# A survey of road feature extraction methods from raster maps

DOI: 10.1111/tgis.12812

REVIEW ARTICLE

Transactions in GIS WILEY

# A survey of road feature extraction methods from raster maps

Chenjing Jiao | Magnus Heitzler | Lorenz Hurni

Institute of Cartography and Geoinformation, ETH Zürich, Zürich, Switzerland

**Correspondence**
Chenjing Jiao, Institute of Cartography and Geoinformation, ETH Zürich, Stefano-Franscini-Platz 5, CH-8093 Zürich, Switzerland.
Email: cjiao@ethz.ch

## Abstract

Maps contain abundant geospatial information, such as roads, settlements, and river networks, to name a few. The need to access this information to carry out analyses (e.g., in transportation, landscape planning, or ecology), as well as advances in software and hardware technologies, have driven the development of workflows to efficiently extract features from maps. The aim of this article is to provide a comprehensive overview of such methods to extract road features from raster maps. The methods are categorized based on the classes of techniques they employ (e.g., line extraction), as well as their subclasses (e.g., line tracing, Hough transform), the amount of user intervention required (e.g., interactive, automatic), the required data (e.g., scanned maps, contemporary vector data) and the produced results (e.g., raster-based predictions, vector-based results, attributes). Additionally, recent road extraction methods from overhead imagery, together with evaluation methods that will possibly benefit road extraction from raster maps, are reviewed. Furthermore, the evolution of this research field is analyzed over the past 35 years and the limitations of the current techniques, as well as possible future directions, are discussed.

## 1 | INTRODUCTION

Nowadays, a large number of scanned raster maps of different ages are available in digital map archives. Yet the wealth of cartographic information, such as roads, building footprints, hydrography, and so forth, is still locked in these maps and therefore cannot be directly analyzed or used in GIS (Chiang, Duan, Leyk, Uhl, & Knoblock, 2020; Chiang, Leyk, & Knoblock, 2011). Therefore, there is a high demand for effective and efficient feature extraction methods.

Cartographic information on road data has a particularly broad application domain. Long-term historical road data are used to analyze the evolution of the road networks. For instance, the development of the Swiss road and railway network during the period of 1950–2000 is investigated by Erath, Löchl, and Axhausen (2009). Swiss road-based accessibility since 1,850 is reported by Axhausen, Fröhlich, and Tschopp (2011) to understand the impact of transport investment on the economy and land use patterns. Strano, Nicosia, Latora, Porta, and Barthélemy (2012) reported the analysis of the evolution of the road networks for almost 200 years in a large area located north of Milan. Masucci, Stanilov, and Batty (2014) studied the growth of London's street networks over 224 years by modeling the networks in dual representation and analyzing their topological properties, such as degree, connectivity, average path length, and so forth. Zhao, Jia, Qin, Shan, and Jiao (2015) statistically analyzed the evolution of the OpenStreetMap (OSM) road network for Beijing. Casali and Heinimann (2019) studied the growth of the road network from 1955 to 2012 in Zürich, Switzerland. Wang et al. (2019) analyzed the evolution of the road network in Changchun, China from 1912 to 2017. Road data extraction from historical maps is a prerequisite of the studies on road evolution that cover a long time span. Furthermore, historical road data are also used to realistically reconstruct streetscapes of the past for education, entertainment, and research purposes (https://ai.googleblog.com/2020/10/recreating-historical-streetscapes.html). Moreover, especially in geoscience, one application is geospatial data integration as road features exist across various geospatial data sources. For example, maps and remote sensing imagery covering the same area can be aligned through utilizing the extracted common road intersection points from these two spatial data sources as control points or "glue" (Chen, Knoblock, & Shahabi, 2008; Chen, Knoblock, Shahabi, Chiang, & Thakkar, 2004). Chiang, Knoblock, Shahabi, and Chen (2009) use road intersection templates containing the information of the positions, connectivity, and orientations of the road intersections extracted from raster maps to extract roads from remote sensing imagery, so that these two data sources are integrated. In addition, the extracted road features can be applied for improving the extraction results themselves. For instance, in the study by Chiang et al. (2009), a template is constructed based on each extracted road intersection. Subsequently, the localized template matching (LTM) method (Chen et al., 2008) is utilized to adjust the intersection points to the precise location, thereby improving the accuracy of the extracted road intersections in terms of position, connectivity, and orientation. Another study reported by Chiang and Knoblock (2009a) used the identified road pixels to generate a road template to find road pixels that cannot be detected by the Hough transform (Ballard, 1987). Furthermore, up-to-date road data are essential to update the existing road database, including fitting the existing road data to the real landscape, improving the planimetric accuracy, and deriving the height of the road centerlines (Eidenbenz, Käser, & Baltsavias, 2000; Fortier, Ziou, Armenakis, & Wang, 2001; Zhang, 2003), and timely road maps are crucial in applications including disaster management, urban planning, car navigation (Baltsavias & Zhang, 2005; Itonaga, Matsuda, Yoneyama, & Ito, 2003), intelligent transportation systems (Zhang, Baltsavias, & O'Sullivan, 2005), and impervious surface extraction (Wang, Song, Chen, & Yang, 2015).

Despite the wide application domains of road data, the extraction of road features from raster maps can be challenging due to the similarity of road symbols to those of other features (e.g., isolines, streams), the long length of some features (e.g., highways spanning whole countries), the adjacency of other map elements with the same color (e.g., buildings), and the marginal differences between the symbols representing different classes of roads (e.g., main roads, country roads) (Herold, 2015). Especially paper maps that have been printed decades or even centuries ago may suffer from poor quality. The reasons might be inaccurate printing technologies as well as chemical and physical deterioration (e.g., bleaching, fractures, paper distortion). Furthermore, blurring and color aliasing may be induced by the scanning process (Leyk, Boesch, & Weibel, 2005; Liu, Xu, & Zhang, 2019). Additionally, map readers interpret the maps by checking the map context, such as road labels or map legends, which is a challenging task for machines (Chiang et al., 2009). To overcome these problems, researchers have developed numerous approaches, such as line tracing, morphological operation, color image segmentation (CIS), machine learning, and so forth.

This article aims to review the studies on road feature extraction and vectorization from raster maps, as well as analyzing the development and progress of these road extraction techniques. Also road extraction methods, as well as evaluation metrics from overhead imagery, that will possibly benefit the task for raster maps are reviewed, aiming to bring inspiration from the field of remote sensing. Both interactive and automatic road extraction methods from raster maps are reviewed in this article. Although there already exist surveys on feature extraction and vectorization from topographic maps (Chiang, Leyk, & Knoblock, 2014; Liu et al., 2019), they aim to cover map feature extraction methods in general and thus lack information on the special case of road feature extraction, such as interactive road extraction methods. Interactive or semi-automatic methods refer to the process in which a user collaborates with a computer. Interactive methods take advantage of a computer's ability to precisely delimit a feature and to combine it with a user's high-level understanding of the map image (Moučka, 2018). Originally, interactive methods worked on paper maps or binary raster map images, and roads were extracted together with other linear features (e.g., counter lines) (Eikvil, Aas, & Koren, 1995; Kennie, 2014; Stevenson, 1994). Recently, interactive methods have generated promising road extraction results from color map images and have the capability to segment road features with other linear features. Until now, interactive road extraction has still played an important role in data acquisition. However, in comparison with interactive methods, automatic approaches nowadays clearly dominate the field of road extraction from raster maps. Automatic methods largely reduce manual intervention during the extraction process. But they require the road features to be consistent within the data source. Moreover, a set of rules or parameters usually must be pre-defined before the automatic extraction and vectorization of road features (Stevenson, 1994). Although the degree of automation and accuracy of feature extraction has been continuously improving in recent decades, fully automatic extraction of road features still cannot be achieved because human inspection of the raster maps is necessary to achieve reliable map interpretation results (Bin & Cheong, 1998; Chiang et al., 2014; Suzuki & Yamada, 1990; Yang, An, & Huang, 2012). For example, an interactive correction step is necessary to extract unrecognized and remove falsely recognized features. Apart from this, in the existing surveys, the evolution of road extraction techniques is rarely investigated. Thus, a brief analysis of the development and progress of road extraction techniques is given to review the research trends over the past decades, as well as to indicate possible future research directions. Furthermore, a novel categorization scheme is proposed to classify the methods, based on the concrete techniques applied and their purposes. In this categorization scheme, detailed technique characteristics of each method are presented. Classical and general techniques for line extraction (e.g., Hough transform, morphological operations), image segmentation (e.g., histogram technique, K-means), noise filtering (e.g., conventional filters like mean-shift), and so forth are listed in Table 1 and expounded in the corresponding text. The remainder of this article is structured as follows. Section 2 first shows the categorization scheme, and reviews road extraction methods from raster maps. Selected road extraction methods and evaluation methods from overhead imagery are reviewed in Section 3. Section 4 reports the analysis of the development and progress of the road extraction methods, followed by a discussion of the current technical limitations and future technology trends. Section 5 concludes.

## 2 | ROAD FEATURE EXTRACTION FROM RASTER MAPS

### 2.1 | Categorization

This section presents a detailed review of the road extraction methods based on raster maps. The methods are categorized in a novel hierarchical categorization scheme shown in Table 1. The methods are first characterized and classified according to the applied techniques, whether user intervention is required, the input data, and the

**TABLE 1** Categorization of road extraction methods based on raster maps

| | Line extraction | | | Image filtering | | | CIS | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Paper | Line tracing | Parallel characteristic extraction | Hough transform | Conventional filters | Morphological operations | CNN | Histogram technique | K-means | Other clustering techniques | SVM | Input | Output |
| Miyatake (1985) | $a^{\dagger}$, $2^{*}$ | a, 1 | | | | | | | | | Raster maps | Road centerlines |
| Nakajima et al. (1985) | a, 2 | a, 1 | | | | | | | | | Raster maps | Road centerlines and the connectivity of intersections |
| Amin and Kasturi (1987) | a, 1 | | | | a, 2 | | | | | | Raster maps | Road centerlines |
| Suzuki et al. (1987) | a, 1 | | | | | | | | | | Raster maps | Road centerlines |
| Alemany and Kasturi (1988) | a, 1 | | | | a, 2 | | | | | | Raster maps | Road centerlines |
| Kasturi and Alemany (1988) | a, 1 | | | | a, 2 | | | | | | Raster maps | Road centerlines |
| Nagao et al. (1988) | | a, 1 | | | | | | | | | Raster maps | Road centerlines |
| Kasturi et al. (1989) | a, 1 | | | | a, 2 | | | | | | Raster maps | Road centerlines |
| Nagao et al. (1990) | | a, 1 | | | | | | | | | Raster maps | Road centerlines |
| Suzuki and Yamada (1990) | a, 1 | | | | | | | | | | Raster maps | Road centerlines |
| Yamada et al. (1990) | | | | | a, 1 | | | | | | Raster maps | Road pixels |
| Yamada et al. (1991) | | | | | a, 1 | | | | | | Raster maps | Road pixels |
| Yamada et al. (1993) | | | | | a, 1 | | | | | | Raster maps | Road pixels |
| Ebi et al. (1994) | | | | | a, 2 | | a, 1 | | | | Raster maps | Road centerlines |

**TABLE 1** Continude

| | Line extraction | | | Image filtering | | | CIS | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Paper | Line tracing | Parallel characteristic extraction | Hough transform | Conventional filters | Morphological operations | CNN | Histogram technique | K-means | Other clustering techniques | SVM | Input | Output |
| Stevenson (1994) | i, 1 | | | | | | | | | | Raster maps | Centerlines |
| Eikvil et al. (1995) | i, 1 | | | | | | | | | | Raster maps | Centerlines |
| Ahn et al. (1997) | | | | | a, 2 | | | | i, 1 | | Raster maps | Road centerlines |
| Bin and Cheong (1998) | | a, 1 | | | | | | | | | Raster maps | Road centerlines |
| Nishijima and Watanabe (1998) | | a, 1 | | | | | | | | | Raster maps | Road centerlines |
| Watanabe and Oshitani (2001) | | a, 1 | | | | | | | | | Raster maps | Road centerlines |
| Yin and Huang (2001) | | | | a, 1 | | | a, 2 | | | | Raster maps | Road centerlines |
| Liu (2002) | a, 4 | a, 3 | | a, 1 | a, 2 | | | | | | Raster maps | Road centerlines |
| Itonaga et al. (2003) | | | | | a, 1 | | | | | | Raster maps | Road centerlines |
| Chiang et al. (2005) | a, 3 | a, 2 | | | a, 4 | | a, 1 | | | | Raster maps | Road width, intersection positions, connectivity and orientations |
| Chiang and Knoblock (2006) | | | | | | | | | | a, 1 | Raster maps | Road pixels |
| Dhar and Chanda (2006) | | a, 3 | a, 4 | | a, 2 | | | a, 1 | | | Raster maps | Road centerlines |
| Chiang and Knoblock (2008) | a, 3 | a, 2 | | | a, 4 | | a, 1 | | | | Raster maps | Road width, intersection positions, connectivity and orientations |
| Chiang et al. (2009) | a, 3 | a, 2 | | | a, 4 | | a, 1 | | | | | |

