LLM-Based Multi-Agent System and Simplicial Self-Supervised Learning Model for Regional Cancer Prevalence Estimation Using Satellite Imagery

Jiue-An Yang
City of Hope Cancer Center
jyang@coh.org

Yuzhou Chen
University of California, Riverside
yuzhou.chen@ucr.edu

Calvin P. Tribby
City of Hope Cancer Center
ctribby@coh.org

Huikyo Lee
NASA Jet Propulsion Laboratory
huikyo.lee@jpl.nasa.gov

Loretta Erhunmwunsee
City of Hope Cancer Center
lorettae@coh.org

Tarik Benmarhnia
University of California San Diego
tbenmarhnia@ucsd.edu

Caroline A. Thompson
UNC Gillings School of Global Public Health
caroline.thompson@unc.edu

Yulia R. Gel
Virginia Tech
ygl@vt.edu

Marta M. Jankowska
City of Hope Cancer Center
mjankowska@coh.org

Abstract

Traditional cancer rate estimations are often limited in spatial resolutions and lack considerations of environmental factors. Satellite imagery has become a vital data source for monitoring diverse urban environments, supporting applications across environmental, socio-demographic, and public health domains. However, while deep learning (DL) tools, particularly convolutional neural networks, have demonstrated strong performance in extracting features from high-resolution imagery, their reliance on local spatial cues often limits their ability to capture complex, non-local, and higher-order structural information. To overcome this limitation, we propose a novel LLM-based multi-agent coordination system for satellite image analysis, which integrates visual and contextual reasoning through a simplicial contrastive learning framework (Agent-SNN). Our Agent-SNN contains two augmented superpixel-based graphs and maximizes mutual information between their latent simplicial complex representations, thereby enabling the system to learn both local and global topological features. The LLM-based agents generate structured prompts that guide the alignment of these representations across modalities. Experiments with satellite imagery of Los Angeles and San Diego demonstrate that Agent-SNN achieves significant improvements over state-of-the-art baselines in regional cancer prevalence estimation tasks.

CCS Concepts

• **Computing methodologies** → **Machine learning approaches**.

Keywords

GeoAI, Large Language Models, Satellite Imagery, Precision Population Health

SIGSPATIAL '25, Minneapolis, MN, USA

ACM ISBN 979-8-4007-2086-4/2025/11
https://doi.org/10.1145/3748636.3763225

ACM Reference Format:
Jiue-An Yang, Yuzhou Chen, Calvin P. Tribby, Huikyo Lee, Loretta Erhunmwunsee, Tarik Benmarhnia, Caroline A. Thompson, Yulia R. Gel, and Marta M. Jankowska. 2025. LLM-Based Multi-Agent System and Simplicial Self-Supervised Learning Model for Regional Cancer Prevalence Estimation Using Satellite Imagery. In *The 33rd ACM International Conference on Advances in Geographic Information Systems (SIGSPATIAL '25), November 3–6, 2025, Minneapolis, MN, USA.* ACM, New York, NY, USA, 4 pages. https://doi.org/10.1145/3748636.3763225

1 Introduction

Accurate assessment of geographically refined cancer rates is essential for effective disease surveillance and planning targeted screening and risk reduction interventions. Traditional cancer surveillance data are often limited in coverage, timeliness, and spatial resolution, especially in regions lacking robust public health infrastructure. In turn, the association between environmental factors and increased cancer incidence and mortality has been well documented [7, 9], but is yet to be utilized in understanding and predicting cancer incidence. Satellite imagery offers a widely available data source for monitoring a broad range of urban geographical contexts related to environmental, socio-demographic, and neighborhood characteristics which are important for evaluating cancer incidence. Recent advancement of deep learning (DL) tools provides a competitive suite of algorithms, e.g., convolutional neural networks (CNNs) or graph neural networks (GNN), to analyze such satellite imagery for urban patterns and features identification [10]. Finally, Large Language Models (LLMs) have also made significant strides in a diverse spectrum of tasks, extending beyond natural language processing into language-centric applications across disciplines [3, 15], opening new path for integration of multi-modal information associated with assessment of cancer rates.

