Research Square

Preprints are preliminary reports that have not undergone peer review.
They should not be considered conclusive, used to inform clinical practice,
or referenced by the media as validated information.

Interpretable prototype-driven detection of rare Hodgkin/Reed- Sternberg cells in multicenter lymph-node fine-needle aspiration whole-slide images

Chong-Wai Io
National Chung Cheng University
Giang Huong Tran
University Medical Center Ho Chi Minh City
Tu Anh Thai
Ho Chi Minh City Oncology Hospital
Tho Ngoc-Quynh Le
University Medical Center Ho Chi Minh City
Uyen Le Ly
Ho Chi Minh City Oncology Hospital
Chun-Chieh Wu
Kaohsiung Medical University Hospital
Hao-Wen Chuang
Kaohsiung Veterans General Hospital
Minh-Xuan Dang
University Medical Center Ho Chi Minh City
Diem Nhu-Thi Pham
University Medical Center Ho Chi Minh City
Toan Dang Pham
Ho Chi Minh City Oncology Hospital
Wei-Chen Tang
National Chung Cheng University
Ming-Jie Zhuo
National Chung Cheng University
Lu-Han Fang
Kaohsiung Veterans General Hospital
Chien-Chin Chen
hlmarkc@gmail.com

Ditmanson Medical Foundation Chia-Yi Christian Hospital

Article

Keywords: Hodgkin lymphoma, Hodgkin and Reed-Sternberg cells, fine-needle aspiration, whole-slide imaging, interpretable image analysis, rare-cell detection, digital cytopathology

Posted Date: November 24th, 2026

DOI: https://doi.org/10.21203/rs.3.rs-11816964/v1

License: This work is licensed under a Creative Commons Attribution 4.0 International License.
Read Full License

Additional Declarations: No competing interests reported.

Interpretable prototype-driven detection of rare Hodgkin/Reed-Sternberg cells in multicenter lymph-node fine-needle aspiration whole-slide images

Chong-Wai Io [1,†], Giang Huong Tran [2,3,†], Tu Anh Thai [4,†], Tho Ngoc-Quynh Le [2], Uyen Le Ly [4], Chun-Chieh Wu [5], Hao-Wen Chuang [6], Minh-Xuan Dang [2], Diem Nhu-Thi Pham [2], Toan Dang Pham [4], Wei-Chen Tang [1], Ming-Jie Zhuo[1], Lu-Han Fang [6], Chien-Chin Chen [7,8,9,10,*]

[1] Department of Physics, National Chung Cheng University, Chiayi, Taiwan.

[2] Department of Pathology, University Medical Center Ho Chi Minh City, Ho Chi Minh City, Vietnam.

[3] Department of Histology, Embryology, and Pathology, School of Medicine, University of Medicine and Pharmacy at Ho Chi Minh City, Ho Chi Minh City, Vietnam.

[4] Department of Pathology, Ho Chi Minh City Oncology Hospital, Ho Chi Minh City, Vietnam.

[5] Department of Pathology, Kaohsiung Medical University Hospital, Kaohsiung, Taiwan.

[6] Department of Pathology and Laboratory Medicine, Kaohsiung Veterans General Hospital, Kaohsiung, Taiwan.

[7] Department of Pathology, Ditmanson Medical Foundation Chia-Yi Christian Hospital, Chiayi, Taiwan.

[8] Doctoral Program in Translational Medicine, National Chung Hsing University, Taichung, Taiwan.

[9] Department of Cosmetic Science, Chia Nan University of Pharmacy and Science, Tainan, Taiwan.

[10] Department of Biotechnology and Bioindustry Sciences, College of Bioscience and Biotechnology, National Cheng Kung University, Tainan, Taiwan.

† C.W.I., G.H.T., and T.A.T. contributed equally to the manuscript.

***Corresponding author:**

Chien-Chin Chen, MD, FIAC

Department of Pathology, Ditmanson Medical Foundation Chia-Yi Christian Hospital,

No. 539, Zhongxiao Rd., East Dist., Chiayi 600, Taiwan.

Tel: +886-5-2765041 ext. 7521; Fax: +886-5-2781961; Email: hlmarkc@gmail.com

Highlights

1. An interpretable prototype-driven method detects rare HRS-like cells.
2. Makes review time manageable by narrowing ~10^6 objects to ~10^2 candidates.
3. Cross-institution intensity calibration addresses staining variation.
4. Positive and negative slides showed clearly separated slide-level PPVs.
5. The system supports prescreening while retaining expert confirmation.

Abstract

Identifying rare Hodgkin and Reed–Sternberg (HRS) cells among 10^5–10^6 background objects in lymph-node fine-needle aspiration (FNA) smears is laborious and restricts labeled datasets needed for data-intensive machine learning. We developed an interpretable, prototype-driven algorithm utilizing expert-selected HRS cells across five staining classes to define quantitative criteria based on RGB intensity, contour morphology, and intranuclear heterogeneity, calibrated via histogram cross-correlation. Evaluated across 78 digitized slides from five institutions (12 training, 38 positive test, and 28 negative-control slides), the method achieved a 91.6% mean prototype-cell recovery rate. Mean candidate-level positive predictive values (PPVs) were 39.2% (training), 50.0% (intra-institutional test), 55.0% (cross-institutional test), and 1.1% (negative controls), yielding non-overlapping positive (minimum 10.3%) and negative (maximum 9.7%) cohort ranges. By reducing 10^6 objects to ~10^2 candidates, this interpretable approach streamlines cytopathologist-led prescreening while preserving human-in-the-loop diagnostic confirmation.

Keywords

Hodgkin lymphoma; Hodgkin and Reed-Sternberg cells; fine-needle aspiration; whole-slide imaging; interpretable image analysis; rare-cell detection; digital cytopathology

Introduction

Classical Hodgkin lymphoma (cHL) is a distinct germinal-center B-cell neoplasm that accounts for approximately 10% of lymphomas. It is characterized by a striking imbalance between the malignant and non-neoplastic compartments: mononuclear Hodgkin cells and multinucleated Reed-Sternberg (HRS) cells usually constitute only approximately 1%-2% of the cellular population and are dispersed within a polymorphous microenvironment of lymphocytes, eosinophils, plasma cells, histiocytes, and stromal elements. The World Health Organization recognizes nodular sclerosis, mixed cellularity, lymphocyte-rich, and lymphocyte-depleted subtypes of cHL. Although these subtypes differ in architecture and background composition, the scarcity and morphological heterogeneity of HRS cells are common obstacles to their recognition in limited cytological preparations [1-5].

Excisional lymph-node biopsy remains the diagnostic reference standard because it preserves tissue architecture and provides sufficient material for integrated histomorphological and immunophenotypic assessment [6,7]. Fine-needle aspiration (FNA), however, is frequently used as a rapid, economical, and minimally invasive initial investigation of lymphadenopathy and can guide patient triage and ancillary testing [8-11]. Its diagnostic performance varies among lymphoma categories. Cytologically high-grade non-Hodgkin lymphomas may yield readily recognizable monotonous large-cell populations, whereas flow cytometry can improve the evaluation of many B-cell lymphomas by demonstrating an aberrant or light-chain-restricted population. In contrast, Hodgkin lymphoma remains a prominent source of false-negative or indeterminate FNA interpretations [8,9,12,13].

This limitation is intrinsic to both disease biology and specimen acquisition. HRS cells may be absent from the sampled area, obscured within a dense reactive background, or difficult to recover from fibrotic lesions; nodular sclerosis cHL is particularly prone to paucicellular aspirates [13-19]. Conventional flow-cytometric workflows may also fail to capture HRS cells because they are rare, large, fragile, and frequently excluded by routine lymphocyte gating. Specialized multiparameter assays can identify HRS cells and HRS-cell-T-cell rosettes in selected small biopsy and cytology specimens, but their performance depends on specimen quality, antibody design, acquisition strategy, and laboratory expertise [20,21]. Consequently, direct morphological recognition of rare HRS-like cells remains an important step in FNA assessment, often followed by cell-block immunocytochemistry or tissue biopsy for definitive classification.

Whole-slide imaging provides a means to assist this visually demanding search by converting glass slides into high-resolution digital images that can be systematically reviewed and computationally analyzed [17,18]. Digital cytology nevertheless presents challenges beyond those encountered in many histological sections, including three-dimensional cell groups, variable focal planes, overlapping cells, uneven cellular distribution, and large sparsely populated slide areas. Systematic reviews and professional task-force recommendations therefore emphasize the need for cytology-specific validation, representative datasets, and careful integration of whole-slide imaging and artificial intelligence into clinical workflows [24,25].