**TABLE 1** Continude

| | Line extraction | | | Image filtering | | | CIS | | | | Input | Output |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Paper | Line tracing | Parallel characteristic extraction | Hough transform | Conventional filters | Morphological operations | CNN | Histogram technique | K-means | Other clustering techniques | SVM | | |
| Chiang and Knoblock (2009a) | | | a, 4 | a, 1 | a, 5 | | | a, 3 | a, 2 | | Raster maps, user labels | Road pixels |
| Chiang and Knoblock (2009b) | | | a, 4 | a, 1 | a, 5 | | | a, 3 | a, 2 | | Raster maps, user labels | Road centerlines, intersection positions |
| Henderson and Linton (2009) | | | | | | | a, 1 | | | | Raster maps | Road pixels, intersections positions |
| Henderson et al. (2009) | | | | | | | a, 1 | a, 2 | | | Raster maps | Road pixels |
| Leyk (2009) | | | | | | | | | a, 1 | | Raster maps | Road centerlines |
| Linton (2009) | | | | | a, 2 | | a, 1 | | | | Raster maps | Road centerlines, intersections positions |
| Pezeshk and Tutwiler (2010) | | | | | a, 1 | | | | | | Raster maps | Road centerlines |
| Pezeshk (2011) | | | | | a, 1 | | | | | | Raster maps | Road centerlines |
| Pezeshk and Tutwiler (2011) | | | | | a, 1 | | | | | | Raster maps | Road centerlines |
| Callier and Saito (2011) | | | | a, 1 | | | a, 2 | | | | Raster maps | Road pixels |
| Chiang et al. (2011) | | | | a, 2 | a, 3 | | | | a, 1 | | Raster maps, user labels | Road centerlines |
| Yang et al. (2012) | i, 3 | | | | a, 2 | | | a, 1 | | | Raster maps | Centerlines |
| Chiang and Knoblock (2013) | a, 5 | | a, 3 | a, 1 | a, 4 | | | a, 2 | | | Raster maps, user labels | Road centerlines, intersection positions and orientations |
| Henderson (2014) | | | | | a, 2 | | a, 1 | | | | Raster maps | Road centerlines, intersections positions |

**TABLE 1** Continude

| | Line extraction | | | Image filtering | | | CIS | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Paper | Line tracing | Parallel characteristic | Hough transform | Conventional filters | Morphological operations | CNN | Histogram technique | K-means | Other clustering | SVM | Input | Output |
| Kennie (2014) | i, 1 | | | | | | | | | | Paper maps | Centerlines |
| Duan et al. (2017) | | | | | | a, 1 | | | | | Raster maps, road vector data | Aligned road data/training data for road extraction |
| Chiang et al. (2020) | | | | | | a, 1 | | | | | Raster maps | Road pixels |
| Duan et al. (2020) | | | | | | | | a, 1 | | | Raster maps, road vector data | Aligned road data/training data for road extraction |
| Saeedimoghaddam and Stepinski (2020) | | | | a, 1 | | a, 2 | | | | | Raster maps | Positions and bounding boxes of road intersections |

† "i" refers to an interactive way to implement the technique, while "a" an automatic way.

* The numbers refer to the orders in which the techniques are applied.

results they produce. Furthermore, the techniques are classified based on their goals, including line extraction, image filtering, and CIS. To clearly show the procedure of each method, numbers are used to indicate the orders of the techniques applied in the method. Whether user intervention is required is indicated by "i" (interactive) or "a" (automatic) (e.g., "i, 1" means the technique is the first one in the method and is used in an interactive way; "a, 2" means the technique is used automatically and secondly). The methods in the table are ordered according to publication year, so the evolution and development trend of the methods can be explicitly shown, based on which the development and progress of the road extraction methods is analyzed in Section 4. Methods with the same techniques and technique orders, as well as the same input and output, are clustered and shown in the same color.

Line extraction directly yields lines and usually operates in the image space. It includes the techniques of line tracing, parallel characteristic extraction, and Hough transform. Image filtering yields pixel-based results and carries out operations in the image space. The image space is defined by the height and width of the image as well as its channels. The channels typically contain three components, one for each of the primary colors red, green, and blue. Sometimes an alpha channel as a fourth component is present.

Image filtering-based road extraction techniques encompass conventional filters, morphological operations, and convolutional neural networks (CNNs). The term "conventional" is used to distinguish from CNNs. A CNN is considered an image filter-based technique since its convolutional layers, which are probably the most important layer type, essentially make use of a multitude of image filters to modify an input image.

CIS yields pixel-based results and operates in the color space. The color space consists of the following dimensions: one for each color red, green, and blue. Sometimes an alpha channel is added as a fourth dimension. In rare cases, these dimensions are transformed into an alternative color space, such as L*u*v. The width and height of the image do not play a role in the color space, which distinguishes it from the image space. CIS-based road extraction techniques comprise the histogram technique, K-means, other clustering techniques, and support vector machine (SVM).

Moreover, the input and output of each method are also included in this categorization scheme. As input to the methods, paper maps refer to hard copies of printed maps and raster maps refer to maps that exist in raster form, which is typically obtained via scanning. As for the output, if a method only extracts road pixels without a subsequent vectorization step, the output in the table is "road pixels." Otherwise, if a method is specifically developed for extracting road centerlines in vector form, the output is "road centerlines." But if the method is developed to extract centerlines of all kinds of linear features (e.g., roads, contour lines, streams), its output is "centerlines." In this way, a detailed and comprehensive characterization of the methods is presented and the techniques are explicitly associated with their goals.

## 2.2 | Line extraction-based road extraction

### 2.2.1 | Line tracing

*Interactive line tracing*

A large portion of the graphical elements on maps is composed of line features. Thus, researchers have implemented many approaches to extract and vectorize linear features, including roads (Kennie, 2014; Stevenson, 1994). Interactive line tracing vectorizes linear features through following the lines and recording line points every few pixels. It is a popular and intuitive approach that gives the operator an opportunity to guide the process and utilize human knowledge (Eikvil et al., 1995).

In line tracing, the operator interactively identifies a section of the scanned raster map image, zooms into the selected area, places the screen cursor on the linear feature, and determines the direction of the line. Then, the software takes over and automatically traces the linear features by following and drawing along the lines.

The $x$ and $y$ coordinates of points on the line are recorded at a preset interval. When some complicated features are reached, such as the end point of a line or a potential intersection, the program waits for operator input. A potential intersection is recognized by a sudden increase in line width. Some line tracing programs have the capability for dashed line recognition, line intersection straightening, line generalization, and so forth (Kennie, 2014; Stevenson, 1994).

More advanced tools and algorithms emerged later to improve the extent of automation for line tracing. Eikvil et al. (1995) implemented an interactive tool for binary map image vectorization. With this tool, linear features are extracted and vectorized through tracing the contours at both sides of the lines. The course of the contours is stored as they are traced. The midpoint between the current two contour points is computed at short intervals, and stored to represent the line. Moreover, curvatures and corners are detected by checking whether the direction of the contour segments changes. For intersections, the type of intersection is determined by the user before subsequent line tracing. The tracer may determine the direction of the traced segments. Otherwise, the tracer stops and asks for user interaction.

Even until recently, interactive line tracing was still widely used and further developed, as it brings line tracing under human control and provides the ability to correct the data immediately if required. Yang et al. (2012) proposed a local adaptive segmentation method based on a sliding window to extract linear features, followed by a sequential line tracing process. Specifically, the starting point and an initial direction of the line to be traced are specified by the user, based on which a sliding window is created. The color image in the sliding window is converted into a gray-scale image. Then, K-means clustering is used to segment the linear feature in a small neighborhood in the center of the sliding window. Subsequently, directional region growth is performed in the whole sliding window to segment the linear feature. A thinning operation is applied to the segmentation results. Sequential line tracing and vectorization are performed by moving the sliding window along the linear feature until an endpoint or an intersection is met. Automatic line tracing stops at an intersection point, and the next point is manually provided. The method may not work for tracing linear features symbolized by dashed lines.

In summary, due to the complexity of linear features (e.g., intersected road lines, dashed lines), interactive line tracing guided by a human operator is a practical and effective way of vectorization. Interactive line tracing is much more efficient and economic than manual line tracing, and more practical than automatic line tracing, as fully automatic extraction and vectorization of linear features is still far from being mature enough to be universally applicable. Usually, interactive line tracing works on binary raster maps. However, Howman and Woodsford (1977) point out that interactive line tracing works more effectively for single lines (e.g., road features symbolized with single lines, contour lines) than for double lines (e.g., road features symbolized with parallel lines).

#### *Automatic line tracing*

Single line tracing methods are developed to trace either the medial lines of linear features (roads, railways, and water areas) or the line segment array generated through a map vectorization process, thereby extracting road features from maps. Suzuki, Kosugi, and Hoshino (1987) traced and extracted the medial lines of road features by assuming that road features have the same width and small curvatures. Later, Suzuki and Yamada (1990) proposed an algorithm to trace the medial lines of road features by searching a line segment array that is obtained through vectorization. The array records the starting and ending points of each line segment, and line branches that contain the segment. Nonetheless, in these two studies, the vectorization process that generated the medial lines through the thinning operation may lead to distortion and lower the accuracy of the medial lines. Moreover, if road features have the same width as other linear features (e.g., railways, contours), these two methods may fail in distinguishing these features. A similar approach that traces and recognizes road features based on the line segment list is designed by Amin and Kasturi (1987). The line tracing algorithm is developed to order line segments constituted of connected pixels into lines based on geometrical constraints like segment length and orientation. Subsequently, road features are recognized through analyzing the line segment list. When a query is requested, the image processing routines are triggered. Then, operations such as skeletonization and line tracking are used to extract and display all the lines representing the queried roads. Similar methods to trace road lines are implemented in Alemany and Kasturi (1988), Kasturi and Alemany (1988), and Kasturi, Fernandez, Amlani, and Feng (1989). Nevertheless, the application of the algorithms is limited to simple maps, and several parameters have to be manually specified. Despite the effectiveness of the single line tracing approaches, the thinning operation that produces the medial lines of linear features often leads to line distortion, especially around road intersections (Chiang & Knoblock, 2008). More importantly, the single line tracing methods usually neglect the parallel characteristic of road lines, which is the most prominent characteristic of road features (Bin & Cheong, 1998; Liu, 2002), as the parallel characteristic has been lost in the vectorization process. In addition, most single line tracing methods did not take advantage of the color information of raster maps.

Many other methods employ the parallel characteristic of road lines to extract road features from map images. Miyatake (1985) proposed a parallel line extraction method to extract road features. Specifically, the spacing width between the parallel lines is computed. The length of each connected component is set as the length of its surrounding rectangle. The pixels with labeled width and length values in a given range are regarded as the space between parallel lines. Next, the pixels adjacent to the space are extracted as parallel lines. Furthermore, road intersections are extracted through an expansion and contraction process. The medial lines of the extracted parallel lines are obtained with the tracing method proposed by Kakumoto, Miyatake, Shimada, and Ejiri (1983). With this method, road features of different widths and road intersections can be extracted. Nevertheless, several parameters—like the range of the spacing width between the parallel lines—have to be set manually. Moreover, an interactive step to correct the road extraction results and to connect the disconnected lines is required. Nakajima, Agui, and Iituka (1985) defined and used parallel vector tracers to extract road features. The tracers have two pairs of connecting vectors that move along the parallel edges of the road features. This algorithm succeeds not only in tracing the parallel edges, but also the intersections. Nonetheless, the starting points of the vector tracers and the initial length of each vector cannot be automatically determined. Liu (2002) introduced an algorithm named the “rolling ball” method, which can recognize and vectorize raster map features simultaneously. Specifically, parallel road curb lines are traced with the rolling ball method, in which a ball that always touches the two parallel road lines moves along the road branches. The center points of the rolling ball are the road centerline points. Namely, the centerlines of road branches are obtained with this method. Furthermore, an adaptive road inter-junction detector is developed to detect road intersections. Subsequently, road networks are constructed based on the road centerlines and the center points of the road intersections. Although this method generates promising road extraction results, it may fail in recognizing road branches symbolized with dashed lines, as the rolling ball always needs continuous road lines. Later, Chiang, Knoblock, and Chen (2005) proposed an approach that can automatically distinguish single- or double-line road features by varying the road width from 0 to 10 pixels for parallel pattern tracing. In this process, the road width can be automatically obtained for double-line road features. The foreground pixels that do not have the parallel properties for the given road width are removed. Moreover, road intersection points are detected by using the interest operator (Chen et al., 2004) and filtered with the criterion of a connectivity (the number of lines intersecting at an intersection) of more than two. Nonetheless, there is distortion around each road intersection due to the morphological operation. Thus, this approach was improved by Chiang and Knoblock (2008) and Chiang et al. (2009). In Chiang and Knoblock (2008), road orientations are computed by tracing the road medial lines. Road intersection templates are then generated based on the road orientations to refine the positions of the road intersections. In Chiang et al. (2009), LTM (Chen et al., 2008) is utilized to enhance the accuracy of the position, connectivity, and orientation of the extracted intersections. Specifically, a double-line road intersection template or a single-line template is constructed for each extracted road intersection. Then, the position of the intersection point is adjusted with this template based on LTM. Although the approaches achieve promising results in various maps, it seems that they may not work for road features symbolized with dashed lines, as the properties employed in parallel pattern tracing did not cover the condition of dashed lines.

