced interpretability in clinical datasets.[28, 29] Feature selection (FS) helps by identifying compact, informative subsets that preserve performance and improve clarity.[27] FS methods are usually grouped into filters, embedded methods, and wrappers.[29] Examples of embedded methods include LASSO-regularized models and tree-based ensembles.[30, 31] Wrapper methods evaluate candidate subsets using a downstream classifier and often perform well by capturing feature interactions, though at higher computational cost.[32]

The main challenge for wrapper FS is the combinatorial search over $2^D - 1$ non-empty feature subsets.[27] Metaheuristic optimization is popular for this setting because it is flexible, derivative-free, and able to escape local optima.[33, 34] Classic examples include Genetic Algorithms and Particle Swarm Optimization.[35, 36] Nature-inspired methods such as the Grey Wolf Optimizer have also been adopted.[37] More recent additions include Harris Hawks Optimization and the Slime Mould Algorithm.[38, 39] These approaches have achieved strong empirical results in biomedical applications.[40, 41] At the same time, standard metaheuristics can suffer from premature convergence, unstable exploration–exploitation balance, and sensitivity to hyperparameters.[42, 43] General limits are also captured by the No Free Lunch theorems for optimization.[44]

Researchers have proposed enhanced and hybrid strategies to improve search quality. Cooperative and multi-population learning are common examples.[45] Comprehensive-learning and

dynamic-neighborhood variants of PSO further encourage exploration.[46, 47] Opposition-based learning can be incorporated directly or via differential evolution schemes.[48, 49] When binary decisions are required, refined transfer functions help map continuous updates to feature inclusion.[50] Despite these advances, methods may still over-rely on a single best solution and lose diversity, especially when each fitness evaluation retrains a classifier. In practice, this can produce unstable subsets and variable performance across folds.

We propose a Multi-dimensional Gradient interaction Search Polar Lights Optimizer (MGSPLO) for wrapper-based feature selection on blood routine and inflammatory markers. MGSPLO integrates mathematical gradient estimation with bio-inspired mechanics: (i) it utilizes a Dynamic Gradient Interaction System (DGIS) inspired by the Adam optimizer, which calculates the first and second moments of the population's position to generate a momentum-guided reference point for stable convergence; (ii) it maintains population diversity via Starfish-inspired hybrid search patterns, which adaptively switch between five-dimensional and unidimensional updates to navigate high-dimensional feature spaces efficiently; and (iii) it applies a binary mapping scheme that turns these momentum-driven continuous trajectories into probability-driven feature inclusion decisions. In our diagnostic pipeline, MGSPLO wraps a support vector machine (SVM), using cross-validated performance as the fitness signal to move toward compact, high-performing subsets.

We study a cohort of patients evaluated for pneumonia, with age, comorbidities, blood routine indices (white blood cell and differential counts; red cell and platelet parameters) and CRP measured at presentation. From these we derive composite markers (e.g., NLR, PLR), creating a feature pool that reflects both routine clinical practice and emerging biomarker evidence. Our goals are to: (1) quantify the diagnostic value of blood routine and inflammatory markers—alone and in combination—for pneumonia detection; (2) test whether MGSPLO-driven wrapper FS can identify small, stable, interpretable subsets that match or improve SVM performance compared with using all markers or standard FS baselines; and (3) relate the selected markers to current biomarker literature.

We list our contribution of this study as follows:

- We introduce MGSPLO, a hybrid metaheuristic integrating **Adam-inspired gradient interaction** and **multi-dimensional starfish search patterns** to balance exploitation stability with exploration diversity, effectively reducing premature convergence in high-dimensional clinical feature spaces.
- We derive a binary version of MGSPLO that maps continuous, **momentum-driven search trajectories** into discrete feature selection decisions, optimizing the inclusion of blood routine and inflammatory markers.
- We build an end-to-end SVM pneumonia pipeline that integrates MGSPLO with cross-validated evaluation and compare against standard FS techniques and alternative wrapper metaheuristics.
- On a real-world cohort, we show that MGSPLO–SVM achieves competitive or superior diagnostic accuracy while selecting compact, interpretable panels of routine laboratory markers, which is helpful where imaging resources are limited.

The rest of the paper is organized as follows: Section 2 reviews related work on pneumonia biomarkers, ML-based diagnosis, and feature selection. Section 3 describes MGSPLO, its binary version, and the SVM framework. Section 4 details the dataset, features, and experiments. Section 5 reports results and ablations. Section 6 discusses limitations, clinical implications, and future work.

References

[1] World Health Organization. *Pneumonia*. Available at: https://www.who.int/health-topics/pneumonia. Accessed 2025.

[2] GBD collaborators. Trends in the global burden of lower respiratory infections: the knowns and unknowns. *Lancet Infect Dis*. 2022. Available at: https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(22)00445-5/fulltext.

[3] Metlay JP, Waterer GW, Long AC, et al. Diagnosis and treatment of adults with community-acquired pneumonia. *Am J Respir Crit Care Med*. 2019;200(7):e45–e67. (ATS/IDSA guideline). https://www.atsjournals.org/doi/10.1164/rccm.201908-1581ST.

