Multiscale Image Enhancement and Visualization Analysis for Historical Map/s

Guilei Ren
Faculty of History and Archaeology, Anyang Normal University, Anyang 455000, China

***Abstract*—Based on the degradation of historical maps, such as fading, stains, creases, blurred lines, and text adhesion, multi-scale image enhancement and spatial element visualization analysis are studied. This study constructs a technical process consisting of sample division, image preprocessing, multi-scale enhancement, quality evaluation, and spatial feature extraction, and further evaluates map registration, vectorization, optical character recognition, and semantic segmentation on the same test maps. The results show that the PSNR, SSIM, EPI, and RCS of the multi-scale enhancement model are 27.94 dB, 0.842, 0.817, and 0.81, respectively, and the comprehensive quality score is 0.872. The road connectivity, water boundary matching, and administrative boundary topology consistency rates reach 87.9%, 85.4%, and 82.6%, while place name and map symbol recognition rates increase to 84.7% and 82.3%. Downstream evaluation shows that registration RMSE decreases from 8.31 to 4.12 px, vectorization F1 increases from 0.713 to 0.871, OCR character accuracy increases from 62.4% to 84.7%, and semantic-segmentation mIoU increases from 0.614 to 0.802. These results indicate that the enhancement improves both visual quality and the reliability of GIS-oriented spatial processing.**

Keywords—Historical maps; multiscale image enhancement; image restoration; spatial feature recognition; visualization analysis

I. INTRODUCTION

Historical maps preserve urban morphology, transport networks, water systems, administrative boundaries and place names, supporting historical geography and cultural-heritage digitization. Long-term storage and scanning introduce fading, stains, creases, blurred linework and text adhesion, reducing both readability and spatial-processing reliability. The task is therefore to restore line, text and symbol structures without amplifying paper noise or altering spatial relationships. The methodological contribution is a degradation-aware, task-coupled workflow rather than a new primitive enhancement operator: quantified degradation characterization guides controlled evaluation, while global, local and edge responses are fused to preserve map topology and support spatial-feature extraction and GIS-downstream tasks.

II. RELATED WORK

A. Historical Map Digitization and Feature Extraction

Existing historical-map studies are largely task-specific. Schlegel [1] proposed semi-automated object extraction, treating preprocessing as support for extraction rather than an independently evaluated stage. Mäyrä et al. [2] used historical maps for long-term land-use interpretation, emphasizing source quality rather than restoration procedures. Ahmed et al. [3] applied unsupervised learning and GEOBIA to cadastral-map digitization, while Mutch and Newsam [4] examined conversion of hand-drawn maps into GIS; both concentrate on structural conversion. Arzoumanidis et al. [5] addressed semantic segmentation, whereas Pai et al. [6] focused on symbol detection. These approaches optimize individual extraction or recognition objectives and leave the effect of restoration quality across multiple spatial tasks insufficiently examined.

B. Multiscale Enhancement and Research Gap

Vijayalakshmi et al. [7] developed a multiscale enhancement framework for low-contrast images, but its formulation is not tailored to historical maps where paper texture, weak linear features, text adhesion, and topological continuity coexist. The contribution claimed here is therefore not a new brightness-correction operator, texture enhancer, edge filter, optimization algorithm, or standalone feature extractor. It lies primarily in a degradation-aware workflow that quantifies map-specific deterioration, couples global, local, and detail responses with line-topology and text-readability requirements, and carries the restored output into spatial extraction and GIS evaluation. The architecture is thus a task-specific organization of established enhancement operations, while the degradation model and downstream evaluation provide the map-oriented constraints that distinguish the workflow from generic image restoration.

III. MATERIALS AND METHODS

A. Historical Map Data Sources and Sample Composition

Historical map samples cover urban blocks, transportation routes, waterway boundaries, administrative divisions, and topographic symbols. The primary dataset contains 120 map sheets from publicly available scanned archives, digitized historical-map datasets, and verifiable geographic sources, with 84 sheets for training, 18 for validation, and 18 for testing; scanning resolution ranges from 300 to 600 dpi and map sheets are uniformly tiled for processing. To assess transfer beyond this collection, external validation is additionally performed on the public Semantic Segmentation Map Dataset (Semap) [8], which contains 1,439 manually annotated real map samples from heterogeneous historical-map collections. The official test partition contains 144 real samples and was used exclusively for external evaluation. Semap defines background, boundary/contour, built, non-built, water, and road-network labels; following its benchmark protocol, mIoU was calculated over the four geographic classes of built, non-built, water, and road network. No Semap image was used for parameter selection, and the complete G6 configuration was fixed before external testing.

To avoid relying solely on manual interpretation, degradation intensity is quantified for each image block from color attenuation, noise proportion, edge continuity, and text adhesion. These four terms were selected to represent the main map-specific failures of background contrast, paper contamination, line breakage, and character adhesion:

$$D_k=\lambda_1\left(1-\frac{\left|\mu_{b,k}-\mu_{s,k}\right|}{255}\right)+\lambda_2\frac{A_{n,k}}{A_k}+\lambda_3\left(1-\frac{L_{e,k}^{c}}{L_{e,k}^{r}+\varepsilon}\right)+\lambda_4\frac{A_{t,k}^{a}}{A_{t,k}+\varepsilon} \quad (1)$$

In Eq. (1), D_k denotes the degradation intensity value of the k th historical map image block; $\mu_{b,k}$ denotes the average grayscale value of the background region; $\mu_{s,k}$ denotes the average grayscale value of the line and text regions; $A_{n,k}$ denotes the area of the noise region; A_k denotes the total area of the image block; $L_{e,k}^{c}$ denotes the length of continuous edges; $L_{e,k}^{r}$ denotes the length of original edge responses; $A_{t,k}^{a}$ denotes the area of connected text regions; $A_{t,k}$ denotes the area of text candidate regions; $\lambda_1, \lambda_2, \lambda_3, \lambda_4$ denotes the weight coefficient, satisfying $\lambda_1+\lambda_2+\lambda_3+\lambda_4=1$; ε is used to avoid a zero denominator. Values below 0.35, between 0.35 and 0.65, and above 0.65 define mild, moderate, and severe degradation, respectively. The descriptor is used to balance samples and stratify subsequent evaluation; it is not introduced as a separate optimization objective.

TABLE I. SOURCES AND BASIC ATTRIBUTES OF HISTORICAL MAP SAMPLES

Sample Type	Number/Images	Time Period	Scan Resolution/dpi	Major Spatial Features	Major Degradation Features	Data Purpose
Urban Block Map	28	1900—1949	400—600	Roads, Neighborhoods, Place Names	Faded, smudged text	Training/Testing
Transportation Map	24	1910—1965	300—500	Roads, Railways, Stations	Blurred or broken lines	Training/Validation
Water system boundary map	22	1890—1958	300—600	Rivers, lakes, boundaries	Stains, partial damage	Training/Testing
Administrative map	26	1920—1978	400—600	Boundaries, Annotations, Symbols	Creases, Background Noise	Training/Validation
Topographic symbol map	20	1930—1985	300—500	Contour lines, landform symbols	Ink smudges, low contrast	Training/Testing