A previous study [2] has experimented with CNN models to predict cancer prevalence with promising results. Here we introduce a novel LLM-based multi-agent system and simplicial neural networks (Agent-SNN) and validate its application in estimating census tract level cancer prevalence. Agent-SNN is the first framework to introduce simplicial complex structures into satellite image

representation learning through a multi-agent coordination system built on LLM-based agents. By jointly contrasting higher-order interactions derived from superpixel graphs with semantically rich prompts generated by LLM-guided agents, Agent-SNN enables more expressive and discriminative representations. This design is particularly well-suited for capturing the heterogeneity and complex structural patterns inherent in urban landscapes, offering a principled approach for learning richer and more nuanced features from satellite imagery. We demonstrate the superior performance of Agent-SNN in application to the analysis of urban forms and their latent relationships with regional cancer prevalence at two urbanized regions, i.e., Los Angeles and San Diego, California.

2 Methodology

Satellite Imagery. Publicly available Sentinel-2 (L2A) satellite images are collected for the study areas. The 10-meter spatial resolution spectral bands 4-3-2 (R-G-B) are used for nature color image composition. Images were split into smaller image tiles (approximately $450m \times 450m$ each), at 45×45 pixels. The center coordinates of each tile were used to georeference image tiles to census tracts.

Superpixel Graph Generation. We now introduce how to construct a superpixel graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ for a satellite image where $\mathcal{V}$ and $\mathcal{E}$ denote the nodes and edges, respectively. We first apply the simple linear iterative clustering (SLIC) function [1] SLIC to the target image $\boldsymbol{I} \in \mathbb{R}^{H \times W \times C}$ (here H, W, and C are height, weight, and the number of channels of the image $\boldsymbol{I}$ respectively), resulting in the corresponding superpixels $\boldsymbol{S}$:

$$\boldsymbol{S}, \mathcal{V} = \text{SLIC}(\boldsymbol{I}; Q), \tag{1}$$

where Q denotes the pre-defined number of superpixels in the segmented image $\boldsymbol{I}$ for SLIC, and $\mathcal{V}$ is a node set containing $\mathcal{N}$ superpixel nodes, i.e., $\mathcal{V} = \{v\}_i^{\mathcal{N}}$, and $|\mathcal{V}| = \mathcal{N}$ is the exact number of superpixels. The edge set $\mathcal{E}$ is constructed based on spatial adjacency of superpixels, i.e., if two superpixels v_i and v_j share a common boundary in the segmented image, then an undirected edge (v_i, v_j) is added to $\mathcal{E}$. Hence, given a source satellite image dataset $\mathcal{I} = \{\boldsymbol{I}_j\}_{j=1}^{N}$, where N is the number of satellite images, we can obtain the corresponding set of graphs $\mathcal{G} = \{\mathcal{G}_j\}_{j=1}^{N}$.

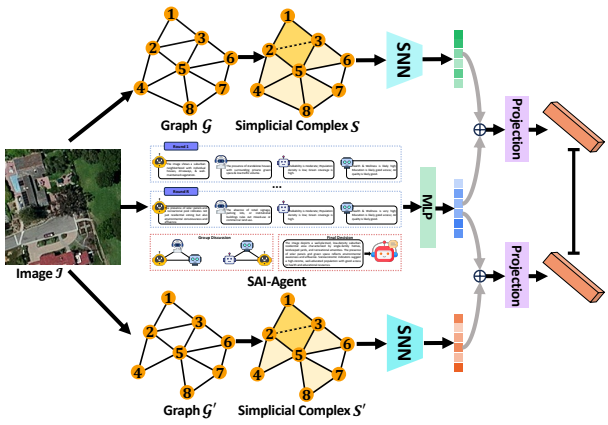

Figure 1: The overview of Agent-SNN model.

Next we turn to the detailed discussion of our proposed Agent-SNN paradigm. Figure 1 presents the Agent-SNN pipeline. We first introduce the image-to-graph conversion and guide LLM-based agent through designed prompt statements, which allows us to distill rich textual information about the satellite imagery via image-to-text generation. We then develop a novel contrastive learning framework which incorporates the complex higher-order interactions described by simplices into the representation learning step, systematically combining both text embedding and simplicial embeddings. As a result, this step allows us to substantially enhance the satellite image latent representation learning.