The studies reviewed above fully employ the parallel characteristic by tracing the parallel road lines and can obtain the width of road features. Nonetheless, many of the methods may not work for road features symbolized with dashed lines. Apart from this, most of them did not employ the color information of raster maps. These are the possible reasons why the parallel line tracing methods have rarely been developed and applied to road extraction in recent years.

*Parallel characteristic extraction*

Some other studies take advantage of the parallel characteristic of road lines to extract road features, although they do not directly trace the parallel lines, as the parallel tracers may stop where other features like characters cut off the road lines (Nagao, Agui, & Nakajima, 1988). To alleviate this problem, Nagao, Agui, and Nakajima (1990) proposed a skip-scan method to automatically extract the medial lines of parallel line road features. Specifically, a map is scanned by horizontal and vertical scan lines of constant intervals, respectively, so all the intersection points of a scan line and parallel lines are obtained. The middle points of these intersection points are connected to form the medial lines of road features. Experiments show that more than 90% of road features can be successfully extracted. Nevertheless, the width of road features has to be known before processing, and the location precision of the extracted medial lines is dependent on the interval of the scan lines. Moreover, it is not accurate to simply use straight lines to connect the broken road lines resulting from the elimination of characters.

Bin and Cheong (1998) developed a system to extract and generate road networks from urban maps. In this system, parallel lines are obtained through parallel grouping. Subsequently, road medial lines are extracted based on the parallel lines and topologically connected to construct the road networks. Despite the successful extraction and generation of road networks, it seems that the method did not automatically remove the medial lines resulting from either the parallel lines of building edges, or the building edges parallel with the road lines. Nishijima and Watanabe (1998) reported a study that applies the generation and verification paradigm of hypotheses to road extraction from urban maps. In this method, road networks are extracted and constructed by searching out pairs of successive parallel line segments. Similarly, in Watanabe and Oshitani (2001), the inferred road fragments are verified by searching pairs of parallel line segments. In the study by Dhar and Chanda (2006), road features represented by parallel lines are extracted by applying the Hough transform. This technique is further elaborated in the next section.

### 2.2.2 | Hough transform

The Hough transform can be used to detect lines in images by detecting intersection points between the sinusoidal curves in the Hough space (Ballard, 1987). In the study demonstrated by Dhar and Chanda (2006), the Hough transform is used in the pre-processing step to compute the orientations of lines obtained from the thinning operation, which contributes to joining the broken lines. Apart from this, the Hough transform is applied to distinguish different road classes represented by a solid single line, two parallel solid lines, and two parallel dashed lines by finding two maxima at the same angle, but at slightly different positions regarding the distance to the image center.

Chiang and Knoblock (2009a) applied the Hough transform to identify Hough lines from the color layers segmented from the user label image through CIS. The user label image is a small rectangle labeled by the user, which covers a road intersection or a road segment on the input map image. Then, the colors in the color layers where the average distance between the detected Hough lines to the image center is within a threshold are selected as road colors. Nevertheless, a limitation of the Hough transform in this method is that its effectiveness relies on the number of detected Hough lines. If a color layer has only a small number of road pixels, the corresponding color will not be selected as road color. Thus, an alternative technique has to be used to extract road pixels from the color layer. The method is also applied in Chiang and Knoblock (2009b, 2013).

## 2.3 | Image filtering-based road extraction

### 2.3.1 | Conventional filters

An image filter or kernel is a small matrix used to carry out a convolution operation on an image in order to change its appearance. In this way, different effects, such as sharpening, blurring, embossing, edge detection, and so forth can be obtained. Conventional image filters can be used to remove noise. Liu (2002) applied median filtering to reduce map image noise in the pre-processing step. In the study by Chiang et al. (2011), noise in the road layer identified based on a user label and CIS is removed by employing proper image processing filters and generating parameter sets identified through incorporating user-provided "noise samples."

Conventional filters play an important role in CIS. Yin and Huang (2001) employed the median filter to smooth the segmented images representing different classes of roads. The images are segmented through a gray-scale histogram. Chiang and Knoblock (2009a, 2009b) applied the mean-shift technique to reduce the number of colors in the input image. The mean-shift filter merges two colors into one if their distance in the RGB color space and the spatial distance of the corresponding pixels in the image space are both within the preset thresholds. Considering that a Gaussian or median filter may result in a substantial loss of information about the position of the roads and their edges, Callier and Saito (2011) selected mean-shift to reduce the number of colors and noise in the input map image. Similarly, in Chiang and Knoblock (2013), mean-shift is used to reduce noise and the number of colors but to preserve the edges of map features (e.g., road lines). Mean-cut is then applied to further reduce the number of colors to 1,024 at most. The color cluster boxes in the HSL space are continuously divided at the median until the total number of boxes is smaller than the desired number of colors. The colors in the same color box are then represented by their median color.

Conventional filters can also be used as auxiliary techniques in machine learning methods to extract road features. In the study reported by Saeedimoghaddam and Stepinski (2020), to enlarge the training dataset, a Gaussian filter is used to provide tiles with low graphical qualities for data augmentation. To analyze the effect of sharpness or blurriness of the input map image to the prediction results, a Laplacian filter is employed to compute the minimal blurriness.

### 2.3.2 | Morphological operations

A morphological operation is conceptually defined by moving a window (e.g., a shifting window of 3 × 3 pixels) over the image, in such a way that it is eventually centered over every image pixel where a local logical operation is performed (Bovik, 2009). The most commonly used morphological operations in image processing tasks are binary operations, to which the input is a binary image. When the image is "scanned" by the shifting window, these morphological operations generate binary results based on the hit-or-miss transform. Eventually, the morphological operation creates a new binary image. Fundamental binary operators include erosion and dilation. The erosion operation removes small-scale details from a binary image but simultaneously reduces the size of regions of interest, too. The dilation operation has the opposite effect to erosion. It adds pixels to both the inner and outer boundaries of regions. Many morphological operations are represented as combinations of the erosion and dilation operators, such as opening and closing. The opening operation can open up a gap between objects connected by a thin bridge of pixels. The closing operation fills holes in the regions while keeping the initial region sizes. Especially the thinning operation is useful for producing the skeleton of a group of foreground pixels, and thus is often applied for extracting the centerlines of road areas (Chiang et al., 2014).

Morphological operations are frequently applied in road feature extraction from map images, in order to get the road centerlines (Ahn, Kim, Rhee, & Lee, 1997; Chiang et al., 2009; Chiang & Knoblock, 2009a, 2009b, 2013; Itonaga et al., 2003), to reduce noise (Ahn et al., 1997; Chiang & Knoblock, 2013; Linton, 2009; Liu, 2002), to reconnect broken road lines (Chiang & Knoblock, 2009a, 2009b, 2013; Chiang et al., 2009; Dhar & Chanda, 2006), and to refine the extracted road areas or road intersections (Ahn et al., 1997; Chiang & Knoblock, 2013). In these methods, morphological operations are used as auxiliary techniques.

The studies that employ a morphological operation as a main technique to extract road features were initiated by Yamada, Yamamoto, Saito, and Matsui (1991). They proposed a concept of multi-angled parallelism (MAP), which unifies the two concepts of non-isotropic neighbors for feature orientation and directional elements. Each pixel of the map image is composed of multiple directional elements (e.g., 8 or 16 directions). The authors then define a set of MAP operations based on conventional erosion and dilation to extract linear features. The operations are performed upon multiple directional elements of the pixels. Specially, a directional erosional operation is performed to extract the road lines. Road areas are extracted using fan-shaped dilation operations, as road areas can be regarded as an overlapped region of expansion from two parallel lines. Nevertheless, the extracted linear features are separated at the curved sections, as in this method "line" was defined as a straight line segment. Thus, Yamada, Yamamoto, and Nakamura (1990) solved this problem by connecting the line segments of long linear features using a set of directional dilation operations. Moreover, the broken parts in the long linear features are restored by using dilation operations. Road features are extracted using the directional dilation of facing directions as they consist of parallel lines. Nonetheless, some roads in the downtown areas are wrongly extracted as hatched regions (e.g., buildings). Therefore, Yamada, Yamamoto, and Hosokawa (1993) improved this method to solve the problem by using directional dilation to restrict the directional operation.

MAP may detect all parts of linear features at the expense of misclassifying segments of characters as linear features (Pezeshk & Tutwiler, 2010). Therefore, Pezeshk and Tutwiler (2011) demonstrated an approach to solve this problem. Specifically, they defined four primary directions instead of eight directions, as the strict directionality imposed by MAP results in over-fragmentation of lines into many short segments (Pezeshk & Tutwiler, 2010). Furthermore, line segment pixels that have neighbors in at least two of the adjacent primary directions are selected as seed pixels. These pixels are located in areas where the local direction of a linear feature is changing. The whole linear feature is obtained through seed growth and linking the line segments. The length criteria are used to separate linear features from character segments. Experiments show that the approach generates promising road extraction results from U.S. Geological Survey (USGS) maps. Based on this approach, a system for automatic extraction of various map features and recognition of the text content from scanned topographic maps is developed by Pezeshk (2011). Nonetheless, some straight segments of large characters were still misclassified as linear features.

Additionally, Itonaga et al. (2003) developed an approach for automatic extraction of road networks from urban planning maps. First, road areas are recognized by stochastic relaxation based on the geometrical properties of the areas. Then, a thinning operation is applied to the recognized areas to extract road centerlines. But the thinning operation results in geometrical distortion in the extracted centerlines, especially around the line intersections. To solve this problem, the position of the intersection point is updated based on the angle of a short line segment connected to the intersection area. Next, piecewise linear approximation is applied based on the corrected road centerlines, and the road network is constructed. Nonetheless, this method may not work for areas where road features are intersected with overpasses, labels, and so forth. Moreover, the quality of the extracted road networks depends on the setting of parameter values, which may make the method map-specific.

In summary, morphological operations are commonly used to modify image contents to facilitate the extraction of linear features. Usually, CIS is carried out on the map image to obtain a set of binary images representing individual map layers before the application of binary morphological operations. Nonetheless, they have some drawbacks. First, morphological operations (e.g., the thinning operation) may distort the geometry of the original linear features, especially around line intersections. Thus, alternative techniques need to be developed to correct the geometric distortion (Chiang et al., 2014). Moreover, some parameter values (e.g., the operation iteration, the length of line segments) usually have to be determined manually based on the attributes of the maps or through experiments. This makes the morphological methods map-specific. Additionally, morphological operations may work less well for extracting road lines symbolized by dashed lines.

### 2.3.3 | Convolutional neural network

Machine learning methods provide an effective way to extract features from raster map images because of their excellent performance in classification. Some early machine learning methods are developed to obtain a set of feature values based on the training images to facilitate the extraction of road features. The road extraction method proposed by Yin and Huang (2001) computed the geometrical feature values of the map title box and the legend index table (e.g., the ratio of the longer side to the shorter side of the legend index table) from the training map images. Recently, CNNs—especially deep CNNs—have exerted their superiority in automated feature extraction from images. Deep CNNs contain tens or hundreds of successive layers that gradually extract complex features from an input image, and then predict the probability with which a certain area (e.g., whole image, single pixel, rectangles within an image) depicts a certain class. A deep CNN architecture usually consists of convolutional layers, different activation functions (e.g., rectified linear units), and pooling layers (e.g., max or average pooling) (Saeedimoghaddam & Stepinski, 2020).

One of the advantages of CNNs lies in their generality compared with other machine learning models for image feature recognition like SVM. CNNs have the capability to recognize different types of map features, or the same type of features represented by different symbols (e.g., different classes of roads represented by different symbols) (Duan et al., 2017). Yet, CNNs require a large amount of training data to perform sufficiently well. Thus, it is laborious and time-consuming to manually create the training data. To alleviate this problem, Duan et al. (2017) proposed an algorithm that automatically generates training data to facilitate the subsequent extraction of railroad features from USGS historical maps using CNNs. The algorithm automatically and accurately aligns contemporary railroad vector data with the corresponding railroad features on the maps. Later, Duan, Chiang, Leyk, Uhl, and Knoblock (2020) presented another automatic vector-to-raster alignment algorithm to generate training data for the extraction of road features from USGS historical maps. This algorithm models the alignment problem using the reinforcement learning framework to precisely annotate the locations of road features on the maps. Nonetheless, the algorithm did not move the adjacent vector road segments as a group, resulting in losing geometric and topological information of intersections, or distorting the road orientation. To compare the impact of CNN architectures on feature extraction accuracy from raster maps, Chiang et al. (2020) presented a set of experiments for railroad extraction from USGS historical maps. Although the railroad features are successfully recognized from the map images, the experiment results show a limitation of CNNs, that is, the convolutional and pooling layers included in CNNs make it difficult to recover the detailed spatial locations of map features, especially the locations of small features.

