A Graph Neural Network-based Map Tiles Extraction Method Considering POIs Priority Visualization on Web Map Zoom Dimension

HUAZE XIE¹, DA LI², YUANYUAN WANG³ AND YUKIKO KAWAI.⁴
¹Division of Frontier Informatics, Kyoto Sangyo University, Kyoto, Japan (e-mail: x.huaze2019@outlook.com)
²Faculty of Engineering, Fukuoka University, Fukuoka, Japan (e-mail: lida@fukuoka-u.ac.jp)
³Graduate School of Sciences and Technology for Innovation, Yamaguchi University, Ube, Japan (e-mail: y.wang@yamaguchi-u.ac.jp)
⁴Division of Frontier Informatics, Kyoto Sangyo University, Kyoto, Japan (e-mail: kawai@cc.kyoto-su.ac.jp)
Corresponding author: Yukiko Kawai (e-mail: kawai@cc.kyoto-su.ac.jp).

This work was partially supported by JSPS KAKENHI Grant Numbers JP19K12240, 19H04118, JP20H00584, and JP21K17862

ABSTRACT
Owing to the tremendous popularity of mobile networks, point-of-interest (POI) data of location-based social networks (LSBN) provide significant geographic information on maps and can be utilized to discuss the dynamic characteristics of map tiles as segmented by city roads. In this study, motivated by the problem of the dynamic characteristic analysis of the map tile, we propose a spatial-zoom graph-attention model (SZ-GAT) based on a global-attention mechanism and 5-category POI attributes for each map tile zoom dimension. Furthermore, a social-media dataset (Twitter with geolocation) is utilized to promote POI visualization at different zoom levels and improve the aggregation efficiency of geographic records in zoom dimensions. In the experiments, we extract POI geo-features from Twitter and display the user’s favorite POI features at each map zooming level with 5-dimensional tweet attributes. We evaluate the accuracy of the POI prediction on Google, OpenStreetMap, Bing, and Yahoo! maps by comparing the tweets’ visit history. The predictive performance of the proposed method is more than 86% for each zoom level on 60 randomly-selected map tiles in Kyoto City.

INDEX TERMS
Map Zooming, Graph Neural Networks, Map Tile Extraction, Dynamic POI Visualization

I. INTRODUCTION
Owing to the increasing popularity of social media services, mobile terminals have enabled the public to share their daily activities online and leave their digital footprint in urban areas. The information provided by social media reflects the real-time characteristics (flows, services, scenery, and parking) of urban facilities such as restaurants, transportation, and attractions [1]. GPS coordinates within urban areas collected by geotagged social-media messages help researchers understand dynamic spatially oriented human activities and spatial urban patterns [2]. For the web map, the zoom-in and zoom-out operations easily observe every country, city, and street. Displayed categories (buildings, placemarks, streets, and boundaries) gradually increase when the map zooms in and disappear when it zooms out [3]. The geographic characteristics displayed from a country to a street-level scale are called the level of detail (LOD) on a map [4]. Particularly, the web map is composed of several tiles, and each map tile displays only a small part of the placemark names and hides most of the geographic information. The dynamic characteristics of map tiles are described by point-of-interest (POI) data, where the location based on users’ visit history records are the underlying pattern of trips and spatial interactions in cities, POI revisited-spatial interaction, and distance decay in spatially embedded networks. In this study, we expect to maintain users’ favorite POIs at multiple map zooming levels and focus on the problem of POI dynamic display (zoom-in) and disappearance (zoom-out) on the map. The Google Map was utilized as the base map in this study, and its merits are as follows: (1) Google Map is the most commonly used map in daily activities, and it shows many user reviews about placemarks; (2) there are few studies on the map zooming layer compared with other map-tile-analysis methods; (3) based on the experience of the area location, map tile dense
Dynamic POI analysis combined with map zooming is a new challenge in the task of user interest recommendation. There are three challenges associated with using machine-learning models for analyzing geographic information data for social media. 

**a) POI matching uncertainty.** The POI attributes are analyzed by place labels, which match users’ preferred attributes to obtain personalized recommendations, and geographic dynamic characteristic predictions are the basic idea of machine learning. However, users’ access records or addresses shared on social media are inaccurate or even unmatched by the actual POI location. Some texts from social-media data are removed as noise because they do not mention POI attributes [5]. Therefore, the extraction of accurate POI categories which match social-media data to POI features is a challenge [6].

**b) Map zooming dynamicity.** Figure 1 introduces the tweets (red spots) and POIs (black spots) from the Kyoto City Road map with LOD processing. We extracted the dynamic POI categories (ex: restaurants, shops, and transportation) in three map scales. The problem is that each level of the map zooming task requires a set of a priori POI data to summarize the user-visit training [7]. Particularly, the results of the supervised training model must regress the zooming feature and discuss the user-preferred POI of each map tile.

**c) Attention mechanism complicacy.** For the graph structure of graph neural network (GNN), the self-attention mechanism aggregates the neighbor nodes, which realizes adaptive matching of the weights of different neighbors, thereby improving the accuracy of the model [8]. The aggregation process works on a map with a POI recommendation [9]. However, in this study, POIs contained multiple categories from tweets information, which required multilayer attention for POI attributes to match the raster units of each map tile.

In this study, we propose a spatial-zoom graph-attention model (SZ-GAT) that can efficiently extract POI geolocation features from social-media data and focus on users’ favorite POIs of each map zooming levels. We extract dense POI cores as 160 map tiles from the Google Map API, including POI attributes, POI coordinates, and POI raster unit information. Tweets with geolocation records are mapped to each POI in the rasterized-unit-map tile. The proposed training model verifies the effectiveness of this rasterized statistical method with GNN, that works on map zooming [10]. The training is completed by 4,716,335 placemarks with five categories on 100 map tiles, and the remaining 1,834,164 marks on 60 map tiles are used to evaluate the model performance. The proposed method promotes POI visualization at different zoom levels and improves the aggregation efficiency of geographic records in zoom dimensions. We propose that the raster distribution be used to predict POI features with user-access records and display users’ favorite POI on different map scales [11]. The main contributions of this paper are as follows,

1) Tile-feature annotation on the map areas containing POI data is completed.
2) Rasterization statistics for map tiles, which is a significant contribution to POI data analysis is completed.
3) We propose a graph neural network model for map zooming and placemark training, including two train-
II. RELATED WORK

Map zooming with POIs priority visualization focus on the user’s prefer attributes of POI. The POI contains various user information and place characteristics [12], which are two research points: (a) from users to placemarks, and (b) from placemarks to users. A study [13] introduced geolocation prediction (GP) and reviewed the two research points as user-profiling-based and content-based geolocation prediction.