[4] National Institute for Health and Care Excellence (NICE). *Pneumonia: diagnosis and management* (NG250). Updated 2025. https://www.nice.org.uk/guidance/ng250.

[5] Musher DM, Thorner AR. Community-Acquired Pneumonia. *JAMA*. 2014;311(20):2197–2206. (See also JAMA 2020 updates). https://jamanetwork.com/journals/jama/fullarticle/2760882.

[6] NICE Quality Standard QS110. Chest X-ray and antibiotic treatment within 4 hours. Updated 2025. https://www.nice.org.uk/guidance/qs110.

[7] Self WH, et al. The clinical utility of chest radiography for identifying pneumonia. *AJR Am J Roentgenol*. 2020;214(6):1208–1212. https://www.ajronline.org/doi/pdf/10.2214/AJR.19.21521.

[8] Okada F, et al. Community-Acquired Pneumonia with Negative Chest Radiography Findings: Clinical and Radiological Characteristics. *Respiration*. 2018;97(6):508–516. https://karger.com/res/article/97/6/508/290847/Community-Acquired-Pneumonia-with-Negative-Chest.

[9] Claessens YE, et al. Early chest CT-scan to assist diagnosis and guide treatment decision for suspected CAP in the ED. 2015 randomized trial. https://research.pasteur.fr/en/publication/early-chest-ct-scan-to-assist-diagnosis-and-guide-treatment-decision-for-suspected-comm

[10] Minnaard MC, et al. The added value of CRP in diagnosing pneumonia in primary care: IPD meta-analysis. *CMAJ*. 2017;189(2):E56–E63. https://www.cmaj.ca/content/189/2/E56.

[11] Schuetz P, et al. Procalcitonin to guide initiation and duration of antibiotics in acute respiratory infections: IPD meta-analysis. *Clin Infect Dis*. 2012;55(5):651–662. https://academic.oup.com/cid/article/55/5/651/350305.

[12] Schuetz P, et al. Effect of procalcitonin-guided antibiotic treatment on mortality in acute respiratory infections. *Lancet Infect Dis*. 2018;18(1):95–107. https://www.thelancet.com/journals/laninf/article/PIIS1473-3099(17)30592-3/fulltext.

[13] Krüger S, et al. Procalcitonin predicts severity and outcome in CAP; comparison with CRP and leukocyte count. *Eur Respir J*. 2008;31(2):349–355. https://publications.ersnet.org/content/erj/31/2/349.full.pdf.

[14] Chalmers JD, et al. CRP, severity of pneumonia and mortality in elderly patients. *Age Ageing*. 2009;38(6):693–697. https://academic.oup.com/ageing/article-abstract/38/6/693/40913.

[15] Menéndez R, et al. Utility of C-reactive protein in assessing severity and outcomes in CAP. *J Infect*. 2008;56(5):355–362. https://www.sciencedirect.com/science/article/pii/S1198743X14617985.

[16] Lee JH, et al. Prognostic value of serial neutrophil-to-lymphocyte ratio in hospitalized CAP. *PLoS One*. 2021;16(4):e0250067. https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0250067.

[17] Enersen CCE, et al. NLR and PLR and association with mortality in CAP: derivation–validation cohort. *Infection*. 2023;51:1339–1347. https://link.springer.com/article/10.1007/s15010-023-01992-2.

[18] Colosi IA, et al. The Neutrophil/Lymphocyte Ratio and Outcomes in Hospitalized CAP. *Biomedicines*. 2024;12(2):260. https://www.mdpi.com/2227-9059/12/2/260.

[19] Simadibrata DM, et al. NLR, MLR, PLR and RDW to predict outcome and differentiate etiologies in pneumonia. *Sci Rep*. 2022;12:16109. https://www.nature.com/articles/s41598-022-20385-3.pdf.

[20] Ye Y, et al. Machine learning–assisted prediction of pneumonia based on symptoms and routine data. *Front Public Health*. 2022;10:938801. https://www.frontiersin.org/articles/10.3389/fpubh.2022.938801/full.

[21] Kang SJ, et al. Machine Learning for CAP diagnosis using only clinical and laboratory data. *Comput Biol Med*. 2024;169:108100. https://www.sciencedirect.com/science/article/pii/S2213716524004090.

[22] Pan J, et al. Mortality prediction in severe CAP in ICU using ML. *Sci Rep*. 2025; https://www.nature.com/articles/s41598-025-85951-x.pdf.

[23] Liu X, et al. ML model for mortality prediction in severe pneumonia using accessible clinical data. *BMJ Open Respir Res*. 2025;12:e001983. https://bmjopenrespres.bmj.com/content/12/1/e001983.

[24] Delaney BC, et al. Applying ML to EHR to predict CAP after RTI consultations in primary care. *J Biomed Inform*. 2022;128:104028. https://www.sciencedirect.com/science/article/pii/S0895435622000154.

[25] Rajpurkar P, et al. CheXNet: Radiologist-Level Pneumonia Detection on Chest X-rays with Deep Learning. arXiv:1711.05225 (2017). https://arxiv.org/abs/1711.05225.

