Article

Identification of Inundation Using Low-Resolution Images from Traffic-Monitoring Cameras: Bayes Shrink and Bayesian Segmentation

Yin Chao Wu ¹, Seong Jin Noh ^{2,*} and Suyun Ham ^{1,*}

¹ Department of Civil Engineering, The University of Texas at Arlington, Arlington, TX 76019, USA; yinchao.wu@mavs.uta.edu

² Department of Civil Engineering, Kumoh National Institute of Technology, Gumi, Gyeongbuk 39177, Korea

* Correspondence: seongjin.noh@kumoh.ac.kr (S.J.N.); s.ham@uta.edu (S.H.); Tel.: +1-817-272-5217 (S.H.)

Received: 14 May 2020; Accepted: 12 June 2020; Published: 17 June 2020

Abstract: This study presents a comparative assessment of image enhancement and segmentation techniques to automatically identify the flash flooding from the low-resolution images taken by traffic-monitoring cameras. Due to inaccurate equipment in severe weather conditions (e.g., raindrops or light refraction on camera lenses), low-resolution images are subject to noises that degrade the quality of information. De-noising procedures are carried out for the enhancement of images by removing different types of noises. For the comparative assessment of de-noising techniques, the Bayes shrink and three conventional methods are compared. After the de-noising, image segmentation is implemented to detect the inundation from the images automatically. For the comparative assessment of image segmentation techniques, k-means segmentation, Otsu segmentation, and Bayesian segmentation are compared. In addition, the detection of the inundation using the image segmentation with and without de-noising techniques are compared. The results indicate that among de-noising methods, the Bayes shrink with the thresholding discrete wavelet transform shows the most reliable result. For the image segmentation, the Bayesian segmentation is superior to the others. The results demonstrate that the proposed image enhancement and segmentation methods can be effectively used to identify the inundation from low-resolution images taken in severe weather conditions. By using the principle of the image processing presented in this paper, we can estimate the inundation from images and assess flooding risks in the vicinity of local flooding locations. Such information will allow traffic engineers to take preventive or proactive actions to improve the safety of drivers and protect and preserve the transportation infrastructure. This new observation with improved accuracy will enhance our understanding of dynamic urban flooding by filling an information gap in the locations where conventional observations have limitations.

Keywords: inundation; road flooding; image processing; de-noising filtering; bayes shrink; traffic monitoring image; CCTV application

1. Introduction

In the face of natural disasters such as flash flooding, prompt information is crucial to establish a mitigation plan and find the best route for first responders. These rains cause unprecedented flooding and cause severe fatalities and hundreds of billions of US dollars in damages. Such an extreme flood not only damages roads and bridges but also cuts off evacuation routes and rescue paths. In many parts of the US, occurrences of “rare” extreme precipitation and flooding events are now a new normal [1].

There are different types of observations to monitor and detect floods in urban areas. Among them, typical measurement methods include in-situ water level sensors in streams, remote sensing

from satellites and airborne drones, on-site images from social media, and traffic-monitoring systems. Each observation type contributes to filling the information gap to grasp a holistic picture of urban flooding with different spatiotemporal scales. Despite the advance in measurement techniques, there are still limitations known for each measurement approach. For instance, in-situ water-level sensors are adapted to only stream monitoring. Instrumenting entire hydrological basins, which can cover hundreds of square kilometers, is practically and economically infeasible. Satellites are still limited to monitoring water levels and flow remotely with a low temporal frequency. Optical measurements from satellites and drones allow measuring only in a short period, and they are impossible during floods with severe weather conditions. For example, thick cloud layers interfere with the observation of satellites, while drones cannot fly when the wind is strong.

The vertical resolution of current synthetic aperture radars (tens of centimeters) with several days of the repeat cycle is insufficient for the task [2,3]. Without promising real-time flooding information, local citizens face the possible danger of tragedies. To address those limitations, effective, inexpensive, and reliable approaches are needed instead of installing additional facilities or equipment to detect flooding or inundation near civil infrastructure (e.g., highways and bridges). On-site images from traffic monitoring systems which are automatically photographed at regular times provide critical information for citizen and government, sometimes as the only reliable and practical sources, to identify the occurrence of flooding under extreme weather conditions in real-time, and protect people from exposure to danger for taking pictures in these conditions [4,5]. Despite these advantages, low-resolution images such as closed-circuit television (CCTV) photos are often corrupted by noises that can degrade image quality due to transmission or capture by inaccurate equipment or natural weather environments (e.g., raindrops or light refraction on camera lenses). The low-resolution images from traffic monitoring cameras or CCTV are one of the few reliable sources to know the outside conditions during extreme natural disasters [6,7].

The low resolution creates a dilemma to be overcome; for example, there are monitors in different places such as intersections or highways, but the resolution of CCTV is not fine enough for decision-makers to use the footage to decide to evacuate people to a safe place or guide a correct route for transportation. In addition, it is challenging to detect inundation from CCTV images using only a deep learning approach, which is one of the common applications of object detection. However, it requires many images sources in the same position but different situations to form the training database [8]. Unless the rain is heavy enough to quickly accumulate water on the road that can be recorded by erected equipment at the same time, it is difficult to obtain water levels changing with obvious differences in the image. From the context of the recognition target, detecting inundation from an image is different from the conventional image processing or target tracking procedure. In the conventional image processing, the water in an image may be regarded as noise or unnecessary background to be removed to better recognize common objects of interest such as cars or humans. For the detection of inundation, on the contrary, the edge of water bodies with changing boundaries and reflectivity is the object of interest to be recognized in while the other objects, but the water needs to be filtered out.