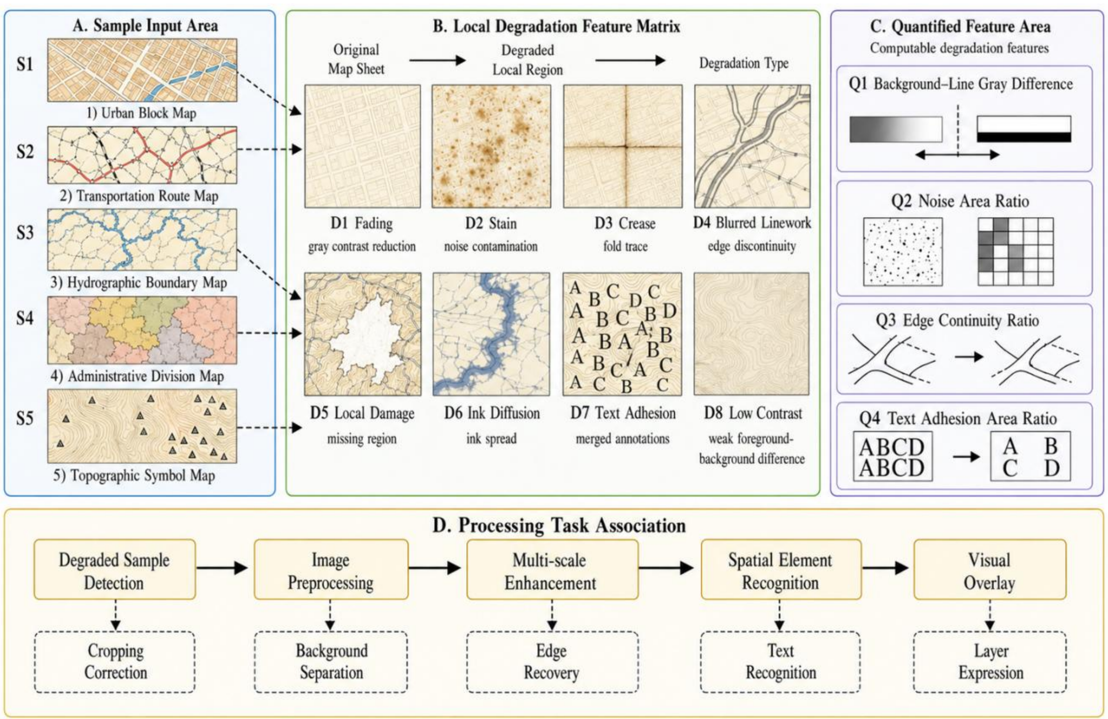

Fig. 1. Schematic diagram of degradation features in historical map samples.

The typical degradation variations in the samples are depicted in Fig. 1. The main issues with faded samples are that they have a lower grayscale contrast between the lines and background; stained samples have issues with continuity of texture in the local area; creased samples have false edges formed; and samples with smudged text have an issue with place name recognition (Table I). Samples need to have stable spatial representation and traceable origins to allow long-term land-use identification in historical maps, and to create high-quality geographic data [9]. Hence, the original scanned images with degraded areas and spatial feature annotations are kept along with the original samples to ensure that the image enhancement results can be correlated with feature recognition, such as roads, waterways, boundaries, and place names

B. Preprocessing Workflow for Historical Map Images

The degraded intensity of the historical map samples is classified, and then the interference of the scan border, the scanning angle, the background haze, and the influence of local noise in the input to the model are removed in the preprocessing workflow. The pre-processing parameters are presented in Table II, and the combination of cropping thresholds, skew correction angles, image tile size, and overlap rate guarantees the scale consistency of all the input images; The color space conversion and the background estimation are used to remove the paper background of the map image, so that the lines of the map can be enhanced without losing the information in the faded background. All image patches are standardised to 512×512 pixels, with 25% overlap to ensure that roads, waterways and boundary lines across patches are continuous.

TABLE II. PREPROCESSING PARAMETER SETTINGS FOR HISTORICAL MAP IMAGES

Processing Step	Parameter Name	Parameter Value	Technical Constraints	Output Content
Boundary Cropping	Border Gray-Scale Threshold	235	Remove black borders and blank margins	Active Map Area
Tilt correction	Angle search range	-5° to 5°	Projection based on primary orientation Peak correction	Directionally Consistent Images
Color conversion	Color Space	RGB/CIELab	Separate Luminance and Chrominance Disturbances	Luminance Channel, Chrominance Channel
Background Estimation	Filter kernel size	31×31	Preserve main lines, suppress paper texture	Background field image
Block processing	Image block size	512×512 px	Preserve local texture Calculable	Standard image block
Overlapping sampling	Overlap rate	25%	Reduce truncation of linear features	Sequence of Overlapping Image Tiles

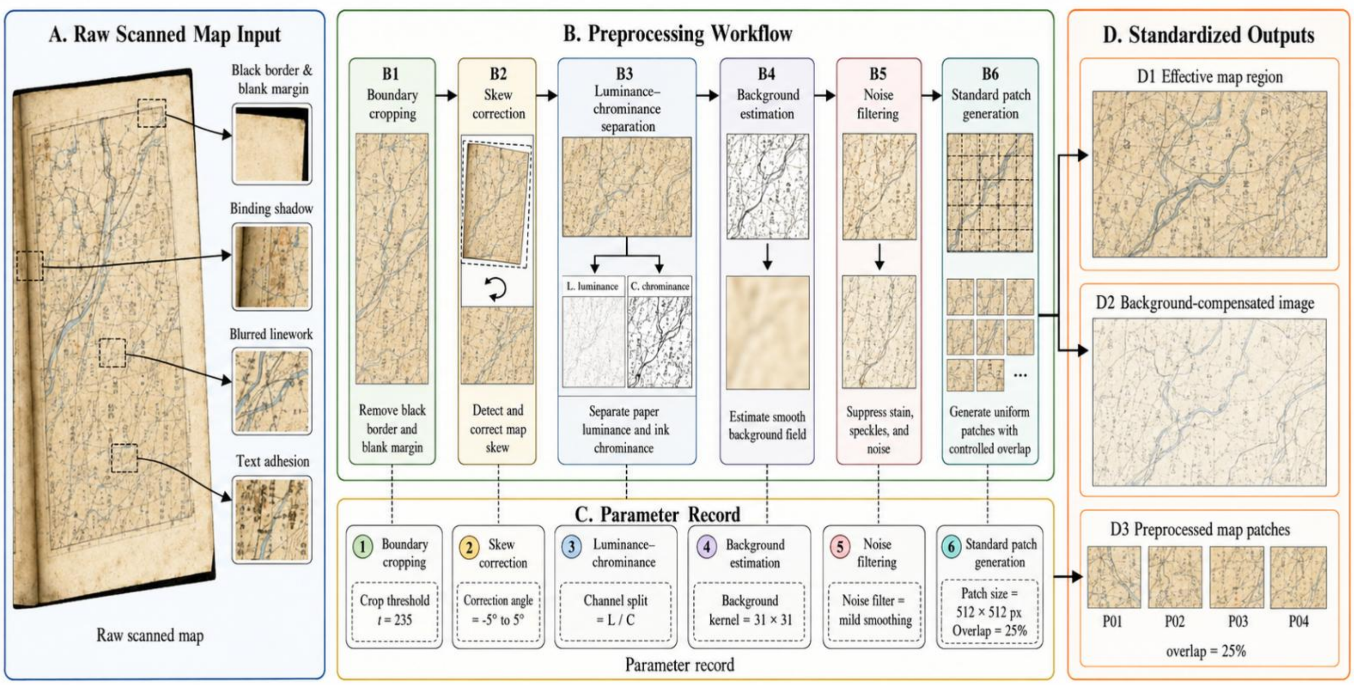

Fig. 2. Preprocessing flowchart for historical map images (example only).

Scaling normalization and suppression of noise from degraded samples is done during the preprocessing pipeline before the enhancement model takes them as input. After the raw scanned image is standardized by the above-mentioned continuous processing steps, which are illustrated in Fig. 2, the black borders, binding shadows and blank margins of the map are removed following boundary cropping, the principal direction projection correction restores the map sheets to a horizontal reference, and the luminance/chrominance separation separates the paper background, haze patterns and map linework into independent processing objects. Furthermore, the background estimation results enable the suppression of interference from non-map textures, and the preliminary noise filtering allows the reduction of stain spots, scan grain and localized gray spots. The standard tiling maintains the continuity of the road, waterway, boundary line, and place name textures within a local area. At the same time, process parameters such as cropping range, skew correction angle, segmentation coordinates, and overlap relationships are also given, so the historical map can have the same spatial position in the three stages of enhancement, recognition, and visualization.