(a). Users to placemarks represent a series of studies based on the text-feature-analysis method containing user location information and extended to placemark attribute analysis on the map [14] [15]. The authors demonstrated the visualization of user-shared placemark map tiles for application to natural events and human-change-predictive tasks [16]. This method is used to construct recommendation algorithms for POIs by extracting user characteristics and inferring user preferences based on object attributes. A previous study on spatio-temporal neural networks achieved excellent results on raster map tiles with user location-related geographic information [17] [18]. The abovementioned studies demonstrate the feasibility of the location information shared by users, and the related geographic placemark attributes are learnable [6].

(b). Placemarks to users utilize the placemark-feature distribution on a map to describe fixed-region attributes. Social-media data can be used to verify the prediction accuracy with different dimensions in a map tile [5]. For a certain map tile, the included multiple POI features depend on the observed attributes and the location features of adjacent POI data, thereby inferring the unknown attributes of a geographic location [19]. Some difficulties are associated with the quality of POI annotations and the relevance of social-media data to geographic locations [20].

For the related works of 2D or 3D flat maps, the spatial problems of fixed distance or fixed position between POIs are always discussed on a basic map with a fixed scale and zoom level. Such as the efficient destination retrieval for a query to POIs [21], incremental POI nodes mapping between large graph structures [9], and traffic flow prediction [22] are all adopted spatial-temporal frameworks with a learnable positional attention mechanism. However, the geographic information in these studies is aggregated from a priori maps with fixed zoom levels. The map zoom level as a controlled variable is currently only introduced in the geographic information systems field. On the one hand, the POI data generalization technique changes the degree of abstraction of map content. Content zooming provides the user with the capability to change the amount and the granularity of foreground information presented while keeping the geometric map scale the same [23]. Content zooming allows overriding the effects of standard map generalization, focusing on optimized content representation to aid the information-seeking task of a mobile user. On the other hand, a matching spatial data at different map scales for integrating, evaluating, and updating spatial data collected and maintained at various scales. They proposed a relaxation labeling technique to match the areal features and/or for a different scale range with a contextual approach [24]. More related to our research is a web map service platform development to browse the behavior modeling and brows interest extraction with multiple map scales [25]. They utilize the user’s zoom-in map operations to indicate interest increasing. In contrast, the zoom-out indicates interest decreasing and proposed a new hierarchical Gaussian mixture model to model the multi-granular spatial structure and extract browsing interests. For our research, the map zooming process is used to reflect the POI visit history with social media and user preference on the raster maps. This is the first time map zoom level has been applied as a map dimension to a graph structure framework.
For POI filtering, we focus on recommending a ranked set of POIs for a user where the POI after map zooming should be the user highly preferred. A model proposed to learn content-aware POI embeddings through user visit sequences and POI textual information [26]. A category-aware gated recurrent unit model is proposed to mitigate the negative impact of sparse check-in data, capture long-range dependence between user check-ins and get better recommendation results of POI category [27]. An interactive multi-task learning framework is proposed to exploit the interplay between activity and location preference with temporal-aware activity encoder [28]. In our previous research, we constructed the POI data set with text categories for regional competitiveness training, which contain the POI data with social media records and user preference distributions [29].

The proposed model exhibits several significant differences. First, we focus on learnable POI graph representations through a graph attention network (GAT) and zooming process analysis. The dynamic-location-space changes in the POI attributes are described by the map-zooming process. Therefore, the proposed method exhibits differences in the optimization goals of the algorithm and zooming dimensions. Second, the graph structure in the proposed method was improved for spatial zooming. Third, the learnable POIs in the zooming-evaluation process was imported from a different resource, and the predictive results of the proposed model effectively described the hiding and appearance of POIs during map zooming.

III. DATA COLLECTION
In this study, the collected data included location information and map-tile data. Location information is a set of POIs with geographical information used to describe the dynamic changes in the map-zooming process. The collection of POIs is tiered according to the zoom levels shown in Figure 2, which all the POIs are displayed on the base map in zoom level 0. The map-tile data are a set of vector maps containing dense POI cores from different map platforms, as shown in Figure 1. The map features were collected from the Google Map API as follows:

1. Base map: from Google Road Map in the area, used to visualize the division of map tiles;
2. POIs: from OpenStreetMap, utilized to determine the scope of the basic POI data set in the area;
3. Placemarks: from Google-map-location reviews and Foursquare, used to count the popularity and user behavior of each placemark;
4. Tweets with geolocation: used to evaluate the user visits and preferences in the map-zooming process of the proposed model.

The dense POI cores are divided into 160 maps using roads for rasterized map feature extraction shown in Figure 2. In order to classify each POI as a trainable POI attribute more efficiently on the base map, we only focus on 5 kinds of POI categories. In Figure 3, the rich POIs and tweets with geolocation are extracted on the Google Base Map, where all the POIs and tweets are divided into 5-dimensional categories to provide the trainable POI attribute vectors.

IV. RASTER PROCESS & FRAMEWORK OVERVIEW
A. RASTER MAP TILE
In 2005, map tile originated from a technology called “slippy map”, which makes maps more accessible and load more quickly [30]. Since 2005, the map was no longer a single image; however, several images were joined together seamlessly. Interaction with the map only requires the necessary sections of the map to be loaded and displayed, and not the entire image [31]. The images were called raster tiles, and all major map applications and software libraries used Google’s method [32].

In this section, the raster tiles divide the target map area into grids and construct a set of square tiles. As shown in Figure 2, we introduced a raster tile on the base map. We divided the base map into a set of square maps using the

![Google Base Map](image1)

![POI Category](image2)

**FIGURE 3.** An example of POIs and tweets with geolocation on the Google Base Map, where the POIs and tweets are divided into 5-dimensional categories to provide the trainable POI attribute vectors.
Mercator projection system. Each grid on the map tile is called a raster unit, and all units share a data-storage scheme on a map tile. The schemes are described as having the same parameters, such as the grid pixel size, tile shape and size, coordinate system origin, tile matrix size, and the number of zoom levels. The raster map tiles obey the Web Map Tile Service (WMTS) standard definition from the Open Geospatial Consortium (OGC).