The de-noising filtering method is used to enhance the quality of an image by removing noises mixed in when the image is digitized and reconstructing a signal in the original image by extracting features in the image [9]. The impact of image noise can be decreased by changing the pixels to adjust brightness and contrast with the de-noising filtering methods (e.g., mean filtering and median filtering) [10]. In recent years, engineering communities have developed de-noising technologies [9,11]. Thresholding discrete wavelet transform (DWT) coefficients have been mainly studied for de-noising [12–14]. The wavelet is usually calculated by using spatial Gaussian variables, while different wavelets are derived from different Gaussian multi-order derivative functions [15,16]. The principal of the wavelet coefficient is to set a processing range by a threshold to achieve de-noising [17]. The wavelet coefficients method has the characteristics of bandpass filtering. Thus, the use of the wavelet decomposition and reconstruction method allows feasible de-noising [18,19].

Compared with other de-noising filtering techniques (e.g., mean filtering and median filtering), the wavelet coefficients method is a customized denoise threshold known as an adaptive threshold that can be more accurate to separate noise than other de-noising filtering techniques with a fixed threshold.

Since wavelet coefficients for de-noising are well-studied, many threshold approaches have been proposed. Among those threshold methods, the wavelet Bayes shrink approach is the most effective wavelet coefficient method [15]. Based on Bayesian theory, the Bayes shrink is changed according to different image information, so the Bayes shrink is also called the adaptive threshold [16,20,21]. The most important concept of de-noising is that there is no best de-noising method, but only the most suitable de-noising method because the noise of each image is different [22]. Thus, it is important to test and choose the most accurate de-noising filtering for a CCTV image to enhance image segmentation, which allows estimating the inundated area.

The image affected by any natural environments (e.g., mist, light refraction) may cause image overexposure and fogginess. There are advanced image process approaches in addition to de-noising processing: dark channel prior and dehaze filtering, which can eliminate mist and light refraction situations. The dark channel prior is operated by the darkest pixels in the image being separated from the image, and the remaining pixels are relatively bright image or called foggy image. By the dark channel prior, the difference between a foggy image and a fog-free image can be calculated with light refraction and brightness of an environment. The process of de-hazing is to obtain the fog-free image by changing the information of the fog image and combining it with the dark channel prior. By dehaze filtering calculated by dark channel and light refraction, the fog can be effectively eliminated [23,24]. After de-noising the image, to further identify the inundation or water area in the image, the edge and contour of the object must be determined first. Image segmentation is an effective way to find the edge of the objects on the image.

Image segmentation is one of the hotspots in image processing and computer vision, which is the basis for image analysis and understanding of image feature extraction and recognition. It refers to dividing the image into several areas based on grayscale, color, texture, and shape. The features divided into the same area are similar, and there are significant differences between different areas. There is a common principle in image segmentation algorithms which can be divided to region-based segmentation, edge detection segmentation, and clustering segmentation. Dijk and Hollander [25] describe each algorithm in unified frameworks that introduce separate clusters and data weight functions. Felzenszwalb and Huttenlocher [26] study two different local neighborhoods in constructing the graph, by which the important characteristic of the method is its ability to preserve detail. We deploy three different image-segmentation methods, including k-means clustering segmentation, Otsu region-based segmentation, and Bayesian threshold segmentation.

A neural network may be another solution to the limitations of CCTV data, and uses a large number of different images of flood and water level conditions to establish a database as a reliable statistical method for identifying the status of flash floods [27]. However, a neural network for flood monitoring requires many images with different views in the same location. It is challenging to meet the requirements because, practically, there are not enough images in the same CCTV location with different flooding conditions to build a suitable database.

Thus, this paper presents an effective image-processing procedure that requires only a single image to detect the inundation area in a CCTV image to overcome limitations on current flooding detection. First, we investigate and propose de-noising approaches to improve the quality of the image. Then we utilize different image segmentation methods, including k-means segmentation, Otsu segmentation, and Bayesian segmentation, to detect flooding areas. The obtained segmentation results are compared to determine which matches the flooded area in the CCTV image best.

2. Methodology

To address the challenges to detect inundation in CCTV images using other approaches including a neural network, the paper proposed a de-noising and image segmentation approach to find the

water area in the image by de-noising and image segmentation. The first step is to find the most suitable de-noising method for CCTV images. The second step is to use image segmentation to find the edge further to find the water area in the image. The flowchart of this study is shown in Figure 1. The effectiveness of de-noising is determined by the peak signal-to-noise ratio (PSNR), which is commonly used for image compression and reconstruction after image de-noising. The higher the PSNR, the better the de-noising effect, and the more original image information is retained.

Figure 1. The basic flowchart of this study for inundation detection. (1) Collect closed-circuit television (CCTV) images from government websites or social media; (2) use four different de-noising filtering to find out which has the best de-noising quality, which evaluated by the peak signal-to-noise ratio (PSNR); (3) use image segmentation methods to understand how the computer interprets images and finds an edge which is the vital part for image object detection; (4) compare segmentation results between each method and detect inundation area.

In image processing, we face various random noises: Gaussian noise, impulse noise, and speckle noise. They are distributed in the CCTV image, caused by digitized transmission compression or equipment, which affects the performance of image processing. There are two requirements for the de-noising filtering: keeping intact important information (e.g., the object edges) and making the image clearer with a better visual impact so the image's information can be clearly seen. We will study several de-noising filtering techniques: mean filtering, and median filtering, which belong to image enhancement. The performance of the de-noising method depends on the type of noise. For example, median de-noising filtering is very effective in smoothing impulse noise while it allows keeping the sharp edges of the image. The results of the image segmentation to find inundation objects edge effectively and accurately by de-noising filtering.

To find out which type of de-noising method is the most suitable for flood identification from CCTV images, we need to understand the type of noise. During image acquisition, encoding, transmission, and processing steps, noise always appears in the digital image. Without prior knowledge of filtering techniques, it is difficult to remove noise from digital images. Image noise is a random change in brightness or color information in the captured image. It is degradation in image signal caused by external sources. We can model a noisy image as $A(x, y) = H(x, y) + B(x, y)$ where, $A(x, y)$ is a function of the noisy image, $H(x, y)$ is a function of image noise, and $B(x, y)$ is a function of the original image. Before de-noising, we need to understand which noises are in the image. There are different

types of image noise. They are typically divided into 3 types, which are Gaussian noise, impulse noise, and speckle noise. Gaussian noise is generated by adding a random Gaussian function to the image, while impulse noise is caused by adding random white and black dots to the image, and speckle noise is a granular noise that inherently exists in an image and reduces its quality. An example of adding noise to the image is shown in Figure 2. Due to the wide variety of image noise, it is necessary to test different de-noising methods separately to determine the most suitable de-noising method for CCTV images.