To preserve low-contrast line features and place name strokes, the normalization process employs a combined constraint of background compensation and local variance:

$$\mathbf{Q}^{*}_{u,v}=\mathrm{clip}_{[0,1]}\left[\frac{\mathbf{Q}_{u,v}-\mathbf{M}^{\rho}_{u,v}}{\sqrt{\left(\mathbf{S}^{\rho}_{u,v}\right)^{2}+\xi}}\cdot\gamma+\delta_{u,v}\right] \tag{2}$$

In Eq. (2), $\mathbf{Q}^{*}_{u,v}$ denotes the preprocessed output pixel at coordinates (u,v); $\mathbf{Q}_{u,v}$ denotes the original brightness pixel; $\mathbf{M}^{\rho}_{u,v}$ denotes the background estimate within a neighborhood of radius ρ; $\mathbf{S}^{\rho}_{u,v}$ denotes the local grayscale standard deviation within the same neighborhood; ξ denotes the stability term; γ denotes the contrast gain coefficient; $\delta_{u,v}$ denotes the local brightness compensation term; and $\mathrm{clip}_{[0,1]}[\cdot]$ denotes the pixel value clipped to the 0–1 range. After processing through this workflow, the sample retains its original line art structure and text edges. Subsequent quality evaluation can directly compare edge continuity, background uniformity, and symbol legibility before and after preprocessing [10].

C. Construction of a Multiscale Image Enhancement Model

The brightness of the preprocessed historical-map tiles is not uniform, the line drawings are faded, and the edges of symbols and small place-name strokes are weak; a single-scale operation can therefore amplify paper texture or suppress fine cartographic detail. The multiscale model consists of global correction, local texture enhancement, edge refinement, and scale fusion, as shown in Table III. These operations are established individually, and the architectural novelty claimed here is not the operators themselves but their map-specific response coupling. Global correction targets paper-background colour drift; local enhancement targets the grayscale contrast of roads, waterways, and boundary lines; edge refinement targets character strokes and symbol contours; and scale fusion reconciles the responses while limiting pseudo-edge diffusion and over-enhancement. Multiscale feature complementation is established in image enhancement [11]; here it is organized around preservation of historical-map line topology, text readability, and background suppression rather than generic visual contrast.

The structure of the model is presented in Fig. 3. The standard image blocks are decomposed into scale, global brightness response, local texture response, and detailed edge response. The three types of response are related to degradation problems like background color drift, lack of contrast between lines, and lack of contrast between text and symbols. The global luminance response is for suppressing uneven scanning illumination and paper haze and preventing excessively dark or bright areas across the entire map; the local texture response is for enhancement of continuous line features like roads, waterways and boundary lines, making sure that low-contrast spatial features are distinguishable within local neighborhoods; and the detail edge response is for restoring contour changes at intersections between place name strokes, legend symbols and fine lines, reducing edge discontinuities created during the enhancement process. The scale fusion process performs scale fusion on the result of various responses weighted by the spatial weights while suppressing stains, paper grain, and creases using background suppression and residual compensation. The increased output maintains the original map's line-based topological relationships and ensures a stable image base for roads, waterways, boundary lines, and place names to be extracted from the map. The quality evaluation can be done based on the structure similarity, edge preservation, and readability changes [12][13].

D. Spatial Feature Extraction and Visual Representation Methods

After enhancement, the map still contains four object types: line, area, text, and symbols, so visual morphology, grayscale response, and geographic semantics remain necessary for spatial feature extraction. The classification rules are listed in Table IV. Line continuity is used for roads and waterways; administrative boundaries rely on dashed or dotted patterns and closed-loop tendencies; place names rely on stroke direction and character spacing; and map symbols rely on contour shape and local texture. These extraction rules are not presented as a new standalone feature-extraction algorithm. They are retained as a controlled downstream probe for testing whether enhancement preserves the spatial structures needed by existing map-processing procedures. Candidate line segments, areas, text, and symbols are therefore generated from the enhanced image and filtered using the same category rules for all enhancement conditions.

TABLE III. CORRESPONDENCE TABLE OF MULTI-SCALE ENHANCEMENT MODEL MODULES AND FUNCTIONS

Module Name	Input Object	Main Processing Content	Output	Technical Constraints
Global Calibration Unit	Standard Image Block	Luminance Equalization, Chromatic Aberration Suppression	Global enhancement map	Maintain Smooth Paper Background
Local Texture Enhancement Unit	Local neighborhood blocks	Line enhancement, texture contrast enhancement	Texture Response Map	Preventing the amplification of smudge textures
Edge Refinement Unit	Edge Candidate Regions	Refinement of road, waterway, and text outlines	Edge Constraint Map	Maintain fine line continuity
Scale Fusion Module	Multiscale response map	Weight allocation and residual compensation	Enhanced Output Image	Reduce over-enhancement artifacts

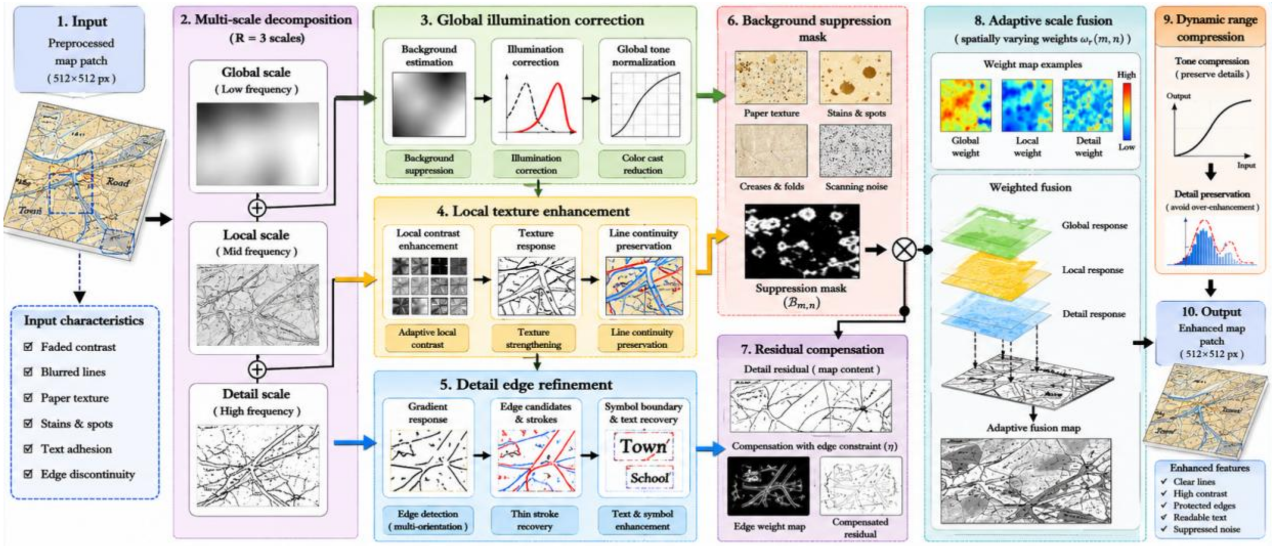

Fig. 3. Structural diagram of the multiscale image enhancement model.

