SKELETON-BASED AUTOMATIC ROAD NETWORK EXTRACTION FROM AN ORTHOPHOTO COLORED POINT CLOUD

Elyta Widyaningrum (1)(2), Roderik C. Lindenbergh (1)

[1] Geoscience and Remote Sensing, Delft University of Technology, the Netherlands
[2] Geospatial Information Agency, Bogor, Indonesia
Email: e.widyaningrum@tudelft.nl; r.c.lindenbergh@tudelft.nl

KEY WORDS: Road network, skeleton, aerial image, colored point cloud, regularization.

ABSTRACT: Reliable and up-to-date road network information is crucial to guarantee efficient logistic distribution, emergency response, urban planning, etc. Road networks in developing urban areas tend to change rapidly. Periodic remapping is necessary to maintain the temporal quality of the road network information. Updating the road network using conventional methods can be a tedious task. This paper presents a methodology to extract road network automatically from an airborne LiDAR point cloud combined with color information from an aerial orthophoto. First, ground points are separated from non-ground points. We then classify the filtered ground points to road and non-road points using the Random Forest (RF) algorithm. Parallel thinning, method for skeletonization of the road segment, is carried out on a binary image extracted by a so-called density map of the classified road points. Finally, road centerline is obtained by our proposed topological order and regularization approach. The proposed method is tested on ISPRS benchmark data of Vaihingen - Germany. Skeleton-based road network extraction is a promising method as more than 95% roads in the study area are extracted. In the future, regularization of the skeleton to obtain smoother line representation is still an essential but challenging research.

1. BACKGROUND

Road networks are an important foundation for predicting and controlling urban development. In developing cities, road networks change rapidly and expand continuously. Updating road network maps is an inevitable task to accommodate many applications such as transportation, land use planning, emergency response, etc. However, obtaining an accurate and reliable road network map periodically can be time consuming and labor intensive. The fact that road network extraction involving complex network topology and has different representation (color, size, shape, and neighboring area) make this research remain challenging, especially when using high resolution remote sensing data.

Airborne Lidar and aerial images are two dominant remote sensing data sources used for mapping. Airborne Lidar point clouds provide robust and dense 3D representation of objects including intensity and multiple echo information. However, such 3D coordinate and intensity information may not be enough to provide semi-automatic and automatic processing such as object classification, detection, or extraction. On the other hand, image provides color information that is also essential for object detection or extraction. Combination of both point cloud and image color information may increase accuracy of the extraction results.

Road networks have specific structural and geometrical characteristics. Skeletonization provides a compact shape representation intuitively while maintaining topology, geometry and scale of the input shape (Stavros and Dickinson, 2017; Elber et al., 2004; Vanajakshi et al., 2010). A skeleton is used to reduce the dimensionality of an object. It should be centered within the object, be thin, have the same topology and geometry as the object and allow for complete recovery of the original object (Saha el at., 2016). Considering benefits of skeletonization, application for road network

extraction is promising. The main goal of this study is to provide a procedure for extracting road networks automatically by using orthophoto colored airborne point clouds. The expected output is road centerline and road polygon in vector format.

2. RELATED WORK

More than two decades of study on the road network extraction have been carried out (Wagner et al., 2013). Wang et al (2016) review works on road classification from remote sensing images using ANN, SVM, Markov Random Field, Mean Shift, mathematical morphology and active contour model. They specify the correctness as well as the disadvantages and dependencies of each method. However, many studies on the automatic road extraction from remote sensing data focus more in the classification processing without considering the extraction of a smooth and GIS-ready data. Alshehhi et al. (2017) propose a single patch-based Convolutional Neural Network (CNN) architecture for roads and buildings extraction from high-resolution remote sensing data. The experiments prove that the method achieves good performance in the localization of urban objects (80.9% correctness for Abu Dhabi area and 91.7% correctness for Massachusetts area), but it required additional processing to outline boundaries more accurately.