The POIs are represented with LOD through data generalization, and the dynamic POI display on different map scales is called map zooming. Typically, the definition of zoom level is described by geographic coordinates. However, for the base map at zoom level 0, the difference in geographic coordinates between the two POIs is a small change. In Figure 2, POIs are defined as different colors with different zoom levels, where the difference of POI$_1$ and POI$_2$ coordinates distance is $(0.00000312886, 0)$. This difference is difficult to describe the relationship between POIs on the base map. Therefore, we utilize raster units to represent the spatial information of each POI. Since the map zooming operation is based on the map zooming rule of nine raster units fused into one raster unit to become a new map in our study, the POI graph structure is a fully connected relationship with other POIs which are in the nine neighbor raster units. This adjacent POI graph structure definition constructs the zoom feature with a map zooming process and raster zooming visualization. An example of POI$_1$ introduced the graph structure construction in which POIs composed the node set and adjacent unit connections are the edge set. And in zoom-out detail, the POI graph in a new raster unit with red mark $(2, 2, 1)$ is zoomed from the raster unit $(2, 8, 1)$ on a base map. In this study, each raster-map-tile extraction contained the details of the POI attributes and zoom levels as the data input. We will introduce the details of the raster units in the subsequent section.

**B. ZOOM DIMENSION**

For an ordinary 2D flat web map, the location attribute of the POI is described by the latitude and longitude, or the plane distance on the X-Y axis. The 3D map adds a height dimension to flat maps to make the web map look more three-dimensional. Previous studies introduced the details of basic rules, definitions, and general scales of map zooming on different platforms of the world map. In this study, we define the POI position of each map tile on the X-Y-Z axis. Unlike the definition of a 3D map, the third dimension in this study was composed of a zoom level, and the flat map was composed of the X-Y and zoom dimensions: Z. In this study, all mathematical symbols related to the zoom levels are denoted as feature z and set Z.

In Figure 1, we introduce the zoom level, which is composed of four levels: base map ($z = 0$) and zoomed-out results $z = 1, z = 2, z = 3$. All POI positions are stored on the base map ($z = 0$), and $z = 3$ map tiles generally contain the maximum map scale and least POIs. The zoom level is imported as one of the POI attributes in the POI raster-unit database, as shown in Figure 2. Figure 4 shows the details of the map-zooming-dimension analysis in this study.

**C. FRAMEWORK OVERVIEW**

The overview of ST-GAT framework is introduced in Figure 4. It consists of three components include data preprocessing, map tile graph attention network block, and output evaluation block.

**FIGURE 4.** The overview of the SZ-GAT framework contains Data Preprocessing Block, Map Tile Graph Attention Network Block, and Output Evaluation Block.
block, map tile graph attention network block, and the output evaluation block. The POIs and tweets geolocation on the Google base map is divided by a set of given POI categories. The 5-dimensional POI categories construct the input POI attributes \( P_A \) as a set of POI vectors \( \vec{P} \) with coordinates \( P_C \) and raster unit information \( P_U \). Then, the POI graph structure contains spatial feature \( P_A \), \( P_C \) and zoom feature \( P_A \), \( P_U \) are embedded into the SZ-GAT training model. The global attention mechanism generates the rasterized geographical awareness with POI spatial information, and user preference zooming aware representation embeddings based on the raster unit attributes and the user’s prefer POIs denote on multiple zoom levels, respectively. After that, each raster unit fuses its corresponding zoom levels \( f_{z_1}, f_{z_2}, f_{z_3} \) with user preference representation POI embeddings. Finally, the output evaluation block computes the matching scores between users’ favorite places and zoomed POIs on different web map platforms including Google Map, Bing Map, Yahoo! Map, and Open Street Map.

V. APPROACH

In this section, we present the spatial-zoom graph-attention network (SZ-GAT). Specifically, we first present the POI data-preprocessing analysis using the map raster statistics method in Figure 2. Subsequently, we discuss the POI place-mark attributes and map the zooming process of each map tile in Figure 4. Finally, the details of the SZ-GAT are presented in Figure 6.

A. RASTERIZATION STATISTICS & GRAPH STRUCTURE

In this study, the POI category of tweets were mapped to raster units and they contributed to learnable graph structures. We introduce a set of graph structure slices \( G_M \) on the rasterized map tiles \( M \), in which POIs comprise the node set and the edge POIs are fully connected with the neighboring grids as the input, as shown in Figure 5.

The process of data preprocessing is introduced in Figure 4. Each user’s post-tweet is mapped to the related POI category, and combined with the POI information, the POI attribute \( P_A \) is defined as a 5-dimensional vector containing the features of restaurants, hotels, transportation, shops, and travel attractions. The output POI vector \( P \) is denoted by POI attributes \( P_A \), POI position \( P_C \), and raster unit information \( P_U \). \( P_C \) contains the geographical location (latitude and longitude) of each POI and user-visit history. \( P_U (x, y, z, d, n) \) is collected to describe the POI raster-unit information, where \( (x, y) \) is the position of the POI in the unit, \( z \) is the zoom level from 0 to 3 (example, \( z = 0 \): base map, \( z = 1, 2, 3 \): zoom level 1, 2, 3), \( d \) is the distance of the nearest-neighbor POI in the same unit, and \( n \) represents the POI dynamic position grid in the map tile (example, base map tile \( z = 0 \) has 81 units, and \( POI_1 \) position is on the 38th grid).

The POI constructed graph of each raster unit \( G = P, E \) comprises a placemark set \( P \) and a link set \( E \) within the grid. The vector of POIs is imported into the raster units, and the POIs are fully connected with neighboring grids to compose the graph structure \( G_P \). For the rasterization map tile shown in Figure 5, \( M \) is denoted by a set of basic raster units \{ \( M_{U1}, M_{U2}, M_{U3} \), and ..., \( M_{U_n} \) \}. We introduced a map tile, which is composed of a set of basic raster units in 50 m side square. Raster units are the basic POI statistics grids in a map tile, and these units help complete the translation of the graph structure and map zooming.