Multi-Agent Coordination System. Vision large language models (VLLMs) have emerged recently as promising tools for visual reasoning tasks [19]. In the context of satellite image analysis, each geographic region can be represented by high-resolution imagery capturing diverse features such as land use, infrastructure, vegetation patterns, and environmental anomalies. By translating these visual signals into descriptive language representations, VLLMs enable semantic understanding of complex geospatial scenes and support tasks such as land cover classification, object detection, and activity inference. However, fundamental challenges remain, including hallucinated descriptions, limited domain-specific knowledge, and a lack of interpretability within standalone models within single VLLM. To address the above limitations, we develop a novel SAtellite Imagery LLM-Agent (SAI-Agent) system for satellite imagery reasoning and analysis.

<Prompt 1: Identifying Visual-Level Urban Features> *You are an expert Earth/Environmental scientist. Given the satellite image: Describe the image and count the number of cars, buildings, trees, roads, waterbodies \n What is the proportion of green space in this image? \n What is the proportion of built environment in this image?*

<Prompt 2: Inferring Land Use> *You are an expert Earth/Environmental scientist. Given the satellite image: What is the major land use of this area? \n Is this a residential neighborhood? \n Is this a commercial area? \n Is this an urban, suburban, or rural region? \n What city or region is this?*

<Prompt 3: Inferring Characteristics from Visual-Level Features> *You are an expert Earth/Environmental scientist. Given the satellite image: Is this area crowded? \n Is this a rich or poor region? \n Is this area good for outdoor activity? \n Is this area good for walking around?*

<Prompt 4: Inferring Characteristics of the Neighborhood> *You are an expert Earth/Environmental scientist. Given the satellite image: What are the socioeconomic characteristics of this region? \n Is this a healthy or unhealthy neighborhood? \n Do you think this area has poor air quality? \n What is the average educational level of the resident in this area?*

Figure 2: A prompt template of satellite image reasoning.

Our SAI-Agent system enables agents to produce answers in each single round, interact multiple rounds, and move on to the final task solving stage. For each round r, we design a zero-shot prompt template and instruct each agent $\mathcal{A}_i$ (where $i \in [2, \ldots, \mathcal{I}]$) to act in capacity of an Earth science expert (or environmental science expert). Fig. 2 shows an example of the prompt $\mathcal{P}$ designed for a node in our dataset. Our expert-designed prompt consists of two parts: (i) the general prompt which introduces the task scenario and (ii) dataset description which characterizes the targeted satellite image. Thus, this designed prompt provides the agent $\mathcal{A}_i$ with the general goal and the reasoning task. Then we use the proposed prompt $\mathcal{P}$ to query the agent $\mathcal{A}_i$ to get the description of the satellite

image. For a node u, the first round process is formally defined by:

$$\mathcal{R}_{i,u}^{(1)} = \mathscr{A}_i(\mathcal{P}, \mathcal{D}(u)), \tag{2}$$

where $\mathcal{R}_{i,u}^{1}$ denotes the explanations during the first round, and $\mathcal{D}(u)$ is the description for the satellite image u which includes geographic and spatial context, land use/land cover features, environmental conditions, and vegetation health and density. Beginning from the second round, for each agent $\mathscr{A}_i$, we allow agents to view the historical explanations and answers from all the previous rounds. Hence, in the r-th round, each LLM-based agent $\mathscr{A}_i$ receives the prompt $\mathcal{P}$ and historical responses and answers, and generates the current round response $\mathcal{R}_{i,u}^{(r)}$, which is written as:

$$\mathcal{R}_{i,u}^{(r+1)} = \mathscr{A}_i(\mathcal{P}, \mathcal{D}(u), \mathcal{R}_{i,u}^{(1):(r)}, \{\mathcal{R}_{j,u}^{(1):(r)}\}_{j\neq i}^{I}), \tag{3}$$

where $\mathcal{R}_{i,u}^{(1):(r)}$ contains the agent $\mathscr{A}_i$'s responses of r rounds $\{\mathcal{R}_{i,u}^{(1)}, \ldots, \mathcal{R}_{i,u}^{(r)}\}$, and $\{\mathcal{R}_{j,u}^{(1):(r)}\}_{j\neq i}^{I}$ are the answers and responses from other agents from previous r rounds, respectively.