Machine-learning and deep-learning methods have achieved strong performance in cell detection, segmentation, and classification in several areas of digital hematology; for example, convolutional neural networks have classified leukocytes and recognized blast cells in peripheral-blood and bone-marrow images [26]. Direct translation of such data-intensive approaches to lymph-

node FNA cytology is difficult, however, because HRS cells are extremely uncommon, morphologically diverse, and embedded among abundant benign inflammatory cells. These properties create severe class imbalance, increase the annotation burden, and can produce large numbers of false candidates during whole-slide screening. The scarcity of multi-institutional HRS-cell datasets further complicates model training and external validation.

We therefore formulated HRS-cell screening as an interpretable rare-object identification problem rather than an autonomous slide-level diagnostic task. The proposed method uses a limited set of expert-selected prototype cells to define transparent color, intensity, texture, and contour criteria. Sequential filtering reduces approximately 10^5–10^6 background cells to a manageable candidate list for cytopathologist review. The feature-engineering strategy was adapted from contour-based object-identification and particle-tracking methods previously used to analyze heterogeneous active-particle systems [27-32]. To accommodate differences in staining, preparation, and imaging among institutions, the framework incorporates multiple prototype classes and cross-correlation-based intensity calibration. This multicenter retrospective study evaluates candidate enrichment in positive FNA slides and the false-candidate burden in negative controls. The resulting system is intended to support prescreening and prioritization while retaining expert review and conventional clinicopathological confirmation as the basis of diagnosis.

Results

Prototype-cell recovery and training-set refinement

Across the 12 training slides, the mean prototype-cell recovery rate was 91.6%, exceeding the predefined 90% criterion (Table 1). Whole-slide screening of the training set was therefore performed, and expert review of the resulting candidates was used to refine the screening intervals. After refinement, the mean slide-level positive predictive value (PPV) in the training set was 39.2% (Table 2).

Table 1. Prototype-cell recovery performance across the 12 annotated training slides. The slide-specific recovery rate is calculated as the number of recovered prototype cells divided by the number of selected prototype cells.

Case ID	Selected prototype HRS cells	Recovered prototype HRS cells	Recovery rate (%)
No. 1	4	3	75.0
No. 2	6	6	100.0
No. 3	5	4	80.0
No. 4	6	6	100.0
No. 5	5	4	80.0
No. 6	8	8	100.0
No. 7	4	4	100.0
No. 8	6	6	100.0
No. 9	4	3	75.0
No. 10	4	4	100.0
No. 11	3	3	100.0
No. 12	6	6	100.0
Mean			**92.5**

Table 2. Slide allocation and slide-level candidate-screening results across the 78-slide multicenter dataset.

No.	Institute	Function	Preparation	BG cell	Candidate	TP	FP	PPV (%)	Mean PPV (%)
1	A	TS	CS	481,612	61	18	43	29.50	39.18
2	A	TS	CS	311,065	126	55	71	43.70	

No.	Institute	Function	Preparation	BG cell	Candidate	TP	FP	PPV (%)	Mean PPV (%)
3	A	TS	CS	1,237,007	63	25	38	39.70	
4	A	TS	CS	916,713	165	33	132	20.00	
5	A	TS	CS	720,863	63	8	55	12.70	
6	A	TS	CS	1,495,102	66	56	10	84.80	
7	A	TS	CS	848,977	61	29	32	47.50	
8	A	TS	CS	605,240	29	3	26	10.30	
9	A	TS	CS	672,230	6	4	2	66.70	
10	A	TS	CS	131,159	41	9	32	22.00	
11	B	TS	CS	1,425,472	15	5	10	33.30	
12	B	TS	CS	1,499,918	55	33	22	60.00	
13	A	IPT	CS	605,369	76	29	47	38.20	
14	A	IPT	CS	1,344,558	54	42	12	77.80	
15	A	IPT	CS	661,200	184	59	125	32.10	
16	A	IPT	CS	474,470	133	64	69	48.10	
17	A	IPT	CS	780,130	19	9	10	47.40	
18	A	IPT	CS	441,353	79	22	57	27.80	
19	B	IPT	CS	1,134,518	5	3	2	60.00	49.96
20	B	IPT	CS	1,069,652	7	3	4	42.90	
21	B	IPT	CS	866,044	35	18	17	51.40	
22	B	IPT	CS	540,687	28	21	7	75.00	
23	B	IPT	CS	1,243,285	10	5	5	50.00	
24	B	IPT	CS	1,354,296	81	26	55	32.10	
25	B	IPT	CS	446,659	15	10	3	66.70	
26	C	CPT	CS	1,419,586	102	45	57	44.10	
27	C	CPT	CS	895,273	100	36	64	36.00	
28	C	CPT	CS	378,706	21	11	10	52.40	
29	C	CPT	CS	516,066	61	20	41	32.80	
30	C	CPT	CS	289,878	41	13	28	31.70	
31	C	CPT	CS	277,074	13	9	4	69.20	
32	C	CPT	CS	915,464	82	25	57	30.50	
33	C	CPT	LBC	18,534	18	9	9	50.00	
34	D	CPT	CS	36,619	5	3	2	60.00	
35	D	CPT	CS	34,029	5	3	2	60.00	53.23
36	D	CPT	LBC	350,623	25	15	10	60.00	
37	D	CPT	LBC	36,604	285	214	71	75.10	
38	D	CPT	LBC	122,716	23	15	8	65.20	
39	D	CPT	LBC	122,716	35	20	15	57.10	
40	D	CPT	CS	82,148	18	8	10	44.40	
41	D	CPT	CS	82,148	70	22	48	31.43	
42	E	CPT	LBC	262,838	20	10	10	50.00	
43	E	CPT	LBC	150,576	8	4	4	50.00	
44	E	CPT	LBC	204,668	12	4	8	33.30	
45	E	CPT	LBC	569,037	28	14	14	50.00	

No.	Institute	Function	Preparation	BG cell	Candidate	TP	FP	PPV (%)	Mean PPV (%)
46	E	CPT	LBC	106425	15	13	2	86.70	
47	E	CPT	LBC	61,403	253	152	101	60.10	
48	E	CPT	LBC	117,371	30	21	9	70.00	
49	E	CPT	LBC	33,472	8	6	2	75.00	
50	E	CPT	LBC	23,454	9	5	4	55.60	
51	C	NC	LBC	366,829	46	2	44	4.30	
52	C	NC	LBC	278,747	31	3	28	9.70	
53	C	NC	LBC	12,387	0	-		0.00	
54	C	NC	LBC	260,643	39	3	36	7.70	
55	C	NC	LBC	71,510	0	-		0.00	
56	C	NC	LBC	40,267	14	0	14	0.00	
57	C	NC	LBC	41,690	9	0	9	0.00	
58	C	NC	CS	569,115	12	1	11	8.30	
59	B	NC	CS	1,594,981	3	0	3	0.00	
60	B	NC	CS	1,792,106	0	0	0	0.00	
61	B	NC	CS	1,628,670	1	0	1	0.00	
62	B	NC	CS	1,849,994	0	0	0	0.00	
63	B	NC	CS	2,033,651	0	0	0	0.00	
64	B	NC	CS	1,493,552	0	0	0	0.00	1.07
65	B	NC	CS	2,053,314	0	0	0	0.00	
66	B	NC	CS	2,949,944	0	0	0	0.00	
67	B	NC	CS	1,297,794	6	0	6	0.00	
68	B	NC	CS	1,901,598	4	0	4	0.00	
69	B	NC	CS	3,428,090	11	0	11	0.00	
70	B	NC	CS	295,000	13	0	13	0.00	
71	B	NC	CS	4,921	1	0	1	0.00	
72	B	NC	CS	1,303,295	1	0	1	0.00	
73	B	NC	CS	2,797,078	5	0	5	0.00	
74	B	NC	CS	1,419,974	2	0	2	0.00	
75	B	NC	CS	805,572	0	0	0	0.00	
76	B	NC	CS	966,548	1	0	1	0.00	
77	B	NC	CS	1,869,074	3	0	3	0.00	
78	B	NC	CS	3,363,257	3	0	3	0.00	

Abbreviations: BG cell: background cell, TS: training set, IPT: intra-institutional positive test set, CPT: cross-institutional positive test set, NC: negative-control set, CS: conventional smear, LBC: liquid-based cytology, positive predictive value (PPV).

Note: PPV was calculated separately for each slide as TP/(TP + FP); reported group means are arithmetic means of slide-specific PPVs, so every slide contributes equally regardless of candidate count.

Intra-institutional evaluation of independent positive slides

The refined algorithm was first applied to 13 independent positive slides obtained from Institutions A and B, both of which were represented in the training set. These slides had not undergone expert pre-annotation and were not used for prototype-cell selection or refinement of the screening criteria.

The arithmetic mean of the 13 slide-specific PPVs was 50.0%, exceeding the corresponding 39.2% mean for the 12-slide training set. Individual slide-level PPVs ranged from 27.8% to 77.8%, reflecting variation in HRS-like cell abundance and morphology and in the complexity of the cytological background (Table 2).