In the data preprocessing, we analyze the tweets with 5 POI categories and mapped them to the POI in each raster unit, where the 5 POI categories are denoted by \( A (A = A_1, A_2, A_3, A_4, A_5) \). The POI category of each user post-tweet \( Tweet(P) \) is

\[
Tweet(P) = \sum_{i=0}^{i} \hat{A} * i(positive) \tag{1}
\]

where \( i \) is the number of the tweet and only "positive" tweets matching the POI category is credited to the corresponding POI attribute. Each tweet contributes POI category \( A \) as an augmented vector shown in Figure 3. Combined with the POI and Tweet geolocation, each POI-related positive tweet position is denoted in each raster unit \( R_{P_A} \)

\[
R_{P_A} = \sum_{n} Tweet(P) * n(Lat, Lng) \tag{2}
\]

where \( n \) is the raster unit number of each tweets position, \( (Lat, Lng) \) is the coordinate of tweets. After importing the social-media data into the map tile, we construct the POI graph structure \( G_M \) on the map tile, as shown in Figure 5.

B. GLOBAL ATTENTION MECHANISM

We constructed a fully connected POI graph structure on map tiles \( G_P = P, E \). In this subsection, we introduce the global attention mechanism to match the multilayer attention to the POI and focus on the 5-category attributes in the map-zooming process. The graph attention network (GAT) [9] introduces a self-attention mechanism to learn the user-preferred visits of \( P_{A_i}, P \in P \cup Tweet(P) \) to every POI node \( P \). Each graph-attention-aggregation module updates the hidden \( H \) representations and aggregates information from every POI neighbor in proportion to the learnable attention coefficients. The aggregation operation of the \( l \)-th graph attention for placemark \( P \) is

\[
\bar{h}^P_l = \sigma \left( \sum_{i,j \in \{P_A\}} \alpha_{ij} h^{l-1}_{P_{ij}} W \right) \tag{3}
\]

where \( \bar{h}^P_l \) is the POI feature embedding, \( \sigma \) is an activation function (ReLU), \( W \) is the trainable parameter matrix of the \( l \)-th layer, and \( \alpha_{ij} \) denotes the attention coefficient between the placemarks \( P_i \) and \( P_j \), which obey the following:

\[
\alpha_{ij} = \text{softmax}_j \left( \text{attn} \left( \bar{h}^P_l, W, \bar{h}^P_l, W \right) \right) \tag{4}
\]
Data Input

| POI Set | POI Attribute | POI Coordinate | POI Raster Unit |
|---------|--------------|---------------|-----------------|
| POI_1  | 0 0 3 0 0    | lat_1 lon_1  | 8.15 7.68 0 0   |
| POI_2  | 0 0 0 0 5    | lat_2 lon_2  | 3.64 4.05 1 2.33|
| POI_3  | 2 0 0 0 0    | lat_3 lon_3  | 3.61 2.08 1 0   |
| POI_4  | 0 3 0 0 0    | lat_4 lon_4  | 2.73 6.72 3 0   |
| POI_n  | ...          | ...           | ...             |

POIs in Raster Unit

Rasterized POI Base Map

Map Rasterization

Map Tile Graph

FIGURE 5. The detail of input POI data set (Attributes, Coordinates, and Raster Unit Information), where A_1 to A_5 donate the POI location attributes in the category of Restaurant, Hotel, Transportation, Attraction, and Shop. lat, lng donates latitude and longitude of the POIs. The map tile base features are donated by each raster unit POI positions: flat position (x, y), zoom level: z, nearest POI distance d, and the number of the raster unit n in tiles.

where $attn$ is the attention mechanism, and $softmax$ is the regularized function of the selected node $P_i$ neighbors.

The user visits POI feature $h_P$ in each raster unit, which is simply described as follows:

$$h_P = \sum_{P_A} attn(Tweet(P_i))W^{(0)}$$

(5)

In the Google base map of zoom level 0 in our research, all POI placemarks with latitude and longitude were collected. Each map tile is divided into a set of $50 \times 50$ m grids as raster units. In Figure 5, we show a $9 \times 9$ raster base map $M$. The grid containing specific POI vectors $h_P$ denotes the raster unit features $h_R$, and each POI position with the X-Y axis, zoom levels, the distance with the nearest POI neighbor in the same unit, and the number of units are recorded as a set of vectors $h_{P_U}$:

$$h_R = \sigma \left( \sum_{U_{P_i}} attn(h_{P_U})W^{(0)} \right)$$

(6)

Same to the user visits POI feature $h_P$, each raster unit vector $h_{P_U}$ contains a base map attention mechanism "attn" and the weight of POI attributes $W^{(0)}$ are shared in the same layers.

Tile feature $F$ contains the POI attributes $h_{P_A}$, raster unit $h_{P_U}$ shown in Figure 6 as follows:

$$F = \sigma \left( \sum_{j=1,2,3} attn(h_{P_U}, h_{P_A})W^{(0)} \right)$$

(7)

where $f_{z(1,2,3)}$ denotes the POI prediction result features contained in each zoom level, respectively. The output of POI features represent the user’s preferences in each map tile.

C. MAP ZOOMING PROCESS

Zoom Dimension is the visualization level of electronic maps in map tiles of different scales. In our work, "zooming" is exploited as a new flat map dimension, so that each POI can be drawn on the map tiles with an X-Y-Z axis. We aim to sequentially export five categories of map tile features as the output of map zooming: zoom levels 1, 2, and 3. To describe the POI attributes contributed by 5-category tweet records in raster units. POIs are mapped to the linear zoom levels:

$$Z_{0,1,2,3} = \exp(logZ),$$

where logZ is tiles linearly interpolate in logarithmic domain as follows:

$$logZ = Z_0 + (Z_1 - Z_3) \times s/(s_1 - 1)$$

(8)

where $s$ is the zoom step, $s_1$ is the step of the first zoom level, and $logZ$ donates the linearly interpolate between map zoom level 3 and level 1, $Z_0$ is base map scales. We extract $P_A$ in spatial attention concatenate as a POI 5-dimensional matrix $R_A$, and map feature $h_M$ with zoom 3 levels as a raster unit convolution process:

$$h_{M+1} = \sigma \left( \tilde{D}^{-k} \hat{A} \tilde{D}^{-k} h_M W^{(U)} \right)$$