(a) Gaussian noise

(b) Impulse noise

(c) Speckle noise

Figure 2. The example images with three different types of noises. (a) The Gaussian noise; (b) the impulse noise; and (c) the speckle noise.

For the following contents in Section 2, several de-noising methods are presented in the first half, which are mean filtering, median filtering, Gaussian filtering, and wavelet coefficients. The image segmentation methods (e.g., k-means segmentation, Otsu segmentation, and Bayesian segmentation) are proposed in the second half.

2.1. De-Noising Method

The image enhancement is performed by changing the pixel number of images with several convolution approaches (e.g., spatial convolution and frequency convolution), which is a mathematical operation to determine a new pixel value from a linear combination of pixel values and its neighboring pixels. The spatial convolution is simply calculated by arithmetic, such as add, minus, multiply, and divide pixel value. The frequency convolution is calculated by the information of the image after the fast Fourier transform (FFT), which converts the information from the spatial domain to a frequency

domain [27]. The principle of image enhancement is to modify pixels by changing the brightness, contrast, and simply de-noising [28,29].

2.1.1. Median Filtering and Arithmetic Filtering

If a signal changes gently, the output value, which are pixels of images, can be replaced by the statistical median value in a certain size neighborhood of this pixel point, and this neighborhood is called a window in the signal-processing field. The larger the window, the smoother the output, but it may also erase useful signal characteristics [30]. In order to keep the useful signal, the size of the window should be determined according to the signal and noise characteristics. Usually, the size of the window is odd because the odd number of data (e.g., pixel number) has a unique median value. The concept of mean filtering is similar to median filtering; the only difference is that the former uses the arithmetic mean as a filter [31].

2.1.2. Gaussian Filtering

Gaussian filtering is commonly used as a linear filtering algorithm. A two-dimensional Gaussian function distribution is used to make a smooth image. The principle of Gaussian filtering is the weighted average of all pixel values in the entire image through the Gaussian distribution. More precisely, Gaussian filtering is the result of convolution operation on pixels by Gaussian normal distribution [32]. The value of each pixel is obtained by a weighted average of the values of itself and nearby pixels. The two-dimensional Gaussian function is:

$$G(x, y) = \frac{1}{2\pi\sigma^2} e^{-\frac{(x^2+y^2)}{2\sigma^2}} \quad (1)$$

where x and y are the number of pixels on the x and y -axis of the image, respectively and σ is the standard deviation of a Gaussian distribution.

2.1.3. Wavelet Coefficients

The discrete wavelet transform (DWT) can be interpreted as signal decomposition in a set of independent, spatially oriented frequency channels. Signal decomposition means that the signal passes through two complementary filters (low-pass and high-pass filters) and appears in the form of approximate and detailed signals as known as wavelet coefficients [33]. The approximate and detailed signals can be assembled back into the original signal without loss of information. The process is called reconstruction. For decomposition, the image is divided into four sub-bands, as shown in Figure 3a. The image is divided into four different sub-band based on their frequency. The four sub-band come from the separable application of vertical and horizontal directions. Each wavelet coefficient represents a spatial area corresponding to approximately a 2×2 area of the original image. Each coefficient in the sub-bands represent a spatial area corresponding to approximately a $2^1 \times 2^1$ area of the original image. The frequencies ω can be divided into two ranges, the low-frequency range ($0 < |\omega| < \frac{\pi}{2}$) and the high-frequency range ($\frac{\pi}{2} < |\omega| < \pi$). The sub-band labeled L or H depends on their frequency. The four sub-band come from the separable application of vertical and horizontal direction. These four sub-band present image information called details: HH_1 is diagonal detail, LH_1 is vertical detail, HL_1 is horizontal detail and LL_1 is the remaining image details, where number 1 means detail in the first scale decomposition [34]. To obtain the next, more critical scale of wavelet coefficients, the sub-band LL_1 is further decomposed, as shown in Figure 3b. The image is divided into four different sub-band based on their frequency, which low-frequency range in second scale is $0 < |\omega| < \frac{\pi}{2^2}$ while high frequency range in second scale is $\frac{\pi}{2^2} < |\omega| < \frac{\pi}{2^1}$. Each coefficient in the sub-bands of the second scale HH_2 , LH_2 , HL_2 , and LL_2 represents a spatial area corresponding to approximately a $2^2 \times 2^2$ area of the original picture. The decomposition process continues until a certain final scale is reached, while the degree of matching between the reconstructed signal and the original signal is 90%. The DWT shows the wavelet

analysis is a measure of similarity between basis wavelets and the signal function. [35]. The wavelet coefficients method for image de-noising is the process of decomposition and reconstruction of details.

Figure 3. The scale decomposition of discrete wavelet transforms: (a) the first scale decomposition of a discrete wavelet transform representing each wavelet coefficient in a spatial area corresponding to approximately a 2×2 area of the original image and (b) an n-scale wavelet decomposition. The image is divided into four different sub-bands based on their frequency. Each coefficient in the sub-bands of the second scale HH_2 , LH_2 , HL_2 , and LL_2 represent a spatial area corresponding to approximately a $2^2 \times 2^2$ area of the original image.

The wavelet threshold is the reference point to divide the frequency of the image sub-band. The image and noise have different characteristics after wavelet transform. After the noisy signal is decomposed in the wavelet scale, the information of the image is mainly concentrated on the low-resolution sub-bands [36], and the noise signal is mainly distributed on each high-frequency sub-bands. Thus, the choice of wavelet threshold directly affects the performance of wavelet de-noising. The wavelet coefficients of each scale are classified according to different threshold algorithms they used [35]. If the wavelet coefficients are smaller than the threshold, set it to zero; otherwise, it maintains or slightly decreases the magnitude [34]. Because of this characteristic of wavelet coefficients, it is very effective in energy compression, which can better save important image features such as edge changes in the image. Finding an optimal threshold is a tedious process. If using a smaller threshold, it produces a poor performance of de-noising, while using a larger threshold also causes image details to be removed as noise [16].