A recent trend is to use machine learning techniques to classify point clouds of urban area (Lodha et al., 2006). Supervised machine learning techniques are based on hand-crafted features and a classifier algorithm. In literature, a variety of supervised classification techniques, including support vector machines and neural networks are discussed. Niemeyer et al. (2012) classified the airborne LiDAR point cloud of ISPRS Vaihingen benchmark by using Conditional Random Fields. Geometric features and intensity value are used to distinguish the five object classes: building, low vegetation, tree, terrain, and asphalt ground using a Conditional Random Field (CRF) approach. However, they only showed and evaluated the result for the building and tree classes. Guo et al. (2010) use a Random Forest (RF) algorithm to classify LiDAR data in urban area in four classes (building, vegetation, artificial ground, and natural ground) from the combination of optical multi spectral and LiDAR data. They stated the best five features are two optical image channels (Blue and Red), height differences, echo-amplitude and echo-cross section. However, the accuracy of the natural ground and vegetation result is still need to be improved (70% and 72% respectively). Many other studies using the RF algorithm to classify LIDAR data can be found in literature (Rodriguez-Galiano et al. 2012; Guan et al., 2013, Ma et al., 2017). Lodha et al. (2006) employed another LiDAR data classification work. The authors used Support Vector Machines (SVM) for classifying LiDAR data into buildings, trees, roads, and grass using five features: height, height variation, normal variation, LiDAR return intensity, and image intensity. They compare ground truth to their classification result resulting in 90% accuracy. Ma et al. (2017) compared the SVM and RF algorithm to classify LiDAR data. The authors classified data in four categories: trees, buildings, farmland, and ground. According to their findings, the RF algorithm gave a better result than the SVM algorithm for the classification of LiDAR point cloud data.

Skeletonization in 2D and 3D has been widely used in many computer vision and computer graphics applications including shape recognition and analyses, motion tracking, segmentation, etc (Saha et al., 2017). Several skeleton algorithms have been presented for different shape and input format (image and point cloud). Broersen et al. (2017) use 3D skeleton, also called as medial axis transform, to identify watercourses in a flat and engineered landscapes from geographic LiDAR point cloud. Peters investigate how the 3D skeleton can be applied for geographical point clouds comprehensively (2018). Haunert and Sester (2008) presented straight skeleton extraction to derive linear representations of buildings and road centerlines from a cadastral map for automatic generalization. Therefore, classification is not necessary for their study. Yirci et al. (2013) applied straight skeleton and medial axis to generate road centerline of pedestrian networks from by using two skeleton operators (straight skeleton by parallel thinning and medial axis by Voronoi diagram) from a road vector that was provided manually.

3. EXPERIMENTS

This study provides an approach to extract road network (centerlines) automatically from a colored airborne point cloud. The general framework of this study is given in Figure 1. Ground filtering of the point cloud data is first applied to separate ground and non-ground points. Road points are then detected by the Random Forest algorithm using four features: intensity, infrared, red and green channels. Outliers in the classified road points data are removed by DBSCAN clustering and a size threshold. A parallel thinning algorithm is used to get the skeleton of the road from a density map. The final road network is obtained by regularization of the skeleton result.

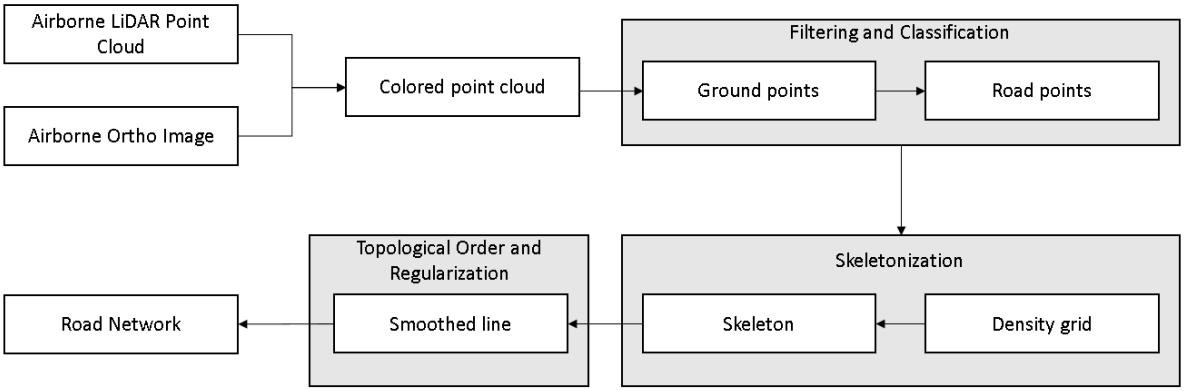


Figure 1. General framework of the study.