(9)

where $\hat{A} = A + R^{k=1,2,3}_A$, $\tilde{D}^{-k} \hat{A} \tilde{D}^{-k}$ is the normalized adjacency matrix of $\hat{G}_M(p_1, p_2)$ in Figure 6, $\tilde{D}$ is the degree matrix of $\hat{A}$, and $W^{(U)}$ denotes the learnable parameter of map zooming unit levels.
Map tiles are zoomed as a new tile with next step of zoom dimension such as the red rectangle shown in Figure 2 comprises the units:

\[
\begin{pmatrix}
(1, 7) & (2, 7) & (3, 7) \\
(1, 8) & (2, 8) & (3, 8) \\
(1, 9) & (2, 9) & (3, 9)
\end{pmatrix}
\]  

The zoom level \( z = 0 \) was ignored in the base map, and in the next stage, it was added to \((1, 7, 0)\). The position of \( POI_1 \) is \((2, 8, 0)\) in the zoom level 0 unit and \((2, 2, 1)\) in the zoom level 1 unit. To describe the unit features concisely, the position of each unit is rewritten as the count of the raster unit on the map tile. For example, unit \((2, 2, 1)\) is the 5th unit, and it is recorded as \( n = 5 \) in the \( 3 \times 3 \) tile with zoom level 1 introduced in Figure 5. After the map tile zooms-out to units \((2, 2, 1)\), \( POI_3 \), \( POI_4 \), and \( POI_5 \) disappear with map zooming, and the learnable parameter \( W^{(U)} \) trains the POI features during the dynamic zooming of the raster unit.

VI. EVALUATION

In this section, we evaluate the performance of the proposed SZ-GAT training model using real-world POI datasets collected from Kyoto City as the area. We compared the performance of the SZ-GAT with that of mainstream baselines, including mathematical statistics and deep neural network models. Additionally, we introduced a set of experiments for POI map tile training and trainable model prediction. The training process fits the POI marks displayed at different zoom levels and learns the regional features in the map tile from the basic units.

The training process includes: First, the POI displayed on 160 Google-map tiles was used as the learnable benchmark. The POI attributes, coordinates, and raster unit information are shown in Figure 5. Second, the POI sets are trained by different models include the variation algorithm of convolutional neural networks (CNN), graph convolutional neural networks (GCN), graph attention networks (GAT). Third, the existing Google-map-zooming results were evaluated using platforms, such as Bing Maps, Yahoo! Maps, and Open Street Maps (OSM) in Figure 4. We introduce the accuracy of the SZ-GAT prediction model on the Google base map and compare it with the remaining three maps.

A. DATASET

We used tweets and user-visit-history datasets on the map tiles, which included three domains: spatial locations, visit attributes and zoom dimensions. The real-world POI attributes \( P_A \) include five categories: restaurants, shops, transportation, hotels, and attractions, as shown in Figure 5, as \( A_1, A_2, A_3, A_4, A_5 \). The domains of the labeled POI include the following:

- **Spatial Location**: learnable POI-placemark coordinates on a base map using the Google Map API: \( P_C \) as shown in Figure 5.
- **Visit Attributes**: the visit attributes are contributed by the social-media data of each POI. Google Maps was utilized to enhance the POI-visit attributes. Specifically, the tweet text usually provides unclear visit records, such as "Pudding" without geolocation. Therefore, we used Google-map review ("This restaurant sells pudding and sweet drinks.") to collect related POI-restaurant attributes.
- **Zoom Dimensions**: the raster-unit set \( P_U \) is collected by grid number \( n \) and zoom level \( z \). The web-map-zooming dimension is visualized by three levels: \( POI_1 : z = 1, 2, 3 \) is \( POI_1 \), shown on all three zoom-level maps in Figure 6. A threshold is given by a map zooming regular,
in which the same attribute POIs in the same raster units compete on map zoom-out processing.

The dataset was divided into training, validation, and test sets. We collected 7,535,137 short texts with geolocation, of which 6,550,499 were Twitter data and 984,638 were from Google-Map reviews. Social-media data were mapped to the POI of each raster unit on the map tiles; 4,716,335 tweet records were utilized for a 100-tile training set, and 1,834,164 POI markers were recorded in a 60-tile test set. The remaining review set was used as the validation set.

B. BASELINES

Our proposed method does not require completing the process of sampling the building in the fitting block with the sub-features, because a set of priority map tile data sets with POI spatial information is utilized to construct the graph structures. In the training groups, we considered the geographical distribution of tweet features on Google base maps, and divided the dense POI core with roads as a set of trainable map tiles. We compare our SZ-GAT model with the baseline models including 3-dimensional dynamic graph convolutional network (3D-GCN), adaptive sequence partitioner with power-law attention (ASPPA), graph-based geographical latent representation model (GGLR), and spatial-temporal-priority-user dimensional graph attention network (STP-UDGAT).

- **3D-GCN** [33]: An GCN variant that model improved from DSTG and consider the POI visit preference on the 3D maps. The semi-supervised node-partition strategy on GCN provides the cross-pixel correlations.

- **ASPPA** [34]: ASPPA contains a stacked recurrent neural network framework Adaptive sequence Partitioner to identify the sequential patterns among semantic subsequences by automatically learning the latent structure in the user’s check-in sequence.

- **GGLR** [35]: This is a graph-based POI recommendation model and considers the directed POI-POI graph constructed from check-in histories.

- **STP-UDGAT** [36]: This method focuses on the next POI recommendation task and the inclusion of user embeddings global STP factors across all users to improve the recommendation task and user experience.

The given POI test sample, for accuracy, if the shown POI is within the map zoom level \( D \) of the Google Base Map scale set, then a positive of 1 is awarded, else 0. And we adopt the root mean squared error (RMSE) and mean absolute percentage error (MAPE) to evaluate the baseline model and SZ-GAT.

\[
RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (a_i - f_i)^2} \quad (11)
\]

\[
MAPE = \frac{100}{n} \sum_{i=1}^{n} \left| \frac{a_i - f_i}{a_i} \right| \quad (12)
\]

A MAPE value of 0% indicates the best model and a MAPE value greater than 100% indicates an inferior model. The performance score is the recall of the test model POI attribute prediction as follows:

\[
Score = \frac{TP_{\text{tweets}}}{TP_{\text{tweets}} + FN_{\text{reviews}}} \times 100 \quad (13)
\]

where \( TP_{\text{tweets}} \) is the positive tweet prediction and \( FN_{\text{reviews}} \) is the negative prediction POI from Google-map reviews.