In this paper, Bayes shrink is used for wavelet coefficients, which has the best performance of de-noising for high-frequency noise [20]. The following is the Bayes shrink algorithm introduction. The Bayes shrink is known to be effective for images with Gaussian noise. The observation model is expressed as follows:

$$Y(i, j) = X(i, j) + V(i, j) \quad (2)$$

where Y is the wavelet transform of the noisy image; X is the wavelet transform of the original image, and V denotes the wavelet transform of the noise components following the Gaussian distribution $N(0, \sigma_v^2)$. Since X and V are mutually independent, the variances σ_y^2 , σ_x^2 and σ_v^2 of Y , X and V is given by:

$$\sigma_y^2 = \sigma_x^2 + \sigma_v^2 \quad (3)$$

It has been shown that the noise variance σ_y^2 can be estimated from the first decomposition level diagonal high-frequency sub-band, HH_1 by the robust and accurate median estimator [37],

$$\sigma_y^2 = \left[\frac{\text{median}(|HH_1|)}{0.6745} \right]^2 \quad (4)$$

The variance of the degraded image can be estimated as:

$$\sigma_y^2 = \frac{1}{M} \sum_{m=1}^M A_m^2 \quad (5)$$

where A_m are the wavelet coefficients of wavelet on every scale; M is the total number of wavelet coefficient. Use of soft threshold which is based on sub-band and level-dependent near-optimal threshold as the equation condition for Bayes shrink thresholding:

$$T_{\text{Bayes}} = \begin{cases} \frac{\sigma_v^2}{\sigma_x} & \text{if } \sigma_v^2 < \sigma_y^2 \\ \max(|A_m|) & \text{otherwise} \end{cases} \quad (6)$$

where

$$\sigma_x = \sqrt{\max(\sigma_y^2 - \sigma_v^2, 0)} \quad (7)$$

The basic framework of the wavelet transform-based image de-noising is shown in Figure 4.

Figure 4. The basic framework of wavelet image de-noising. It shows the process of wavelet de-noising. There are three main steps. (1) Apply wavelet transform to image data and calculate the wavelet coefficients. (2) Find the optimum value for threshold and applying a soft threshold. (3) Calculate the de-noised signal and reconstruct the image.

2.2. Image Segmentation

The goal of this paper is to identify the inundation or water area in CCTV images. In order to achieve this goal, the edge and contour of the object must be determined first using an image segmentation, which is an effective way to find the edge of the objects on the image. The classification of image segmentation is based on grayscale, color, texture, and shape to divide the image into several areas. The features that have been divided into the same area are similar, while there are significant differences between different areas. Moreover, this is the basis for image analysis and understanding of

image feature extraction and object detection. There several image segmentation approaches are studied (e.g., region-based segmentation, clustering segmentation) [26]. The region-based segmentation divides the image into two regions of the target and the background by a single threshold. With different threshold calculation methods, the region-based segmentation presents different results. The clustering method is used to segment the image with the corresponding feature pixel points. According to their features in the cluster, the image is segmented into several different clusters in which each cluster has similar features. A global threshold can effectively segment different targets and backgrounds with different grayscales. However, when the grayscale difference of the image is not obvious, the local threshold or adaptive threshold method should be used.

To be able to understand how the computer interprets images to detect an object edge and find which image segmentation has the best performance to detect the water area, we use three different image-segmentation methods, which are k-means clustering segmentation, Otsu region-based segmentation, and Bayesian threshold segmentation.

2.2.1. K-Means Segmentation

Each pixel in a color image is a point in three-dimensional space; k-means segmentation uses pixels of the image as data points according to the specified number of clusters, replacing each pixel with its corresponding cluster center to reconstruct the image. K-means clustering minimizes the sum of the squared errors of the data in the cluster and the center of the cluster [38]. The purpose is to find a similar cluster in the data so that members in the same subset have similar attributes. Assume there is a set of n -dimensional data:

$$x_i \in R^d, i = 1, 2, \dots, n; \{S_1, S_2, \dots, S_k\}, k \leq n \quad (8)$$

where x_i is a set of i data points as the data to be clustered and d is a number of dimension of data points i ; S_k is the number of clusters from data points x_i . By using the formula of Euclidean distance to calculate the sum of least squares between clusters center and pixel points x_i is the minimum value to define the number of clusters:

$$\arg_{\mu} \min \sum_{c=1}^k \sum_{i \in S_k} \|x_i - \mu_c\|^2 \quad (9)$$

where μ_c is the center of k clusters and $\arg_{\mu} \min$ is the value of the variable μ_c reaches the minimum value in the following formula.

The image segmentation based on k-means uses the pixels as data points, using Equation (9) to calculate the number of clusters then replace each pixel with its corresponding cluster center to reconstruct the image. The different clusters present different colors and other characteristics, while the pixel points in the same cluster have similar characteristics.

2.2.2. Otsu Segmentation

The most commonly used threshold segmentation algorithm is the most substantial interclass variance method (Otsu), which selects threshold by maximizing the variance between classes. According to the grayscale characteristics of the image, Otsu assumes that the image is composed of two parts: the foreground and background. By calculating the variance of the foreground and background of the segmentation result under different thresholds, the threshold with the largest variance is the Otsu threshold [39]. The larger the between-class variance between the background and foreground, the better the effect to distinguish these two parts. The main calculated between-class variance equation is:

$$g = \omega_1 \times \omega_2 \times (\mu_1 - \mu_2)^2 \quad (10)$$

where ω_1 and ω_2 are the ratio of background pixels and foreground pixels in the image, respectively; μ_1 and μ_2 are the average grayscale value of background and foreground.