A study that employs deep CNNs for road intersection extraction is reported by Saeedimoghaddam and Stepinski (2020). They adopted the faster region-based deep convolutional neural network (RCNN) framework to extract road intersections from USGS historical maps. Specifically, the faster RCNN framework first uses a deep CNN to extract the feature maps of the map image, followed by an implementation of a region proposal network (RPN) in order to select excellent candidates from the feature maps. Then, the selected candidates are fed into two fully connected layers to compute the probability of the candidates being a road intersection and to refine their bounding boxes. In this study, the authors used a pre-trained deep CNN to reduce the training time. Moreover, as data size is the key factor in deep CNN performance, the authors enlarged the training dataset using data augmentation techniques. Experimental results demonstrate that road intersections represented as both single lines and double lines can be successfully extracted. Nonetheless, road branches cannot be extracted.

Despite the rapid development and superiority in feature recognition and extraction of CNNs, there exist up to now only limited numbers of studies that apply CNNs to road feature extraction from raster maps. It should be explored how CNNs can be fully applied to the problem of road extraction, including how the parameters of the CNN architecture affect the road extraction results, how the characteristics of the map images (e.g., color diversity, blurriness) impact the road extraction accuracy (Saeedimoghaddam & Stepinski, 2020), and how to use CNNs for generating large amounts of high-quality training data.

## 2.4 | CIS-based road extraction

CIS separates thematic map layers based on homogeneous color information, as thematic layers in maps—such as road networks, hydrography, vegetation, and so forth—are normally represented by a distinct color (Leyk & Boesch, 2010). CIS is of critical importance since the outcome directly determines the image processing methods to be applied in all subsequent stages of map feature extraction (Chiang et al., 2011; Leyk, 2009).

### 2.4.1 | Histogram technique

The histogram technique can be used in automatic CIS, aimed at separating the different color layers in a map without user intervention (Chiang et al., 2014). In a study by Ebi, Lauterbach, and Anheier (1994), the scanned raster maps are segmented into color layers before the recognition of map features. First, the RGB data of the map image are transformed into the u′v′ chromaticity plane (L*u*v* color space), and the u′v′ histogram is generated. The peaks and ridges are detected in the histogram, from which the color cluster centers are derived. The map image is then segmented using the cluster centers based on chromaticity and lightness criteria. Subsequently, region growth techniques are applied to correct the defects caused by overprinting with other layers (e.g., black tree symbols printed over a green forest region). After obtaining the color-homogeneous layers, the geometric properties of map features are used to detect whether a layer contains mainly region or line structures. The line-based layers (e.g., the road layer) are thinned and vectorized to produce the medial line. The methods succeed in distinguishing road features from contour lines, as road features are in the black layer, while contour lines are in the brown layer. Nevertheless, if road features and contour lines have similar color and are segmented in the same layer in the CIS process, the methods may fail in distinguishing them.

Another approach that utilizes histogram-based CIS is demonstrated by Yin and Huang (2001). In this approach, based on the histogram of the gray-level map image, the gray-level distributions of map features are analyzed using the multilevel thresholding technique, so different classes of roads (e.g., national highways, county roads) are segmented into separate layers. The extracted road features are vectorized, followed by a post-processing to restore the broken road lines by analyzing their slopes and the endpoint locations.

Histogram analysis is used to separate the foreground from the background pixels in the input map image (Chiang et al., 2005, 2009; Chiang & Knoblock, 2008). The authors analyze the shape of the grayscale histogram. First, the largest luminosity cluster in the histogram is identified as the background cluster. Then, other clusters are classified as either background clusters or foreground clusters based on the number of pixels in the clusters.

A color histogram segmentation approach is reported by Henderson and Linton (2009) to extract road pixels from USGS maps. First, the color usage information (e.g., the number of pixels of the same color) is retrieved from the map legend, based on which different color layers are segmented. Furthermore, geometric properties—including spatial proximity, continuity, and closure—are employed in a tensor voting method to find roads and intersections in the segmented layers. Nevertheless, this approach requires specific knowledge about the use of colors in the maps (e.g., the RGB value of each color) as the basis for a color histogram segmentation. Moreover, the parameter used in tensor voting depends on the size of road features, which seems too map-specific. The approach only analyzed a set of 200 × 200 sub-images from USGS raster maps. Similar approaches are presented by Linton (2009) and Henderson, Linton, Potupchik, and Ostanin (2009). Another similar color histogram-based map image segmentation approach is demonstrated by Henderson (2014). In order to generate an initial feature extraction result, the histogram model of the feature is created as a set of sample histograms representative of the feature class from the map legend. This is a pre-processing step for the subsequent analysis by tensor voting. More importantly, this study made a detailed comparison between the techniques of histogram-based classification as well as the techniques for extracting road features from the road curve map produced by the tensor voting process. The curve map gives the likelihood of the presence of a road curve passing through each pixel. The study also discussed how to approximate ideal parameter values for tensor voting.

Another study that employs histogram-based CIS is reported by Callier and Saito (2011) to extract road features from tourism maps. First, the mean-shift is used to reduce the total number of colors in the raster map. The method then searches for possible lines for each pixel in the 15 × 15 shifting window centered at the pixel. If a line is found to have almost the same color, then it is considered as a possible linear feature. Pixels with high probability of being a road are selected as seed points and seeded region growing is applied to find other possible road pixels. To extract complete road features, two three-dimensional histograms of the colors of the detected road pixels and the background pixels are created. The colors corresponding to the main peaks in the road pixel histogram are selected as road color. Based on this color, the undetected road pixels in the previous steps are retrieved. But it can be challenging to find reliable parameters for determining the same color and selecting pixels with high probability of being a road, and the parameter values vary between different types of maps.

The histogram technique is frequently used in automatic CIS that contributes to the successful extraction of road features in subsequent image processing and feature recognition steps, but assumes high levels of homogeneity within color layers (Chiang et al., 2014). Automatic CIS techniques may be map-specific, as it can be challenging to choose the number of color clusters and the values of other parameters. Moreover, the histogram technique usually requires combination with other image processing techniques.

### 2.4.2 | K-means

K-means is an unsupervised machine learning method that is aimed at classifying the dataset into *k* pre-defined clusters. K-means repeatedly divides the data into *k* clusters according to a certain distance function until an optimization function reaches convergence.

In the study by Dhar and Chanda (2006), K-means is applied to Indian survey maps to separate different color layers (e.g., red layer for man-made structures like roads). Specifically, K-means clusters the RGB colors of the image pixels. Nonetheless, in the clustering process, *k* is preset to five, which can be map-specific. Henderson et al. (2009) reported a study that employs K-means to segment semantic classes (e.g., roads) in raster maps. In the method proposed by Chiang and Knoblock (2009a, 2009b), K-means is applied to the map image processed by mean-shift to further reduce the number of colors to not larger than the preset *k* value. Similarly, Chiang and Knoblock (2013) applied K-means to the image processed by mean-shift and median-cut to further merge similar colors. Yang et al. (2012) applied K-means clustering to a small neighborhood (5 × 5) to segment foreground pixels from background pixels. To facilitate the alignment of contemporary vector data to the features on historical maps, Duan et al. (2020) used K-means to group the pixel colors into clusters to detect the dominant pixel colors overlapping with vector segments. The clusters are used to formulate the reward function in the reinforcement learning framework. If a cluster center is not within the color range of the target map feature (e.g., road, water body), the reward for the segment is 0; otherwise, it is 1.

K-means is a simple and unsupervised technique in CIS. Sometimes, it requires other techniques as pre-processing steps, like median-cut to reduce the runtime (Chiang & Knoblock, 2013) and image enhancement to reduce color variations (Dhar & Chanda, 2006). Nonetheless, limitations of K-means include that it may be difficult to forecast the number of clusters, namely the value of *k*, and that the clustering result is highly influenced by the original input (e.g., the value of *k*). With a small *k* value, K-means may merge different semantic features (e.g., roads and text), as it considers only the color space (Chiang & Knoblock, 2009b).

### 2.4.3 | Other clustering techniques

Other techniques can also be used in CIS, and CIS can be developed as an interactive procedure, because it requires user input to indicate the colors of road pixels. Ahn et al. (1997) demonstrated a road extraction method based on interactive CIS for Korean topographical maps. The user needs to manually specify points that contain the colors of road features, and the center of the color cluster is calculated. The pixels that have the shortest distance to the center in the color space are segmented from the map image into a separate layer, which is a binary image. For example, contour lines and roads are colored in red. Thus, they are segmented into one layer. Next, opening and thinning operations are applied to remove contour lines and obtain centerlines of road features, respectively. Vectorized road data are obtained by tracing the centerlines. Nevertheless, this method did not correct the distortion around road intersections resulting from the thinning operation.

To alleviate the parameterization problem in CIS methods and the color variation problem in USGS maps, Leyk (2009) proposed a two-stage color sampling approach. The first stage is implemented for the derivation of the color value centroid in the color layer based on color value sampling. With these color centroids, homogeneous regions are extracted based on their minimum distance to the centroid in the color space. The black layer contains thematic objects such as road infrastructure, buildings, administrative boundaries, and characters. But the color centroids for black areas can be very different from those of black linear features. For example, buildings and road lines both belong to the black layer but appear in different color tones due to color bleaching and aliasing of the historical maps. Thus, the second stage is aimed at classifying the parts that suffer from color deviation by resampling of color values for the adjustment of the color centroids obtained in the first stage. Then, a post-processing step is implemented to generate a cartographic representation of road features.

Chiang and Knoblock (2009a) developed a supervised method that requires user input for extracting road pixels from raster map images. After the number of colors in the input image is reduced, the user needs to provide a user label for each road color in the map. Subsequently, each user label is processed by employing the Hough transform and a template matching technique, so that a color filter with all identified road colors is generated. Next, all the road pixels are extracted using the identified road colors. Despite the successful extraction of road pixels, the method is considered incomplete, as the geometry of road features is not extracted and vectorized. Thus, the method is improved in Chiang and Knoblock (2009b). In the improved method, morphological operations are applied to generate road centerlines based on the extracted road pixels. A problem of using the morphological operators is that the thinning operator usually distorts the lines near the intersections. For correcting the distortion and generating accurate road vector data, the authors detect the intersections of the thinned lines and trace the lines outside the distorted areas to generate accurate road orientations and intersection positions. Similarly, in the method reported by Chiang et al. (2011), a user provides a “road sample” through labeling a sample area centered on a road line, so that the features in the same color as the labeled road lines are automatically recognized. The recognized results are refined by employing user labels that provide samples of road and noise pixels to remove the non-road pixels through image processing filters. Subsequently, the refined road features are vectorized. Nonetheless, the refined road features still contain some undesired pixels (e.g., grid lines, characters) and broken road lines, which indicates that a further refinement step or manual post-processing is needed. Moreover, the quality of the CIS results relies on the parameter value of the number of desired map layers.

Chiang and Knoblock (2013) demonstrated a general road vectorization approach by integrating and improving the approaches reported in their earlier work (Chiang et al., 2009; Chiang & Knoblock, 2009a, 2009b). Specifically, road pixels are extracted using interactive CIS (Chiang & Knoblock, 2009a). The authors improved the time complexity of the parallel pattern tracing algorithm reported by Chiang et al. (2009) and developed a single-pass parallel pattern tracing algorithm to detect the road width and road format. Next, morphological operations are used to generate road centerlines based on the detected road width and road format. As the thinning operator usually distorts the lines near the intersections, the method then traces the road lines outside the distorted areas (Chiang & Knoblock, 2009b). The locations of the road intersections are updated using the traced lines. Subsequently, to generate the road vector data, the road centerlines are traced based on the accurate positions of the road intersections (Chiang & Knoblock, 2009b). Nevertheless, one limitation of the method is that the vectorization process relies on the width of the majority of road features. As a result, some small road branches are eliminated by mistake, as their width is smaller than the width of the majority of roads. Moreover, a post-processing step is required to reconnect the broken road lines.

The clustering techniques in CIS generate a color filter for the subsequent map feature extraction. Usually, other image processing techniques (e.g., Hough transform, morphological operations) are required in the subsequent steps to refine the results of CIS and extract the geometry of target features.

### 2.4.4 | Support vector machine

SVM is a machine learning method which is used for the separation of map layers based on color information or the separation of foreground and background pixels. SVM was first proposed for classification and regression analysis. SVM solves classification problems by finding an optimal hyperplane for linearly separable data, and is extended to non-linearly separable data by transforming the original data to a higher-dimensional space with kernel functions (Ben-Hur, Horn, Siegelmann, & Vapnik, 2001).

Chiang and Knoblock (2006) reported a study that employs SVM in combination with discrete cosine transformation (DCT) coefficients to extract road pixels from raster maps. First, foreground pixels are separated from the background through generating the DCT coefficients for pixels on the input raster map. Based on the property of consistent color in the background, pixels with low DCT coefficients are classified as background. The second stage is to classify road and character pixels among the foreground pixels. As characters are generally more complex than road lines, the DCT coefficients of character pixels are higher than those of road lines. Thus, the authors generate the DCT coefficients for each foreground pixel and send them to the SVM for classification. The authors use two street maps from Google Maps and one from ViaMichelin as road training data and manually remove the characters from them. The misclassified pixels are corrected by performing connected component analysis. Nonetheless, as the foreground pixels include other map features than road lines and characters, the approach still needs refinement.