The number of candidates identified on individual intra-institutional positive slides ranged from 5 to 184. Thus, whole-slide screening reduced the review space from approximately 10^5–10^6 background cells in a conventional smear to a candidate list on the order of 10^2 objects.

Cross-institutional evaluation of independent positive slides

To evaluate generalizability, the algorithm was subsequently applied to 25 independent positive slides obtained from three external institutions, designated Institutions C, D, and E. Before screening, the intensity-related criteria were calibrated using the cross-correlation-based shifting procedure described in the Methods section, whereas the morphological criteria remained unchanged.

Mean slide-level PPVs were 43.3% for the eight slides from Institution C, 62.1% for the eight slides from Institution D, and 59.0% for the nine slides from Institution E. The arithmetic mean of the 25 slide-specific PPVs was 55.0% (Table 2); each slide contributed equally, irrespective of its candidate count.

The mean slide-level PPV for the cross-institutional positive set was slightly higher than the corresponding mean of 50.0% for the intra-institutional positive set. LBC slides generally exhibited higher slide-level PPVs than conventionally prepared smears. However, preparation method and institution were not evaluated as independent variables in the present dataset.

Evaluation of negative-control slides

Clinically negative control slides were evaluated to determine whether the screening criteria indiscriminately retained background objects. Among the 28 slides from Institutions B and C, 22 produced fewer than 10 candidates and eight produced none. The arithmetic mean of the slide-specific PPVs was 1.1%, substantially lower than those of the positive evaluation sets (Table 2).

Nine candidates from four clinically negative slides were accepted by expert review as exhibiting HRS-like cytomorphology at the candidate-object level. These slides nevertheless remained clinically negative and were not reclassified on the basis of isolated HRS-like objects. The remaining negative-control slides either produced no candidates or produced candidates that were rejected during expert review.

Overall performance and exploratory separation of positive and negative slides

The highest slide-level PPV among negative controls was 9.7%, whereas the lowest among the 38 independent positive slides was 10.3%; therefore, the distributions did not overlap in this cohort. Threshold-dependent true-positive rate (TPR) and false-positive rate (FPR) curves show the exploratory slide-level classification performance across candidate-level PPV thresholds, and the inset receiver operating characteristic (ROC) curve summarizes the corresponding relationship (Fig. 1).

Any PPV threshold strictly between 9.7% and 10.3% separated all positive test slides from all negative-control slides, yielding a slide-level TPR of 100% and an FPR of 0%. This threshold range was identified retrospectively in the study cohort.

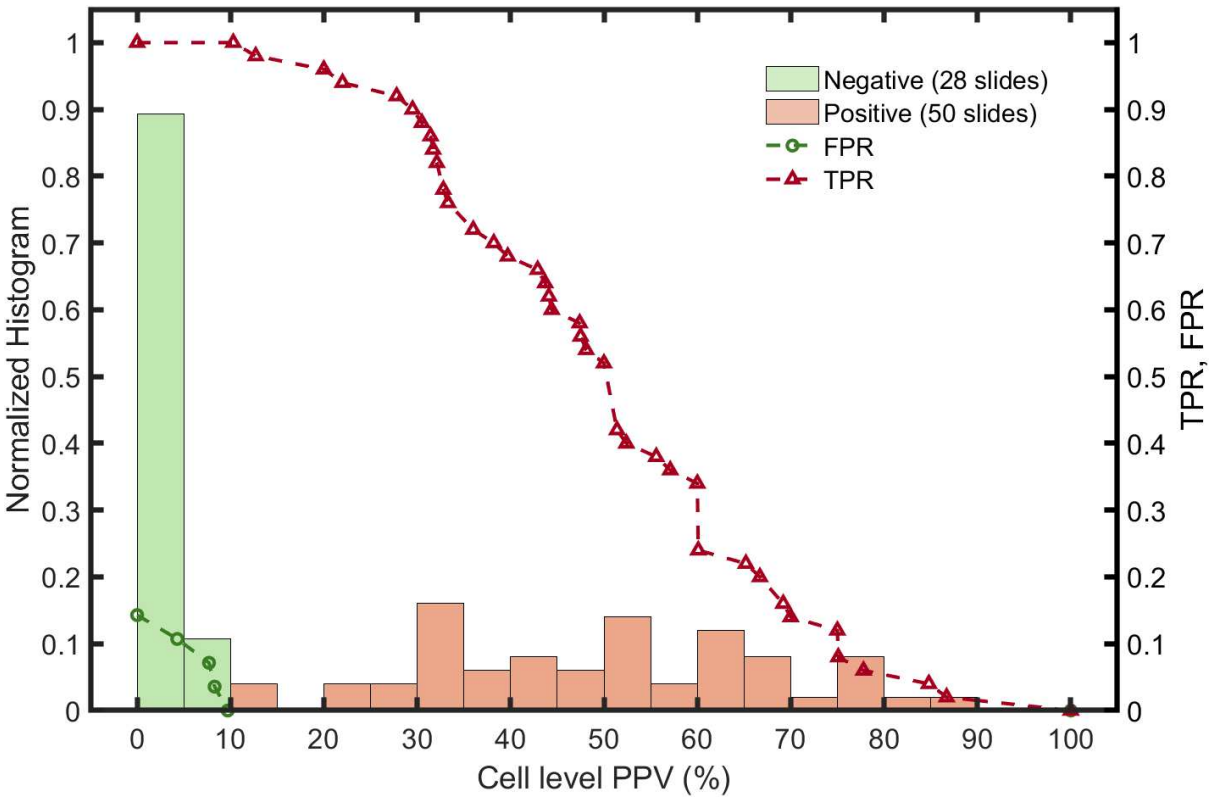

Fig. 1. Threshold-dependent exploratory slide-level classification performance based on candidate-level PPV. True-positive rate (TPR) and false-positive rate (FPR) are plotted against the positive predictive value (PPV) threshold used to classify slides. TPR = TP/(TP + FN) and FPR = FP/(FP + TN), where TP, FN, FP, and TN denote slide-level outcomes. The inset shows the corresponding receiver operating characteristic (ROC) curve across the evaluated thresholds.

Discussion

Principal findings

This multicenter study demonstrates that an interpretable prototype-driven algorithm can enrich rare HRS-like candidates across heterogeneous lymph-node FNA preparations. Mean slide-level PPVs increased from 39.2% in the training set to 50.0% and 55.0% in the intra- and cross-institutional positive test sets, respectively, whereas the negative-control mean was 1.1%. The non-overlapping PPV distributions in this cohort support the feasibility of candidate-based slide prioritization, although the retrospectively identified threshold requires independent prospective validation.

Clinical context and intended role

Excisional lymph-node biopsy remains the reference standard for diagnosing cHL because it preserves tissue architecture and permits integrated histomorphological, immunophenotypic, and molecular evaluation [5,6]. However, this procedure is invasive and may require surgical resources, whereas lymph-node FNA is less invasive, rapid, and readily repeatable [7-10]. FNA can therefore provide a useful initial assessment and help prioritize patients requiring definitive biopsy. Its principal limitation in cHL is that diagnostically relevant HRS cells are sparse, morphologically variable, and dispersed within a substantially larger reactive cellular background [11-16]. The present algorithm was developed specifically to address this rare-cell search problem.

The difficulty of assembling the present cohort further illustrates the rarity of suitable cytological material. Despite retrospective retrieval of archival slides spanning more than two

decades across five institutions, only 78 analyzable slides were available, including 50 positive slides and 28 negative controls. This limited availability, together with substantial variation in staining, specimen preparation, and imaging conditions, presents a major challenge for data-intensive machine-learning approaches that require large, balanced, and extensively annotated datasets. A prototype-driven, interpretable, feature-based approach was therefore considered particularly suitable for this rare-event setting because its screening criteria could be established from a limited number of expert-selected cells and adjusted transparently across staining conditions.

The proposed method should be regarded as an assistive prescreening system rather than an autonomous diagnostic or slide-classification tool. Its primary purpose is to reduce the manual search burden by narrowing approximately 10^5–10^6 background cells in a whole-slide image to a manageable candidate set on the order of 10^2 objects, typically fewer than 100 on a positive slide. Cytopathologists can then focus their review on algorithm-identified objects, potentially shortening the visual search pathway while retaining expert interpretation as the final decision-making step. Candidate-level PPV and any threshold derived from the present cohort should therefore be interpreted as measures for candidate enrichment and slide prioritization, not as independent diagnostic criteria. Definitive diagnosis must continue to rely on cytopathological assessment together with histological, immunophenotypic, and clinical evidence.