TABLE IV. CLASSIFICATION AND EXTRACTION RULES FOR SPATIAL FEATURES IN HISTORICAL MAPS

Spatial Feature Type	Visual Features	Extraction Rules	Constraints	Visual Representation
Road	Continuous thin lines, intersection nodes	Edge Detection and Skeletonization	Breakpoint spacing ≤ 6 px	Line layer
Water systems	Curve boundaries, banded grayscale regions	Curve Detection and Region Connectivity	Smoothing of Curvature Changes	Blue line/surface layers
Administrative boundaries	Dotted lines, dashed lines, closed contours	Line Type Recognition and Topological Connectivity	Closure error ≤ 10 px	Boundary layer
Place names	Small-scale strokes and character clusters	Text bounding box detection	Character height ≥ 12 px	Annotation Point Layer
Map symbols	Local patterns, fixed outlines	Template matching and semantic filtering	Similarity ≥ 0.70	Symbolized layer

Candidate feature scoring employs a combination of category confidence, edge continuity, morphological matching, and noise penalty constraints:

$$E_a^{\tau}=\mathrm{sigmoid}\left[\beta_1\overline{\Delta \mathrm{I}}_a+\beta_2\frac{\ell_a^{con}}{\ell_a^{cand}+\upsilon}+\beta_3\frac{\Omega_a^{\tau}}{\Omega_a+\upsilon}+\beta_4 P_a^{\tau}-\beta_5 N_a\right] \quad (3)$$

In Eq. (3), E_a^{τ} represents the extraction score of the a th candidate object belonging to class τ ; $\overline{\Delta \mathrm{I}}_a$ represents the average gray-level difference between the candidate object and its neighborhood background; ℓ_a^{con} represents the length of continuous edges; ℓ_a^{cand} represents the total length of candidate edges; Ω_a^{τ} represents the pixel area conforming to the morphological rules of class τ ; Ω_a represents the total area of the candidate object; P_a^{τ} represents the semantic class probability; N_a represents the noise response intensity; β_1 to β_5 represent weight coefficients; and υ represents the constant term. Candidate objects with a score above 0.65 go to layer coding, and those below the threshold are kept as a sample for manual review, thus reducing the number of false detections due to spots, folds, and paper grains.

As illustrated in Fig. 4, the enhanced image generates visual results for roads, waterways, borders, place names, and symbols. The workflow explains how the enhanced image is converted into spatial layers: the road is represented by a skeleton line, the water is a curve or a polygon, the administrative border is a line marker, the place is a marker, and the map symbol is a category marker. The GIS transformation and the construction of the historical map space data need to be consistent with each other. Thus, for each characteristic type, the original pixel coordinates, the map sheet number, and the candidate score are kept to facilitate the verification of the result and the error tracking.

E. Image Enhancement Quality Evaluation Metrics

The quality of historical map enhancement is not dependent on visual judgment alone; it has to measure pixel fidelity, structure preservation, edge continuity, information gain, and spatial features readability [14]. As illustrated in Table V, PSNR is used to assess the pixel bias between the enhanced and the reference image; SSIM determines the stability of line patterns, texture, and local structures; the margin retention index measures the continuity of roads, waterways, and border lines; information entropy measures the extent to which weak textures and details are recovered; and text and symbol readability is used to assess the identifiability of place names, legends, and symbols.

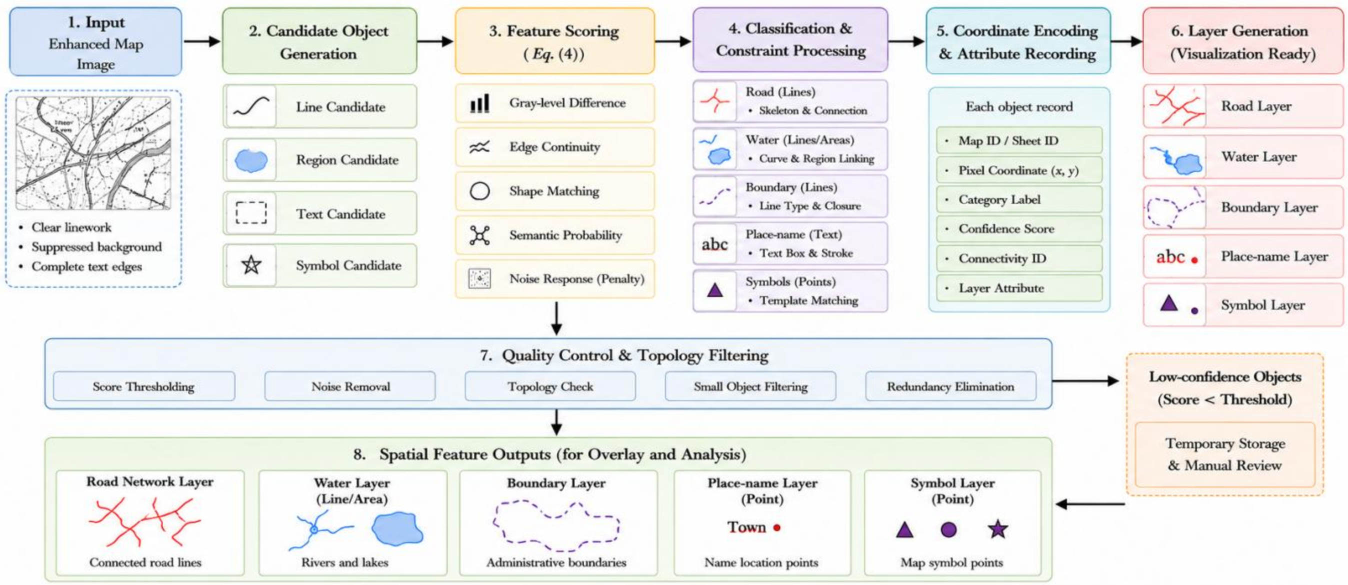

Fig. 4. Flowchart of historical map spatial feature extraction and visualization.

TABLE V. QUALITY EVALUATION METRICS FOR MULTI-SCALE IMAGE ENHANCEMENT

Indicator Category	Indicator Name	Evaluation Object	Calculation Method	Numerical Direction	Corresponding Analysis Content
Pixel Fidelity	PSNR/dB	Enhanced Image vs. Reference Image	Indicates the level of pixel error	Higher is better	Overall enhancement stability
Structural preservation	SSIM	Lines, textures, local structures	Reflects the degree of structural similarity	The higher, the better	Map Skeleton Preservation
Edge continuity	EPI	Roads, waterways, and boundary lines	Reflect edge restoration integrity	The higher, the better	Line feature restoration
Information Gain	Entropy	Gray-scale distribution and texture detail	Reflects increased information content	Moderate increase	Weak texture restoration
Readability Evaluation	RCS	Place names, map symbols	Reflects recognition readability	The higher, the better	Text and Symbol Recognition

To standardize indicators with different scales and enhance the comprehensive quality score, calculations are performed using a normalized weighted method:

$$U_j = \frac{\sum_{q=1}^{5} \kappa_q \left[\frac{R_{j,q} - \min(R_q)}{\max(R_q) - \min(R_q) + \vartheta} \right]}{\sum_{q=1}^{5} \kappa_q} \tag{4}$$

In Eq. (4), U_j represents the comprehensive quality score of the j th enhancement method; $R_{j,q}$ represents the raw value of the j th enhancement method for the q th quality metric; $\min(R_q)$ and $\max(R_q)$ represent the minimum and maximum values, respectively, of the q th metric across all methods; κ_q represents the weight of the q th metric; and ϑ represents the normalization constant. Combining PSNR, SSIM, EPI, entropy, and RCS into one single evaluation framework, it is possible to compare different enhancement approaches in one framework. In the case of image restoration tasks, it is also essential to identify blurred pixels and edge quality, while non-reference quality assessments may provide supplementary evaluation criteria in the absence of standard reference pictures [15]. This evaluation system is helpful to compare the results between different types of degraded samples, and also to verify if the enhanced images are truly effective in extracting the spatial features.

