Control Point Detection for Network Alignment Using Large Language Model

Müslüm Hacar(✉) and Özge Öztürk Hacar

Yildiz Technical University, Esenler, Istanbul 34210, Türkiye
mhacar@yildiz.edu.tr

Abstract. Control point (CP) selection is a crucial preprocessing step in network alignment workflows in GIS and spatial data integration. CP-based registration commonly depends on selecting corresponding intersections between a source network and a target network, after which a geometric transformation such as similarity, affine, projective, or locally adaptive warping can be estimated. This paper evaluates GPT-5.2 "thinking" mode, a multimodal Large Language Model (LLM) for CP detection and correspondence establishment between two network inputs that represent the same geographic extent. We analyze twelve experimental runs organized by three factors: input modality as rasterized network imagery versus vector networks, availability of explicit intersection node markers and identifier labels versus line-only inputs, and distortion level of the target network including regular affine-like distortions, globally irregular distortions, and local distortions. The results indicate that regular affine-like deformations enable strong LLM-assisted CP detection, particularly when nodes are provided and the raster/vector representation is chosen appropriately. This advantage diminishes under irregular distortions (global or local), where achieving consistently good performance becomes more difficult in vector-based settings. Finally, when deformation exists by localized detail-level differences, neither raster nor vector yields reliable results, making local distortions a modality-agnostic limitation for the current LLM-assisted CP detection setup.

Keywords: Large Language Model · Alignment · Georeferencing · Road Network · Coordinate Transformation · GeoAI

1 Introduction

Map alignment is central to a wide range of geospatial tasks, including map conflation of multi-source data, integration of utility graphs, map updating, and harmonization of digitized, aerial-based, remotely sensed, and surveyed datasets. When two maps represent the same geographic region but differ by georeferencing error or deformation, a common practical approach is to select corresponding control points (CPs) and estimate a transformation that aligns one map to the other. Corners of map sheets, intersections at cartesian grid networks, and junctions at street networks are natural CPs since they tend to be semantically stable and often distinctive in local structure. The identification and matching of such points remains manual, especially when representations differ across raster and vector formats or when distortions are spatially complex.

© The Author(s), under exclusive license to Springer Nature Switzerland AG 2027
O. Gervasi et al. (Eds.): ICCSA 2026 Workshops, LNCS 16766, pp. 365–373, 2027.
https://doi.org/10.1007/978-3-032-30539-8_28

CP-based registration and georeferencing has long been a standard mechanism in GIS and cartography, relying on human selection of tie points and the fitting of parametric transforms such as similarity and affine transformations or higher-order and spline-based warps when distortions are complex. In network contexts, intersections are favored CPs since they represent topological events where multiple edges meet, and their local neighborhood can, in principle, provide additional cues beyond raw coordinates. Nevertheless, CP selection remains a manual step in many production environments, particularly in cases of representation mismatch or local deformation where geometric heuristics break down.

Automated network conflation and graph matching methods address the same underlying problem by attempting to compute correspondences between edges, nodes, or subgraphs [1–5]. These approaches range from geometry-driven techniques that use proximity, overlap, and distance measures to topology-aware methods that incorporate degree distributions, adjacency patterns, and subgraph signatures. While these methods can be effective, they often require careful tuning and may be sensitive to missing edges, inconsistent sampling density, or distortions that vary spatially. Local distortion can defeat nearest-neighbor matching and can complicate the use of a single global model for correspondence.

In parallel, computer vision and point-set registration offer families of methods for matching under noise and deformation, including robust estimation and outlier rejection [6–10]. Their performance, however, depends on the availability of stable features in both input maps (or spatial data) and on whether the deformation difference can be captured by the chosen model. When raster linework is warped or aliased, key point salience can degrade and the definition of a "point" representation for junctions becomes ambiguous.

Multimodal LLMs are increasingly used for visual reasoning tasks that involve associating symbols, reading labels, and describing spatial relationships in images and diagrams. Their potential contribution to GIS workflows was evaluated recently in the studies [11–21] lies less in replacing classical geometric computation and more in providing interactive assistance, triaging candidate matches, and adapting to heterogeneous representations through language-guided reasoning. However, the reliability of such models in CP detection and correspondence under geometric distortion has not been widely established. This paper contributes empirical evidence for that setting, emphasizing the conditions under which LLM assistance should be constrained and combined with classical candidate generation. In this study, we use a multimodal LLM, GPT-5.2 "thinking" mode in a zero-shot setting, developed by OpenAI [22]. This work extends our earlier assessment of LLM-based map conflation, in which GPT-5 "thinking" mode was evaluated for line-to-line road correspondence under strict 1:1 and flexible M:N matching strategies [23]. While that study addressed edge-level conflation on topologically structured networks, the present paper focuses on a junction-level CP detection and correspondence under controlled geometric distortions. This paper investigates the reliability of LLM-assisted CP detection under controlled conditions that resemble real-world data heterogeneity: different input modalities, the presence or absence of explicit node layers with identifiers, and different distortion regimes that stress both detection and matching. The study is based on twelve experimental tests runs in which the same general task is executed under different combinations of modality, node availability,