HRS-like candidates in negative controls

The presence of nine HRS-like candidates among clinically negative slides was not unexpected. Reactive lymph nodes may contain immunoblasts, activated lymphoid cells, histiocytes, and other inflammatory cells with enlarged nuclei, prominent nucleoli, and heterogeneous chromatin. These cytomorphological mimics can satisfy the algorithm's color, size, contour, and intensity criteria and may also be accepted as HRS-like during candidate-level expert review. Their detection indicates morphological resemblance rather than malignancy. An isolated HRS-like object is insufficient for a positive slide-level diagnosis, which must integrate the overall cytological background, tissue architecture when available, clinical findings, and appropriate immunophenotypic evidence.

Influence of specimen preparation

The observed differences between preparation methods also provide insight into the performance of the algorithm. The quantitative screening criteria were derived primarily from selected prototype HRS cells that were well isolated from surrounding background cells and exhibited clearly visible nuclear and nucleolar features. Consequently, the algorithm may be less effective when HRS cells are embedded within dense tissue fragments, obscured by background cells, or overlapped with other cells in multilayered regions. Under these conditions, cellular boundaries, contour morphology, and local intensity distributions cannot be measured as reliably. This limitation likely contributed to the higher PPV observed for liquid-based cytology preparations, including slides obtained from institutions not represented in the training set. Liquid-based preparations generally provide a more uniform cellular distribution, less background material, and reduced cellular overlap, allowing the color, intensity, and morphological features of HRS-like cells to be more clearly resolved. Thus, specimen preparation and cellular arrangement may influence algorithm performance independently of institutional staining differences.

Limitations and future work

Several additional limitations should be considered. First, although each patient contributed only one slide, FNA sampling is inherently dependent on needle placement and may not always accurately target the involved or diagnostically representative region of a lymph node. An aspirate may therefore contain few or no HRS cells or may predominantly sample surrounding reactive or nonlesional tissue. A low candidate count or negative screening result may consequently reflect sampling inadequacy rather than failure of the detection algorithm. Second, the archival slides were collected over more than two decades, and prolonged storage may have altered their staining appearance through fading, discoloration, or deterioration of the mounting medium. These storage-related changes may contribute to variations in color intensity and contrast in addition to the original differences in staining protocols and specimen preparation. Although the use of multiple criterion classes and cross-correlation-based intensity calibration improved adaptability across institutions, simple intensity shifting may not fully compensate for nonlinear staining differences, uneven illumination, focus variation, cellular overlap, and preparation- or storage-related artifacts. Unusual, degenerative, weakly stained, or low-contrast HRS cells may consequently be missed.

Finally, true positive (TP) and false positive (FP) cells were operationally defined at the candidate-object level by expert cytomorphological review. Because every HRS cell in each whole-slide image was not exhaustively annotated, candidate-level false negatives and object-level sensitivity could not be determined. Candidate-level PPV therefore measures candidate enrichment and potential review workload rather than diagnostic sensitivity or slide-level accuracy. Expert review may also introduce interobserver variability. The modest, institutionally imbalanced cohort and retrospective derivation of the PPV threshold further limit generalizability. Future work should prospectively evaluate cross-institutional calibration and PPV-based slide prioritization in larger cohorts containing broader spectra of benign mimics and hematolymphoid malignancies, with exhaustive annotation of representative slides, interobserver assessment, and measurement of review-time reduction.

Conclusions

This multicenter retrospective study demonstrates that an interpretable, prototype-driven feature-based algorithm can enrich rare HRS-like candidates in lymph-node FNA whole-slide images while maintaining a low candidate burden in negative controls. Cross-institutional intensity calibration supported application across heterogeneous staining conditions, and the marked slide-level PPV separation observed between positive and negative slides warrants prospective validation. The method is intended to shorten the visual search pathway and support, rather than replace, expert cytopathological assessment.

Materials and methods

Study design and overall workflow

This multicenter retrospective study developed and evaluated an interpretable, feature-based computer-assisted algorithm for detecting HRS-like cells in lymph-node FNA smears. The workflow (Fig. 2) comprised sample collection, whole-slide digitization, expert pre-annotation, prototype selection, algorithm development, prototype-cell recovery testing, whole-slide screening, expert review, criteria refinement, independent testing, and expert confirmation. The finalized criteria were first evaluated on slides from institutions represented in the training set and then applied, with intensity calibration, to slides from external institutions.

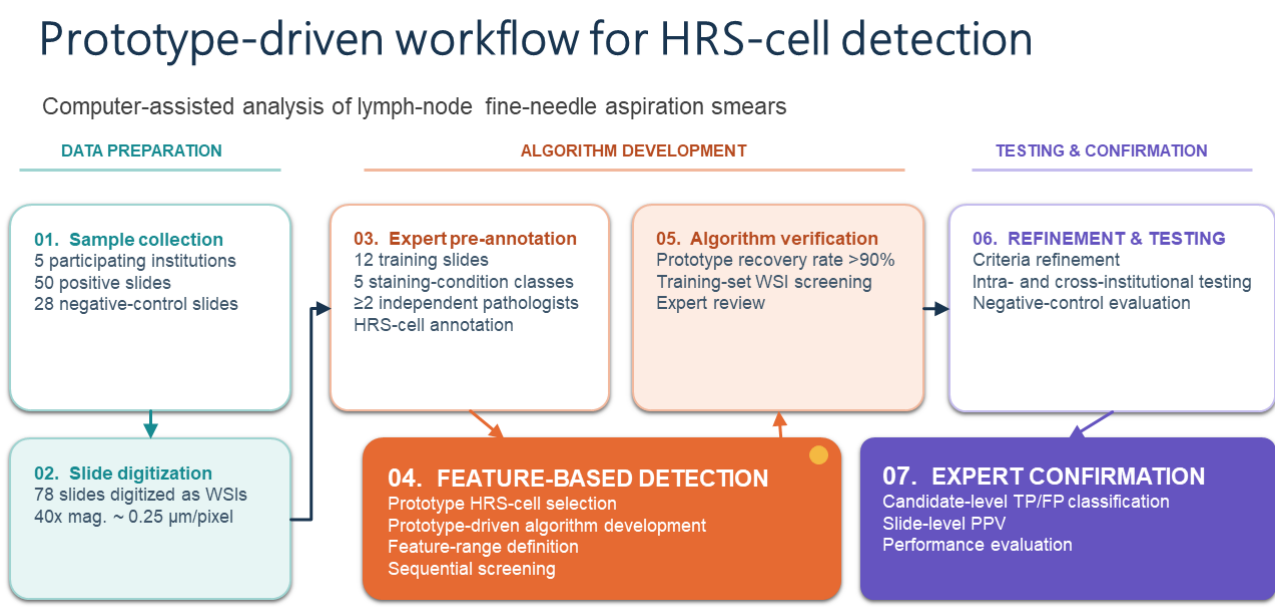

Fig. 2. Overall workflow for multicenter slide collection, whole-slide digitization, expert annotation and prototype selection, interpretable feature-based detection, prototype-cell recovery testing, criteria refinement, independent testing, and expert confirmation.

Study cohort and dataset allocation

Retrospective lymph-node FNA smear slides were collected from five collaborating medical institutions, Chia-Yi Christian Hospital (CYCH, Taiwan), Kaohsiung Medical University Hospital (KMUH, Taiwan), Kaohsiung Veterans General Hospital (KVGH, Taiwan), Oncology Hospital (OH, Vietnam), and University Medical Center (UMC, Vietnam). The participating institutions represented diverse clinical and laboratory settings, enabling evaluation of algorithm robustness across variations in staining protocols, specimen preparation, and background cellular composition. All specimens were de-identified before analysis and handled in accordance with the applicable institutional ethical requirements.

A total of 78 slides were included and allocated to three functionally distinct datasets (Table 2): a 12-slide annotated training set, a 38-slide independent positive test set, and a 28-slide negative-control set. The training slides from Institutions A and B were selected to encompass a broad range of staining characteristics and expand the feature-distribution domain used for algorithm development. These slides underwent expert pre-annotation and were used to define representative HRS-cell characteristics and the initial quantitative feature criteria.

The independent positive test set comprised 38 slides that were not used for prototype selection or algorithm development. Of these, 13 slides originated from Institutions A and B and were used for intra-institutional evaluation. The remaining 25 slides originated from Institutions C-E and were used to assess cross-institutional generalizability. An additional 28 FNA slides with a corresponding benign histological diagnosis were retrieved from Institutions B and C and formed the negative-control set used to quantify the false-candidate burden and evaluate whether the screening criteria were excessively sensitive.

The clinicopathological characteristics of all HL cases (cases 1–50) are detailed in Table 3. Patient ages ranged from 18 to 74 years without gender predisposition. Aspiration sites covered primary anatomical nodal basins, predominantly the neck (23 cases) and supraclavicular region (18

cases), alongside mediastinal (5 cases), abdominal (1 case), and inguinal (2 cases) presentations. Histological subtypes encompassed the full spectrum of HL, including nodular sclerosis (22 cases), lymphocyte-rich (10 cases), mixed cellularity (8 cases), lymphocyte-depleted (3 cases), classical type (4 cases), and nodular lymphocyte-predominant HL (2 cases), spanning clinical stages 1 to 4. Notably, original FNA cytological interpretations were reported as benign in 30 cases (60.0%), atypical in 11 cases (22.0%), and suspicious for malignancy/malignant in 9 cases (18.0%), underscoring the diagnostic difficulty and high rate of false-negative or indeterminate interpretations inherent to routine manual cytological evaluation.