Through multiple rounds, to make SAI-Agent system converge on a single and consensus answer, we consider using the meta-judge framework. In the setting of meta-judge framework, we create a new agent which is assigned a role as the meta-judge coordinator assistant denoted as $\mathscr{A}_J$. Instead of comparing voting behaviors, the agent $\mathscr{A}_J$ aggregates all agents' responses and answers, and then produces a definitive descriptive result for each image. The meta-judge evaluates these outputs by considering factors including agent-specific confidence weights and historical performance.

$$\mathcal{R}_{u}^{Final} = \mathscr{A}_J(\mathcal{P}, \mathcal{D}(u), \{\mathcal{R}_{i,u}^{(1):(R)}\}_{i=1}^{I}). \tag{4}$$

Thus, by using the SAI-Agent system, we can obtain descriptions of all images, i.e., $\vec{\mathcal{R}}^{Final} = [\mathcal{R}_1^{Final}, \ldots, \mathcal{R}_N^{Final}]$. Then, for each image u, we can generate its corresponding latent embedding, i.e.,

$$\boldsymbol{H}_{\text{SAI-Agent},u} = \text{MLP}(\text{LM}(\mathcal{R}_u^{Final})), \tag{5}$$

where MLP denotes a multi-layer perception (MLP), LM is a pre-trained language model (e.g., BERT and GPT-2), and $\boldsymbol{H}_{\text{SAI-Agent},u}$ denotes the latent embedding of satellite image u.

Superpixel Simplicial Neural Networks. To better capture the spatial dependencies between higher-order substructures, we leverage the concepts of the Hodge theory [6]. In particular, we extend the notion of a standard combinatorial graph Laplacian, which addresses diffusion from node to node of $\mathcal{G}$ through edges, to diffusion over higher-order substructures of $\mathcal{G}$, described by k-simplices of $\mathcal{G}$ (where $k = \{0, 1, 2, \cdots\}$). The generalization of graph Laplacians enables us to account for complex multi-node interactions in $\mathcal{G}$. The standard graph Laplacian $\boldsymbol{L}_0 \in \mathbb{R}^{\mathcal{N}\times\mathcal{N}}$ is a special case of the above k-th combinatorial Hodge Laplacian [4, 11] and the matrix $\boldsymbol{L}_1 \in \mathbb{R}^{\mathcal{M}\times\mathcal{M}}$ is the Hodge 1-Laplacian. We then define the k-simplex convolution operation as:

$$\boldsymbol{Z}_{\text{SC}_k}^{(\ell+1)} = \sigma(\boldsymbol{L}_k \boldsymbol{Z}_{\text{SC}_k}^{(\ell)} \boldsymbol{\Theta}_k^{(\ell)}), \tag{6}$$

where $\boldsymbol{Z}_{\text{SC}_k}$ is the input activation matrix to the ℓ-th hidden layer (here $\boldsymbol{Z}_{\text{SC}_k}^{(0)} = \boldsymbol{X}_k$), $\boldsymbol{\Theta}_k^{(\ell)}$ is the ℓ-th layer trainable weight matrix, and $\sigma(\cdot)$ is the non-linear activation function such as ReLU. Also, inspired by [18], we use layer-wise aggregation of learned features on all layers to enhance the learned k-simplex representations:

$$\tilde{\boldsymbol{Z}}_{\text{SC}_k} = [\boldsymbol{Z}_{\text{SC}_k}^{(1)}, \boldsymbol{Z}_{\text{SC}_k}^{(2)}, \cdots, \boldsymbol{Z}_{\text{SC}_k}^{(\ell+1)}], \tag{7}$$

where $[\cdot,\cdot]$ denotes the concatenation of $(\ell+1)$ outputs of k-simplex convolution operation. Then, we employ global pooling layer READOUT to obtain the graph-level feature: Ξ_k

$$\boldsymbol{H}_{\text{SC}_k} = \sigma([\text{GAP}(\tilde{\boldsymbol{Z}}_{\text{SC}_k}), \text{GMP}(\tilde{\boldsymbol{Z}}_{\text{SC}_k})]), \tag{8}$$