and distortion level, and the outcomes are recorded as accuracy along with qualitative observations of error behavior.

2 Study Data, Experimental Design, and LLM-Assisted CP Detection

Figure 1 defines the experimental design through six source–target input combinations used for CP detection. In every combination, the left panel represents the source (original) network and the right panel represents the target network obtained by applying a controlled deformation operator. These six combinations are treated as distinct processes because they change the underlying difficulty of the CP task by independently varying the deformation regime applied to the target and whether intersections are provided as explicit node candidates. Each of the six processes is then instantiated in two data modalities: raster and vector, yielding twelve total experimental processes.

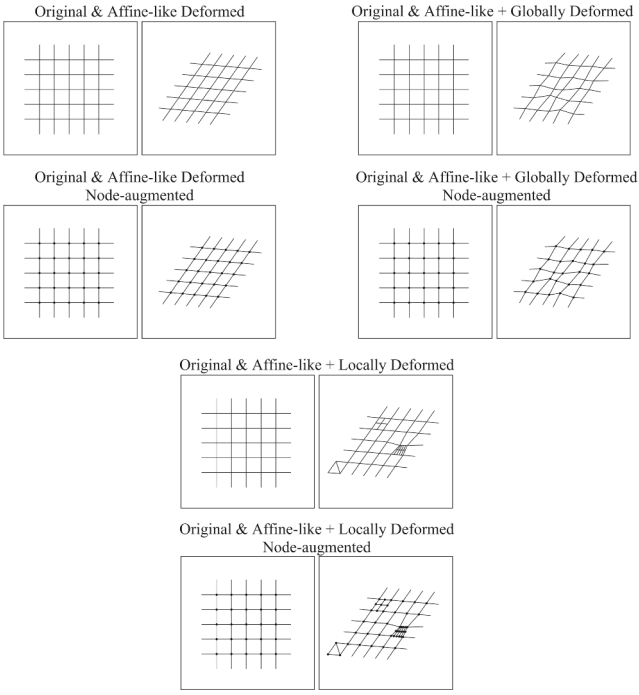

Fig. 1. Input combinations for CP detection.

The first axis of variation is the deformation regime. The "Original & Affine-like Deformed" combination represents globally consistent misalignment that can be approximated by a single parametric mapping; in practice this regime is associated with shifting, shear, scale, and rotation effects that preserve a coherent transformation across the entire

extent (Fig. 2). The "Original & Affine-like + Globally Deformed" combination represents a more challenging target in which a global affine-like component is present but additional globally acting deformation is superimposed, reducing global consistency and weakening the reliability of purely distance-based matching logic (Fig. 3). The "Original & Affine-like + Locally Deformed" combination represents spatially heterogeneous deformation concentrated within subregions; this regime is characterized by locally inconsistent displacements that may alter the apparent neighborhood structure of intersections within the deformed area while leaving other regions relatively coherent (Fig. 3). These three regimes allow the study to separate performance under globally explainable misalignment from performance under irregular and spatially varying deformation.

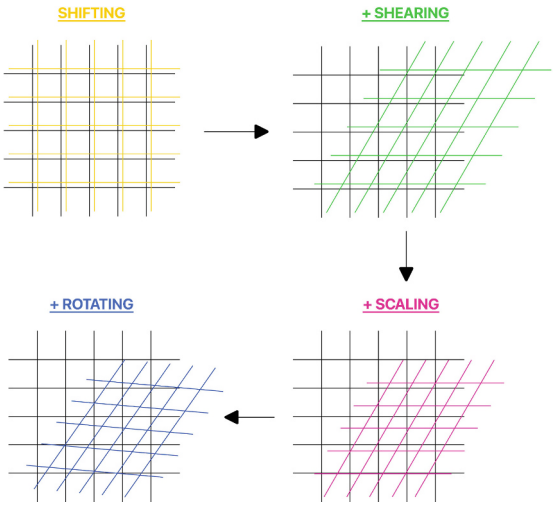

Fig. 2. Original (black), shifted (yellow), sheared (green), scaled (pink), and rotated (blue) networks.