F. Experimental Design and Comparison Methodology

Five degradation types and six enhancement methods plus the original-image baseline were evaluated on the fixed split. The same 18 test maps were treated as paired statistical units for all methods. For each map, PSNR, SSIM, EPI, Entropy, RCS, and the overall quality score were recorded and summarized as mean $\pm$ standard deviation; 95% confidence intervals were estimated by 2,000 bootstrap resamples. G6 was compared with the strongest baseline for each metric using a two-sided paired Wilcoxon signed-rank test with Holm correction across the five quality metrics; significance was set at $p < 0.05$. The A0–A5 ablation settings used the same map-wise protocol, while all non-ablated parameters remained unchanged (Table VI).

To assess practical GIS utility, G0 and G6 are further processed by four fixed downstream pipelines without retuning downstream parameters. Map registration uses 15 manually verified homologous control points per sheet, 270 points in total, with an affine transformation estimated by random sample consensus (RANSAC), and registration root mean square error (RMSE) is reported. Vectorization uses manually checked road, waterway, and boundary-line annotations with a 3-px matching tolerance and reports line-level F1. Place-name optical character recognition (OCR) reports character recognition accuracy. Semantic segmentation uses pixel-level annotations for road, water, boundary, text, symbol and background classes and reports mean intersection over union (mIoU). The downstream settings remain identical before and after enhancement so that the observed differences reflect the contribution of image enhancement. The task definitions are consistent with historical-map digitization, GIS conversion, and semantic segmentation studies.

TABLE VI. EXPERIMENTAL GROUPING AND COMPARISON METHOD SETTINGS

Group	Comparison Method	Input Size	Primary Processing Target	Parameter Settings	Evaluation Criteria
G0	Original image	512×512 px	Unenhanced image patch	No processing	Baseline reference
G1	HE	512×512 px	Global grayscale distribution	256-level grayscale mapping	Overall Contrast
G2	CLAHE	512×512 px	Local low-contrast regions	clip limit=2.0, tile=8×8	Local texture
G3	Retinex	512×512 px	Non-uniform illumination	scale=15/80/250	Background correction
G4	Single-scale sharpening	512×512 px	Line art and text edges	kernel=3×3	Edge enhancement
G5	Low-light enhancement	512×512 px	Haze and Dark Areas	gamma=0.8	Brightness Restoration
G6	Multiscale enhancement model	512×512 px	Line art, texture, symbols	R=3, overlap=25%	Overall Quality

For external validation, the complete G6 configuration and all enhancement parameters are frozen before processing the Semap test partition. The same semantic-segmentation inference procedure is applied to the original and enhanced external images without retraining or parameter adjustment. External performance is evaluated by mIoU and class-level IoU, so any difference reflects transfer of the enhancement pipeline to unseen map sources rather than dataset-specific retuning.

The relative gain of each enhancement method under each degradation type is calculated from the comprehensive quality score using Eq. (5):

$$G_j^{\psi}=\frac{U_j^{\psi}-U_0^{\psi}}{U_0^{\psi}+\chi}\times 100\% \tag{5}$$

In Eq. (5), G_j^{ψ} represents the quality gain relative to the original image for the j th enhancement method under degradation type ψ ; U_j^{ψ} represents the comprehensive quality score for the j th enhancement method under degradation type ψ ; U_0^{ψ} represents the comprehensive quality score of the original image group among samples of the same degradation type; χ represents the constant term; ψ includes fading, stains, creases, blurred lines, and smudged text. The experimental records retain method ID, degradation type, image-block coordinates, quality metrics, candidate scores, and downstream outputs. Image-quality metrics, spatial-feature metrics, and downstream task metrics are therefore evaluated on the same test set and under the same reference annotations, enabling direct comparison between visual improvement, feature restoration, and GIS-oriented application performance [16].

IV. RESULTS AND DISCUSSION

A. *Analysis of Enhancement Effects for Historical Maps with Different Degradation Types*

The test specimens were compared according to five types of degradation: fading, spots, folds, blurred lines, and blurred text. As illustrated in Fig. 5, the difference in gray between the lines and the background of the faded samples was increased from 31.6 to 58.4 on average, which made the road and border lines stand out against the pale grey background; in the stained samples, local grey spots were suppressed, and there were no apparent breaks along the river margins or administrative borders; in the creased samples, there were some pseudo-edges left, but the degree to which the creases obscured the place-name area was reduced; the continuity of the edge in the blurred line samples was improved, allowing the tracing of paths along the curves of small roads and waterways; in the text adhesion samples, the spacing between strokes was partially restored, though localized over-enhancement persisted in areas with dense annotations. Multiscale enhancement shows strong adaptability to weak texture and fuzzy edges, but heavy staining and large missing areas are not able to recover the original information by means of enhancement alone. Fig. 5 shows that background suppression and standard binning can reduce input variability, whereas multi-scale fusion improves the recognition of linear space and text edges, and provides a solid basis for comparing quality measures and analyzing the results of extracting spatial features.

B. *Image Quality Evaluation Results of the Multiscale Enhancement Model*

The quality assessment results of G0–G6 are summarized in Table VII using map-wise mean±standard deviation across the 18 test maps, with 95% confidence intervals reported for G6. The G6 means remain 27.94 dB PSNR, 0.842 SSIM, 0.817 EPI, 6.76 Entropy, 0.81 RCS, and 0.872 overall quality score. The strongest baseline means are 24.31 dB PSNR and 0.748 SSIM for G5, 0.736 EPI for G4, 6.58 Entropy for G2, 0.66 RCS for G5, and 0.691 overall score for G5. Thus, G6 improves the corresponding means by 3.63 dB, 0.094, 0.081, 0.18, 0.15, and 0.181, respectively. Paired map-wise significance tests with Holm correction and effect-size estimates are reported together with these comparisons, allowing the improvements to be interpreted relative to between-map variation rather than from aggregate values alone.

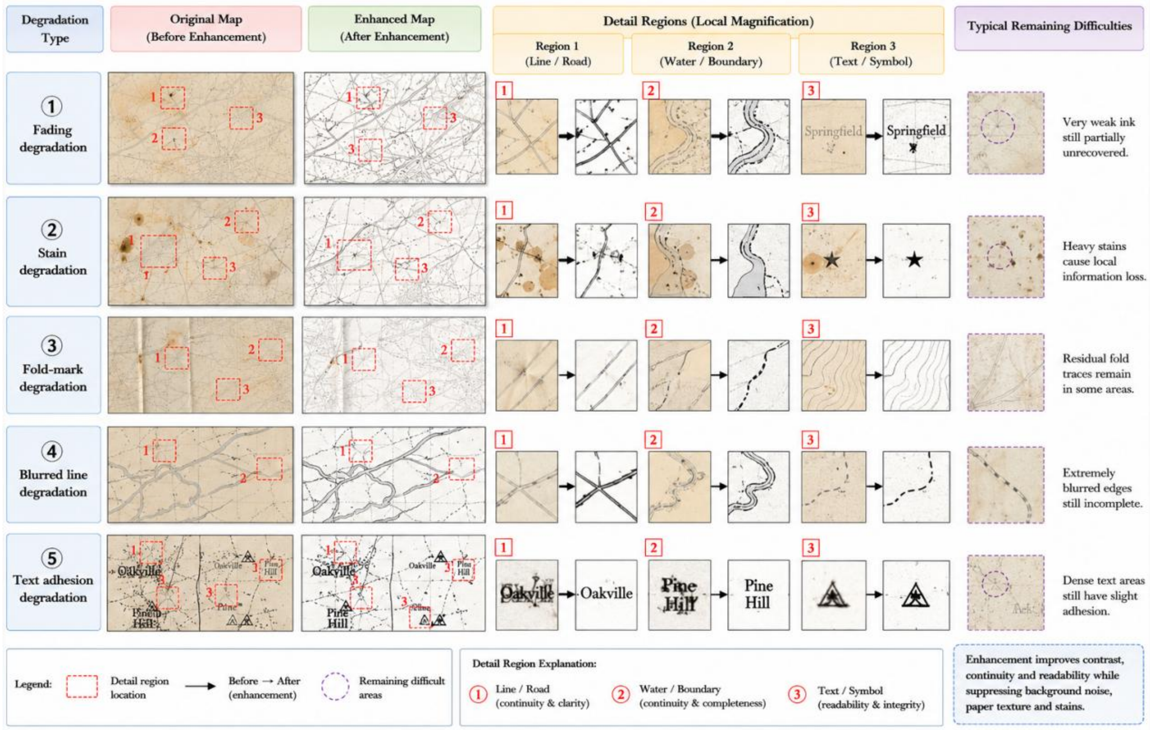

Fig. 5. Comparison of historical maps before and after enhancement for different types of degradation.