3.1 Study Area

We evaluate our approach on an ISPRS benchmark dataset of Vaihingen – Germany. The size of the study area is about 658 by 656 m and consists of residential buildings surrounded by trees. The Vaihingen dataset we used consists of two main data: an airborne LiDAR point cloud and a true orthoimage. The point density of the airborne LiDAR point cloud of the study area (Figure 2.a) is 6 – 8 points per meter. The total number of points of the study area is 3 332 278 points. The intensity image of the airborne point cloud is shown in Figure 2.b . It can be seen that in general, road points have low intensity value. A true orthoimage of the study area at 9 cm spatial resolution is presented in Figure 2.c. This false color images combines the three available bands, Near-Infrared, Red, and Green. In the orthoimage, road points typically have a grey color.

a. Airborne point cloud colored by height

b. LiDAR intensity

c. Airborne true orthoimage (NIR, Red, Green)

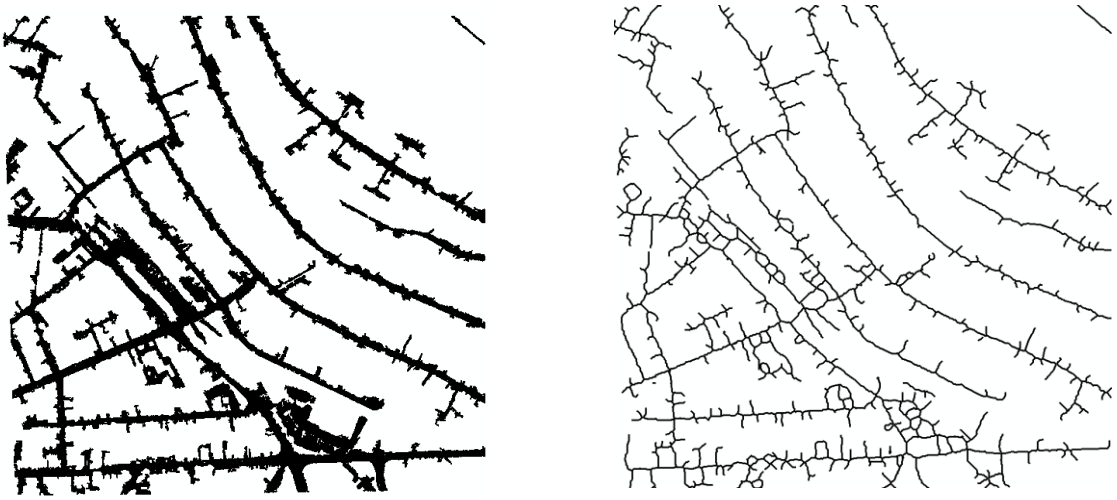
Figure 2. Dataset of the study area. ©ISPRS.

3.2 Ground Filtering and Road Classification

Prior to classification, a filtering procedure is applied to separate ground points from non-ground points. Ground filtering is performed using LAStools with default settings for town or flats.
Random Forest (RF) classification is reported to be less sensitive to the quality of training samples and to overfitting than other machine learning methods [Belgiu and Dragut, 2016; Zhou et al, 2019]. The RF algorithm consists of many decision trees (Breiman, 2001). This algorithm uses two parameters: the number of trees *T* in the forest and the number of randomly selected features *M* at each node. When building a tree, training data points are sampled randomly from the full training dataset. Furthermore, only a random subset of predictive features *M* is chosen for splitting each node in each decision tree *DT*. This will decrease the strength of individual trees, but will lead to a lower correlation between trees and smaller generalization error. RF uses bootstrapping that allows several samples are used multiple times in a single tree during training. Stability is achieved as the entire forest has low variance, which increase the classification accuracy. The final prediction is determined by a majority vote of all individual trees.
Reducing *M* reduces both the correlation and the strength. But, the generalization error converges as the number of trees *T* increases. Therefore, RF tends not to over fit the data.
We implement a point-based classification using the RF algorithm using four features, LiDAR intensity and three image channels (NIR, Red, and Green). The result is a set of points divided into two classes (road and non-road points). The classifier has been trained with manually labeled points using $T = 50$ decision trees. The number of features *M* at each node was set to 2, which is the square root of the number of input variables (4 in total). Labeled reference points for training are obtained manually. The ground truth (of 600000 points) data was divided randomly into 2/3 and 1/3 for training and testing, respectively.
Next, DBSCAN clustering followed by size filtering is applied to remove outliers and small road segments.