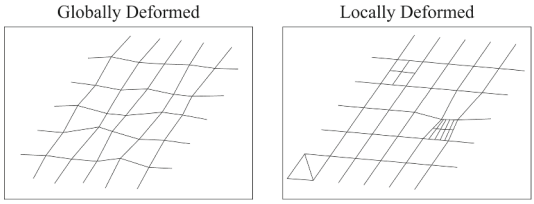

Fig. 3. Irregular deformations: Globally (left) and locally (right) distorted networks.

The second axis of variation is the availability of explicit intersection nodes. For each deformation regime, the experiment includes a line-only and a node-augmented variants (Fig. 1 and Table 1). In the line-only variant, intersections are just crossing lines and must be inferred from network geometry, which makes the task a joint detection

and matching problem. This means that the model must first decide where valid CPs might exist and then establish correspondences across the source and target. In the node-augmented variant, intersections are explicitly marked as points, which constrains the task to selecting and pairing from an explicit candidate set of junction points. Availability of nodes is assumed reducing the ambiguity in CP definition by shifting the task toward correspondence establishment under deformation. Using these variants is central to the evaluation because it isolates errors due to control point discovery from errors due to correspondence reasoning.

Table 1. Assisting prompts used to conduct experiments in LLM.

Input combination	Network structure	Assisting prompt
Original & Affine-like Deformed	Node-augmented	Possibility of one-to-one alignment At least 4 CPs
	Line-only	At least 4 CPs
Original & Affine-like + Globally Deformed	Node-augmented	Possibility of one-to-one alignment At least 4 CPs
	Line-only	N/A
Original & Affine-like + Locally Deformed	Node-augmented	At least 4 CPs
	Line-only	N/A

Each of the six processes, shown in Fig. 1, is executed in two modalities. In the raster setting, inputs are provided as network renderings in .tiff format, in which linework and, when applicable, node marks are interpreted as image structures. In this modality, intersection salience depends on pixel-level cues such as stroke crossings and junction shapes. In the vector setting, inputs are provided as line-based network geometry in .gpkg format. In node-augmented variants, an explicit point representation of intersections is provided as a separate .gpkg file. Comparisons between the processes carried out with raster and vector inputs isolate modality effects, including whether intersection salience in raster is preserved under deformation and whether vector line-only inputs introduce ambiguity in the absence of explicit node candidates.

In all experiments, the LLM was first given a set of standard task instructions requiring it to read the provided files, identify corresponding CPs for network alignment, and store the resulting CP associations in a downloadable output file. In addition to these baseline instructions, the assisting prompts listed in Table 1 were incorporated in the relevant tests to further guide the model's behavior. These auxiliary prompts were used selectively depending on the input combination and network structure, for example by explicitly framing the task as a one-to-one alignment problem in node-augmented settings and by enforcing a minimum number of CPs when appropriate.

3 Experimental Results

The results of raster tests in Table 2 and vector tests in Table 3 indicate that affine-like distortion is the regime in which LLM assistance is most reliable, particularly when explicit node markers with identifiers are available. In the node-augmented condition under affine-like distortion, both raster and vector runs achieve perfect correctness, consistent with the expectation that labeled node anchors make correspondence largely deterministic. In the line-only condition under the same distortion, raster achieves perfect accuracy for the requested minimum set of four CPs, whereas vector drops to half accuracy. This suggests that when the task is restricted to producing a small set of intersections under a globally consistent warp, raster imagery may offer cues that are easier for the LLM model to exploit than vector linework.

Table 2. Performance statistics in the CP detection scenarios using raster sources.

Input combination	Network structure	Total CP	Correct CP	%
Original & Affine-like Deformed	Node-augmented	25	25	100
	Line-only	4	4	100
Original & Affine-like + Globally Deformed	Node-augmented	25	25	100
	Line-only	20	2	10
Original & Affine-like + Locally Deformed	Node-augmented	18	9	50
	Line-only	25	8	32

Table 3. Performance statistics in the CP detection scenarios using vector sources.