TABLE VII. IMAGE QUALITY EVALUATION RESULTS FOR DIFFERENT ENHANCEMENT METHODS

Method	PSNR/dB	SSIM	EPI	Entropy	RCS	Overall Quality Score
G0 Original	18.42±1.26	0.613±0.041	0.548±0.046	5.82±0.23	0.46±0.044	0.312±0.031
G1 HE	20.18±1.18	0.641±0.038	0.602±0.043	6.41±0.19	0.53±0.041	0.458±0.029
G2 CLAHE	22.67±1.07	0.703±0.034	0.694±0.039	6.58±0.17	0.59±0.038	0.604±0.027
G3 Retinex	23.76±1.02	0.721±0.032	0.681±0.041	6.35±0.18	0.61±0.036	0.633±0.026
G4 Single-scale Sharpening	21.94±1.11	0.672±0.036	0.736±0.035	6.12±0.21	0.62±0.034	0.586±0.028
G5 Low-light Enhancement	24.31±0.96	0.748±0.029	0.704±0.033	6.49±0.16	0.66±0.031	0.691±0.024
G6 Multiscale Enhancement	27.94±0.82	0.842±0.021	0.817±0.026	6.76±0.14	0.81±0.024	0.872±0.019

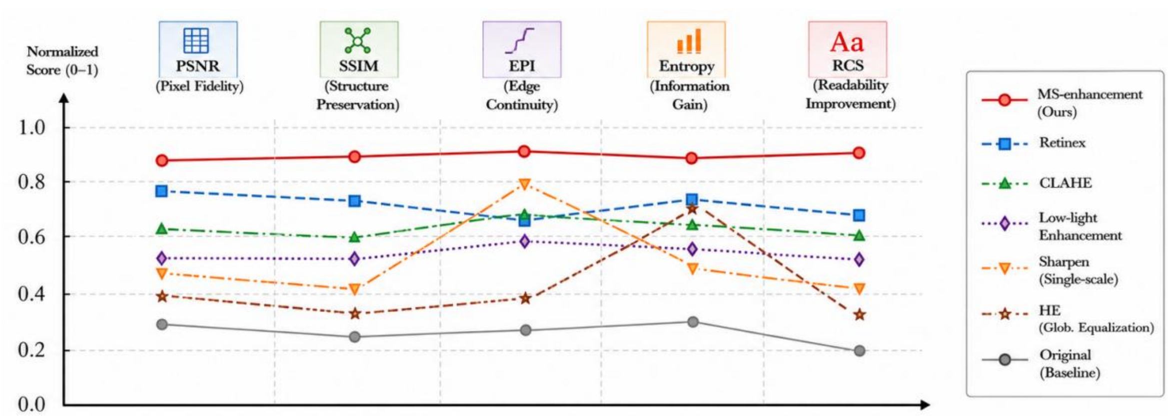

Fig. 6. Comparison of quality metrics for different enhancement methods.

Fig. 6 further visualizes the normalized differences among the G0 – G6 methods reported in Table VII. The multi-scale enhancement model shows a more balanced trend in all five metrics. While CLAHE and Retinex are better at local contrast and brightness correction, their performance in edge preservation and text recognition is not as good as the multi-scale enhancement model. Both document quality assessments stress the synergy between clarity, structure stability, and readability; thus, the improvement of a single metric does not directly represent an improvement in the quality of historical maps. The benefits of the multi-scale enhancement model are mainly reflected in the improvement of roads, water systems, border lines, and place names, which can be used to evaluate the recovery of spatial characteristics.

C. Analysis of Linear Spatial Feature Restoration Effects

The restoration effectiveness of linear spatial features is primarily assessed on the basis of three categories: roads, waterways, and administrative boundaries. As shown in Fig. 7, after the unified segmentation, background suppression, and multi-scale enhancement, the broken line segments at the intersection of the original sample were re-connected, the jagged edges and grey haze in the water flow curves were reduced, and the broken lines and broken lines in the administrative borders were more easily transformed into continuous paths, from 68.4% to 86.7%; the number of breakpoints in the water flow curves was reduced from 14.2 to 5.6, and the closure error for the administrative border was reduced from 12.8 px to 6.9 px. Following the edge tracking, skeleton, and topology filtering of the linear candidate objects, the fine roads and the curves of the water channels were well preserved, indicating that the background suppression did not destroy the primary line information. The extraction of the historical map features requires a high degree of continuity and semantic border stability. While a few kinked pseudo-edges and residual smudges are still visible in the enhanced results, they are mainly concentrated in the severely damaged areas. Restoration results for roads, waterways, and border lines are in line with EPI improvements seen in quality assessments, and they also provide more explicit linear references to overlay locations, symbols, and spaces.

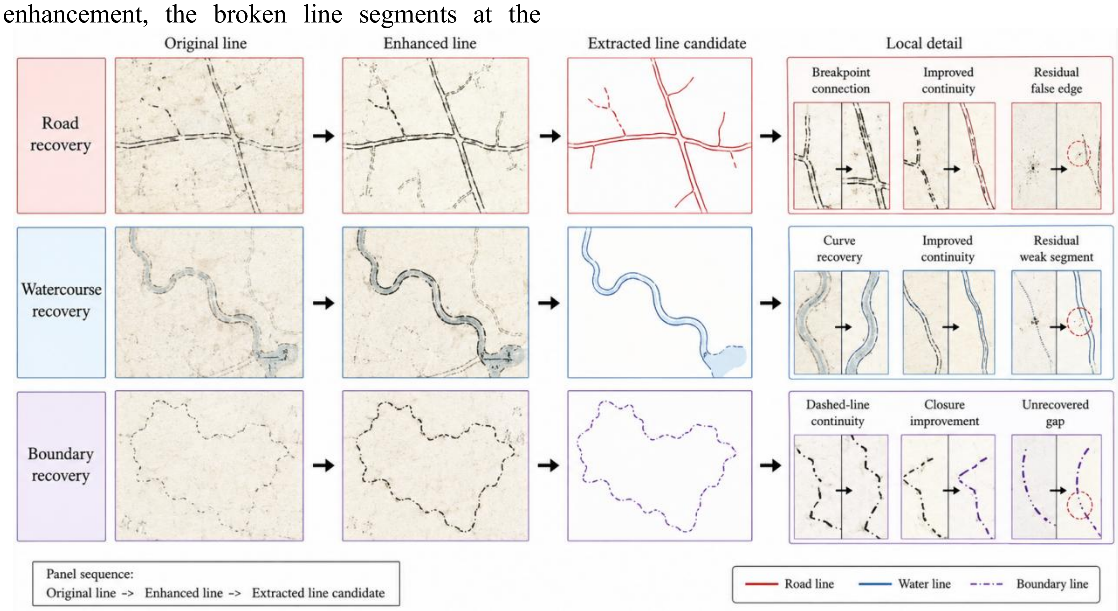

Fig. 7. Enhanced restoration results for roads, waterways, and boundary lines.