Table 3. Clinicopathological characteristics of the 50 HL cases with lymph-node FNAs.

No.	Institute	Gender	Age	Cytology site	Original cytology diagnosis	Histological diagnosis	Stage	EBER	Treatment	Outcome
1	A	M	68	Right supraclavicular	Benign	HL, mixed cellularity	3	NA	C/T	NA
2	A	F	36	Left supraclavicular	Atypical	HL, lymphocyte-depleted	2	NA	C/T	NA
3	A	M	68	Left neck	SFM	HL, classical type	2	NA	C/T	NA
4	A	M	68	Right supraclavicular	Benign	HL, classical type	3	NA	C/T	NA
5	A	F	55	Right inguinal	Atypical	HL, mixed cellularity	2	NA	C/T	NA
6	A	F	42	Abdominal	SFM	NLPHL	2	NA	C/T	NA
7	A	M	70	Right neck	SFM	HL, nodular sclerosis	3	NA	C/T	NA
8	A	F	18	Left neck	Atypical	HL, lymphocyte-rich	2	NA	C/T	NA
9	A	F	47	Right neck	Benign	HL, nodular sclerosis	2	NA	C/T	NA
10	A	F	18	Right neck	Atypical	HL, lymphocyte-rich	2	NA	C/T	NA
11	B	F	33	Right supraclavicular	Benign	HL, nodular sclerosis	2A	NA	IO+C/T	Alive
12	B	F	33	Right supraclavicular	Benign	HL, nodular sclerosis	2A	NA	IO+C/T	Alive
13	A	M	68	Right supraclavicular	Benign	HL, classical type	3	NA	C/T	NA
14	A	F	18	Left neck	Atypical	HL, lymphocyte-rich	2	NA	C/T	NA
15	A	F	74	Left neck	Benign	HL, mixed cellularity	3	NA	C/T	NA
16	A	F	36	Left supraclavicular	Atypical	HL, lymphocyte-depleted	2	NA	C/T	NA
17	A	M	68	Left neck	SFM	HL, classical type	2	NA	C/T	NA
18	A	F	55	Right inguinal	Atypical	HL, mixed cellularity type	2	NA	C/T	NA
19	B	F	26	Right supraclavicular	Atypical	HL, nodular sclerosis	2A	NA	IO+C/T	Alive
20	B	F	33	Left supraclavicular	Benign	HL, nodular sclerosis	2A	NA	IO+C/T	Alive
21	B	F	33	Left supraclavicular	Benign	HL, nodular sclerosis	2A	NA	IO+C/T	Alive
22	B	F	28	Right supraclavicular	Benign	HL, mixed cellularity	3	NA	IO+C/T	Alive
23	B	M	34	Right supraclavicular	Benign	HL, lymphocyte-rich	3B	NA	IO+C/T	Alive
24	B	M	34	Right supraclavicular	Benign	HL, lymphocyte-rich	3B	NA	IO+C/T	Alive
25	B	F	28	Right supraclavicular	Benign	HL, mixed cellularity	3	NA	IO+C/T	Alive
26	C	F	31	Left neck	Benign	HL, lymphocyte-rich	2	Negative	IO+C/T	Alive
27	C	F	31	Left neck	Benign	HL, lymphocyte-rich	2	Negative	IO+C/T	Alive
28	C	M	48	Right neck	Benign	HL, nodular sclerosis	3	Positive	IO+C/T	Alive
29	C	M	48	Right neck	Benign	HL, nodular sclerosis	3	Positive	IO+C/T	Alive
30	C	M	28	Left neck	Benign	HL, nodular sclerosis	2	Negative	Refuse treatment	Alive
31	C	M	28	Left neck	Benign	HL, nodular sclerosis	2	Negative	Refuse treatment	Alive
32	C	M	47	Left neck	Benign	HL, mixed cellularity	3	Positive	IO+C/T	Alive
33	C	M	64	Right neck	Benign	HL, lymphocyte-rich	4A	Positive	IO+C/T	Alive
34	D	M	26	Mediastinal	Benign	HL, lymphocyte-rich	4	NA	C/T	Alive

No.	Institute	Gender	Age	Cytology site	Original cytology diagnosis	Histological diagnosis	Stage	EBER	Treatment	Outcome
35	D	M	26	Mediastinal	Benign	HL, lymphocyte-rich	4	NA	C/T	Alive
36	D	F	26	Left neck	SFM	HL, nodular sclerosis	2	Negative	C/T	Alive
37	D	M	29	Right neck	SFM	HL, nodular sclerosis	2	Negative	C/T	Alive
38	D	F	26	Right neck	Benign	HL, nodular sclerosis	2	Negative	C/T+RT	Alive
39	D	M	30	Left neck	SFM	HL, nodular sclerosis	2	Positive	NA	NA
40	D	F	27	Mediastinal	SFM	HL, nodular sclerosis	2	Negative	C/T	Alive
41	D	M	18	Mediastinal	Atypia	HL, nodular sclerosis	3	Negative	C/T	Alive
42	E	M	43	Right neck	Atypical	HL, nodular sclerosis	NA	NA	NA	Alive
43	E	F	23	Right neck	Benign	HL, nodular sclerosis, syncytial variant	3B	Negative	IO+C/T	Alive
44	E	M	55	Left neck	Benign	NLPHL	1A	Negative	C/T+RT	Alive
45	E	M	54	Left neck	Benign	HL, lymphocyte-rich	1	Negative	C/T	Alive
46	E	F	29	Left neck	Benign	HL, nodular sclerosis	NA	Negative	Refuse treatment	Alive
47	E	M	39	Left neck	Malignant	HL, lymphocyte-depleted	4	NA	C/T	Alive
48	E	F	30	Left neck	Atypical	HL, nodular sclerosis	NA	NA	NA	Alive
49	E	M	37	Left neck	Benign	HL, lymphocyte-rich	2A	Negative	C/T+RT	Alive
50	E	M	33	Left supraclavicular	Benign	HL, nodular sclerosis	3	Negative	IO+C/T	Alive

Abbreviations: HL: Hodgkin lymphoma, FNA: fine needle aspiration, M: male, F: female, SFM: Suspicious for malignancy, NLPHL: nodular lymphocyte-predominant Hodgkin lymphoma, NA: not available, C/T: chemotherapy, IO: immunotherapy, RT: radiotherapy.

Whole-slide imaging, expert annotation, and prototype selection

All FNA smear slides were digitized using a standardized whole-slide imaging scanner (Soranus-300; Song-Yi System Corp., Ltd., Taiwan) equipped with a 40× objective lens, providing a spatial resolution of 0.25 μm/pixel. Multifocus scanning was employed to minimize local focus variations and ensure adequate visualization of diagnostically relevant nuclear and nucleolar features. For each scanned field, images were acquired at nine focal planes. The three planes exhibiting the highest image contrast were selected and combined by focus stacking to generate the final whole-slide image. The resulting whole-slide images were stored in a pyramidal tiled format to facilitate efficient processing of the high-resolution cytological images.

The 12 training slides encompassed diverse staining and specimen-preparation conditions and underwent expert pre-annotation by experienced cytopathologists. Based on staining characteristics, the slides were categorized into five representative staining-condition classes, broadening the feature distributions available for algorithm development and adaptive criterion selection (Fig. 3). Conceptually, these classes served as representative basis elements spanning the relevant staining and morphological feature space. Each annotated HRS cell was independently reviewed by at least two cytopathologists, and the annotated population was selected to reflect the morphological variability encountered in routine practice rather than only ideal textbook appearances.

From the pre-annotated HRS cells, a subset with well-defined nuclear and nucleolar characteristics was further selected as prototype cells. These prototype cells served as anchor points for establishing representative morphological and intensity-based feature distributions and defining the initial quantitative screening criteria. Representative prototype HRS cells from the five staining-condition classes are shown in Fig. 4. Accordingly, the prototype set represented a refined subset of the pre-annotated HRS-cell population rather than an exhaustive ground truth.