2.2.3. Bayesian Segmentation

Similar to Otsu segmentation, the image is divided into the foreground and background by Bayesian segmentation. The Bayesian theorem calculates the posterior probability with the smallest Bayesian risk as Bayesian shrink, which is defined as the probability distribution of the expected values. Image segmentation is a conditional assumption question in which the decisions are usually based on probability to select value [40,41]. If $P\left(\frac{H_0}{z}\right) > P\left(\frac{H_1}{z}\right)$, H_0 is selected; if $P\left(\frac{H_0}{z}\right) < P\left(\frac{H_1}{z}\right)$, then H_1 be chosen where P is a probability, H_0 and H_1 are decisions, and z is the independently distributed Gaussian variables. For an image $I(m, n)$, the segmentation by using Bayesian theorem can be presented as:

$$I(m, n) < \lambda : I(m, n) \in H_0 \quad (11)$$

$$I(m, n) \geq \lambda : I(m, n) \in H_1 \quad (12)$$

where λ is the Bayesian threshold of image and satisfies the following formula:

$$\frac{P\left(\frac{\lambda}{H_0}\right)}{P(H_0)} = \frac{P\left(\frac{\lambda}{H_1}\right)}{P(H_1)} \quad (13)$$

Assume that $P(z)$ is the probability density function with the expected Bayesian threshold values of image $I(m, n)$, which is defined as the probability distribution based on Equations (11) to (13):

$$P(z) = P\left(\frac{z}{H_0}\right)P(H_0) + P\left(\frac{z}{H_1}\right)P(H_1) \quad (14)$$

For the image is divided into a background part ω_1 and target part ω_2 by a threshold, which their probability is $P(\omega_1)$ and $P(\omega_2)$, separately. The posterior probability can be presented by the Bayesian theorem:

$$P(\omega_i|x) = \frac{P(x|w_i)P(w_i)}{\sum_{j=1}^2 P(x|w_j)P(w_j)} \quad (15)$$

The threshold with minimum Bayesian risk has the maximum expectation of posterior probability represented by Equation (15), and can be written as:

$$T = \{x | P(w_1|x_0) = P(w_2|x_0)\} \quad (16)$$

Based on threshold T presented on Equation (16), which has the maximum expectation of threshold based on the Bayesian theorem, the image segmentation results can get the minimum error, which means the distortion of the image after segmentation is minor.

2.3. Data Collection and Assessment Approach

2.3.1. Data Collection

In this study, 14 CCTV images which are collected from public websites on transportation information government website such as TranSTAR or downloaded from public social media have been tested to determine which de-noising method is the best for CCTV images. Six of the CCTV images are shown in Figure 5.

Figure 5. Example images used for inundation detection experiments which are No. 1 to No. 6 in Table 1 collected from CCTV and social media used in this paper. (a,b) are flooded roads during Hurricane Harvey collected from CCTV; others are collected from social media.

2.3.2. Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR)

The quality and information of the image after compression or reconstruction are usually different from the original image. Image de-noising is also a process of compression and reconstruction, which can eliminate most image noise while maintaining image information. However, the differences are difficult to identify the performance of de-noising by the human eye. The criteria for the quality of de-noising filtering are determined by mean squared error (MSE) and peak signal-to-noise ratio (PSNR). MSE in mathematical statistics refers to the expected value of the squared difference between the estimated values and the true value, which can evaluate the degree of change of data. The smaller the value of MSE, the better the accuracy of the experimental data. PSNR is a measurement method to quantify the impact of image processing, which is commonly used for image compression and reconstruction after image de-noising. The higher the PSNR, the better the de-noising effect, and the more original image information is retained.

$$\text{MSE} = \frac{1}{mn} \sum_{i=0}^{m-1} \sum_{j=0}^{n-1} [I(i, j) - K(i, j)]^2 \quad (17)$$

$$\text{PSNR} = 10 \log_{10} \frac{\text{MAX}_I^2}{\text{MSE}} = 20 \log_{10} \frac{\text{MAX}_I}{\sqrt{\text{MSE}}} \quad (18)$$

where m, n is a resolution of image; $I(i, j)$ is an image after de-noise; $K(i, j)$ is a noisy image; MAX_I is the maximum of resolution (i.e., 8-bits image is $2^8 = 256$ resolution). According to Equation (9), the better performance of the de-noising method, the higher PSNR since MAX_I is a fixed value of resolution, and MSE is the error between de-noised image and noisy image. In theory, the de-noising method can only accurately remove image noise and retain the details of the image. The reconstructed image after de-noising must be consistent with the original image, except that it contains noise.

3. Results and Discussion

3.1. The Efficiency of De-Noising Methods

The results of CCTV images via different de-noising methods are shown in Figure 6. According to this comparison Figure 6, the results demonstrate (b) the Bayes shrink is the best of these de-noising methods for CCTV images. It not only removes the noise but also retains important information, including brightness, color, and resolution of the original image. A reasonable explanation for the blurring of images after de-noising filtering like (c), (d) and (e) is the threshold applied in these methods is fixed, which means the threshold does not change with the important information of an image; therefore, the most important information is removed when the noise causes distortion of filtered images.

(a) Original CCTV image (Case No.1)

(b) After Bayes shrink (Case No.1)

(c) After Arithmetic filtering(Case No.1)

(d) After Median filtering (Case No.1)

(e) After Gaussian filtering (Case No.1)

Figure 6. The results of a CCTV image (Case No.1 in Figure 6) with different de-noising methods: (a) original CCTV image; (b) an image by using Bayes shrink; (c) an image by using Arithmetic filtering; (d) an image by using median filtering and (e) an image by using Gaussian filtering. (b) Bayes shrink is the best of these de-noising methods for CCTV images. It not only removes the noise but also retains important information, including brightness, color, and resolution of the original image.