where a READOUT function yields a permutation-invariant graph-level representation via aggregating node representations into a graph-level representation. Specifically, we adopt the concatenation of global average pooling (GAP) and global max pooling (GMP) to read out the graph representation. Therefore, we obtain the final embedding $\boldsymbol{H}_O$ by combining two embeddings:

$$\boldsymbol{H}_O = \alpha_{\text{SC}_k} \times \boldsymbol{H}_{\text{SC}_k} + \alpha_{\text{SAI-Agent}} \times \boldsymbol{H}_{\text{SAI-Agent}}, \tag{9}$$

where α_{SC_k} and $\alpha_{\text{SAI-Agent}}$ are hyperparameters (in this work, we set $\alpha_{\text{SC}_k} = \alpha_{\text{SAI-Agent}} = 0.5$).

Contrastive Learning on Agent-SNN. Using Eq. 9, we can generate latent SAI-Agent based simplicial complex representations $\tilde{\boldsymbol{H}}_{i,O}$ and $\tilde{\boldsymbol{H}}'_{i,O}$ from two perturbed superpixel graphs $\tilde{\mathcal{G}}_i$ and $\tilde{\mathcal{G}}'_i$ (i.e., two augmented views from the same superpixel graph $\mathcal{G}$), respectively. For every latent SAI-Agent based simplicial complex representation $\tilde{\boldsymbol{H}}_{i,O}$ being the anchor instance, its positive sample is the latent SAI-Agent based simplicial complex representation $\tilde{\boldsymbol{H}}'_{i,O}$ of the another augmented view. Naturally, negative pairs are latent extended topological representations (e.g., $\tilde{\boldsymbol{H}}_{\gamma,O}$) generated from other augmented superpixel graphs (e.g., $\tilde{\mathcal{G}}_\gamma \in \tilde{\boldsymbol{\mathcal{G}}} \setminus \{\tilde{\mathcal{G}}_i, \tilde{\mathcal{G}}'_i\}$, where $\tilde{\mathcal{G}}_\gamma$ is a augmented superpixel graph of $\mathcal{G}_\gamma$ and $\tilde{\boldsymbol{\mathcal{G}}}$ is a set of 2Υ augmented superpixel graphs). The loss of function is then defined to enforce maximizing the consistency between positive pairs $(\tilde{\boldsymbol{H}}_{i,O}, \tilde{\boldsymbol{H}}'_{i,O})$ compared with negative pairs, which is formulated as

$$\ell_i(\tilde{\mathcal{G}}_i, \tilde{\mathcal{G}}'_i) = -\log \frac{\exp\left(\text{sim}(\tilde{\boldsymbol{H}}_{i,O}, \tilde{\boldsymbol{H}}'_{i,O})/\zeta\right)}{\sum_{\gamma,\gamma\neq i}^{2\Upsilon} \exp\left(\text{sim}(\tilde{\boldsymbol{H}}_{i,O}, \tilde{\boldsymbol{H}}_{\gamma,O})/\zeta\right)},$$

where $\text{sim}(\tilde{\boldsymbol{H}}_{i,O}, \tilde{\boldsymbol{H}}'_{i,O}) = \tilde{\boldsymbol{H}}_{i,O}^{\top}\tilde{\boldsymbol{H}}'_{i,O}/||\tilde{\boldsymbol{H}}_{i,O}||||\tilde{\boldsymbol{H}}'_{i,O}||$, $\tilde{\boldsymbol{H}}_{\gamma,O}$ is the latent SAI-Agent based simplicial complex representation of $\tilde{\mathcal{G}}_\gamma$ learnt by our proposed, and ζ is the temperature hyperparameter.

3 Experiments

Census Tract Cancer Prevalence. The estimated crude and age-adjusted cancer prevalence among adults was extracted from the 2022 PLACES release for census tracts in the City of Los Angeles (LA; $n = 959$) and San Diego (SD; $n = 280$). PLACES data provides model-based estimates based on data from Behavioral Risk Factor Surveillance System (BRFSS), Census decennial population counts and annual county population estimates, and the American Community Survey 5-year estimates.