3.3 Skeletonization

To get the skeleton of the classified road points, a masked image of the road network is created first. A masked image is a binary image discriminating road and non-road pixels based on a density grid. A density grid is a method to visualize scattered point data by transforming irregular points into a regular grid as a function of nearby points. This means, each grid has a value determined by the density of neighboring points. For transforming road points into a raster, we apply a Epanechnikov kernel density estimation with 1 meter radius search. This kernel was introduced by Epanechnikov (1996) and is described as the most efficient kernel function in terms of minimizing the mean square error. Although the difference compared to other kernels may be relatively small, the ability of the Epanechnikov kernel to adapt position of the point to other data points can greatly increase the accuracy of the estimation (Riksen, 2014).
A raster-based skeletonization is preferred as it directly results in a skeleton in a line format. Thus rasterization is applied. The kernel density map is used instead of interpolation to exclude the outside area of the road. Besides transforming points into a raster and regularizing the boundary, a density grid is useful to fill in gaps on the road due to some obstructions such as cars or trees. Figure 3.a shows a binary image of road (black) and non-road (white) pixels resulting from a kernel density gridding.



a. Rasterized road points from the density grid

b. Skeleton result of the study area

Figure 3. Road area and skeleton in a binary image.

The skeleton of the road network is obtained using the fast parallel thinning algorithm introduced by Zhang and Suen (1984). The basic principle of the algorithm is to remove all pixels starting from the boundary of the shape except those pixels belonging to the skeleton. The algorithm requires a binary input image and preserves the skeleton connectivity by using sub-iteration for each boundary removal iteration.

Figure 3 illustrates an example of the thinning algorithm. The black pixels, as shown in Figure 4.a, represent the object to be processed. First, the algorithm searches for all black pixel neighbors of query pixel P1 in a 3×3 window as arranged in Figure 4.b.

The first sub-iteration checks any black pixel P1 to be removed or set to 0 if it satisfies all conditions as follows:

C.1 Pixel P1 is black and has eight neighbors;

C.2 $2 \leq S(P1) \leq 6$

C.3 Number of transition from white to black ($0 \rightarrow 1$) in clockwise ordered circular sequence (P2,P3,P4,P5,P6,P7,P8,P9,P2) = 1

C.4 at least one of (P2, P4, P6) is white ($P2 \times P4 \times P6 = 0$)

C.5 at least one of (P4, P6, P8) is white ($P4 \times P6 \times P8 = 0$)

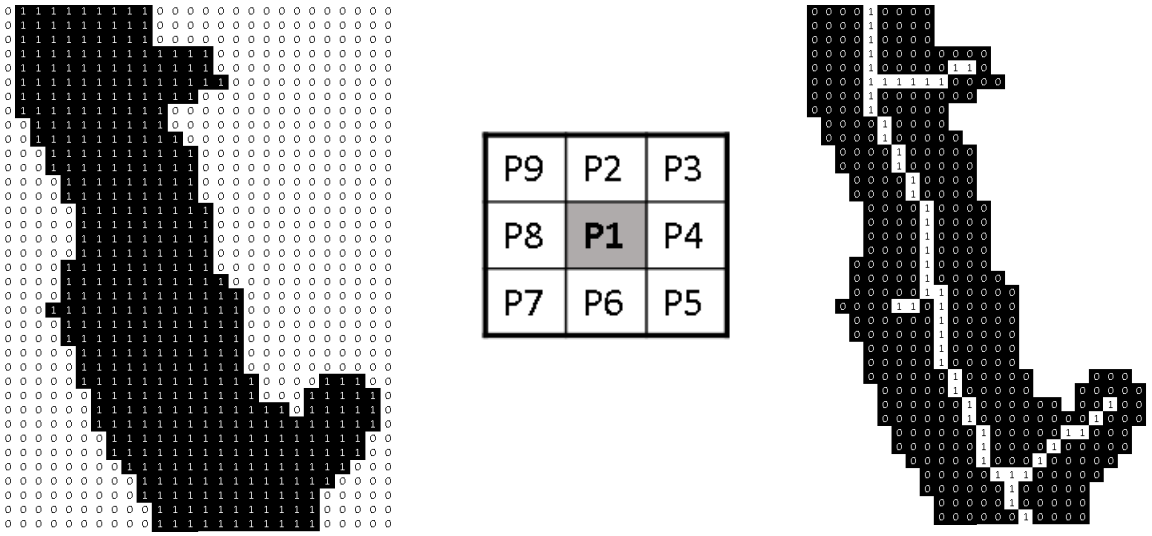
The second sub-iteration re-examines each pixel and set its value to 0 that satisfies all the C.1 to C.3 and C.4' and C.5'.