[26] Khan A, et al. Deep learning for pneumonia detection in CXR: a review of 2012–2023. *J Imaging*. 2024;10(8):176. https://www.mdpi.com/2313-433X/10/8/176.

[27] Guyon I, Elisseeff A. An introduction to variable and feature selection. *J Mach Learn Res*. 2003;3:1157–1182. https://jmlr.org/papers/volume3/guyon03a/guyon03a.pdf.

[28] Chandrashekar G, Sahin F. A survey on feature selection methods. *Computers & Electrical Engineering*. 2014;40(1):16–28. https://www.sciencedirect.com/science/article/pii/S0045790613003066.

[29] Saeys Y, Inza I, Larrañaga P. A review of feature selection techniques in bioinformatics. *Bioinformatics*. 2007;23(19):2507–2517. https://academic.oup.com/bioinformatics/article/23/19/2507/185254.

[30] Tibshirani R. Regression shrinkage and selection via the LASSO. *J R Stat Soc Ser B*. 1996;58(1):267–288.

[31] Breiman L. Random Forests. *Machine Learning*. 2001;45(1):5–32.

[32] Kohavi R, John GH. Wrappers for feature subset selection. *Artificial Intelligence*. 1997;97(1–2):273–324. https://www.sciencedirect.com/science/article/pii/S000437029700043X.

[33] Xue B, Zhang M, Browne WN. A survey on evolutionary computation approaches to feature selection. *Pattern Recognit*. 2016;53:121–143.

[34] Sayed GI, Hassanien AE, et al. A comprehensive survey on recent metaheuristics for feature selection. *Neurocomputing*. 2022;509:99–126. https://www.sciencedirect.com/science/article/pii/S092523122200474X.

[35] Holland JH. *Adaptation in Natural and Artificial Systems*. University of Michigan Press; 1975.

[36] Kennedy J, Eberhart R. Particle swarm optimization. In: *Proc. IEEE Int. Conf. Neural Networks*. 1995:1942–1948.

[37] Mirjalili S, et al. Grey Wolf Optimizer. *Advances in Engineering Software*. 2014;69:46–61.

[38] Heidari AA, et al. Harris Hawks Optimization. *Future Gener Comput Syst*. 2019;97:849–872.

[39] Li S, et al. Slime mould algorithm. *Future Gener Comput Syst*. 2020;111:300–323.

[40] Wazirali R, et al. A systematic review on metaheuristic optimization techniques for disease diagnosis. *Arch Comput Methods Eng*. 2023;30:1–34. https://link.springer.com/article/10.1007/s11831-022-09853-1.

[41] Alrefai N, Ibrahim O. PSO-based feature selection with ensemble learning for cancer microarray classification. *Neural Comput Appl*. 2022;34:13513–13528. https://link.springer.com/article/10.1007/s00521-022-07147-y.

[42] Milano M, et al. A better balance in metaheuristic algorithms: Does it exist? *Appl Soft Comput*. 2020;93:106381. https://www.sciencedirect.com/science/article/pii/S2210650219304080.

[43] Ibrahimov M, et al. Why tuning the control parameters of metaheuristic algorithms is so important. *MENDEL*. 2014;20(1):7–14. https://mendel-journal.org/index.php/mendel/article/download/120/141/.

[44] Wolpert DH, Macready WG. No Free Lunch Theorems for Optimization. *IEEE Trans Evol Comput*. 1997;1(1):67–82. https://www.cs.ubc.ca/~hutter/earg/papers07/00585893.pdf.

[45] van den Bergh F, Engelbrecht AP. A cooperative approach to particle swarm optimization. *IEEE Trans Evol Comput*. 2004;8(3):225–239. https://phoenixwilliams.github.io/PersonalWebsite/LargeScaleEA/A_Cooperative_approach_to_particle_swarm_optimization.pdf.

[46] Liang JJ, Qin AK, Suganthan PN, Baskar S. Comprehensive Learning PSO. *IEEE Trans Evol Comput*. 2006;10(3):281–295. https://tiezhongyu2005.github.io/resources/popularization/CLPSO_2006.pdf.

[47] Zhang Y, et al. A dynamic neighborhood learning based PSO (DNLPSO). *Inf Sci*. 2012;190:19–34. https://www.sciencedirect.com/science/article/pii/S0020025512002927.

[48] Tizhoosh HR. Opposition-Based Learning: A new scheme for machine intelligence. CIMCA 2005. https://www.researchgate.net/publication/4242497_Opposition-Based_Learning_A_New_Scheme_for_Machine_Intelligence.

[49] Rahnamayan S, Tizhoosh HR, Salama MMA. Opposition-Based Differential Evolution. *IEEE Trans Evol Comput.* 2008;12(1):64–79. https://www.researchgate.net/publication/3419015_Opposition-based_differential_evolution_IEEE_Trans_Evol_Comput.

[50] Mirjalili S, Lewis A. S-shaped vs V-shaped transfer functions for binary PSO. *Swarm Evol Comput.* 2013;9:1–14. https://www.sciencedirect.com/science/article/pii/S2210650212000648.