In order to further verify the performance of de-noising methods, the PSNR of each de-noising method is shown in Table 1. The PSNR chart is shown in Figure 7, where the x-axis is the number of cases, and the y-axis is the percentage of PSNR. Based on the results, Bayes shrink has the best de-noising efficiency, in which mostly PSNR is over 80 dB. In some cases, PSNR can reach over 85 dB. The results imply the similarity between the de-noised image and noisy image is 85%. Most image details can be preserved while the noise is accurately eliminated. For other filtering methods, which are median filtering, arithmetic filtering, and Gaussian filtering, their PSNR show around 20 dB to 30 dB, which means they have poor performance for de-noising. The results are so different between Bayes shrink and other methods because the threshold of the latter three methods is fixed instead of calculated by the image information such as Bayes shrink.

Table 1. The peak signal-to-noise ratio (PSNR) of each de-noising filtering.

NO.	Bayes Shrink Wavelet Coefficients (dB)	Median Filtering (dB)	Arithmetic Filtering (dB)	Gaussian Filtering (dB)
1	83.60	36.05	22.64	23.13
2	80.94	32.87	20.83	20.70
3	83.36	34.70	21.58	20.98
4	79.99	31.74	21.15	20.30
5	79.92	32.39	22.22	20.87
6	79.03	32.06	20.68	19.16
7	87.15	39.31	26.68	27.34
8	81.07	32.80	23.93	24.38
9	80.25	28.73	21.36	20.92
10	81.30	29.47	21.04	20.99
11	86.71	33.92	20.65	21.88
12	85.59	33.45	20.28	20.42
13	89.15	34.25	22.54	22.76
14	89.68	33.52	20.92	19.69

Figure 7. The chart of PSNR for each method with different 14 CCTV images. High PSNR indicates that the reconstructed image after de-noising retains the similar information of the original image, while most of the noise has been removed. Bayes shrink has the best de-noising performance since its PSNR average over 85 dB, which implies most image details can be preserved while the noise is accurately eliminated, as shown in Figure 6b. On the other hand, the PSNR of other methods is only around 20 dB to 40 dB, which means that most details of the original image have been deleted as noise. The de-noised image is distorted, as shown in Figure 6c–e.

3.2. Detection of Inundation by Image Segmentation

Based on the results of de-noising shown in Table 1, the Bayes shrink has the best de-noising performance for CCTV images. The next step is to use image segmentation to determine whether these de-noised images are clear and can be used for computer analysis to detect object edge perfectly. This is one of the important keys to affect the subsequent image detection work and to identify the edge and contour of inundated water areas in the CCTV image. First, we use k-means segmentation, Otsu segmentation, and Bayesian segmentation methods to analyze images. A computer recognizes pixel information and changes in pixels, which is very difficult unless computers can clearly distinguish areas with similar attributes before the object detection procedure. Image segmentation helps us understand how computers parse images. The comparison of before and after image de-noising filtering is performed using three different image-segmentation methods and is shown in Figure 8.

Figure 8. The comparison of before and after image de-noising filtering using different image segmentation methods. The result (b,c) presents using the k-means; the result (d,e) presents using the Otsu; and the result (f,g) presents using the Bayesian of the noisy image and the de-noised image, separately. The results (h,i) indicates the water area surrounded by the edge of Figure 8e,g, separately.

Firstly, as shown in Figure 8b,c, the image using k-means segmentation is divided into different colors according to its attributes. It is clearly found that k-means cannot treat the water area as the same object with or without de-noising filtering. Thus, it indicates one part of water belongs to yellow, and the other part belongs to blue. The result of the Otsu image segmentation is shown in Figure 8d,e. Although it can detect mostly the water area, Otsu treats road parts as the same as the water area so that all the images are mostly in the same color, which is white. The results of Bayesian segmentation are shown in Figure 8f,g. There are many black spots in the image segmentation before de-noising filtering shown in Figure 8f due to the noise. The result for Bayesian segmentation with the de-noised image shown in Figure 8g shows that this method can segment important information in the image (i.e., water and roads), which has the best performance among the three methods.

Secondly, based on the comparison results of Figure 8f,g, the importance of de-noising for image recognition can be determined. The Bayesian segmentation result with no de-noising image shown in Figure 8f, which has a lot of black dots, means there is pepper noise, and Figure 8g shows the Bayesian segmentation result with de-noised image, which shows a perfect contour and edge of water area.

By comparing two cases shown in Figure 8f,g, without noise filtering, even in advanced approaches such as the Bayesian method, the result likely presents a poor image.

In addition, to further compare the performance of each segmentation method, the inundation area detection is performed based on the approach of Otsu segmentation and Bayesian segmentation shown in Figure 8h,i in which gray indicates water area and blue indicates other parts out of the area, while k-means already segment the image in Figure 8c into colors—comparing the inundation area detection results with different segmentation methods, which are k-means, Otsu, and Bayesian segmentation shown in Figure 8c,h,i, respectively. The inundation area detection based on Bayesian segmentation is the closest to the original image (a), which shows the perfect edge of the water area, while the other two results in Figure 8c,h cannot display the correct inundation area corresponding to the original image Figure 8a. Summarizing the results shown in Figure 8, de-noising is important for image processing, which may affect the following processing results. Bayesian segmentation has the best performance to find inundation edges and using these edges to find the inundation area corresponding to the CCTV image.

Figure 9. Inundation detection based on the results of Bayesian segmentation. To avoid detection errors, the region of interest (ROI) should be chosen carefully; cutting off the sky part of the CCTV image (Case No.3) allows for providing the correct ROI for accurate inundation detection.

Consequently, there are two other images that use the edge based on Bayesian segmentation results to calculate inundation to achieve object detection. The region of interest (ROI) for this study is the water area. The inundation detection results are shown in Figure 9, which presents the water area only. Case No.2 is shown in Figure 9a,d. The gray area in Figure 9b indicates inundation ROI, while blue implies background outside the inundation area. Case No.3 is shown in Figure 9c,d. Figure 9d presents an incorrect ROI of inundation; there are two different inundation areas at the top and bottom of (d). The gray at the top represents the sky, while the blue represents the background of the building, and the gray at the bottom represents the inundation. Compared with Case No. 2, the sky part in Case No. 3 is detected as the inundation area. In this case, choosing the appropriate ROI is important, which controls the location of CCTV images to avoid the sky part in the image to make inundation detection more accurate, as shown in Figure 9e,f.