C.4 ' at least one of (P2, P4, P8) is white ($P2 \times P4 \times P8 = 0$)

C.5' at least one of (P2, P6, P8) is white ($P2 \times P6 \times P8 = 0$)

S(P1) is the number of black (non-zero) neighboring pixels of P1.

Figure 4.c shows the result of the parallel thinning algorithm on a binary image. The non-skeleton pixels (black pixels with 0) are removed or, in this case their value has been changed from 1 to 0, until the skeleton pixels (white pixel with 1) remain.



a. binary image b. neighbors of P1 in 3 × 3 window c. Skeleton obtained by parallel thinning

Fig 4. Skeletonization obtained by parallel thinning of the binary shape (black: road pixel; white: skeleton result).

3.4 Topological order and regularization

Resulting skeleton lines shown in green in Figure 5.a may still be noisy and jaggy and not meet the criteria for representing the road network on a map. Therefore, a regularization step to get a smoother line from the resulted skeleton is required. This smoother line is used later as final road centerline. In this study, we consider a road network hierarchy consisting of main roads and branch roads. We propose a topological network approach based on connectivity to generalize the skeleton lines by extracting network points consisting of end-points, branch points, and middle-points. Typically, end-points are extracted at both ends of a line segment. A point located at the end of a line that is not intersecting another line is defined as end-point. If a network point intersects another line, it is called a branch-point. A middle-point is a point on a line halfway between two network points (either end-point or branch-point).

For preserving the road topology and generalizing the skeleton, we define a hierarchical road order based on the network connectivity. An order-1 road is the main road. A main road is defined as a line connecting at least two branch-points. Order-2 roads, the branch roads, are lines connecting a branch-point to an end-point.

The algorithm extracts the main road centerline by connecting its ordered middle-points. Then, a branch road centerline is created by connecting end-points and the line is extended until it connects to the main road.

Figure 5 shows an example of a road network topology. End-points $E3_{33,34,35}$, $E3_{35,36,37}$ and $E3_{37,38,39}$ are intersections of 3 lines, L33-L34-L35, L35-L36-L37 and L37-L38-L39 respectively.

$E1_{34}$, $E1_{36}$ and $E1_{38}$ are end-points. In this case, the main (Order-1) road is made by connecting M39-M37-M35-M33. The branch road (order-2) centerline is then extracted by extending a line from its end-point to the main road. The three branch roads in this example are: $E1_{34}$-M34, $E1_{36}$-M36, and $E1_{38}$-M38.

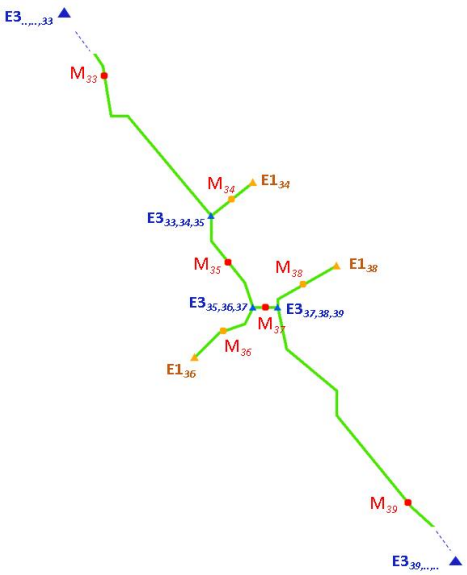
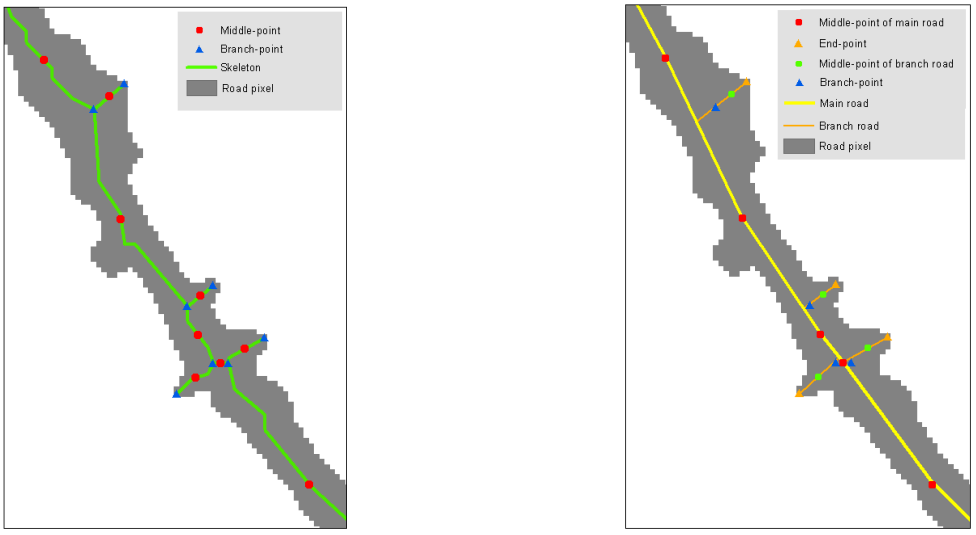