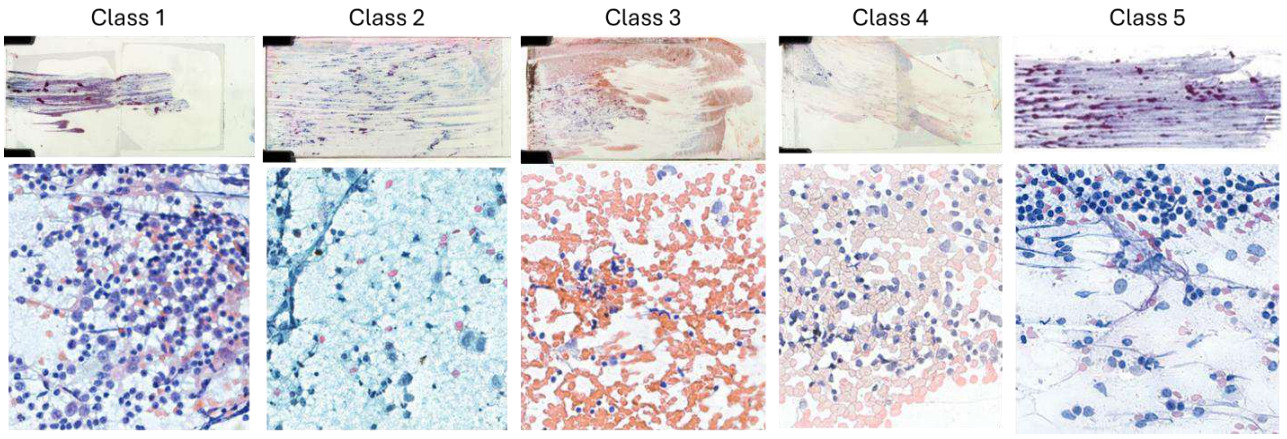

Fig. 3. Five representative staining-condition classes identified from the annotated training slides and used to define class-specific feature distributions.

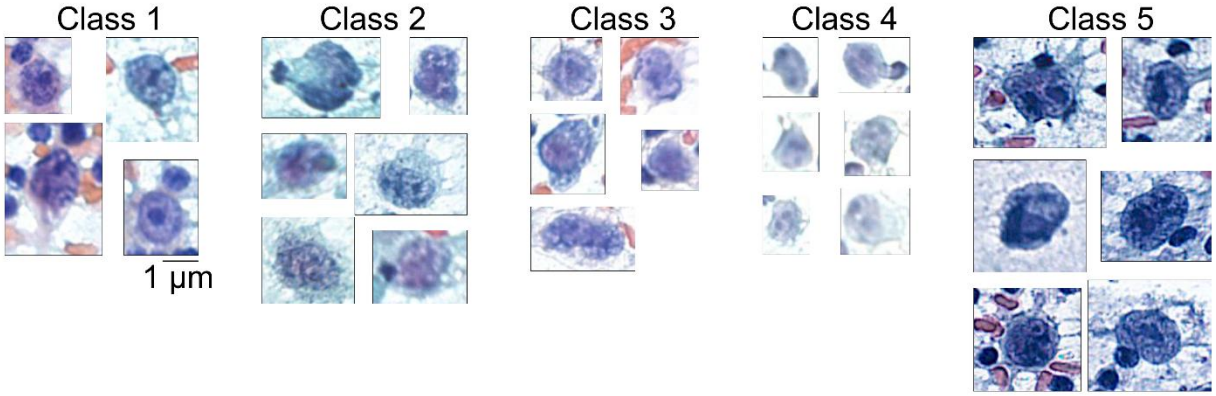

Fig. 4. Representative prototype Hodgkin and Reed–Sternberg (HRS) cells from the five staining-condition classes used to define class-specific screening criteria.

Prototype-driven feature-based detection

A single staining-condition class was selected for each whole-slide image, and its screening criteria were applied throughout the slide. Following RGB segmentation and candidate-object extraction, the algorithm evaluated two complementary features of each candidate: contour morphology, which characterized size, elongation, and contour regularity, and intranuclear intensity heterogeneity, which characterized staining variation within the nuclear contour. These features were evaluated sequentially using the criteria for the selected class.

RGB segmentation and candidate-object extraction

Each 24-bit RGB image tile was separated into red (R), green (G), and blue (B) intensity maps (Fig. 5). Initial channel-specific intensity intervals were derived from prototype-cell distributions. Pixels within the corresponding interval for each channel were retained to generate binary masks. Candidate regions were obtained by intersecting the three masks, after which connected-component and contour analyses identified individual candidate objects. This initial color-filtering stage excluded obvious background regions and reduced the number of objects requiring more computationally intensive analysis of contour morphology and intranuclear intensity heterogeneity.

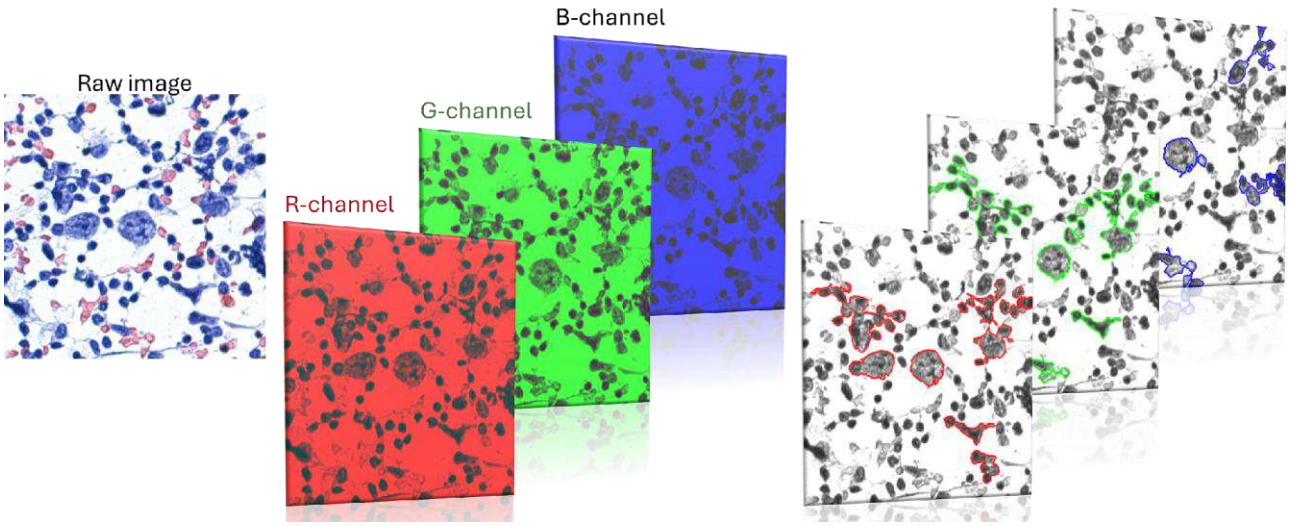

Fig. 5. RGB-based segmentation and contour detection for candidate-object identification. Each RGB tile was separated into red (R), green (G), and blue (B) intensity maps. Channel-specific filtering generated binary masks, which were intersected before connected-component and contour analysis identified candidate objects for quantitative feature evaluation.

Morphological feature extraction

For each candidate contour, the projected area (A), major axis (a), minor axis (b), aspect ratio (α), and spherical order parameter (S) were calculated, as illustrated in Fig. 6. The major and minor axes were obtained by fitting an ellipse to the detected contour, and the aspect ratio was defined as:

$$\alpha \equiv \frac{a}{b} \qquad \text{(eq. 1)}$$

The aspect ratio quantifies the elongation of a candidate object.

To characterize contour regularity independently of its absolute size, the contour was transformed from Cartesian coordinates to a polar representation centered at the contour centroid. The spherical order parameter was then defined as:

$$S \equiv 1 - \sqrt{\frac{\sum_i^N (r_i - \langle r \rangle)^2}{N \langle r \rangle^2}} \qquad \text{(eq. 2)}$$

where r_i is the radial distance between the centroid and the i-th contour point, $\langle r \rangle$ is the mean radial distance, and N is the number of contour points. The value of S approaches 1 for a perfectly circular contour and decreases as radial fluctuations, contour irregularity, or elongation increase.

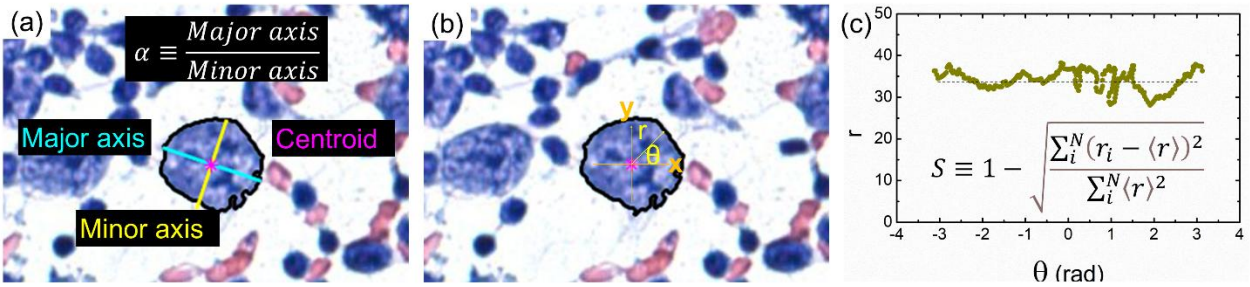

Fig. 6. Contour-based morphological features of a candidate cell. An ellipse fitted to the nuclear contour provided the major and minor axes and aspect ratio. Transformation to the centroid-based polar representation r(θ) enabled calculation of the spherical order parameter from radial fluctuations.