4. Conclusions

In this study, we comparatively studied image-processing methods, such as de-noising methods and image segmentation, to automatically detect the flooded areas from the low-resolution images. The inundation detection results indicate that a series of methods are important and necessary to achieve detection. According to this research, the most effective de-noising method for a CCTV image is the Bayes shrink adaptive wavelet threshold. By using Bayes shrink and segmentation as a pre-processing procedure, future classification and object detection in CCTV images are expected to be more successful. The key findings are summarized below.

- First, by comparing the most recently used de-noising methods, Bayes shrink with adaptive wavelet coefficients shows the best de-noising performance of all indicating this by the minimum MSE and maximum PSNR for CCTV images. The PSNR of CCTV images, by using the Bayes shrink approach, mostly exceeds 85 dB, which means that at least 85 % of the image details are retained after de-noising.
- Second, for image-segmentation techniques, Bayesian segmentation has the best performance to find the inundation edge. The results present the most important part of following object detection. Bayesian segmentation allows identifying the inundation edges correctly in a grayscale image.
- Last, use of the edge based on Bayesian segmentation enabled us to calculate inundation to achieve object detection. We notice the importance of the ROI, which controls the location of CCTV images to avoid the sky part, which has similar features to the inundation part. In this study, the inundation in CCTV can be identified accurately, which is important for following work like water-level detection by using the coordinate of the image.

The image processing presented in this paper is to estimate the inundation from images to assess flooding risks in the vicinity of the local flooding locations. Such information will help traffic engineers to take preventive or proactive actions to improve the safety of the drivers and to protect and preserve the transportation infrastructure.

For further research by using the concept of image processing presented in this paper, which defines the edge of the inundation area, the depth of water can be calculated by the coordinate relationship between image and the real world. It is possible to monitor the inundation status and calculate the water level in real-time by using a traffic-monitoring camera in the future. This research demonstrates the other economical option for people to detect flooding conditions such as the location and water level of the inundation area to provide people with more and faster information.

Author Contributions: Y.C.W. and S.H. designed and investigated the study; Y.C.W. carried out image processing, analyzed the collected, and write a manuscript. S.J.N. acquired the funding, reviewed, and improved the manuscript draft. S.H. acquired the funding, supervised the study, reviewed and improved the manuscript draft. All authors have read and agreed to the published version of the manuscript.

Funding: This research was supported by the U.S. Department of Transportation, Tran-SET under grant number 19SAUTA04 and 19SAUTA03 and by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) under grant number NRF-2020R1C1C1005099.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Blessing, R.; Sebastian, A.; Brody, S.D. Flood risk delineation in the united states: How much loss are we capturing? *Nat. Hazards Rev.* **2017**, *18*, 04017002. [[CrossRef](#)]
2. Mousa, M.; Zhang, X.; Claudel, C. Flash flood detection in urban cities using ultrasonic and infrared sensors. *IEEE Sens. J.* **2016**, *16*, 7204–7216. [[CrossRef](#)]
3. Alfieri, L.; Velasco, D.; Thielen, J. Flash flood detection through a multi-stage probabilistic warning system for heavy precipitation events. *Adv. Geosci.* **2011**, *29*, 69–75. [[CrossRef](#)]
4. Noh, S.J.; Lee, J.H.; Lee, S.; Seo, D.J. Retrospective dynamic inundation mapping of Hurricane Harvey flooding in the Houston metropolitan area using high-resolution modeling and high-performance computing. *Water* **2019**, *11*, 597. [[CrossRef](#)]
5. Wang, Y.; Chen, A.S.; Fu, G.; Djordjević, S.; Zhang, C.; Savić, D.A. An integrated framework for high-resolution urban flood modeling considering multiple information sources and urban features. *Environ. Model. Softw.* **2018**, *107*, 85–95. [[CrossRef](#)]
6. Lo, S.W.; Wu, J.H.; Lin, F.P.; Hsu, C.H. Visual sensing for urban flood monitoring. *Sensors* **2015**, *15*, 20006–20029. [[CrossRef](#)]
7. Van Ackere, S.; Verbeurgt, J.; De Sloover, L.; Gautama, S.; De Wulf, A.; De Maeyer, P. A review of the internet of floods: Near real-time detection of a flood event and its impact. *Water* **2019**, *11*, 2275. [[CrossRef](#)]
8. Xie, J.; Xu, L.; Chen, E. Image de-noising and inpainting with deep neural networks. *Adv. Neural Inf. Process Syst.* **2012**, *1*, 1–9.
9. Buades, A.; Coll, B.; Morel, J.M. A review of image de-noising algorithms, with a new one. *Multiscale Model. Simul.* **2005**, *4*, 490–530. [[CrossRef](#)]
10. Zhu, Y.; Huang, C. An improved median filtering algorithm for image noise. *Phys. Procedia* **2012**, *25*, 609–616. [[CrossRef](#)]
11. Kazubek, M. Wavelet domain image de-noising by thresholding and Wiener filtering. *IEEE Signal. Process Lett.* **2003**, *10*, 324–326. [[CrossRef](#)]
12. Donoho, D.L.; Johnstone, I.M.; Kerkycharian, G.; Picard, D. Wavelet shrinkage: Asymptopia. *J. R. Stat. Soc.* **1995**, *57*, 301–337. [[CrossRef](#)]
13. Donoho, D.L.; Johnstone, L.M. Adapting to unknown smoothness via wavelet shrinkage. *Fundam. Pap. Wavelet Theory* **1995**, *90*, 1200–1224. [[CrossRef](#)]
14. Donoho, D.L. De-Noising by soft thresholding. *IEEE Trans. Inf. Theory* **1995**, *41*, 631–637. [[CrossRef](#)]
15. Portilla, J.; Strela, V.; Wainwright, M.J.; Simoncelli, E.P. Adaptive Wiener de-noising using a Gaussian scale mixture model in the wavelet domain. *IEEE Int. Conf. Image Process* **2001**, *2*, 37–40.
16. Chang, S.G.; Member, S.; Yu, B.; Member, S.; Vetterli, M. Adaptive wavelet thresholding for image de-noising and compression. *IEEE Trans. Image Process* **2000**, *9*, 1532–1546. [[CrossRef](#)]
17. Nowak, R.D. Wavelet-based Rician noise removal for magnetic resonance imaging. *IEEE Trans. Image Process* **1999**, *8*, 1408–1419. [[CrossRef](#)]
18. Mallat, S.G. A theory for multiresolution signal decomposition: The wavelet representation. *Fundam. Pap. Wavelet Theory* **2009**, *11*, 674–693.
19. Mallat, S.; Hwang, W.L. Singularity detection and processing with wavelets. *IEEE Trans. Inform. Theory* **1992**, *38*, 617–643. [[CrossRef](#)]
20. Elyasi, I.; Zarmehi, S. Elimination noise by adaptive wavelet threshold. *World Acad. Sci. Eng. Technol.* **2009**, *56*, 462–466.
21. Tharani, K.; Mani, C.; Arora, I. A comparative study of image de-noising methods using wavelet thresholding techniques. *Int. J. Eng. Res. Appl.* **2016**, *6*, 73–77.
22. Talebi, H.; Milanfar, P. Global image denoising. *IEEE Trans. Image Process* **2014**, *23*, 755–768. [[CrossRef](#)] [[PubMed](#)]
23. Fattal, R. Single image dehazing. *ACM Trans. Graph.* **2008**, *27*, 1–9. [[CrossRef](#)]
24. He, K.; Sun, J.; Tang, X. Single image haze removal using dark channel prior. *IEEE Trans. Pattern Anal. Mach. Intell.* **2011**, *33*, 2341–2353.