In summary, CIS plays a key role as a pre-processing step in the whole workflow of road extraction. However, CIS may fail in separating thematic layers if different map features share the same color. For example, brown pixels in the USGS topographic maps are used for both the contour lines and roads. Thus, many of the road extraction results include contour lines (Chiang & Knoblock, 2009a). Moreover, semi-automatic CIS requires the user to provide enough user labels to cover each road color in the raster map. Usually, the user label has to cover a road segment or intersection and should be located at the center of—or just a few pixels from the center of—a road segment or intersection (Chiang et al., 2011; Chiang & Knoblock, 2009a, 2009b, 2013).

## 3 | ROAD FEATURE EXTRACTION FROM OVERHEAD IMAGERY

In this section, road extraction methods from overhead imagery which can possibly benefit road extraction from raster maps but have not yet been applied to raster maps are reviewed. As a reference for enhancing road extraction performance from raster maps, this section mainly focuses on the new methods and achievements from 2014, as since 2014 special attention has been shifted to deep learning in the remote sensing community (Ma et al., 2019). Furthermore, as criteria to evaluate road extraction results, evaluation methods are also reviewed. Compared with the monotonous evaluation metrics for road extraction from raster maps, there are various metrics for the overhead imagery, which will inspire researchers to possibly apply and tailor the metrics to raster maps.

Roads in overhead imagery and in raster maps are long, extended slender areas (Chen, Papandreou, Schroff, & Adam, 2017; Sun, Di, Che, Liu, & Wang, 2019; Wang et al., 2016; Zhou, Zhang, & Wu, 2018). Ideally, roads should be extracted as continuous, connected, and intersected long lines that can form a network. However, it is usually challenging to preserve the continuity and topology of road features, as in the overhead imagery, roads are occluded by trees and shadows (Wang et al., 2016), and in the raster maps, roads are interrupted by labels and other features (e.g., water bodies, railroads). Thus, different strategies are proposed and applied to address this issue.

### 3.1 | Machine learning architectures

Some studies adapt machine learning architectures in order to enhance their abilities to solve the discontinuity issue of road extraction. Zhou et al. (2018) propose an encoder–decoder neural network structure, named D-LinkNet, which inserts dilated convolution layers between the encoder and the decoder to address the challenge of road extraction from high-resolution satellite imagery by using dilated convolution to enlarge the receptive field of filters. The dilated convolution layers are stacked both in cascade mode and parallel mode in order to take advantage of and combine multi-resolution features. ResNet34 (He, Zhang, Ren, & Sun, 2016) pre-trained on the ImageNet (Deng et al., 2009) dataset is deployed as the encoder, as it is found that transfer learning can accelerate the convergence of the network and make it perform better. D-LinkNet achieved top performance in the CVPR DeepGlobe 2018 Road Extraction Challenge. Nonetheless, the discontinuity issue of the extracted road features is not completely addressed (Abdollahi, Pradhan, Shukla, Chakraborty, & Alamri, 2020).

Inspired by the U-net (Ronneberger, Fischer, & Brox, 2015) and atrous spatial pyramid pooling (ASPP) (Chen et al., 2017) approach, He, Yang, Wang, Wang, and Li (2019) integrated ASPP into U-net in order to grasp multi-scale road characteristics such as local corners, textures, macroscopic lines, and global network structures, as atrous convolution is capable of adjusting the receptive field of the filter. Specifically, the ASPP module used in this article consists of one 1 × 1 convolution and three parallel 3 × 3 convolutions with atrous rates of 6, 12, and 18, respectively, in combination with an image-level pooling layer. Placed after the bottleneck of the encoder–decoder network, ASPP is applied to the feature map produced by the encoder, and the resulting feature map is fed into the decoder. Nonetheless, the pooling layer may reduce the resolution of center feature maps and drop spatial information.

Tao, Qi, Li, Wang, and Li (2019) designed a spatial information inference structure (SIIS), enabling them to extract and transmit not only local road characteristics, but global and contextual road information in four directions. The SIIS is inserted after the bottleneck of the DeepLabV3+ network (Chen, Zhu, Papandreou, Schroff, & Adam, 2018). Specifically, the input of SIIS is a set of feature maps produced by the encoder. The feature map set is split into chunks along two dimensions, and the chunks are fed into a 3D convolutional recurrent neural network (Conv3d-RNN) one by one. The Conv3d-RNN is developed by replacing all the matrix operations in the traditional RNN unit with 3D convolution. Despite the effectiveness and robustness of the SIIS-Net, it may fail in extracting some very narrow country roads.

### 3.2 | Alternative loss functions

Another way to preserve the continuity and topology of roads is to adapt and improve the loss function, as the normal loss functions—such as cross-entropy—assign equal weights to each pixel, thus ignoring the spatial and topological information when evaluating the similarity between the predictions and the ground truth (He et al., 2019; Mosinska, Marquez-Neila, Koziński, & Fua, 2018).

Wei, Wang, and Xu (2017) propose a road structure-based loss function that incorporates the geometric information of road features in cross-entropy loss through imposing a large penalty of loss on the pixels close to road regions while imposing a small penalty on pixels far from road regions.

Mosinska et al. (2018) propose the topology loss as a supplementary term to the binary cross entropy (BCE) loss, which is aware of the higher-order topological characteristics (e.g., connectivity, continuity) of linear features. The feature maps obtained from several layers of a pre-trained VGG19 network (Simonyan & Zisserman, 2014) are used as a description of the higher-order characteristics. The topology loss tries to minimize the differences between the VGG19 description of the ground truth and the corresponding prediction. Experiments show that the prediction performance is increased by using this new loss function without having to change the network architecture.

He et al. (2019) proposed the structural similarity loss that evaluates the similarity between two images by comparing their luminance, contrast, and structure. The luminance is compared based on the mean intensity of pixel values, the contrast compared by using the standard deviation, and the structure information compared by computing the correlation (inner product) of the normalized pixel values. The statistics of the local structural similarity are calculated by using a square sliding window. The structural similarity is added as a term to the BCE loss and the obtained loss is used to train the network. Nonetheless, the hand-designed metrics and the topology loss computed based on the pre-trained VGG19 may be hard to generalize.

## 3.3 | Data fusion

One common problem of machine learning methods lies in the fact that it is usually time-consuming and tedious to manually label and sample the training data. Thus, some studies fuse other data as the complement to the overhead imagery for training. Wang et al. (2015) used a vector-guided sampling strategy for generating training data for road extraction. Specifically, in the step of preparing the training data, cubic B-splines are employed to refine the vector road lines, so the road polylines visually match with the roads on the aerial imagery. Sample points, centered on which the image patches are to be extracted, are classified according to whether they locate in road areas and the angle of the road segment. A local image patch is extracted based on the location of a sample point and the angle information of the point. Then, the labeled image patches are fed into a DNN for road pattern recognition and a finite state machine (FSM) is used for tracking the roads on the imagery. Nonetheless, the tracking process is manually triggered by selecting an initial position and orientation for the tracker.

Sun et al. (2019) proposed fusing GPS data with aerial imagery for road extraction. Specifically, the GPS data are rendered as new input layers and fed together with the RGB channels of the imagery into the encoder–decoder network. To overcome over-fitting, a novel way of data augmentation is applied to the GPS data, including subsampling, resolution reducing, random perturbation, and data omitting. Instead of the conventional $3 \times 3$ transpose convolution filters, 1D transpose convolution is used in the decoder, as the 1D filters are more aligned with road shapes, thus contributing to reducing gaps in the road extraction results.

Zhang, Hu, Li, and Ai (2020) used GPS trajectories of floating car data as training data to extract roads from high-resolution remote sensing imagery, which avoids the tedious and time-consuming manual labeling process. Specifically, the GPS trajectory data are first rasterized and then denoised with morphological operations. Next, the trajectory data are matched to the road in the remote sensing imagery in terms of resolution and road width. The results show that roads occluded by buildings or trees are extracted, verifying that the method is able to preserve the continuity of road features to some extent.

Nevertheless, machine learning methods still cannot completely tackle the challenges in road extraction. Highly accurate road extraction results are still not achieved. Issues like fuzzy boundaries as well as small and dispersed false positives still remain challenging, as CNNs mainly count on texture and spectral features, and the mixed pixels in road borders lead to misclassification. Moreover, many roads extend a far distance in the imagery, which demands high-level semantic information (e.g., multi-scale features) to preserve their completeness and continuity. Furthermore, pre- and post-processing operations are still necessary to achieve satisfactory results (Abdollahi et al., 2020).

### 3.4 | Shape descriptors

Shape descriptors can be used to describe accurately road-specific geometric properties (e.g., narrowness, the parallel characteristic of road edges), thereby effectively recognizing roads from the overhead imagery. Nonetheless, there are only limited numbers of studies that utilize shape descriptors to extract roads from raster maps. Thus, the methods that use shape descriptors for road extraction from overhead imagery are reviewed.

For many years, shape descriptors have played an important role in road extraction from overhead imagery (e.g., Mayer & Steger, 1998; Steger, 1998, 2000). Recently, shape descriptors have served as a supplementary technique in machine learning. Li, Hu, and Ai (2018) define two shape descriptors to automatically recognize roads and filter outliers from the results obtained through superpixel segmentation. Specifically, deviation of parallelism (DoP) is defined as the deviation of the width of a superpixel to describe the parallel characteristic of road edges, and narrow rate (NR) as the ratio of the length and width of the superpixel to describe the long and narrow characteristic of roads. Despite the effectiveness of this method, it may be difficult to obtain proper parameter values. Moreover, it may be challenging to precisely describe roads so as to quickly and accurately extract roads (Wang et al., 2016). Nonetheless, usually raster maps are less complicated than remote sensing imagery. Therefore, shape descriptors can still play a role in road extraction from raster maps.

### 3.5 | Evaluation metrics

Standard evaluation metrics for road segmentation results include recall, precision, quality, and so forth. The computation of the metric values is based on the number of correctly or wrongly segmented pixels, namely true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). Recall that the percentage of TP over all target pixels, namely TP plus FN, describes how completely roads are extracted. Precision, defined as the percentage of TP over all extracted pixels, namely TP plus FP, evaluates how correctly roads are detected. If the predicted road pixels as well as the ground truth are skeletonized or vectorized as road centerlines, completeness and correctness are more suitable than recall and precision, as they take into account the buffer of the centerlines and are therefore regarded as relaxed variants of recall and precision (Mosinska et al., 2018). If the ground-truth centerline lies within a buffer of the predicted centerline, it is deemed TP, and FN otherwise. The lengths of TP and FN are used to compute completeness, which is $\mathrm{TP}/(\mathrm{TP}+\mathrm{FN})$. If the predicted centerline lies within a buffer of the ground-truth centerline, it is TP, or FP otherwise; and $\text{correctness} = \mathrm{TP}/(\mathrm{TP}+\mathrm{FP})$ (Cardim, Silva, Dias, Bravo, & Gardel, 2018; Wang et al., 2015; Wegner, Montoya-Zegarra, & Schindler, 2013; Zhang et al., 2020). Quality and F1 score, regarded as combinations of recall and precision or completeness and correctness, reflect the overall performance (He et al., 2019; Li et al., 2018; Mosinska et al., 2018; Saeedimoghaddam & Stepinski, 2020; Tao, Qi, Li, Wang, & Li, 2019). Quality is estimated as $\mathrm{TP}/(\mathrm{TP}+\mathrm{FP}+\mathrm{FN})$, and F1 score as $2\mathrm{TP}/(2\mathrm{TP}+\mathrm{FP}+\mathrm{FN})$.

Accuracy is an intuitive metric that refers to the ratio of the number of correctly classified pixels to the number of all pixels (Wei et al., 2017; Wulamu, Shi, Zhang, & He, 2019). But if there are only a few target pixels on the image, the value of accuracy may not coincide with the effectiveness of the segmentation results. An alternative for accuracy is "intersection over union" (IoU), which, also known as the Jaccard index, refers to the intersection of the prediction and the ground truth divided by the union of the prediction and the ground truth (Sun et al., 2019; Tao et al., 2019; Wulamu et al., 2019; Zhang et al., 2020; Zhou et al., 2018).

The above mentioned metrics may fail in evaluating the continuity and topology of the extracted road features. Thus, to evaluate the continuity, Tao et al. (2019) compute the number of road breaks, which represents the number of false fractures in the predictions compared with the ground truth. To evaluate topology, Mosinska et al. (2018) and Wegner, Montoya-Zegarra, and Schindler (2015) randomly and repeatedly sample pairs of connected points in the ground truth as well as in the predicted road network, and compare their path lengths. Incorrect gaps in the extracted network cause too long paths, or they disconnect the network into disjoint parts with no connection at all. Incorrect shortcuts result in too short paths. A tolerance parameter is pre-defined to account for geometric uncertainty. Paths with length difference smaller than the tolerance are regarded as correct. The point pairs are sampled until the percentages of these three error types have converged.