Intranuclear intensity heterogeneity

In addition to contour morphology, intensity-based features were evaluated within each candidate region to characterize intranuclear staining patterns. Hematoxylin-stained nuclear and nucleolar structures appear blue to purple and therefore generate substantial contrast in the red and green intensity channels. Hence, the normalized heterogeneity index H was defined as:

$$H \equiv \sqrt{\frac{\sum_i^N \ (R_i - \langle R \rangle)^2}{N \langle R \rangle^2}}, \qquad \text{(eq. 3)}$$

where R_i denotes the red-channel intensity of the ith pixel, $\langle R \rangle$ is the mean red-channel intensity within the contour, and N is the number of pixels. This dimensionless index captures relative intranuclear texture nonuniformity, including prominent or multiple nucleoli features of HRS cell, that cannot be quantified by size and shape alone.

The mean square variation of grayscale intensity was additionally calculated as:

$$\sigma^2 \equiv \frac{1}{N} \sum_i^N \ |I_i - \langle I \rangle|^2 \qquad \text{(eq. 4)}$$

where I_i is the grayscale intensity of the ith pixel and $\langle I \rangle$ is the mean grayscale intensity within the contour. Unlike the heterogeneity as a normalized index, the mean square variation retains the magnitude of grayscale contrast and is sensitive to the absolute separation between dark nuclear or nucleolar regions and the surrounding nucleoplasm. Uniform regions produce low values, whereas prominent nucleoli, heterogeneous chromatin, or multiple high-contrast intranuclear structures produce higher values. Thus, H describes relative heterogeneity after normalization for mean brightness, whereas the grayscale variance quantifies absolute intranuclear contrast.

Sequential multistage screening

Candidate objects were screened sequentially using the RGB intensity, contour morphology, and intranuclear intensity criteria for the staining-condition class selected for the slide (Table 4). At each stage, only objects within the specified range proceeded to the next stage, progressively reducing the number of background objects. Candidates that met all screening criteria were retained as the suspicious HRS cells.

Table 4. Class-specific screening parameters derived from the five representative staining-condition classes.

Screening parameter	Class 1	Class 2	Class 3	Class 4	Class 5
Red-channel intensity	78–149	57–169	137–190	130–198	33–123
Green-channel intensity	81–166	95–170	149–190	147–208	34–130
Blue-channel intensity	123–209	135–215	184–231	151–231	111–211
Projected area (A)	1361–2115	1595–3117	1470–2761	849–1562	2022–3103
Aspect ratio (α)	1.15–1.33	1.18–1.63	1.04–1.87	1.15–1.49	0.09–1.52
Heterogeneity index (H)	0.095–0.125	0.125–0.145	0.115–0.145	0.155–0.215	0.090–0.130
Grayscale variance (σ^2)	383–461	420–599	425–550	234–380	714–877

Prototype-cell recovery and training-set refinement

Following algorithm development, initial verification determined whether the expert-selected prototype cells were successfully recovered. The slide-specific and mean prototype-cell recovery rates are summarized in Table 1. After the mean recovery rate exceeded the predefined 90% criterion, whole-slide screening was performed on the training set. Algorithm-generated candidates were independently reviewed by experienced cytopathologists and classified as HRS-like (true positive) or non-HRS-like (false positive). These results guided refinement of the screening criteria.

Cross-institutional intensity calibration

For slides obtained from different institutions, intensity-related screening criteria were calibrated using histogram cross-correlation to determine the appropriate intensity shift associated with variations in staining and imaging conditions. A representative slide from one of the five staining-condition classes in the training set (Fig. 3) was selected as the reference according to its similarity to the target slide. For each RGB color channel, the cross-correlation between the intensity histogram of the reference slide, R(I), and that of the target slide, T(I), was defined as:

$$C(\Delta I) = \int R(I)T(I + \Delta I)dI \qquad \text{(eq. 5)}$$

where ΔI represents the relative intensity displacement between the two histograms. The value of ΔI that maximized C(ΔI) was used to shift the corresponding channel-specific screening intervals. This procedure compensated for systematic color-level differences across institutions. The RGB intensity histograms of the reference and target images and the resulting cross-correlation function are illustrated in Fig. 7.

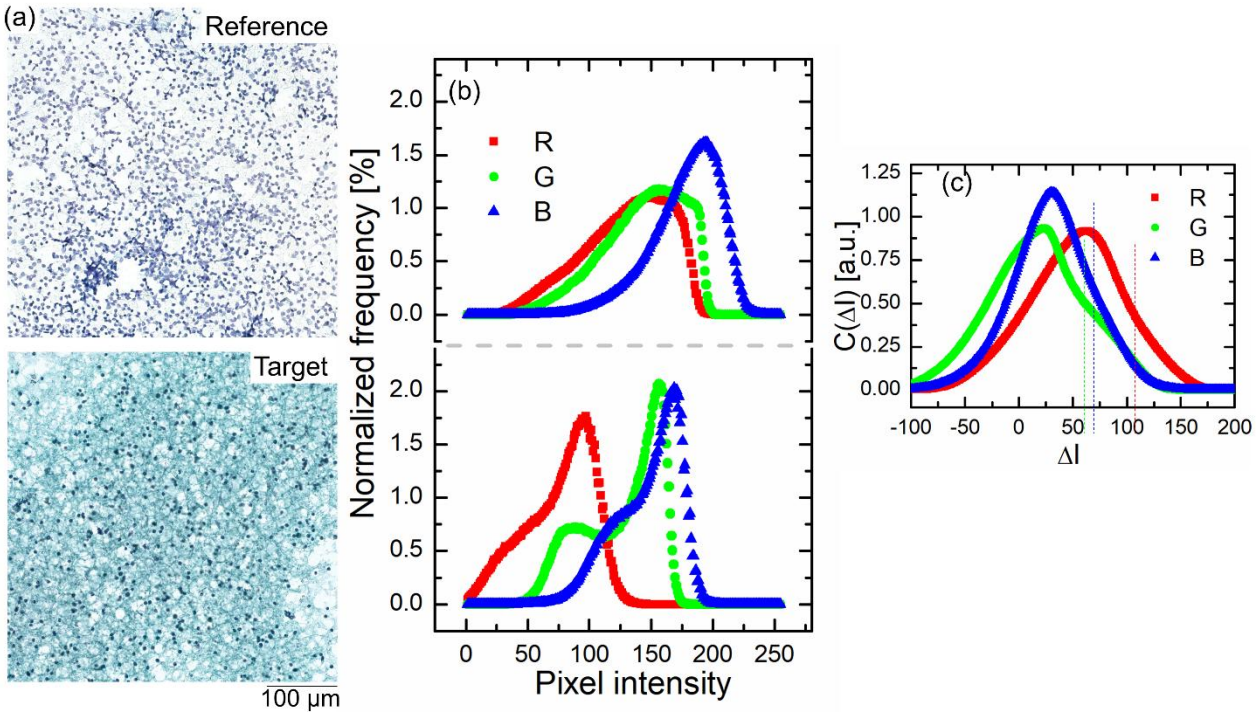

Fig. 7. Cross-institutional intensity calibration by histogram cross-correlation. (a) Representative sub-images from the reference slide (Case 10, staining-condition class 4) and the target slide (Case 36), respectively. (b) Normalized intensity histograms for the red, green, and blue channels of the reference and target images, calculated using only pixels within cell-occupied regions. (c) Cross-correlation functions for the three color channels. For each channel, the lag at which cross-correlation is maximal determines the intensity shift applied to the corresponding screening interval for the target slide.

Independent evaluation, outcome measures, and statistical analysis

The finalized algorithm was applied first to an independent positive test set from the institutions represented in the training set and subsequently to positive and negative-control slides from external institutions, using the cross-institutional intensity calibration described above. No further modification of the screening criteria was performed during independent testing.

During whole-slide screening, objects satisfying all predefined color, morphological, and intensity criteria were retained as candidates and independently reviewed by experienced cytopathologists. A candidate confirmed to exhibit HRS-like cytomorphological features was classified as a true-positive (TP) candidate, whereas a rejected candidate was classified as a false-positive (FP) candidate. For each slide, candidate-level PPV was calculated as TP/(TP + FP). This metric quantifies enrichment within the algorithm-generated candidate list and does not represent diagnostic sensitivity.