Input combination	Network structure	Total CP	Correct CP	%
Original & Affine-like Deformed	Node-augmented	25	25	100
	Line-only	4	2	50
Original & Affine-like + Globally Deformed	Node-augmented	25	11	44
	Line-only	16	5	31.3
Original & Affine-like + Locally Deformed	Node-augmented	16	5	31.3
	Line-only	32	12	37.5

Under global and irregular distortion, performance diverges sharply across conditions. The experiment with node-augmented network in raster format remains perfect, while vector test drops substantially, indicating that even with labeled vector nodes the model can mis-associate correspondences when irregular deformation disrupts the spatial relationships the model expects (Tables 2 and 3). In the test with line-only rasterized network, the performance collapses to a very low accuracy and the qualitative

record indicates that the model selected control points along network edges rather than at intersections, which constitutes a primary detection failure.

Under local distortion, correctness percentages are consistently degraded across all configurations, with no single representation dominating. The qualitative observations highlight that locally distorted regions attract incorrect selections and that proximity-based matching remains a frequent default strategy, which is brittle when distortions vary spatially. Consequently, the study shows three significant results: (1) Regular affine-like deformations can yield strong LLM-assisted CP detection performance, especially when supported by explicit node usage and, where beneficial, an appropriate raster/vector representation choice. (2) the same advantage does not carry over when the discrepancy between the two inputs becomes irregular (whether globally irregular or locally irregular), where achieving consistently good performance becomes more difficult in vector-based settings. (3) When the mismatch is dominated by localized detail-level deformation differences, neither raster nor vector inputs produced reliable results. As a results, local distortions constitute a modality-agnostic limitation for the current LLM-assisted CP detection setup.

4 Discussion and Conclusion

The experiments demonstrate that LLM-assisted CP detection is strongly conditioned by both deformation regularity and the availability of explicit intersection nodes. Under regular affine-like distortions, performance is consistently high, and the node-augmented setting yields perfect correctness in both raster and vector inputs, confirming that labeled junction anchors effectively constrain the correspondence problem. When nodes are not provided, performance becomes more representation-dependent: raster linework supports accurate selection of a small minimum set of intersections under globally consistent warps, whereas vector line-only inputs show reduced correctness, suggesting that the absence of an explicit node representation increases ambiguity in how intersections are operationalized and paired.

When the discrepancy between inputs becomes irregular, the reliability observed in affine-like cases does not generalize. In particular, the correctness of vector node-augmented case decreased significantly, remarking that even with labeled nodes the LLM-assisted detection may fall back on proximity-driven reasoning and mis-associate correspondences when irregular deformation violates expected spatial relationships. In the raster line-only irregular case, correctness collapses primarily due to a detection failure, where the model selects points along edges rather than true intersections. These outcomes indicate that irregular deformation stresses both stages of the problem: identifying valid junction CPs and establishing one-to-one correspondence. In that case, the dominant error mode shifts from candidate ambiguity to systematic mis-selection and mis-association.

Local distortions produce consistently degraded correctness across all configurations, with only modest differences between raster and vector and between node-augmented and line-only settings. The qualitative pattern suggests that locally deformed regions attract incorrect selections and that proximity-based matching becomes brittle when deformation varies spatially, limiting the benefit of representation choice.

Consequently, while regular affine-like misalignments can be handled effectively through node augmentation and, in some cases, careful modality selection, irregular and especially locally varying distortions represent the principal limitation of the current LLM-assisted CP detection setup, as neither raster nor vector inputs provide consistently reliable performance. In the future, the LLM-assisted CP detection may benefit from using additional assisting prompts leading the LLM actions with topology-aware matching cues retrieved from source network structures.

Acknowledgments. This study was supported by the Scientific and Technological Research Council of Türkiye (TÜBİTAK) under the 2224-A - Grant Program for Participation in Scientific Meetings Abroad. The authors gratefully acknowledge this support.

Disclosure of Interests The author declares no conflict of interest.