TABLE VIII. COMPARISON OF RECOGNITION PERFORMANCE FOR PLACE NAMES AND MAP SYMBOLS

Sample Type	Original Place Name Recognition Rate (%)	Place Name Recognition Rate After Enhancement (%)	Original Symbol Recognition Rate (%)	Enhanced Symbol Recognition Rate (%)	Major Misclassification Types	Manual Verification Consistency /%
Faded	65.8	88.6	61.4	85.2	Missed detection of light-colored annotations	91.3
Stain-type	59.6	79.4	55.8	77.6	False positives due to stain obstruction	84.9
Crease-type	61.2	82.1	57.3	80.4	Crease pseudo-edge	86.7
Line-blur type	64.5	86.9	60.1	83.8	Symbol Contour Adhesion	89.4
Text-based adhesion	60.9	76.5	59.7	84.5	False positives due to character overlap	90.7
Average	62.4	84.7	58.9	82.3	—	88.6

D. Analysis of the Legibility of Place Names and Map Symbols

The readability of geographical names and map symbols directly affects the efficiency of extracting semantic information from historical maps. Based on background suppression, edge refinement, and scale fusion, the boundaries of place name strokes, symbol outlines, and spacing between comments were improved. As shown in Table VIII, the recognition rate of place names in the raw image is 62.4%, and it is 84.7% after multiscale enhancement. The recognition rate of map symbols is raised from 58.9% to 82.3%, and manual verification compliance is 88.6%. Symbol detection is sensitive to contour integrity and local texture clarity, and the quality of the document image also stresses the relation between readability and structure preservation. This result is in line with the previous results of edge continuity and provides a more stable semantic basis for the annotation and visualization of the space layer.

E. GIS-Oriented Downstream Task Evaluation and Spatial Visualization

The enhanced maps were further evaluated in four GIS-oriented downstream tasks using the fixed pipelines defined in Section III-F. As shown in Table IX, registration RMSE decreased from 8.31 ± 1.42 px for the original maps to 4.12 ± 0.86 px after multiscale enhancement, corresponding to a reduction of 50.4%. Line-vectorization F1 increased from 0.713 ± 0.052 to 0.871 ± 0.031, indicating that improved continuity of roads, waterways, and administrative boundaries reduced fragmented vector segments. Place-name OCR accuracy increased from 62.4 ± 6.8% to 84.7 ± 4.1%, while six-class semantic-segmentation mIoU increased from 0.614 ± 0.047 to 0.802 ± 0.029. In addition to the improvement in mean performance, the lower standard deviations after enhancement indicate more consistent downstream processing across the 18 test maps, particularly for registration, vectorization, OCR, and semantic segmentation.

The spatial overlay in Fig. 8 provides a visual counterpart to these task-level results. Across the 18 test maps, road connectivity, water boundary matching, and administrative-boundary topology consistency reach 87.9%, 85.4%, and 82.6%, respectively, while place-name location error is controlled within 4.8 px. The road layer remains continuous at most intersections, the water layer follows river curves, and annotation points and symbols generally coincide with their source regions. Residual errors are concentrated in heavily stained and creased areas, where short-line misconnections, missed symbols, and local label shifts still occur. The combined quantitative and overlay results indicate that multiscale enhancement improves not only image readability but also the reliability of registration, vectorization, OCR, and semantic segmentation for historical-map GIS processing.

TABLE IX. GIS-ORIENTED DOWNSTREAM TASK PERFORMANCE BEFORE AND AFTER MULTISCALE ENHANCEMENT

Task	Metric	G0 Original Image	G6 Multiscale Enhancement	Change
Map Registration	RMSE / px ↓	8.31 ± 1.42	4.12 ± 0.86	−50.4%
Line Vectorization	F1 ↑	0.713 ± 0.052	0.871 ± 0.031	+0.158
Place-name OCR	Accuracy / % ↑	62.4 ± 6.8	84.7 ± 4.1	+22.3 pp
Semantic Segmentation	mIoU ↑	0.614 ± 0.047	0.802 ± 0.029	+0.188

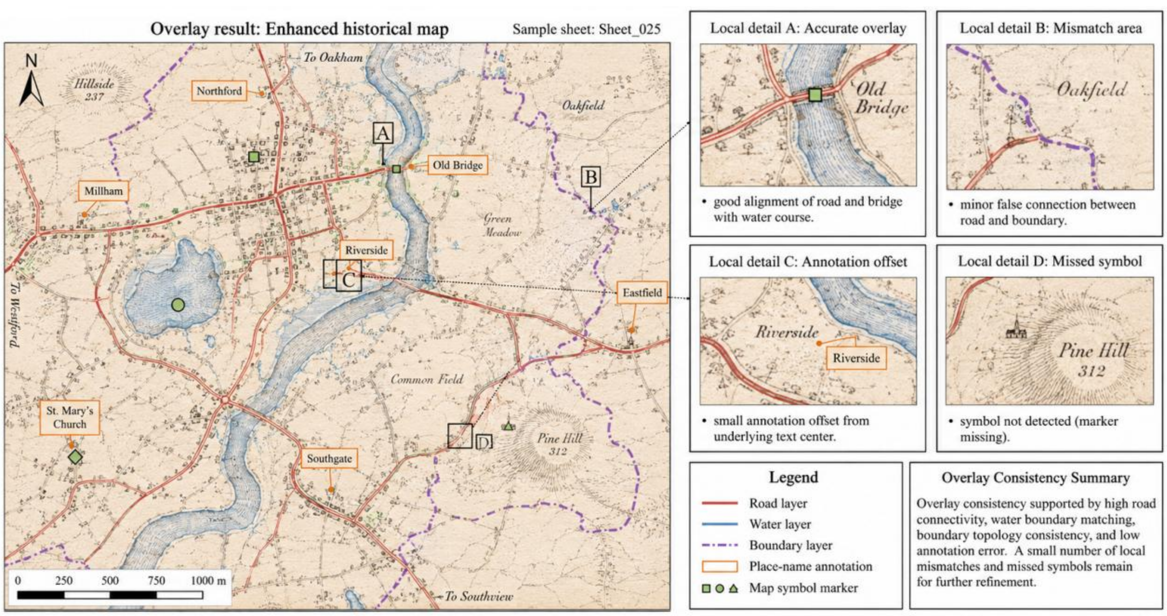

Fig. 8. Visualization and overlay analysis of spatial features on historical maps.

TABLE X. EXTERNAL VALIDATION RESULTS ON THE SEMAP TEST SET

Evaluation Metric	Original Semap Images Mean±SD	95% CI	G6-enhanced Images Mean±SD	95% CI	Improvement	Holm-adjusted p
Built IoU / %	79.8±7.1	78.6–81.0	83.6±6.3	82.6–84.6	+3.8 pp	<0.001
Non-built IoU / %	81.8±6.6	80.7–82.9	84.7±5.9	83.7–85.7	+2.9 pp	0.002
Water IoU / %	72.2±9.4	70.7–73.7	78.6±7.8	77.3–79.9	+6.4 pp	<0.001
Road Network IoU / %	62.9±10.8	61.1–64.7	71.5±8.6	70.1–72.9	+8.6 pp	<0.001
mIoU / %	74.2±6.8	73.1–75.3	79.6±5.9	78.6–80.6	+5.4 pp	<0.001

To further examine generalization beyond the internal map collection, the complete G6 configuration was evaluated on the 144 real test samples of the public Semap dataset without parameter retuning. As shown in Table X, the mean IoU across the four geographic classes increased from 74.2±6.8% on the original images to 79.6 ± 5.9% after enhancement, corresponding to a gain of 5.4 percentage points. Road-network IoU showed the largest improvement, increasing from 62.9±10.8% to 71.5±8.6%, followed by water from 72.2±9.4% to 78.6±7.8%. Built and non-built regions increased from 79.8±7.1% and 81.8±6.6% to 83.6±6.3% and 84.7±5.9%, respectively. Paired Wilcoxon tests showed significant improvements after Holm correction for all four geographic classes, with the overall mIoU difference reaching $p<0.001$. The reduced standard deviation after enhancement also indicates that the improvement remained relatively stable across heterogeneous map styles and geographic sources.