Analyses were descriptive and were performed at the slide level unless otherwise stated. Candidate counts and slide-specific PPVs were reported for individual slides. Group PPV was summarized as the arithmetic mean of the slide-specific PPVs so that each slide contributed equally, irrespective of its cellularity or candidate count. For exploratory slide-level discrimination, a slide was classified as positive when its candidate-level PPV met or exceeded the evaluated threshold. The TPR was the proportion of clinically positive slides classified as positive, and the FPR was the proportion of clinically negative slides classified as positive. TPR and FPR were evaluated across PPV thresholds to generate the ROC curve. Because threshold selection and evaluation used the same retrospective cohort, these analyses were considered exploratory; no inferential hypothesis testing was performed.

Data availability

The datasets generated and/or analyzed during the current study are not publicly available due to containing individual patient data and being under a license agreement with the providing center, but are available from the corresponding author upon reasonable request. Requests for access to these data should be made to Chien-Chin Chen, hlmarkc@gmail.com.

Author contributions

C.C.C. conceived the idea of this work. C.W.I. designed the methodology and the software of this work. G.H.T., T.A.T., T.N.Q.L., U.L.L., C.C.W., H.W.C., M.X.D., D.N.T.P., T.D.P., L.H.F., and C.C.C. provided the resources and participated in the data curation. W.C.T. and M.J.Z. performed the investigation and formal analysis. C.C.C. and C.W.I. carried out the validation and performed visualization and statistical analysis. C.W.I., G.H.T., and T.A.T. prepared and wrote the original draft. C.C.C. and C.W.I. reviewed and edited the manuscript. C.C.C. supervised this work and acquired funding for this work. All authors have read and agreed to the published version of the manuscript.

Acknowledgements

This study is supported by Ditmanson Medical Foundation Chia-Yi Christian Hospital (Project No. R114-14) and the National Science and Technology Council, Taiwan (Grant No. NSTC 115-2112-M-194-003). We are particularly grateful to Professor William C. Faquin (Harvard Medical School) for his valuable insights and inspiration that contributed to the development of this study. We also acknowledge the technical assistance provided by Chun-Chieh Chiang and Songyi System Corporation.

Competing interests

The authors declare no competing interests.

Ethics approval

This multicenter study was approved by the Institutional Review Board of Ditmanson Medical Foundation Chia-Yi Christian Hospital (CYCH-IRB-2025094; September 4, 2025). The requirement for informed consent was waived because all data were fully anonymized and the study involved no more than minimal risk.

References

1. Connors, J. M. et al. Hodgkin lymphoma. *Nat. Rev. Dis. Primers* **6**, 61 (2020).
2. Kuppers, R. The biology of Hodgkin's lymphoma. *Nat. Rev. Cancer* **9**, 15–27 (2009).
3. Alaggio, R. et al. The 5th edition of the World Health Organization Classification of Haematolymphoid Tumours: Lymphoid Neoplasms. *Leukemia* **36**, 1720–1748 (2022).
4. Campo, E. et al. The International Consensus Classification of Mature Lymphoid Neoplasms. *Blood* **140**, 1229–1253 (2022).
5. Matsuki, E. & Younes, A. Lymphomagenesis in Hodgkin lymphoma. *Semin. Cancer Biol.* **34**, 14–21 (2015).
6. Wang, H. W., Balakrishna, J. P., Pittaluga, S. & Jaffe, E. S. Diagnosis of Hodgkin lymphoma in the modern era. *Br. J. Haematol.* **184**, 45–59 (2019).
7. Takahara, T., Satou, A., Tsuzuki, T. & Nakamura, S. Hodgkin lymphoma: biology and differential diagnostic problem. *Diagnostics* **12**, 1507 (2022).
8. Schmidt, R. L., Hall, B. J., Wilson, A. R. & Layfield, L. J. A systematic review and meta-analysis of the usefulness of fine-needle aspiration and core needle biopsy in diagnosing lymphoma. *Am. J. Clin. Pathol.* **139**, 16–28 (2013).
9. Frederiksen, J. K., Sharma, P., Casulo, C. & Burack, W. R. Systematic review of the effectiveness of fine-needle aspiration and/or core needle biopsy for subclassifying lymphoma. *Arch. Pathol. Lab. Med.* **139**, 345–351 (2015).
10. Meda, B. A. et al. Diagnosis of Hodgkin's disease by fine needle aspiration cytology: cyto-histologic correlation. *Acta Cytol.* **44**, 623–627 (2000).
11. Dong, H. Y., Harris, N. L., Preffer, F. I. & Pitman, M. B. Fine-needle aspiration biopsy in the diagnosis and classification of primary and recurrent lymphoma. *Mod. Pathol.* **14**, 472–481 (2001).
12. Ha, H. J. et al. Utility and limitations of fine-needle aspiration cytology in the diagnosis of lymphadenopathy. *Diagnostics* **13**, 728 (2023).
13. Chhieng, D. C., Cangiarella, J. F., Symmans, W. F. & Cohen, J. M. Fine-needle aspiration cytology of Hodgkin disease: a study of 89 cases with emphasis on the false-negative cases. *Cancer Cytopathol.* **93**, 52–59 (2001).
14. Handa, U. et al. Pitfalls in the cytological diagnosis of nodal Hodgkin lymphoma. *Diagn. Cytopathol.* **52**, 715–721 (2024).
15. Zhang, J. R., Raza, A. S., Greaves, T. S. & Cobb, C. J. Fine-needle aspiration diagnosis of Hodgkin lymphoma using current WHO classification. *Diagn. Cytopathol.* **34**, 397–402 (2006).
16. Garg, S. et al. Fine-needle aspiration diagnosis of lymphoma based on cytomorphology alone: how accurate is it? *J. Cytol.* **38**, 164–170 (2021).
17. Das, D. K. et al. Hodgkin's lymphoma: diagnostic difficulties in fine-needle aspiration cytology. *Diagn. Cytopathol.* **37**, 564–573 (2009).
18. Hehn, S. T., Grogan, T. M. & Miller, T. P. Utility of fine-needle aspiration as a diagnostic technique in lymphoma. *J. Clin. Oncol.* **22**, 3046–3052 (2004).
19. Jogai, S., Dey, P., Al Jassar, A., Amanguno, H. G. & Adesina, A. O. Role of fine needle aspiration cytology in nodular sclerosis variant of Hodgkin's lymphoma. *Acta Cytol.* **50**, 507–512 (2006).
20. Fromm, J. R., Thomas, A. & Wood, B. L. Flow cytometry can diagnose classical Hodgkin lymphoma in lymph nodes with high sensitivity and specificity. *Am. J. Clin. Pathol.* **131**, 322–332 (2009).
21. Chan, A. et al. Role of flow cytometric immunophenotyping for classic Hodgkin lymphoma in small biopsy and cytology specimens. *Arch. Pathol. Lab. Med.* **146**, 462–468 (2022).
22. Niazi, M. K. K., Parwani, A. V. & Gurcan, M. N. Digital pathology and artificial intelligence. *Lancet Oncol.* **20**, e253–e261 (2019).
23. Pantanowitz, L. et al. Twenty years of digital pathology. *J. Pathol. Inform.* **9**, 40 (2018).

24. Antonini, P. et al. Relevance of the College of American Pathologists guideline for validating whole slide imaging for diagnostic purposes to cytopathology. *Cytopathology* **34**, 5–14 (2023).
25. Kim, D. et al. Digital cytology part 2: artificial intelligence in cytology: a concept paper with review and recommendations from the American Society of Cytopathology Digital Cytology Task Force. *J. Am. Soc. Cytopathol.* **13**, 97–110 (2024).
26. Matek, C., Schwarz, S., Spiekermann, K. & Marr, C. Human-level recognition of blast cells in acute myeloid leukaemia with convolutional neural networks. *Nat. Mach. Intell.* **1**, 538–544 (2019).
27. Io, C. W., Chen, T. Y., Yeh, J. W. & Cai, S. C. Experimental investigation of mesoscopic heterogeneous motion of laser-activated self-propelling Janus particles in suspension. *Phys. Rev. E* **96**, 062601 (2017).
28. Io, C. W. & Lee, Y. Y. Mesoscopic clustering behavior of thermophoretic-type active particle suspension under quasi-one-dimensional microconfinement. *Phys. Rev. E* **104**, 014607 (2021).
29. Io, C. W., Shane, Y. X. & Tang, Y. S. The collective motion of thermophoretic-type active particle suspension under spatial-temporal modulation. *AIP Adv.* **11**, 125028 (2021).
30. Marchetti, M. C. et al. Hydrodynamics of soft active matter. *Rev. Mod. Phys.* **85**, 1143 (2013).
31. Bechinger, C. et al. Active particles in complex and crowded environments. *Rev. Mod. Phys.* **88**, 045006 (2016).
32. Palacci, J., Sacanna, S., Steinberg, A. P., Pine, D. J. & Chaikin, P. M. Living crystals of light-activated colloidal surfers. *Science* **339**, 936–940 (2013).

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- Supplementarydata.pdf