C. EXPERIMENTAL SETTINGS

We found that the visual information of the POI with the map zooming process could reflect the interest of the POI itself. This POI recommended method focuses on the dynamic changing with map zooming, just like the majority of user interest placemarks are displayed with map zoom-in. We utilize the zoom level of map tiles to discuss the related user interest features of POI, in order to realize the map zooming visual function based on the location information. The POI features were imported to different platforms including Google Map, OpenStreetMap, Bing Map, and Yahoo! Map. The interest positions need to contain as few review redundant details as possible in order to reduce the cognitive load of processing raster units.

The SZ-GAT model was implemented using the PyTorch geometric platform. All POI positions were mapped to the 50 \( \times \) 50 m grids of 160 map tiles by latitude and longitude. The zoom dimension of all map tiles is four levels (\( z = 0 \) for the base map, \( z = 1, 2, 3 \) for map zooming), the input POIs are 12-dimensional vectors, and the representation vectors are fixed to 108 dimensions. The cross-entropy loss utilized a batch size of one and ran the experiments for 100 epochs. We set the regularization coefficient to 0.001 and the initial learning rate to 0.001, followed by a decay of 0.0001 at the 10th epoch. The slope in the LeakyRelu activation function was 0.2, weight coefficient was 0.4. We employed a fully connected neural network for POI-feature extraction from raster units, and the number of neurons in each layer corresponds to map-zooming levels of 225, 144, and 81. The threshold of the user favoring the POI number of each category with map zooming was set to 5.

D. EVALUATION RESULTS

We randomly selected 60 map tiles from the test set divided into map tile sets A (30 tiles) and B (60 tiles), which contained five categories of user-visit POI attributes. We generated user-preferred POIs using the baseline method and SZ-GAT for three map zooming. The results are summarized...
in three tables. Table 1 introduces the prediction performance of each model, Table 2 shows the user-preference scores, and Table 3 shows the prediction accuracy for different web-map platforms.

Table 1 lists the performance of the 3D-GCN, ASPPA, GGLR, STP-UDGAT, and SZ-GAT models. We compared the POI-processing ability of regression data of four models and evaluated them by RMSE and MAPE. The proposed method, SZ-GAT, outperformed the baseline models in user-vist visit and prediction accuracy at three zoom levels. Specifically, the average prediction accuracy of the raster unit in the 5-category attributes in set A improved from 44.87% of ASPPA to 58.62% of SZ-GAT. The RMSE performance is improved from 27.96% of ASPPA to 20.36% of SZ-GAT in test set A zoom level 3. SZ-GAT achieved the best performance of 59.04% on the third zoom level of the Kyoto map. For set B, the results of the four training models were consistent with those of set A, where SZ-GAT obtained the best results from the average value. The RMSE and MAPE value show that the SZ-GAT has a good POI regression ability. Additionally, we will explore the dynamic changing in 5-categories POI features in future work.

In Table 2, compared with the user-preferred POI database from Google-Map reviews, we list the POI-performance scores of the test models that work on map tile zoom-out. The users’ preference score improved from 64.37% to 85.32% for ASPPA and SZ-GAT in set A zoom-out map level 2. Compared with set A, the performance score reduced to 82.10% in set B zoom level 2 and obtained the best performance score of 83.61% in zoom level 3.

The results in Table 3 show the prediction accuracy of the POI mapping of the SZ-GAT model from the Google base map to OSM, Bing Map, and Yahoo! Map. Compared with the Google base map, the performance of OSM contained the best result of 58.57% and 58.40% in zoom level 1. In the second zoom level, Bing Map obtains the best performance of 57.98% and 58.98% in the test set A and B. For zoom level 3, Yahoo! Map has the best result of 58.64% in test set A, and Bing Map obtains the best performance of 58.37% in test set B.

A case visualization of SZ-GAT prediction results is based on a Google Map tile and evaluated with the Open Street Maps, Bing Maps, and Yahoo! Maps shown in Figure 7. For map zooming visualization, only POIs visited by more than 50 tweets are displayed on this map tile. POIs are divided into positive prediction results (orange dots), negative prediction results (green dots), and lost POI prediction results (blue dots). The prediction accuracy of the SZ-GAT training model is computed in three zoom levels of four web map platforms. Our proposed SZ-GAT model is learned from the zooming rules of Google Maps. Therefore, compare with the results of the other three web map platforms in Table 3, the performance of Google Map contains the best prediction results and has fewer losses of POIs after map zooming. Compare with the performance of the SZ-GAT model on the test set, the case of the map tile shows better performance at all zoom levels. The predicted results of this map tile case are better than the average of the test set. Discuss from the POI data composition of the case map tile, which aggregates a large number of business centers and tourist attractions, and is one of the most tweets accessed areas in our tile set. We have reason to believe that the performance of this case is a positive prediction and meets our training model expectations.

VII. DISCUSSION

In this study, we discovered that the visual information of the POI with the map-zooming process could reflect user preferences. This method focuses on the dynamic POI change with map zooming, similar to users’ favorite POIs that are always displayed preferentially while the map is zoomed in. We utilized the zoom level of map tiles to discuss the rela-
FIGURE 7. A case visualization of SZ-GAT prediction results combined with multiple web map zoom levels, where the orange marks are the positive POIs prediction of the SZ-GAT model, the green marks are the negative POIs prediction results, and the blue marks are the prediction results with lost POIs. The accuracy of the prediction results based on the Google Map test set is evaluated by the other web map platforms.

The evaluation results show that the prediction accuracy of test set B is slightly better than that of set A, as shown in Table 1. We examined the tweet attributes as test sets in different data sizes found that the bigger the data size, the greater dispersion of the prediction result and the actual user-preferred POI data. We compared the category composition of the two test sets and found that Set A has numerous "Transportation" and "Restaurant," while Set B has 30% of "Attraction" in the POI-dense cores. Combined with the user-visit frequency, most of the attractions in Kyoto are temples with large areas. Therefore, the higher the frequency of visits to attractions, the more they can be expressed as users' favorite POI during the map-zooming process. We believe that this is related to the user-preference map tiles with different urban facility functionality characteristics and that frequent user access is more representative of the attributes of this map tile.