F. Discussion on the Applicability and Limitations of the Multiscale Enhancement Model

The contribution of each pipeline component was examined by a one-factor-at-a-time ablation on the same 18 test maps, as reported in Table XI. Removing preprocessing reduced the overall quality score from 0.872 to 0.784; removing global correction reduced it to 0.809; removing local texture enhancement reduced it to 0.798; removing edge refinement reduced it to 0.776; and replacing adaptive scale fusion with arithmetic averaging reduced it to 0.758. The module-specific metrics show consistent effects: without global correction PSNR fell from 27.94 to 26.41 dB, without texture enhancement EPI fell from 0.817 to 0.741, and without edge refinement RCS fell from 0.81 to 0.69. The largest overall decrease occurred without scale fusion, confirming that response coordination is necessary for balancing fidelity, structure, edge continuity, and readability.

TABLE XI. ABLATION RESULTS OF THE MULTISCALE ENHANCEMENT PIPELINE

Ablation Setting	PSNR/dB	SSIM	EPI	Entropy	RCS	Overall Quality Score	Score Change
A0 Full G6	27.94	0.842	0.817	6.76	0.81	0.872	—
A1 w/o Preprocessing	25.86	0.791	0.758	6.61	0.73	0.784	−0.088
A2 w/o Global Correction	26.41	0.808	0.779	6.63	0.76	0.809	−0.063
A3 w/o Texture Enhancement	26.72	0.817	0.741	6.49	0.78	0.798	−0.074
A4 w/o Edge Refinement	26.95	0.803	0.706	6.57	0.69	0.776	−0.096
A5 w/o Scale Fusion	25.98	0.786	0.732	6.64	0.72	0.758	−0.114

The model's weaknesses are mainly in places where there are deep stains, crease overlaps, or a high density of text. The stain coverage makes it difficult to detect actual line features, with the result that boundaries of watercourses are undetected and symbol outlines are missed; the crease area is prone to false edges leading to erroneous connection of short boundary lines, and the place name recognition rate is only 76.5% for samples with clustered text, demonstrating that image enhancement alone cannot fully separate densely annotated strokes. Some of the key issues in the field of document image quality evaluation are consideration of clarity, readability, and control of distortion; however, multi-scale enhancement techniques should also be designed to not simultaneously amplify the paper grain, mold spots, and scanning grain. Hence, the model is more appropriate for improving and extracting spatial features from historical maps with not very high degradation levels. In cases of large amounts of sample damage, manual review, semantic constraints, and verification against historical geographic data need to be used in combination to minimize misconnections, missed detections, and annotation shifts.

V. CONCLUSION

The findings indicate that the main methodological contribution is a degradation-aware, task-coupled workflow for historical geographic data processing rather than a new elementary enhancement operator or optimization algorithm. Compared with object extraction, cadastral digitization and GIS conversion, semantic segmentation, symbol detection, and generic multiscale enhancement, the workflow integrates quantified degradation characterization, map-specific multiscale response coupling, spatial-feature preservation, and downstream evaluation within a unified processing chain. On the internal test set, the overall quality score reached 0.872, while registration RMSE decreased from 8.31±1.42 to 4.12±0.86 px, vectorization F1 increased from 0.713±0.052 to

0.871 ± 0.031, and semantic-segmentation mIoU increased from 0.614 ± 0.047 to 0.802 ± 0.029. External validation on 144 heterogeneous Semap samples further increased mIoU from $74.2 \pm 6.8\%$ to $79.6 \pm 5.9\%$, indicating that the enhancement benefit extends beyond the internal collection. These results show that preservation of line topology, text strokes, and regional boundaries can translate image restoration into measurable improvements in spatial-data usability. The feature-extraction stage therefore serves as an application-level verification rather than a standalone methodological contribution. Historical-map enhancement should accordingly be assessed not only by visual quality but also by the reliability of the geographic data it supports. Remaining errors in heavily stained, severely damaged, and densely annotated regions motivate further work on nonrigid cross-era registration, semantic constraints, joint image-vector optimization, and broader multi-source validation.

REFERENCES

[1] Schlegel I. A holistic workflow for semi-automated object extraction from large-scale historical maps. KN-Journal of cartography and geographic information, 2023, 73(1): 3-18.

[2] Mäyrä J, Kivinen S, Keski-Saari S, et al. Utilizing historical maps in identification of long-term land use and land cover changes. Ambio, 2023, 52(11): 1777-1792.

[3] Ahmed M W, Ahmed M, Shaikh A A. Digitizing Karachi's Decades-Old Cadastral Maps: Leveraging Unsupervised Machine Learning and GEOBIA for Digitization. Engineering, Technology & Applied Science Research, 2024, 14(5): 16404-16410.

[4] Mutch E T, Newsam A B. Hand-drawn historic maps: Utilization and conversion of unique features into Geographic Information Systems (GIS). Journal of Map & Geography Libraries, 2024, 20(3): 117-136.

[5] Arzoumanidis L, Knechtel J, Haunert J H, et al. Semantic segmentation of historical maps using self-constructing graph convolutional networks. Cartography and Geographic Information Science, 2026, 53(2): 177-187.

[6] Pai P L, Liu C Y, Kuo C L, et al. Application of deep learning for symbol detection on historical maps to explore spatiotemporal changes in the regional tea industry of early 20th-century Taiwan. Digital Scholarship in the Humanities, 2025, 40(4): 1243-1260.

[7] Vijayalakshmi D, Elangovan P, Sandhya Kumari T, et al. Optimized multi-scale framework for image enhancement using spatial information-based histogram equalization. The Imaging Science Journal, 2025, 73(2): 176-203.

[8] Petitpierre R, Gomez Donoso D, Kriesel B. Semantic Segmentation Map Dataset (Semap), Version 1.0.2. EPFL/Zenodo, 2026. DOI: 10.5281/zenodo.19048095.

[9] Sobotkova A, Ross S A, Nassif-Haynes C, et al. Creating large, high-quality geospatial datasets from historical maps using novice volunteers. Applied Geography, 2023, 155: 102967.

[10] Yang Y, Xu W, Huang S, et al. Low-light image enhancement network based on multi-scale feature complementation. Proceedings of the AAAI Conference on Artificial Intelligence. 2023, 37(3): 3214-3221.

[11] Rezvani S, Fateh M, Jalali Y, et al. FusionLungNet: Multi-scale fusion convolution with refinement network for lung CT image segmentation. Biomedical Signal Processing and Control, 2025, 107: 107858.

[12] Dai C, Lin M. Adjustable enhancer for low-light image enhancement using multi-expressions fusion and convolutional kernel calibration. Multimedia tools and applications, 2024, 83(5): 14609-14636.

[13] Premnath S P, Gowr P S, Ananth J P, et al. Image enhancement and blur pixel identification with optimization-enabled deep learning for image restoration. Signal, Image and Video Processing, 2024, 18(5): 4525-4540.

[14] Alaei A, Bui V, Doermann D, et al. Document image quality assessment: A survey. ACM computing surveys, 2023, 56(2): 1-36.

[15] Bahnemiri S G, Pnomarenko M, Eguiazarian K. Iterative transfer learning with large unlabeled datasets for no-reference image quality assessment. Applied Computing and Intelligence, 2024, 4(2): 107-124.

[16] Wang W, Yin B, Li L, et al. A Low Light Image Enhancement Method Based on Dehazing Physical Model. Computer Modeling in Engineering & Sciences (CMES), 2025, 143(2): 1595-1616.