Baselines and Experiment Settings. Sentinel-2 images at 10-meter spatial resolution were used for the experiment, creating 6,020 tiles for LA and 4,238 tiles for SD. For the CNN-based model baseline, we extract embeddings using the pre-trained ResNet model (ResNet-152) [5], which yields a 2048-dimension representation for each image. For the GNN-based model comparison, we use the InfoGraph [17], i.e., an unsupervised whole graph representation

learning model. Hyperparameters of GNN, LLM-SNN (i.e., the ablated variant of Agent-SNN by using a single LLM), and Agent-SNN are tuned in the range of embedding size {64, 128, 256}, dropout is selected from $[0, 0.5]$, and layer number $\{1, 2, 3\}$. We run our experiments on 4 NVIDIA RTX A5000 GPU cards with 24GB memory. We use the Adam optimizer and grid search to tune the learning rate within the range $[1e^{-3}, 2e^{-3}, \ldots, 1e^{-2}]$, and a batch size of 32 for 1,000 epochs. For LLMs, we use GPT-4o [14] and its temperature parameter is set to be 1.0. Each census tract is represented by a single vector of the mean embeddings from all image tiles within its boundary. Estimation of cancer prevalence for census tracts are through an optimization regularized linear regression model (Elastic Net) following the pipeline illustrated in previous studies [2, 8, 12]. Performance metrics include the coefficient of determination (i.e., R^2), mean squared error (MSE), and mean absolute error (MAE), with 95% confidence intervals.

City	Model	R^2	MSE	MAE
LA	CNN	41.74 ± 5.01	1.43 ± 0.22	0.90 ± 0.06
LA	GNN	2.05 ± 4.03	2.40 ± 0.26	1.23 ± 0.06
LA	LLM-SNN	21.07 ± 4.89	1.93 ± 0.22	1.08 ± 0.05
LA	**Agent-SNN (ours)**	**48.81 ± 4.30**	**1.25 ± 0.16**	**0.83 ± 0.04**
SD	CNN	24.94 ± 12.74	3.05 ± 1.37	1.23 ± 0.21
SD	GNN	−2.82 ± 2.80	4.04 ± 1.21	1.45 ± 0.17
SD	LLM-SNN	10.13 ± 9.43	3.61 ± 1.37	1.35 ± 0.19
SD	**Agent-SNN (ours)**	**28.88 ± 10.79**	**2.92 ± 1.32**	**1.22 ± 0.20**

Table 1: Evaluation of cancer prevalence estimation using satellite imagery features learned from DL models.

Experiment Results. In the application of census tract level cancer prevalence estimation in the two most populated cities of California, Agent-SNN outperforms baseline models of CNN, GNN, and an implementation of Agent-SNN (see Table 1). Embeddings learned from Agent-SNN explain more variances in both study areas and result in better regression model fits. Visualizations of the predicted regional cancer prevalence (Figure 3) demonstrate consistent spatial patterns compared to the cancer prevalence from PLACES data.

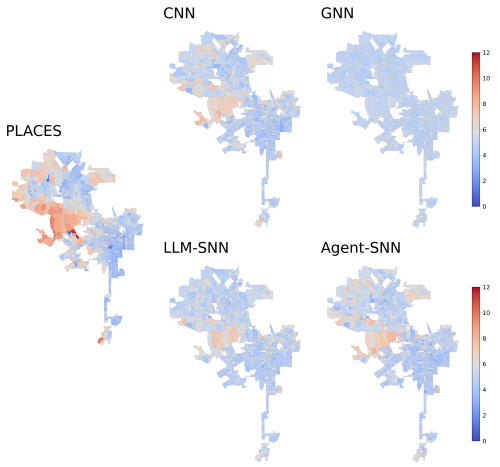

Figure 3: Census tract level cancer prevalence estimations using features from baseline models and Agent-SSN in LA.

4 Conclusion and Future Directions

We have introduced the LLM-Based Multi-Agent System and Simplicial Self-Supervised Learning Model and tested its utility in estimating census tract level cancer prevalence. Our experiments have shown that the new method achieves better estimation results than competing DL tools. Continued efforts are needed to further improve performance, quantify uncertainties, and and gain explainability of the learned features. One promising approach in this direction is to develop tractable multi-resolution visualization, alongside the extracted embeddings [13, 16]. More generally, the obtained results suggest that the DL-based high-resolution modeling has a potential not only to fill gaps where public surveillance data are absent, but also to support precision population health by enabling interventions tailored to specific neighborhoods or even smaller geographic units.