References

1. Saalfeld, A.: Conflation automated map compilation. Int. J. Geogr. Inf. Syst. **2**, 217–228 (1988). https://doi.org/10.1080/02693798808927897
2. Rosen, B., Saalfeld, A.: Match criteria for automatic alignment. In: Proceedings of the 7th International Symposium on Computer-Assisted Cartography (Auto-Carto 7), Washington, D.C. (1985)
3. López-Vázquez, C., Manso Callejo, M.A.: Point- and curve-based geometric conflation. Int. J. Geogr. Inf. Sci. **27**, 192–207 (2013). https://doi.org/10.1080/13658816.2012.677537
4. Doytsher, Y., Filin, S., Ezra, E.: Transformation of datasets in a linear-based map conflation framework. Surv. Land Inf. Syst. **61**, 159–169 (2001)
5. Li, L., Goodchild, M.F.: An optimisation model for linear feature matching in geographical data conflation. Int. J. Image Data Fusion. **2**, 309–328 (2011). https://doi.org/10.1080/19479832.2011.577458
6. Besl, P.J., McKay, N.D.: A method for registration of 3-D shapes. IEEE Trans. Pattern Anal. Mach. Intell. **14**, 239–256 (1992). https://doi.org/10.1109/34.121791
7. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM. **24**, 381–395 (1981). https://doi.org/10.1145/358669.358692
8. Bookstein, F.L.: Principal warps: thin-plate splines and the decomposition of deformations. IEEE Trans. Pattern Anal. Mach. Intell. **11**, 567–585 (1989). https://doi.org/10.1109/34.24792
9. Myronenko, A., Song, X.: Point set registration: coherent point drift. IEEE Trans. Pattern Anal. Mach. Intell. **32**, 2262–2275 (2010). https://doi.org/10.1109/TPAMI.2010.46
10. Ma, J., Zhao, J., Tian, J., Yuille, A.L., Tu, Z.: Robust point matching via vector field consensus. IEEE Trans. Image Process. **23**, 1706–1721 (2014). https://doi.org/10.1109/TIP.2014.2307478
11. Tao, R., Xu, J.: Mapping with ChatGPT. ISPRS Int. J. Geo Inf. **12**, 284 (2023). https://doi.org/10.3390/ijgi12070284
12. Mooney, P., Cui, W., Guan, B., Juhász, L.: Towards understanding the geospatial skills of ChatGPT: taking a geographic information systems (GIS) exam. In: Proceedings of the 6th ACM SIGSPATIAL International Workshop on AI for Geographic Knowledge Discovery, New York, NY, USA, pp. 85–94 (2023)
13. Mansourian, A., Oucheikh, R.: ChatGeoAI: enabling geospatial analysis for public through natural language, with large language models. ISPRS Int. J. Geo Inf. **13**, 348 (2024). https://doi.org/10.3390/ijgi13100348

14. Zhang, Y., Wei, C., He, Z., Yu, W.: GeoGPT: an assistant for understanding and processing geospatial tasks. Int. J. Appl. Earth Obs. Geoinf. **131**, 103976 (2024). https://doi.org/10.1016/j.jag.2024.103976
15. Zhang, Y. et al.: Geospatial large language model trained with a simulated environment for generating tool-use chains autonomously. Int. J. Appl. Earth Obs. Geoinf. **136**, 104312 (2024). https://doi.org/10.1016/j.jag.2024.104312
16. Gulnerman, A.G.: AI-driven urban perception mapping: comparing Gemini and LLaVA. In: International Conference on Computational Science and its Applications, Istanbul, Türkiye, pp. 339–349 (2025). https://doi.org/10.1007/978-3-031-97663-6_30
17. Memduhoğlu, A.: Towards AI-assisted mapmaking: assessing the capabilities of GPT-4o in cartographic design. ISPRS Int. J. Geo Inf. **14**, 35 (2025). https://doi.org/10.3390/ijgi14010035
18. Pierdicca, R., Muralikrishna, N., Tonetto, F., Ghianda, A.: On the use of LLMs for GIS-based spatial analysis. ISPRS Int. J. Geo Inf. **14**, 401 (2025). https://doi.org/10.3390/ijgi14100401
19. Chen, G. et al.: GeoJSEval: an automated evaluation framework for large language models on JavaScript-based geospatial computation and visualization code generation. ISPRS Int. J. Geo Inf. **14**, 382 (2025). https://doi.org/10.3390/ijgi14100382
20. Wu, H. et al.: AutoGEEval: a multimodal and automated evaluation framework for geospatial code generation on GEE with large language models. ISPRS Int. J. Geo Inf. **14**, 256 (2025). https://doi.org/10.3390/ijgi14070256
21. Wu, H. et al.: AutoGEEval++: a multi-level and multi-geospatial-modality automated evaluation framework for large language models in geospatial code generation on Google Earth Engine. Big Earth Data. **9**, 747–796 (2025). https://doi.org/10.1080/20964471.2025.2581425
22. OpenAI: GPT-5.2 System Card Addendum (2025). https://openai.com/index/gpt-5-system-card-update-gpt-5-2/
23. Hacar, M., Öztürk Hacar, Ö.: LLM-based map conflation: performance assessment on matching embedded road lines. ISPRS Int. J. Geo Inf. **15**, 144 (2026). https://doi.org/10.3390/ijgi15040144