Figure 5. Regularization of a skeleton-based line.

By connecting the points based on their topological order, a smoother line can be obtained. Figure 6.a shows a skeleton-based line network (green) of a subset area including the initial network points (red points marked middle-points and blue triangles mark end-points). After the topological order and regularization process, a smoother road line representation is achieved as indicated as yellow line in Figure 6.b. As the final step, we build a polygon of the main road (order 2) based on the smoothed centerline by assuming that the main road has the same road width. The road width is determined by a prior measurement of the average width in the study area.

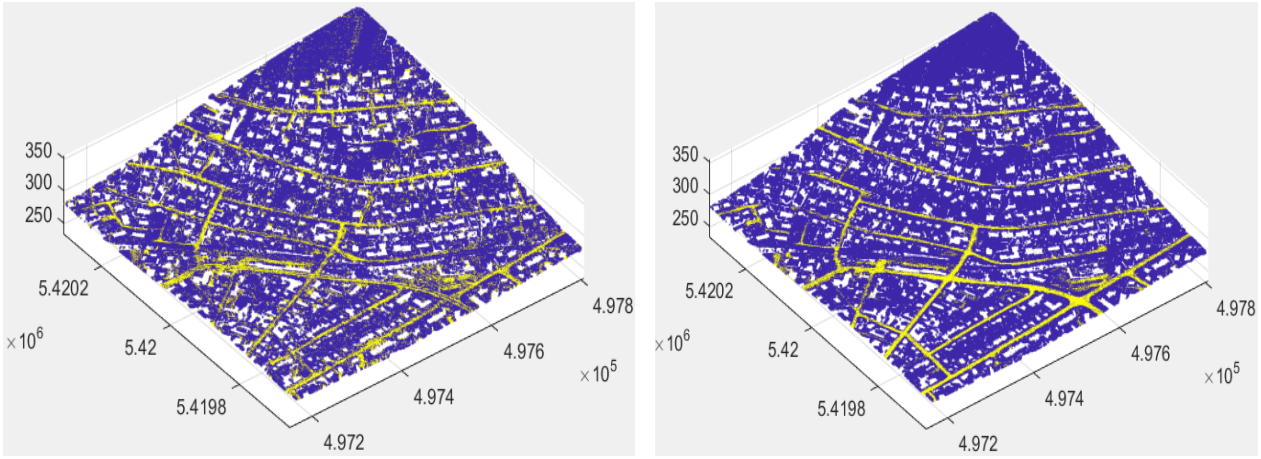


a. Skeleton (green) b. Skeleton-based centerline result (yellow)

Figure 6. Regularization of skeleton-based road line.