There exist similarities and differences between road extraction from overhead imagery and raster maps. A raster map is an electronic map image made up of pixels, which serves as a symbolic depiction of the geographical objects on the land surface. Raster maps can be made by scanning paper maps or by cartographic software. Overhead imagery refers to the images of the earth surface captured by imaging sensors. Features on raster maps are depicted by abstract symbols with various shapes, colors, and so forth, while features on overhead imagery are presented in their natural form. Texture, spectral, and (potentially) 3D information is captured in overhead imagery. Especially, roads are slender, long, and intersected areas in overhead imagery. Roads are recognized and distinguished from other geographical objects mainly by the shape and spectral information. Roads can be long, extended, and intersected areas or lines on raster maps, and are distinguished by symbolic and color information. Roads can be occluded by trees, shadows, and interchanges on overhead imagery, while they can be interrupted by labels and point symbols (e.g., triangulation points). Line tracing works for road extraction from both data sources (Chiang & Knoblock, 2013; Wang et al., 2015). Multi-scale spatial information will benefit both tasks, and the same evaluation metrics, as well as vectorization methods, can be shared by them.

## 4 | DISCUSSION

### 4.1 | An analysis of the development and progress of the road extraction methods of raster maps

The problem of road extraction from raster maps has gained much attention, and the relevant studies and methods have experienced steady development over the past 35 years. Up to 2020, 47 different road extraction methods based on raster maps have been identified. The total citation count of the 47 papers is 1,415 in Google Scholar. Figures 1a,b present yearly paper counts and yearly citation counts, respectively. Figure 1 shows consistency between the paper counts and the citation counts. The paper count experienced two peaks in the late 1980s and around 2010, respectively. Accordingly, the citation count has two peaks in the 1990s and 2010s, as the citation count is lagging behind the paper count.

Interactive line tracing methods dominated the road extraction methods before the emergence of automatic methods. From the order of the techniques applied in the early methods, it is found that line tracing or parallel characteristic of road lines was directly applied without pre-processing like color segmentation, as in early times raster maps were black and white or gray-scale images. For example, the methods that only use interactive line tracing are clustered and colored pink in Table 1. Later, automatic line tracing methods were intensively studied in the late 1980s and early 1990s. For example, the cluster of the methods that first applies line tracing and then morphological operations is colored yellow in Table 1. In these methods, connected pixels are traced, so that line segments are obtained. Then, roads are recognized through an analysis of the line segment list. Especially, the authors implemented a road query function, in which the thinning operation is used to skeletonize the roads, so that the thinned line is displayed as the query result (Alemany & Kasturi, 1988; Amin & Kasturi, 1987; Kasturi & Alemany, 1988; Kasturi et al., 1989). In recent years, line tracing methods were not frequently studied, probably due to the complexity of the line tracing process as well as the emergence and development of other automatic methods. Interestingly, the studies that use morphological operations to extract road features from raster maps emerged in the early 1990s, with the concept of MAP proposed by Yamada's research group. This method cluster is colored orange in Table 1. Based on conventional erosion and dilation, the authors define a set of MAP operations to extract linear features (Yamada et al., 1990, 1991, 1993). These methods were further improved by Pezeshk and Tutwiler in the early 2010s, mainly to tackle the problem of misclassifying character segments as

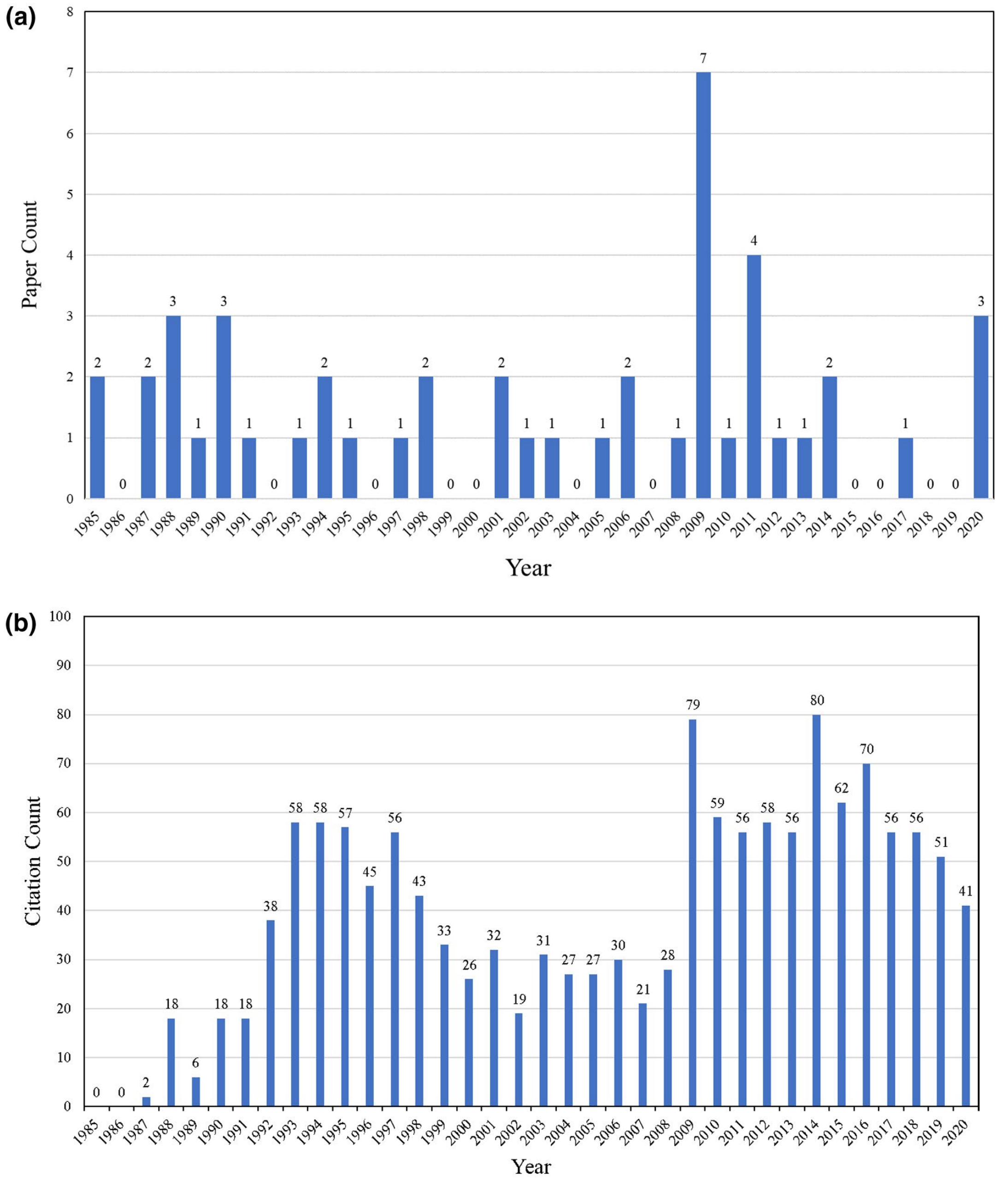


**FIGURE 1** (a) Yearly counts of papers on road extraction from raster maps; and (b) citation counts of these papers. The search was conducted on September 27, 2020 using Google Scholar

linear features by reducing the number of primary directions of MAP as well as selecting pixels where the direction of a line changes as seed pixels. The whole linear feature is then obtained through seed growth and linking the line segments. Subsequently, the length criteria are used to filter linear features (Pezeshk, 2011; Pezeshk & Tutwiler, 2010, 2011). This cluster of methods is colored purple in Table 1. On the one hand, the progress of the methods indicates that the previous studies provide directions and lay foundations for the later studies. On the other hand, the studies that mainly use morphological operations seem to be limited to MAP. Importantly, morphological operations are frequently used in road extraction because of their versatility. They can be used for noise removal in pre-processing, for reconnecting broken road lines, for skeletonizing the recognized road areas, and so forth. Conventional filters, like the median filter, were used for pre-processing the images in the early 2000s, as they work for noise removal and image smoothing. Later, CIS, especially the histogram technique and K-means, became more popular for pre-processing than conventional filters. For example, a set of CIS methods was presented by Henderson and Linton, in which different colored layers are first separated based on the color usage information retrieved from the map legend, and then geometric properties (spatial proximity, continuity, and closure) are used to detect roads and intersections in the separated layers (Henderson, 2014; Linton, 2009). This cluster is colored bright green in Table 1. With the emergence and development of modern algorithms, various techniques are applied in one method to get promising road extraction results. For instance, the methods proposed by Chiang's group involve four different techniques, including the histogram technique as a pre-processing step to segment the foreground pixels from the background pixels, parallel pattern tracing to detect road pixels, morphological operations to reconnect adjacent road pixels, and skeletonizing the road areas (Chiang et al., 2005, 2009; Chiang & Knoblock, 2008). This cluster of methods is colored gray in Table 1. In recent years, machine learning has become a study hotspot and dominates the road extraction methods. Machine learning methods are used either for road feature extraction or automatically producing training data that facilitate subsequent road feature extraction. Usually, CNNs are applied without any pre-processing. CNNs stand out in these methods, owing to their superiority in feature recognition and extraction from images, as well as the wide and intensive attention paid to CNNs. In early times, there were often clusters of very similar road extraction methods. From 2010, there are rarely clusters of methods with the emergence of more modern techniques, indicating a trend of road extraction methods developing in a diverse way. Although machine learning has become a research hotspot, different techniques are applied, different inputs are required, and different goals (e.g., to extract road pixels, to generate road training data) are achieved.

Notably, more attention has been paid to historical maps since around 2009. Historical maps have become an important data source of the study on map interpretation and spatial feature extraction. However, the poor quality of historical maps and the demand for increasing feature extraction accuracy make it urgent to propose more advanced methods for recognizing road features and small objects (e.g., dashed line segments) (Chiang et al., 2020). For instance, due to the spatial distortion inherent to historical maps, there usually exists a shift between the same road feature on the map sheets of different years. How to automatically recognize the same road with flexible tolerance remains unsolved. Additionally, historical map sheets are published every few years. The current methods did not address the problem of predicting the road features in the time between publication years. Also, further and wider applications of machine learning to road feature extraction from historical maps deserve exploration. For example, due to the spatial distortion, the extracted road data need correcting. How to apply machine learning to correct the extracted road data remains unsolved. Further, the corrected road data can be used to correct other map features that have already been extracted, like the alignment of buildings to road lines, as buildings are usually located along roads. The process of alignment could be modeled using a machine learning framework.

## 4.2 | Current limitations and technology trends

There still exist limitations in the current techniques. Roads are represented by long, slender, and intersected linear features on raster maps. It will benefit road extraction performance by employing high-level semantic information, like multi-scale spatial information. However, current techniques may fail in recognizing and taking advantage of high-level semantic information (e.g., road network structures) (He et al., 2019). Moreover, despite the superiority of machine learning methods, up to now it still remains challenging to automatically generate large amounts of high-quality training data to replace the tedious manual labeling tasks. Usually, machine learning methods are developed only to extract road pixels. Thus, vectorization is performed in a following separate step. It deserves exploration to develop a complete machine learning workflow to directly obtain road vector data from raster maps [e.g., the method that combines the DNN and FSM proposed by Wang et al. (2015), the multi-scale machine learning framework developed by Lu et al. (2019)]. Moreover, the obtained road vector data should be topologically correct to construct road networks. These limitations, however, point to future technology trends, like combining different data sources or using machine learning to automatically generate training data (Duan et al., 2020). In addition, image inpainting involves filling in missing regions of an image (Nazeri, Ng, Joseph, Qureshi, & Ebrahimi, 2019). Thus, it can be used to address the quality defects in the historical maps. Furthermore, evaluation metrics that can accurately estimate the topology and continuity of the extracted road features are to be designed. Some methods from the remote sensing domain are hopefully beneficial to address these limitations [e.g., to enable machine learning models to grasp global contextual information by adjusting the receptive field of filters using dilated convolution (Zhou et al., 2018), ASPP (He et al., 2019), or SIIS (Tao et al., 2019); to automatically generate road training data from GPS trajectories (Zhang et al., 2020); to evaluate the connectivity and topology of the extracted roads by comparing the lengths of the corresponding paths in the ground truth and the predictions (Wegner et al., 2015)].

## 5 | CONCLUSIONS

This article provides a detailed review of the studies on road feature extraction from raster maps, which helps to gain a thorough understanding of the existing methods. A novel categorization scheme is proposed to classify the road extraction methods, which associates the image processing techniques with their goals. Moreover, the road extraction methods are characterized by the techniques. Particular attention has been paid to interactive road extraction methods, as they have still been used until recently, but are largely neglected by existing surveys. Furthermore, it has been clarified how the methods in the existing studies have been developed and improved with the continuous advancement of technologies, as well as with the diverting of research hotspots. The issues and directions of further development of road extraction methods are discussed. Machine learning dominates the recent studies. Future research endeavors need to be made to retrieve the accurate and detailed spatial locations of road features. Moreover, it is necessary to pay special attention to the recent progress in road extraction from overhead imagery, as many such approaches can be tailored and applied to extract features from historical maps.