In addition, the POI features were imported into different platforms, including Google Maps, OpenStreetMap, Bing Map, and Yahoo! Map. The results in Table 3 introduce the four web-map platform prediction accuracy and show that there are only minor differences. The purpose of SZ-GAT training with POIs shown by the multi-platform web-map results is to adopt the proposed model to different map-platform zoom rules. The results of map-zooming levels 1, 2, and 3 are trained on the base map to avoid iterative errors. From the results of zoom level 3, because the POI proportion displayed on the tile is higher than the remaining levels, the training requires minimal cost and resources. In future research, we will consider the relationship between the POI display proportion and scaling results.

VIII. CONCLUSION

In this study, we proposed and developed a spatial-zoom graph-attention model that mapped tweet preference to POI position features and retained map-tile-zooming characteristics using the raster statistics method. We applied the user-visit-history data from Twitter and Google Maps to enhance the POI attributes of each raster unit. We completed the training of the proposed SZ-GAT model and evaluated it using other baseline methods. In the 160 pieces of a priori dense map tiles, we completed a set of user-access POI-zooming experiments by analyzing 5-category POI attributes in the urban facilities.

Compared with the baselines, the advantages of the proposed SZ-GAT model are as follows:
1) The SZ-GAT model has several advantages in terms of user-preference regression with tile zooming. The raster unit random sampling results show that the SZ-GAT model obtained better performance for tweet-attribute mapping.
2) The SZ-GAT model connected the POI features as nodes corresponding to all raster unit attributes and had a stronger ability to learn a priori map tile data.
3) The global attention mechanism focused on the given 5-category tweet attributes, POI attributes, raster unit
4) From the different test sizes, the performance of the SZ-GAT model was more stable, and as the test set increased, the prediction convergence of the SZ-GAT was better than that of the remaining baselines.

The experiments show that the SZ-GAT model achieved better performance than the baseline models. We confirmed our result can match the users’ preference related map tiles with different urban facility functionality characteristics. The map zooming dimension is a new field of map visualization and POI-data analysis. In future work, we plan to apply the SZ-GAT model to dynamic social data changing under the influence of social events.

REFERENCES

[1] Shuangli Li, Jingbo Zhou, Tong Xu, Hao Liu, Xinjiang Lu, and Hui Xiong. Competitive analysis for points of interest. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, pages 1265–1274, 2020.

[2] Yanhong Wang, Teng Wang, Ming-Hsiang Tsou, Hao Li, Wei Jiang, and Feng Qin Guo. Mapping dynamic urban land use patterns with crowd-sourced geo-tagged social media (sina-weibo) and commercial points of interest collections in beijing. china. Sustainability, 8(11):1202, 2016.

[3] JLP Barker and Christopher JAC Macleod. Development of a national-scale real-time twitter data mining pipeline for social geodata on the potential impacts of flooding on communities. Environmental modelling & software, 115:213–227, 2019.

[4] Tae Wook Kang and Chang Hee Hong. lcif-citygml lod mapping automation using multiproccessing-based screen-buffer scanning including mapping rule. KSCE Journal of Civil Engineering, 22(2):373–383, 2018.

[5] Xiongfei Yan, Tinghua Ai, Min Yang, and Hongmei Yin. A graph convolutional neural network for classification of building patterns using spatial vector data. ISPRS Journal of photogrammetry and remote sensing, 150:259–273, 2019.

[6] Jiayue Yuan, Pranab K Roy Chowdhury, Jacob McKee, Husien Lexie Yang, Jeannette Weaver, and Budhendra Bhat. Exploiting deep learning and volunteered geographic information for mapping buildings in kano, nigeria. Scientific data, 5:180217, 2018.

[7] Ting Zhong, Tianliang Wang, Fan Zhou, Goce Trajcevski, Kunpeng Zhang, and Yi Yang. Interpreting Twitter user geolocation. In Proceedings of the 58th Annual Meeting of the Association for Computational Linguistics, pages 853–859, Online, July 2020. Association for Computational Linguistics.

[8] Petar Velicković, Guillem Cucurull, Arantxa Casanova, Adriana Romero, Pietro Lio, and Yoshua Bengio. Graph attention networks. arXiv preprint arXiv:1710.10903, 2017.

[9] Xinyan Xia, Jun Gao, and Bin Cui. imap: Incremental node mapping between large graphs using gnn. In Proceedings of the 30th ACM International Conference on Information & Knowledge Management, pages 2191–2200, 2021.

[10] Weijie Shi and Raj Rajkumar. Point-gnn: Graph neural network for 3d object detection in a point cloud. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pages 1711–1719, 2020.

[11] Jon Froehlich and John Krumm. Route prediction from trip observations. Technical report, SAE Technical Paper, 2008.

[12] Xinyan Zhu and Chunhui Zhou. Poi inquiries and data update based on lbs. In 2009 International Symposium on Information Engineering and Electronic Commerce, pages 730–734. IEEE, 2009.

[13] Muhammad Nur Yasir Utomo, Teguh Bharata Adji, and Igi Ardiyanto. Geolocation prediction in social media data using text analysis: A review. In 2018 International Conference on Information and Communications Technology (ICOACT), pages 84–89. IEEE, 2018.

[14] Michael Crutcher and Matthew Zook. Placemarks and waterlines: Racialized cyberscapes in post-katrina google earth. Geoform, 40(4):523–534, 2012.

[15] Mark Graham and Matthew Zook. Visualizing global cyberscapes: Mapping user-generated placemarks. Journal of Urban Technology, 18(1):115–132, 2011.

[16] Hongzhi Yin, Xiaofang Zhou, Bin Cui, Hao Wang, Kai Zheng, and Quoc Viet Hung Nguyen. Adapting to user interest drift for poi recommendation. IEEE Transactions on Knowledge and Data Engineering, 28(10):2566–2581, 2016.

[17] Jean-Benoît Griesner, Tалел Abdessalem, and Hubert Naacke. Poi recommendation: Towards fused matrix factorization with geographical and temporal influences. In Proceedings of the 9th ACM Conference on Recommender Systems, pages 301–304, 2015.

[18] Wei Liu, Hanjiang Lai, Jing Wang, Geyang Ke, Weirei Yang, and Jian Yin. Mix geographical information into local collaborative ranking for poi recommendation. World Wide Web, 23(1):131–152, 2020.