4. RESULTS AND DICUSSION

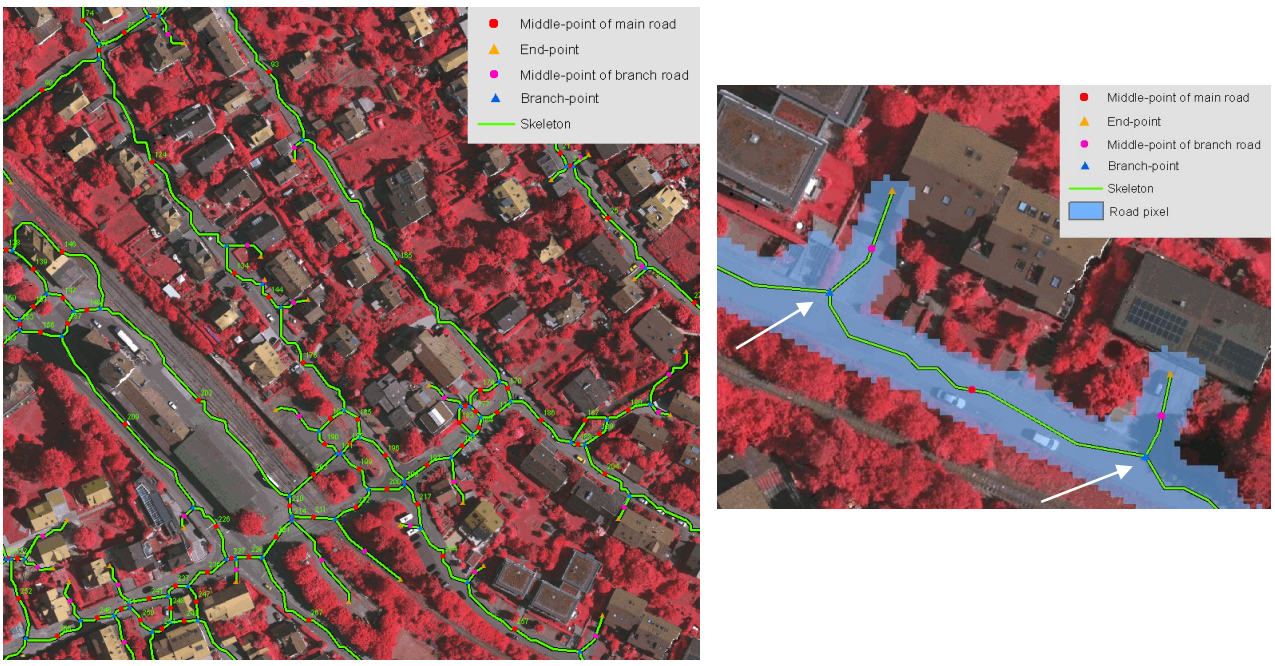
We use a set of labelled ground points for classifying road and non-road points. Our RF classification results show that the combination of RGB and intensity features provides better accuracy than using RGB only. Overall accuracy of RF classification result of the study area using only intensity is 81.03% while combination with RGB features is 91.91%. Figure 7 shows the comparison of the RF classification results using RGB and RGBI features. The RF ability to achieve decent accuracy combined with efficient computation makes it suitable for point cloud classification.



a. RGB b. RGBI

Figure 7. Comparison of the Random Forest classification result (yellow marks road points and blue marks non-road points).

As seen in Figure 8, all major roads in the study area are detected. However, the resulting road network has more branches than the ground truth. These branches include several areas of similar characteristics a main road such as parking areas, house yards, and small alleys. Although such roads connecting to yards or other free space may not be considered a road, we keep these as branch road. As seen in Figure 9.a, several house yards connecting to the road are also detected as road (inside the white ellipses). The skeleton shortcomings due to methods sensitivity to noise made the centerline result not as straight as it should be. This fuzzy line could also have shifted end-points and branch-points. As indicated by white arrows in Figure 8.b, branch point location deviate from the road centerline. In this case, the branch-point offset is caused by the estimated thickness of the road due to variations of surrounding area, e.g. trees or open yards. On the other hand, middle-points in our study area are in general located at a correct location. Therefore, the use of middle-points for regularization is appropriate.



a. The skeleton-based road and the network points b. Branch-points deviate from road centerline (showed by white arrows)

Figure 8. Result of skeleton-based road and network points of the study area

Extracting all network points of the skeleton-based line segment is essential for determining the road topologic order. As shown in Figure 9.b, the main road (yellow line) and branch roads (orange lines) are detected and regularized. For this road segment, the maximum difference of the centerline result to the centerline reference is 1.75 meter (as indicated by green circle in Figure 9.c). However, the regularization approach for the road network presented in this study is not implemented yet for the full complex network of the study area.

a. Skeleton as road centerline before regularization

b. Road centerline after regularization

c. Comparison of road centerline result to the ground truth

Figure 9. Zoomed in skeleton-based road result.

5. CONCLUSION AND FURTHER RESEARCH

This study presents a procedure to extract a road network from a colored point cloud using a skeleton-based approach. The colored point cloud is first filtered and classified using a Random Forest algorithm. Next, the classified road points are gridded using kernel density to obtain a binary image is carried out. This binary image is skeletonized by applying a parallel thinning algorithm. In the end, a connectivity-based approach is carried out to define the road network topological order and to regularize of the skeleton to obtain a smoother and schematic road line representation.
The proposed method shows promising results for extracting a road network automatically from the combination of point clouds and image. However, the proposed skeleton-based road network may not always be accurate especially for complex roads connecting to areas of similar characteristics (such as parking lots or shopping or railroad yards). In such areas, the skeletal shape may deviate from reality. Inaccurate classification results also affect the resulting skeleton. Therefore, line regularization considering topological order within the network will be studied further.