### CONFLICT OF INTEREST

No potential conflict of interest was reported by the authors.

### ORCID

*Chenjing Jiao* https://orcid.org/0000-0002-7619-5540
*Magnus Heitzler* https://orcid.org/0000-0002-9021-4170

### REFERENCES

Abdollahi, A., Pradhan, B., Shukla, N., Chakraborty, S., & Alamri, A. (2020). Deep learning approaches applied to remote sensing datasets for road extraction: A state-of-the-art review. *Remote Sensing*, *12*(9), 1444. https://doi.org/10.3390/rs12091444

Ahn, C., Kim, K. S., Rhee, S. B., & Lee, K. Y. (1997). A road extraction method from topographical map images. In *Proceedings of the 1997 IEEE Pacific Rim Conference on Communications, Computers and Signal Processing*, Victoria, BC, Canada (Vol. *2*, pp. 839–842). Piscataway, NJ: IEEE. https://doi.org/10.1109/PACRIM.1997.620390

Alemany, J., & Kasturi, R. (1988). A computer-vision system for interpretation of paper-based maps. *Proceedings of SPIE*, *829*, 125–139. https://doi.org/10.1117/12.942118

Amin, T. J., & Kasturi, R. (1987). Map data processing: Recognition of lines and symbols. *Optical Engineering*, *26*(4), 264354. https://doi.org/10.1117/12.7974079

Axhausen, K. W., Fröhlich, P., & Tschopp, M. (2011). Changes in Swiss accessibility since 1850. *Research in Transportation Economics*, *31*(1), 72–80. https://doi.org/10.1016/j.retrec.2010.11.010

Ballard, D. H. (1987). Generalizing the Hough transform to detect arbitrary shapes. In M. A. Fischler & O. Firschein (Eds.), *Readings in computer vision* (pp. 714–725). San Francisco, CA: Morgan Kaufmann. https://doi.org/10.1016/0031-3203(81)90009-1

Baltsavias, E., & Zhang, C. (2005). Automated updating of road databases from aerial images. *International Journal of Applied Earth Observation and Geoinformation*, *6*(3–4), 199–213. https://doi.org/10.1016/j.jag.2004.11.002

Ben-Hur, A., Horn, D., Siegelmann, H. T., & Vapnik, V. (2001). Support vector clustering. *Journal of Machine Learning Research*, *2*(12), 125–137. https://doi.org/10.5555/944790.944807

Bin, D., & Cheong, W. K. (1998). A system for automatic extraction of road network from maps. In *Proceedings of the IEEE International Joint Symposia on Intelligence and Systems*, Rockville, MD (pp. 359–366). Piscataway, NJ: IEEE. https://doi.org/10.1109/IJSIS.1998.685476

Bovik, A. C. (Ed.). (2009). *The essential guide to image processing*. San Diego, CA: Academic Press.

Callier, S., & Saito, H. (2011). Automatic road extraction from printed maps. In *Proceedings of the 12th IAPR Conference on Machine Vision Applications*, Nara, Japan (pp. 243–246).

Cardim, G. P., Silva, E. A. D., Dias, M. A., Bravo, I., & Gardel, A. (2018). Statistical evaluation and analysis of road extraction methodologies using a unique dataset from remote sensing. *Remote Sensing*, *10*(4), 620. https://doi.org/10.3390/rs10040620

Casali, Y., & Heinimann, H. R. (2019). A topological analysis of growth in the Zurich road network. *Computers, Environment and Urban Systems*, *75*, 244–253. https://doi.org/10.1016/j.compenvurbsys.2019.01.010

Chen, C. C., Knoblock, C. A., & Shahabi, C. (2008). Automatically and accurately conflating raster maps with orthoimagery. *GeoInformatica*, *12*(3), 377–410. https://doi.org/10.1007/s10707-007-0033-0

Chen, C. C., Knoblock, C. A., Shahabi, C., Chiang, Y. Y., & Thakkar, S. (2004). Automatically and accurately conflating orthoimagery and street maps. In *Proceedings of the 12th ACM International Workshop on Geographic Information Systems*, Washington, DC (pp. 47–56). https://doi.org/10.1145/1032222.1032231

Chen, L. C., Papandreou, G., Schroff, F., & Adam, H. (2017). *Rethinking atrous convolution for semantic image segmentation*. Preprint, arXiv:1706.05587.

Chen, L. C., Zhu, Y., Papandreou, G., Schroff, F., & Adam, H. (2018). Encoder–decoder with atrous separable convolution for semantic image segmentation. In *Proceedings of the 2018 European Conference on Computer Vision*, Munich, Germany (pp. 801–818).

Chiang, Y. Y., Duan, W., Leyk, S., Uhl, J. H., & Knoblock, C. A. (2020). *Using historical maps in scientific studies: Applications, challenges, and best practices*. Cham, Switzerland: Springer.

Chiang, Y. Y., & Knoblock, C. A. (2006). Classification of line and character pixels on raster maps using discrete cosine transformation coefficients and support vector machine. In *Proceedings of the 18th International Conference on Pattern Recognition*, Hong Kong (Vol. *2*, pp. 1034–1037). Piscataway, NJ: IEEE. https://doi.org/10.1109/ICPR.2006.368

Chiang, Y. Y., & Knoblock, C. A. (2008). Automatic extraction of road intersection position, connectivity, and orientations from raster maps. In *Proceedings of the 16th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Irvine, CA (pp. 1–10). New York, NY: ACM. https://doi.org/10.1145/1463434.1463463

Chiang, Y. Y., & Knoblock, C. A. (2009a). A method for automatically extracting road layers from raster maps. In *Proceedings of the 10th International Conference on Document Analysis and Recognition*, Barcelona, Spain (pp. 838–842). Piscataway, NJ: IEEE. https://doi.org/10.1109/ICDAR.2009.274

Chiang, Y. Y., & Knoblock, C. A. (2009b). Extracting road vector data from raster maps. In J. M. Ogier, W. Liu, & J. Lladós (Eds.), *Graphics recognition: Achievements, challenges, and evolution* (Lecture Notes in Computer Science, Vol. *6020*, pp. 93–105). Berlin, Germany: Springer. https://doi.org/10.5555/1875532.1875541

Chiang, Y. Y., & Knoblock, C. A. (2013). A general approach for extracting road vector data from raster maps. *International Journal on Document Analysis and Recognition*, *16*(1), 55–81. https://doi.org/10.1007/s10032-011-0177-1

Chiang, Y. Y., Knoblock, C. A., & Chen, C. C. (2005). Automatic extraction of road intersections from raster maps. In *Proceedings of the 13th ACM International Workshop on Geographic Information Systems*, Bremen, Germany (pp. 267–276). New York, NY: ACM. https://doi.org/10.1145/1097064.1097102

Chiang, Y. Y., Knoblock, C. A., Shahabi, C., & Chen, C. C. (2009). Automatic and accurate extraction of road intersections from raster maps. *GeoInformatica*, *13*(2), 121–157. https://doi.org/10.1007/s10707-008-0046-3

Chiang, Y. Y., Leyk, S., & Knoblock, C. A. (2011). Efficient and robust graphics recognition from historical maps. In Y. B. Kwon & J. M. Ogier (Eds.), *Graphics recognition: New trends and challenges* (Lecture Notes in Computer Science, Vol. *7423*, pp. 25–35). Berlin, Germany: Springer. https://doi.org/10.1007/978-3-642-36824-0_3

Chiang, Y. Y., Leyk, S., & Knoblock, C. A. (2014). A survey of digital map processing techniques. *ACM Computing Surveys*, *47*(1), 1–44. https://doi.org/10.1145/2557423

Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009). Imagenet: A large-scale hierarchical image database. In *Proceedings of the 2009 IEEE Conference on Computer Vision and Pattern Recognition*, Miami, FL (pp. 248–255). Piscataway, NJ: IEEE. https://doi.org/10.1109/CVPR.2009.5206848

Dhar, D. B., & Chanda, B. (2006). Extraction and recognition of geographical features from paper maps. *International Journal of Document Analysis and Recognition*, *8*(4), 232–245. https://doi.org/10.5555/2722900.2723093

Duan, W., Chiang, Y. Y., Knoblock, C. A., Jain, V., Feldman, D., Uhl, J. H., & Leyk, S. (2017). Automatic alignment of geographic features in contemporary vector data and historical maps. In *Proceedings of the First Workshop on Artificial Intelligence and Deep Learning for Geographic Knowledge Discovery*, Los Angeles, CA (pp. 45–54). New York, NY: ACM. https://doi.org/10.1145/3149808.3149816

Duan, W., Chiang, Y. Y., Leyk, S., Uhl, J. H., & Knoblock, C. A. (2020). Automatic alignment of contemporary vector data and georeferenced historical maps using reinforcement learning. *International Journal of Geographical Information Science*, *34*(4), 824–849. https://doi.org/10.1080/13658816.2019.1698742

Ebi, N., Lauterbach, B., & Anheier, W. (1994). An image analysis system for automatic data acquisition from colored scanned maps. *Machine Vision and Applications*, *7*(3), 148–164. https://doi.org/10.1007/BF01211660

Eidenbenz, C., Käser, C., & Baltsavias, E. P. (2000). ATOMI: Automated reconstruction of topographic objects from aerial images using vectorized map information. *International Archives of Photogrammetry and Remote Sensing*, *33*(B3), 463–472. https://doi.org/10.1.1.37.8714

Eikvil, L., Aas, K., & Koren, H. (1995). Tools for interactive map conversion and vectorization. In *Proceedings of Third International Conference on Document Analysis and Recognition*, Montreal, Canada (Vol. *2*, pp. 927–930). Piscataway, NJ: IEEE. https://doi.org/10.5555/839278.840159

Erath, A., Löchl, M., & Axhausen, K. W. (2009). Graph-theoretical analysis of the Swiss road and railway networks over time. *Networks and Spatial Economics*, *9*(3), 379–400. https://doi.org/10.1007/s11067-008-9074-7

Fortier, M. F. A., Ziou, D., Armenakis, C., & Wang, S. (2001). Automated correction and updating of road databases from high-resolution imagery. *Canadian Journal of Remote Sensing*, *27*(1), 76–89. https://doi.org/10.1080/07038992.2001.10854922

He, H., Yang, D., Wang, S., Wang, S., & Li, Y. (2019). Road extraction by using atrous spatial pyramid pooling integrated encoder–decoder network and structural similarity loss. *Remote Sensing*, *11*(9), 1015. https://doi.org/10.3390/rs11091015

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Las Vegas, NV (pp. 770–778). Piscataway, NJ: IEEE.

Henderson, T. C. (2014). *Road and road intersection extraction*. New York, NY: Springer.

Henderson, T. C., & Linton, T. (2009). Raster map image analysis. In *Proceedings of the 10th International Conference on Document Analysis and Recognition*, Barcelona, Spain (pp. 376–380). Piscataway, NJ: IEEE. https://doi.org/10.1109/ICDAR.2009.31

Henderson, T. C., Linton, T., Potupchik, S., & Ostanin, A. (2009). Automatic segmentation of semantic classes in raster map images. In *Proceedings of the Eighth IAPR International Workshop on Graphics Recognition*, La Rochelle, France (pp. 1–10). Stroudsburg, PA: ACL.

Herold, H. (2015). *An evolutionary approach to adaptive image analysis for retrieving and long-term monitoring historical land use from spatiotemporally heterogeneous map sources* (Unpublished PhD dissertation). Dresden, Germany: Technische Universität Dresden.

Howman, C., & Woodsford, P. A. (1977). The Laser-scan Fastrak automatic digitising system. In *Proceedings of the Ninth International Cartography Conference*, Washington, DC (pp. 1–7). Washington, DC: ICA.

Itonaga, W., Matsuda, I., Yoneyama, N., & Ito, S. (2003). Automatic extraction of road networks from map images. *Electronics and Communications in Japan (Part II: Electronics)*, *86*(4), 62–72. https://doi.org/10.1002/ecjb.10144

Kakumoto, S., Miyatake, T., Shimada, S., & Ejiri, M. (1983). Development of auto-digitizer using multi-color drawings recognition method. *IEEE Conference on Computer Vision and Pattern Recognition*, Washington, DC (pp. 114–120). Piscataway, NJ: IEEE.

Kasturi, R., & Alemany, J. (1988). Information extraction from images of paper-based maps. *IEEE Transactions on Software Engineering*, *14*(5), 671–675. https://doi.org/10.1109/32.6145

Kasturi, R., Fernandez, R., Amlani, M. L., & Feng, W. C. (1989). Map data processing in geographic information systems. *Computer*, *22*(12), 10–21. https://doi.org/10.1109/2.42028

Kennie, T. J. (2014). *Engineering surveying technology*. Boca Raton, FL: CRC Press.