25. Dijk, J.; den Hollander, R.J.M. Image enhancement for noisy color imagery. In Proceedings of the International Society for Optical Engineering (SPIE), Cardiff, Wales, UK, 2 October 2008; Volume 71131A. [[CrossRef](#)]
26. Felzenszwalb, P.F.; Huttenlocher, D.P. Efficient graph-based image segmentation. *Int. J. Comput. Vis.* **2004**, *59*, 167–181. [[CrossRef](#)]
27. Jain, V.; Sciences, C.; Seung, H.S. Natural image denoising with convolutional networks. In Proceedings of the 21st International Conference on Neural Information Processing Systems, Vancouver, BC, Canada, 8–10 December 2008; pp. 769–776.
28. Kim, J.Y.; Kim, L.S.; Hwang, S.H. An advanced contrast enhancement using partially overlapped sub-block histogram equalization. *IEEE Trans. Circuits Syst. Video Technol.* **2001**, *11*, 475–484.
29. Abdullah-Al-Wadud, M.; Kabir, M.H.; Dewan, M.A.A.; Chae, O. A dynamic histogram equalization for image contrast enhancement. *IEEE Trans. Consum. Electron.* **2007**, *53*, 593–600. [[CrossRef](#)]
30. Huang, T.; Yang, G.; Tang, G. A fast two-dimensional median filtering algorithm. *IEEE Trans. Image Process.* **1979**, *27*, 13–18.
31. Hwang, H.; Haddad, R.A. Adaptive median filters. *IEEE Trans. Image Process.* **1995**, *4*, 7–10. [[CrossRef](#)] [[PubMed](#)]
32. Deng, G. An adaptive gaussian filter for noise reduction and edge detection. In Proceedings of the IEEE Conference Record Nuclear Science Symposium and Medical Imaging Conference, San Francisco, CA, USA, 31 October–6 November 1993.
33. Sardy, S.; Tseng, P.; Bruce, A. Robust wavelet denoising. *IEEE Trans. Image Process.* **2001**, *49*, 1146–1152. [[CrossRef](#)]
34. Shapiro, J.M. Embedded image coding using zerotrees of wavelet coefficients. *IEEE Trans. Signal Process.* **2009**, *41*, 861–878.
35. Mallat, S. *A Wavelet Tour of Signal Processing*; Academic Press: Burlington, MA, USA, 1999.
36. Chang, S.G.; Member, S.; Yu, B.; Member, S.; Vetterli, M. Spatially adaptive wavelet thresholding with context modeling for image De-noising. *IEEE Trans. Image Process.* **2000**, *9*, 1522–1531. [[CrossRef](#)] [[PubMed](#)]
37. Kimlyk, M.; Umnyashkin, S. Image De-noising using discrete wavelet transform and edge information. In Proceedings of the 2018 IEEE Conference of Russian Young Researchers in Electrical and Electronic Engineering (ElConRus), Moscow and St. Petersburg, Russia, 29 January–1 February 2018; pp. 1823–1825.
38. Ng, H.P.; Ong, S.H.; Foong, K.W.C.; Goh, P.S.; Nowinski, W.L. Medical image segmentation using k-means clustering and improved watershed algorithm. In Proceedings of the IEEE Southwest Symposium on Image Analysis and Interpretation, Denver, CO, USA, 26–28 March 2006; pp. 61–65.
39. Kurita, T.; Otsu, N.; Abdelmalek, N. Maximum likelihood thresholding based on population mixture models. *Pattern Recognit.* **1992**, *25*, 1231–1240. [[CrossRef](#)]
40. Pereyra, M.; McLaughlin, S. Fast unsupervised bayesian image segmentation with adaptive spatial regularization. *IEEE Trans. Image Process.* **2017**, *26*, 2577–2587. [[CrossRef](#)] [[PubMed](#)]
41. Bouman, C.A.; Lafayette, W.; Box, P.O. A multiscale random field model for bayesian image segmentation. *IEEE Trans. Image Process.* **1996**, *3*, 162–177. [[CrossRef](#)] [[PubMed](#)]