Acknowledgments

This work is supported by the National Cancer Institute Cancer Center Support Grant P30CA033572. Y.C. has been supported by NSF DMS-2523484, DMS-2533985, OAC-2530469, TIP-2333703. Y.R.G. has been supported by NSF OAC-2530471, DMS-2533984, and NASA AIST 21-AIST21_2-0020.

References

[1] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. 2012. SLIC superpixels compared to state-of-the-art superpixel methods. *TPAMI* (2012).
[2] J.-E. Bibault, M. Bassenne, H. Ren, and L. Xing. 2020. Deep Learning Prediction of Cancer Prevalence from Satellite Imagery. *Cancers* 12, 12 (2020).
[3] F. Busch et al. 2025. Current applications and challenges in LLM for patient care: a systematic review. *Commun. Med.* 5, 1 (2025), 26.
[4] Y. Chen, Y. R. Gel, and H. V. Poor. 2022. BScNets: Block simplicial complex neural networks. In *AAAI*, Vol. 36. 6333–6341.
[5] K. He, X. Zhang, S. Ren, and J. Sun. 2016. Deep Residual Learning for Image Recognition. In *CVPR*.
[6] W. V. D. Hodge. 1989. *The theory and applications of harmonic integrals.* CUP.
[7] J. S. Jagai, L. C. Messer, K. M. Rappazzo, C. L. Gray, S. C. Grabich, and D. T. Lobdell. 2017. County-level cumulative environmental quality associated with cancer incidence. *Cancer* (2017).
[8] N. Jean, M. Burke, M. Xie, W. M. Alampay Davis, D. B. Lobell, and S. Ermon. 2016. Combining satellite imagery and machine learning to predict poverty. *Science* 353, 6301 (2016), 790–794.
[9] A. M. A. Kamarei, O. Blanton, and S. M. Q. Hussaini. 2025. Association between environmental burden and cancer incidence rates across population subgroups in the United States. *Journal of Clinical Oncology* (2025).
[10] W. Li and C.-Y. Hsu. 2022. GeoAI for Large-Scale Image Analysis and Machine Vision: Recent Progress of AI in Geography. *ISPRS IJGI* 11, 7 (2022).
[11] L.-H. Lim. 2020. Hodge Laplacians on graphs. *Siam Review* (2020).
[12] A. Maharana and E. O. Nsoesie. 2018. Use of Deep Learning to Examine the Association of the Built Environment With Prevalence of Neighborhood Adult Obesity. *JAMA Network Open* 1, 4 (2018), e181535–e181535.
[13] E. Munoz, A. D. VanHelene, N. T. Yang, and A. G. Ramirez. 2025. CancerClarity app: Enhancing cancer data visualization with AI-generated narratives. *Preventive Oncology & Epidemiology* 3, 1 (2025), 2431501.
[14] OpenAI. 2024. GPT-4o Technical Overview. https://openai.com/index/gpt-4o. Accessed June 2025.
[15] M. A. K. Raiaan et al. 2024. A review on LLM: Architectures, applications, taxonomies, open issues and challenges. *IEEE Access* 12 (2024), 26839–26874.
[16] L. L. Ramírez-Ramírez, Y. R. Gel, M. Thompson, E. de Villa, and M. McPherson. 2013. A new surveillance and spatio-temporal visualization tool SIMID: SIMulation of infectious diseases using random networks and GIS. *Comput. Methods Programs Biomed.* 110, 3 (2013), 455–470.
[17] F.-Y. Sun, J. Hoffman, V. Verma, and J. Tang. 2019. InfoGraph: Unsupervised and Semi-supervised Graph-Level Representation Learning via Mutual Information Maximization. In *ICLR*.
[18] K. Xu, C. Li, Y. Tian, T. Sonobe, K. Kawarabayashi, and S. Jegelka. 2018. Representation learning on graphs with jumping knowledge networks. In *ICML*.
[19] J. Zhang, J. Huang, S. Jin, and S. Lu. 2024. Vision-language models for vision tasks: A survey. *TPAMI* (2024).