Leyk, S. (2009). Segmentation of colour layers in historical maps based on hierarchical colour sampling. In J. M. Ogier, W. Liu, & J. Lladós (Eds.), *Graphics recognition: Achievements, challenges, and evolution: GREC 2009* (Lecture Notes in Computer Science, Vol. *6020*, pp. 231–241). Berlin, Germany: Springer. https://doi.org/10.1007/978-3-642-13728-0_21

Leyk, S., & Boesch, R. (2010). Colors of the past: Color image segmentation in historical topographic maps based on homogeneity. *GeoInformatica*, *14*(1), 1. https://doi.org/10.1007/s10707-008-0074-z

Leyk, S., Boesch, R., & Weibel, R. (2005). A conceptual framework for uncertainty investigation in map-based land cover change modelling. *Transactions in GIS*, *9*(3), 291–322. https://doi.org/10.1111/j.1467-9671.2005.00220.x

Li, J., Hu, Q., & Ai, M. (2018). Unsupervised road extraction via a Gaussian mixture model with object-based features. *International Journal of Remote Sensing*, *39*(8), 2421–2440. https://doi.org/10.1080/01431161.2018.1425563

Linton, T. F. (2009). *Semantic feature analysis in raster maps* (Unpublished PhD dissertation). Salt Lake City, UT: University of Utah.

Liu, T., Xu, P., & Zhang, S. (2019). A review of recent advances in scanned topographic map processing. *Neurocomputing*, *328*, 75–87. https://doi.org/10.1016/j.neucom.2018.02.102

Liu, Y. (2002). An automation system: Generation of digital map data from pictorial map resources. *Pattern Recognition*, *35*(9), 1973–1987. https://doi.org/10.1016/S0031-3203(01)00154-6

Lu, X., Zhong, Y., Zheng, Z., Liu, Y., Zhao, J., Ma, A., & Yang, J. (2019). Multi-scale and multi-task deep learning framework for automatic road extraction. *IEEE Transactions on Geoscience and Remote Sensing*, *57*(11), 9362–9377. https://doi.org/10.1109/TGRS.2019.2926397

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A. (2019). Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, *152*, 166–177. https://doi.org/10.1016/j.isprsjprs.2019.04.015

Masucci, A. P., Stanilov, K., & Batty, M. (2014). Exploring the evolution of London's street network in the information space: A dual approach. *Physical Review E*, *89*(1), 012805. https://doi.org/10.1103/PhysRevE.89.012805

Mayer, H., & Steger, C. (1998). Scale-space events and their link to abstraction for road extraction. *ISPRS Journal of Photogrammetry and Remote Sensing*, *53*(2), 62–75. https://doi.org/10.1016/S0924-2716(97)00040-3

Miyatake, T. (1985). Extraction of roads from topological maps using a parallel extraction algorithm. *Systems and Computers in Japan*, *16*(6), 77–86. https://doi.org/10.1002/scj.4690160609

Mosinska, A., Marquez-Neila, P., Koziński, M., & Fua, P. (2018). Beyond the pixel-wise loss for topology-aware delineation. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, Salt Lake City, UT (pp. 3136–3145). Piscataway, NJ: IEEE.

Moučka, B. M. (2018). *Semi-automatic tools for image segmentation* (Unpublished MS thesis). Brno, Czech Republic: Masaryk University.

Nagao, T., Agui, T., & Nakajima, M. (1988). An automatic road vector extraction method from maps. In *Proceedings of the Ninth International Conference on Pattern Recognition*, Rome, Italy (pp. 585–587). Piscataway, NJ: IEEE. https://doi.org/10.1109/ICPR.1988.28299

Nagao, T., Agui, T., & Nakajima, M. (1990). Automatic extraction of roads denoted by parallel lines from 1/25,000-scaled maps utilizing skip-scan method. *Systems and Computers in Japan*, *21*(11), 96–105. https://doi.org/10.1002/scj.4690211110

Nakajima, M., Agui, T., & Iituka, H. (1985). A graphical structure extracting method from an urban map using parallel vector tracers. *Systems and Computers in Japan*, *16*(6), 39–48. https://doi.org/10.1002/scj.4690160605

Nazeri, K., Ng, E., Joseph, T., Qureshi, F. Z., & Ebrahimi, M. (2019). *Edgeconnect: Generative image inpainting with adversarial edge learning*. Preprint, arXiv:1901.00212.

Nishijima, M., & Watanabe, T. (1998). A cooperative inference mechanism for extracting road information automatically. In R. Chin & T. C. Pong (Eds.), *Computer vision: ACCV 1998* (Lecture Notes in Computer Science, Vol. *1352*, pp. 217–224). Berlin, Germany: Springer. https://doi.org/10.1007/3-540-63931-4_218

Pezeshk, A. (2011). *Feature extraction and text recognition from scanned color topographic maps* (Unpublished PhD dissertation). University Park, PA: The Pennsylvania State University.

Pezeshk, A., & Tutwiler, R. L. (2010). Improved multi angled parallelism for separation of text from intersecting linear features in scanned topographic maps. In *Proceedings of the 2010 IEEE International Conference on Acoustics, Speech and Signal Processing*, Dallas, TX (pp. 1078–1081). Piscataway, NJ: IEEE. https://doi.org/10.1109/ICASSP.2010.5495342

Pezeshk, A., & Tutwiler, R. L. (2011). Automatic feature extraction and text recognition from scanned topographic maps. *IEEE Transactions on Geoscience and Remote Sensing*, *49*(12), 5047–5063. https://doi.org/10.1109/TGRS.2011.2157697

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-net: Convolutional networks for biomedical image segmentation. In N. Navab, J. Hornegger, W. Wells, & A. Frangi (Eds.), *Medical image computing and computer-assisted intervention: MICCAI 2015* (Lecture Notes in Computer Science, Vol. *9351*, pp. 234–241). Cham, Switzerland: Springer.

Saeedimoghaddam, M., & Stepinski, T. F. (2020). Automatic extraction of road intersection points from USGS historical map series using deep convolutional neural networks. *International Journal of Geographical Information Science*, *34*(5), 947–968. https://doi.org/10.1080/13658816.2019.1696968

Simonyan, K., & Zisserman, A. (2014). *Very deep convolutional networks for large-scale image recognition*. Preprint, arXiv:1409.1556.

Steger, C. (1998). An unbiased detector of curvilinear structures. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *20*(2), 113–125. https://doi.org/10.1109/34.659930

Steger, C. (2000). Subpixel-precise extraction of lines and edges. *International Archives of Photogrammetry and Remote Sensing*, *33*(3), 141–156.

Stevenson, P. J. (1994). *Scanning for automated data conversion of cartographic documents*. Columbus, OH: Department of Geodetic Science and Surveying, Ohio State University.

Strano, E., Nicosia, V., Latora, V., Porta, S., & Barthélemy, M. (2012). Elementary processes governing the evolution of road networks. *Scientific Reports*, *2*, 296. https://doi.org/10.1038/srep00296

Sun, T., Di, Z., Che, P., Liu, C., & Wang, Y. (2019). Leveraging crowdsourced GPS data for road extraction from aerial imagery. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Long Beach, CA (pp. 7509–7518). Piscataway, NJ: IEEE.

Suzuki, S., Kosugi, M., & Hoshino, T. (1987). Automatic line drawing recognition of large scale maps. *Optical Engineering*, *26*(7), 267642. https://doi.org/10.1117/12.7974129

Suzuki, S., & Yamada, T. (1990). MARIS: Map recognition input system. *Pattern Recognition*, *23*(8), 919–933. https://doi.org/10.1016/0031-3203(90)90137-A

Tao, C., Qi, J., Li, Y., Wang, H., & Li, H. (2019). Spatial information inference net: Road extraction using road-specific contextual information. *ISPRS Journal of Photogrammetry and Remote Sensing*, *158*, 155–166. https://doi.org/10.1016/j.isprsjprs.2019.10.001

Wang, J., Song, J., Chen, M., & Yang, Z. (2015). Road network extraction: A neural-dynamic framework based on deep learning and a finite state machine. *International Journal of Remote Sensing*, *36*(12), 3144–3169. https://doi.org/10.1080/01431161.2015.1054049

Wang, S., Yu, D., Kwan, M. P., Zhou, H., Li, Y., & Miao, H. (2019). The evolution and growth patterns of the road network in a medium-sized developing city: A historical investigation of Changchun, China, from 1912 to 2017. *Sustainability*, *11*(19), 5307. https://doi.org/10.3390/su11195307

Wang, W., Yang, N., Zhang, Y., Wang, F., Cao, T., & Eklund, P. (2016). A review of road extraction from remote sensing images. *Journal of Traffic and Transportation Engineering*, *3*(3), 271–282. https://doi.org/10.1016/j.jtte.2016.05.005

Watanabe, T., & Oshitani, T. (2001). Parallel recognition of roads from urban maps on generation/verification paradigm of hypotheses. In *Proceedings of Sixth International Conference on Document Analysis and Recognition*, Seattle, WA (pp. 1225–1234). Piscataway, NJ: IEEE. https://doi.org/10.1109/ICDAR.2001.953978

Wegner, J., Montoya-Zegarra, J., & Schindler, K. (2013). A higher-order CRF model for road network extraction. In *Proceedings of the 26th Conference on Computer Vision and Pattern Recognition*, Portland, OR (pp. 1699–1706).

Wegner, J. D., Montoya-Zegarra, J. A., & Schindler, K. (2015). Road networks as collections of minimum cost paths. *ISPRS Journal of Photogrammetry and Remote Sensing*, *108*, 128–137. https://doi.org/10.1016/j.isprsjprs.2015.07.002

Wei, Y., Wang, Z., & Xu, M. (2017). Road structure refined CNN for road extraction in aerial image. *IEEE Geoscience and Remote Sensing Letters*, *14*(5), 709–713. https://doi.org/10.1109/LGRS.2017.2672734

Wulamu, A., Shi, Z., Zhang, D., & He, Z. (2019). Multiscale road extraction in remote sensing images. *Computational Intelligence and Neuroscience*, *2019*, 2373798. https://doi.org/10.1155/2019/2373798

Yamada, H., Yamamoto, K., & Hosokawa, K. (1993). Directional mathematical morphology and reformalized Hough transformation for the analysis of topographic maps. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *15*(4), 380–387. https://doi.org/10.1109/34.206957

Yamada, H., Yamamoto, K., & Nakamura, M. (1990). Increasing the performance of MAP (multi-angled parallelism) erosion-dilation for feature extraction through unary/binary operations. In *Proceedings of the 10th International Conference on Pattern Recognition*, Atlantic City, NJ (Vol. *2*, pp. 524–528). Piscataway, NJ: IEEE. https://doi.org/10.1109/ICPR.1990.119419

Yamada, H., Yamamoto, K., Saito, T., & Matsui, S. (1991). MAP: Multi-angled parallelism for feature extraction from topographical maps. *Pattern Recognition*, *24*(6), 479–488. https://doi.org/10.1016/0031-3203(91)90015-W

Yang, Y., An, X., & Huang, L. (2012). Vectorization of linear features in scanned topographic maps using adaptive image segmentation and sequential line tracking. *International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences*, *39*(B4), 103–108. https://doi.org/10.5194/isprsarchives-XXXIX15B4-103-2012

Yin, P. Y., & Huang, Y. B. (2001). Automating data extraction and identification on Chinese road maps. *Optical Engineering*, *40*(5), 663–673. https://doi.org/10.1117/1.1357831

Zhang, C. (2003). *Updating of cartographic road databases by image analysis* (Unpublished PhD dissertation). Zürich, Switzerland: ETH Zürich.

Zhang, C., Baltsavias, E., & O'Sullivan, L. (2005). Performance evaluation of ATOMI system for road database updating from aerial film, ADS40, IKONOS and Quickbird orthoimagery. *International Archives of Photogrammetry, Remote Sensing, and Spatial Information Sciences*, *36*(3/W24), 165–170. https://doi.org/10.1.1.222.5047

Zhang, J., Hu, Q., Li, J., & Ai, M. (2020). Learning from GPS trajectories of floating car for CNN-based urban road extraction with high-resolution satellite imagery. *IEEE Transactions on Geoscience and Remote Sensing*, *59*(3), 1836–1847. https://doi.org/10.1109/TGRS.2020.3003425

Zhao, P., Jia, T., Qin, K., Shan, J., & Jiao, C. (2015). Statistical analysis on the evolution of OpenStreetMap road networks in Beijing. *Physica A: Statistical Mechanics and its Applications*, *420*, 59–72. https://doi.org/10.1016/j.physa.2014.10.0767
Zhou, L., Zhang, C., & Wu, M. (2018). D-linknet: Linknet with pre-trained encoder and dilated convolution for high resolution satellite imagery road extraction. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Salt Lake City, UT (pp. 182–186). Piscataway, NJ: IEEE.

**How to cite this article:** Jiao, C., Heitzler, M., & Hurni, L. (2021). A survey of road feature extraction methods from raster maps. *Transactions in GIS*, 25, 2734–2763. https://doi.org/10.1111/tgis.12812

## APPENDIX

Word cloud of technique terms included in the reviewed papers.