[19] Di Zhu, Fan Zhang, Shengyin Wang, Yaoi Wang, Ximeng Cheng, Zhou Huang, and Yu Liu. Understanding place characteristics in geographic contexts through graph convolutional neural networks. Annals of the American Association of Geographers, 110(2):408–420, 2020.

[20] Renato Andrade, Ana Alves, and Carlos Bento. Poi mining for land use classification: A case study. ISPRS International Journal of Geo-Information, 9(9):493, 2020.

[21] Zixuan Yuan, Hao Liu, Yanchi Liu, Denghui Zhang, Fei Yi, Nengjun Zhu, and Hui Xiong. Spatio-temporal dual graph attention network for query-poi matching. In Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval, pages 628–638, 2020.

[22] Xiaoyang Wang, Yao Ma, Yiqi Wang, Wei Jin, Xin Wang, Jiliang Tang, Caiyan Jia, and Jian Yu. Traffic flow prediction via spatial temporal graph neural network. In Proceedings of The Web Conference 2020, pages 1082–1092, 2020.

[23] Pia Buereteur, Robert Weiibel, and Dirk Burghardt. Content zooming and exploration for mobile maps. In Multiscale-Corpusary Research on Geographic Information in Europe and Beyond Proceedings of the AGILE’2012 International Conference on Geographic Information Science, Avi-gnon, pages 74–80, 2012.

[24] Xiang Zhang, Tinghua Ai, Jiantien Stoter, and Xi Zhao. Data matching of building polygons at multiple map scales improved by contextual information and relaxation. ISPRS Journal of Photogrammetry and Remote Sensing, 92:147–163, 2014.

[25] Zongheng Dong, Rui Li, Huayi Wu, Wenjing Chen, Wei Huang, and Hongping Zhang. Browsing behavior modeling and browsing interest extraction in the trajectories on web map service platforms. Expert Systems with Applications, 195:115690, 2022.

[26] Buru Chang, Yonggyu Park, Donghyeon Park, Seongsoon Kim, and Jae-woo Kang. Content-aware hierarchical point-of-interest embedding model for successive poi recommendation. In IJCAI, pages 3301–3307, 2018.

[27] Yuxun Liu, Xianpei Fan, Xiong Wang, Xihong Yan, Kuyun Zhang, Hao Wang, Hongning Dai, Lianyong Qi, and Rui Ma. An attention-based category-aware gru model for the next poi recommendation. International Journal of Intelligent Systems, 36(7):3174–3189, 2021.

[28] Yang Li, Tong Chen, Hongzhi Yin, and Zi Huang. Discovering collaborative signals for next poi recommendation with iterative seq2graph augmentation. arXiv preprint arXiv:2106.15814, 2021.

[29] Huazhou Xie, Da Li, Yuanrui Wang, and Yukiko Kawai. Visualization of poi competitiveness using extracted map tiles from social media response since covid-19. In IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology, pages 641–646, 2021.

[30] E Stefanakis. Map tiles, and cached map services. Go-Go matics. Magazine of GoGeomatics Canada, 2015.

[31] Rostislav Netek, Jan Masopust, František Pavlicek, and Vílem Pechanec. Performance testing on vector vs. raster map tiles—comparative study on load metrics. ISPRS International Journal of Geo-Information, 9(2):101, 2020.

[32] Matthias Kuhnert, Alexey Voinov, and Ralf Seppelt. Comparing raster map comparison algorithms for spatial modeling and analysis. Photogrammetric Engineering & Remote Sensing, 71(8):973–984, 2005.

[33] Tong Xia, Junjie Lin, Yong Li, Jie Feng, Pan Hui, and Xiaoyang Wang. poi competitiveness using extracted map tiles from social media response. In Proceedings of the 35th ACM International Conference on Information & Knowledge Management, pages 135–144, 2020.

[34] Zixuan Yuan, Hao Liu, Yanchi Liu, Denghui Zhang, Fei Yi, Nengjun Zhu, and Hui Xiong. Spatio-temporal dual graph attention network for query-poi matching. In Proceedings of the 43rd international ACM SIGIR conference on research and development in information retrieval, pages 628–638, 2020.

[35] Michael Crutcher and Matthew Zook. Placemarks and waterlines: Racialized cyberscapes in post-katrina google earth. Geoform, 40(4):523–534, 2012.

[36] Mark Graham and Matthew Zook. Visualizing global cyberscapes: Mapping user-generated placemarks. Journal of Urban Technology, 18(1):115–132, 2011.
[36] Nicholas Lim, Bryan Hooi, See-Kiong Ng, Xueou Wang, Yong Liang Goh, Renrong Weng, and Jagannadan Varadarajan. Stp-udgat: spatial-temporal-preference user dimensional graph attention network for next poi recommendation. In Proceedings of the 29th ACM International Conference on Information & Knowledge Management, pages 845–854, 2020.

HUAZE XIE is currently pursuing the PhD degree in the School of Information Science and Engineering at Kyoto Sangyo University. He received the MS degree in information science and technology from the Yamaguchi University, Japan, in 2019. His research interests include natural language processing, graph convolution networks and social media data mining with GIS.

DA LI is an assistant professor at the Faculty of Engineering, Fukuoka University. He received the PhD degree in information science and technology from the Hokkaido University, Japan, in 2020. He has worked as a researcher with Hokkaido University, in 2020. He has worked as a postdoctoral researcher with the Faculty of Information Science and Engineering, Kyoto Sangyo University, in 2021. His research interests include natural language processing, machine learning, sentiment analysis, data mining, and social media.

YUANYUAN WANG is an associate professor at the Graduate School of Sciences and Technology for Innovation, Yamaguchi University in Japan. She received her PhD degree in Human Science and Environment from the University of Hyogo, Japan in 2014. She has worked as a researcher at Kyoto Sangyo University in 2014 and at Nagoya University in 2015. Her research interests include e-learning systems, data mining, and human-computer interaction.

YUKIKO KAWAI is a professor at the Faculty of Information Science and Engineering, Kyoto Sangyo University. She received her MS and PhD degrees in Information Science and Technology from the Nara Institute of Science and Technology, in 1999 and 2001, respectively. She then worked as postdoc at the National Institute of Information and Communications Technology (NICT) in Japan. Her research interests include data mining, visualization, Web mining, GIS, and information retrieval.

***