Acknowledgement:
The authors acknowledge Indonesian Geospatial Agency (www.big.go.id) for providing the Makassar datasets in this research. The authors gratefully acknowledge support from the Indonesia Endowment Fund for Education (LPDP), Ministry of Finance of Republic of Indonesia for a scholarship support to the first author.

References:

Alshehhi, R., Marpu, P. R., Woon, W. L., Mura, M. D., 2017. Simultaneous extraction of roads and buildings in remote sensing imagery with convolutional neural networks. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 130, pp. 139-149.

Belgiu, M., and Drăguţ, L., 2016. Random forest in remote sensing: A review of applications and future directions. ISPRS Journal of Photogrammetry and Remote Sensing, 114, 24-31.

Breiman, L., 2001. Random Forests, Machine Learning, Vol. 45, pp. 5-32.

Broersen, T., Peters, R., and Ledoux, H., 2017. Automatic identification of watercourses in flat and engineered landscapes by computing the skeleton of a LiDAR point cloud**.** Computers & Geosciences, 106, pp. 171–180.

Epanechnikov, V., 1966. Nonparametric Estimation of a Multidimensional Probability Density, Theory of Probability and its Applications, 14, pp.153-158.

Guan, H., Li, J., Chapman, M., Deng, F., Ji, Z. and Yang, X., 2013. Integration of orthoimagery and LiDAR data for object-based urban thematic mapping using random forests. International Journal of Remote Sensing, vol. 34, Issue 14, pp. 5166-5186

Guo, L., Chehata, N., Mallet, C., and Boukir, S., 2011. Relevance of airborne LiDAR and multispectral image data for urban scene classification using random forests. ISPRS Journal of Photogrammetry and Remote Sensing, Vol. 66, pp. 56–66.

Haunert, J. H., and Sester, M., 2008. Area collapse and road centerlines based on straight skeletons, GeoInformatica, 12(2), pp. 169-191.

Ma L., Zhou M.and Li C., 2017. Land covers classification based on Random Forest method using features from full-waveform LiDAR data. The International Archives of the Photogrammetry Remote Sensing and Spatial Information Sciences, Volume XLII-2/W7, ISPRS Geospatial Week 2017, 18–22September 2017, Wuhan, China.

Niemeyer, J., Rottensteiner, F., and Soergel, U., 2012. Conditional random fields for LiDAR point cloud classification in complex urban areas, ISPRS Annals Photogrammetry Remote Sensing and Spatial Information Science, I-3, pp. 263-268.

Peters, R., 2018. Geographical point cloud modelling with the 3D medial axis transform. PhD thesis, Delft University of Technology.

Saha, P. K., Borgefors, G., and di Baja, G. S., 2016. A survey on skeletonization algorithms and their applications. Pattern Recognition Letters, 76, pp. 3-12.

Saha, P. K., Borgefors, G., and di Baja, G. S., 2017. Skeletonization: Theory, methods and applications. Academic Press.

Stavros, T., and Dickinson, S., 2017. AMAT: Medial axis transform for natural images. Proceedings of the IEEE International Conference on Computer Vision.

Vanajakshi, B., Sujatha, B., and Sri Rama Krishna, K., 2010. An analysis of thinning and skeletonization for shape representation. International Journal of Computer Communication and Information System (IJCCIS), Vol. 2, No. 1.

Wegner, J. D, Montoya-Zegarra, J. A., and Schindler, K., 2013. A higher-order CRF model for road network extraction. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 23 – 28 June 2013, Portland, USA.

Yirci, M., Bredif, M., Perret, J., and Paparoditis, N., 2013. 2D Arrangement-based hierarchical spatial partitioning: An application to pedestrian network generation. Proceedings of the Sixth ACM SIGSPATIAL International Workshop on Computational Transportation Science, pp.31–36.

Zhang, T. Y., and Suen, C. Y., 1984. A fast parallel algorithm for thinning digital patterns. Communications of the ACM, Volume 27, Number 3.

Zhou, K., Lindenbergh, R., and Gorte, B., 2019. Automatic shadow detection in urban very-high-resolution images using existing 3D models for free training. Remote Sensing 11(1), 72.